

**Améliorations et développements d'une méthode de
prévision probabiliste des pluies par analogie.
Application à la prévision hydrologique sur les grands
bassins fluviaux de la Saône et de la Seine.**

Aurélien Ben Daoud

► **To cite this version:**

Aurélien Ben Daoud. Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie. Application à la prévision hydrologique sur les grands bassins fluviaux de la Saône et de la Seine.. Hydrologie. Université de Grenoble, 2010. Français. tel-00483080

HAL Id: tel-00483080

<https://tel.archives-ouvertes.fr/tel-00483080>

Submitted on 12 May 2010

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

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

Université de Grenoble

THESE

pour obtenir le grade de

Docteur de l'Université de Grenoble

Spécialité : Océan, Atmosphère, Hydrologie

préparée dans l'**Unité de Recherche Hydrologie-Hydraulique, Cemagref (Lyon)**

dans le cadre de l'**Ecole Doctorale Terre - Univers - Environnement**

par

Aurélien BEN DAOUD

Le 10 février 2010

**Améliorations et développements d'une méthode de prévision probabiliste
des pluies par analogie.**

*Application à la prévision hydrologique sur les grands bassins fluviaux de la
Saône et de la Seine.*

Directeur de thèse : Charles OBLED

Co-directeurs de thèse : Michel LANG et Eric SAUQUET

JURY

M. Eric BLAYO	LJK-UJF	Président
Mme Carmen LLASAT	University of Barcelona	Rapporteur
M. Eric MARTIN	Météo-France	Rapporteur
Mme Isabelle LELEU	SCHAPI	Examineur
Charles OBLED	LTHE/INPG	Examineur
Eric SAUQUET	Cemagref	Examineur
Guillaume BONTRON	CNR	Examineur

Résumé

Cette thèse porte sur l'amélioration d'une méthode de prévision probabiliste des précipitations adaptée par analogie, appliquée aux grands bassins versants de la Saône et de la Seine. En considérant un jour cible dont on souhaite prévoir les précipitations, cette méthode permet de rechercher dans une archive les situations météorologiques passées les plus similaires. Les cumuls pluviométriques des journées analogues sont ensuite exploités pour estimer de manière probabiliste la quantité de précipitations attendue pour le jour cible. Une étape préliminaire consiste à construire et à choisir les archives nécessaires à l'application de ce type de méthode. A partir de l'algorithme de sélection d'analogues développé par Bontron (2004) sur des bassins du Sud-Est de la France, la méthode de prévision est ensuite améliorée en introduisant deux nouvelles variables : la température, qui permet de prendre en compte les effets saisonniers ; et la vitesse verticale, qui vise à mieux caractériser les mouvements atmosphériques verticaux. Finalement, la méthode est appliquée en contexte de prévision parfaite (situations cibles extraites de réanalyses) sur la plus longue période possible, puis en contexte opérationnel (situations cibles générées par un modèle de prévision) sur une durée de trois ans. Les scénarios de prévision de précipitations sont également exploités par le modèle pluie-débit GR3P pour la prévision des débits sur les bassins de la Saône et de la Seine. De manière générale, les résultats montrent qu'avec ce type d'approche, il est possible d'obtenir une prévision de débit utile, sur les grands bassins tests, pour une échéance de l'ordre de la semaine.

Mots clés : prévision probabiliste quantitative des précipitations, adaptation par analogie, réanalyses ERA-40, prévision hydrologique, bassin de la Saône, bassin de la Seine.

Abstract

This thesis aims at improving a probabilistic quantitative precipitation forecasting method based on analogs, formerly developed on small basins located in South-Eastern France (Bontron, 2004). The study area covers two French large basins of the Saône and the Seine. For a given target situation, this method consists in searching for the most similar situations observed in a historical meteorological archive. Precipitation amounts observed during analogous situations are then collected to derive the empirical predictive distribution function, i.e. the probabilistic estimation of the precipitation amount expected for the target day. A preliminary step is to build the databases necessary to apply such a method. The former forecasting method (Bontron, 2004) is improved by introducing two variables: temperature allows to take seasonal effects into account and vertical velocity enables a better characterization of the vertical atmospheric motion. The new algorithm is then applied in a perfect prognosis context (target situations come from a meteorological reanalysis) and then in an operational forecasting context (target situations come from weather forecasts) for a three years period. Precipitation scenarios are finally introduced in the rainfall-runoff model GR3P to provide discharge forecasts. Results for the two basins under study show that this approach yields useful forecasts up to one week ahead.

Key words : probabilistic quantitative precipitation forecasting, statistical adaptation, analogs, ERA-40 reanalyses, hydrological forecasting, Saône river basin, Seine river basin.

Remerciements

S'il y a bien un domaine dans lequel l'expression « *On n'est jamais mieux servi que par soi-même* » n'est pas vérifiée, c'est dans la construction d'un projet doctoral : un thésard ne saurait se débrouiller seul... C'est donc avec un grand plaisir que j'ai pu travailler avec de nombreuses personnes.

Avant toute chose, je tiens à remercier les membres du Jury extérieurs qui ont eu l'amabilité d'évaluer mes travaux de thèse : M. Eric Martin et Mme Carmen Llasat, rapporteurs, pour le courage et la patience de lire ce long manuscrit ; M. Eric Blayo, examinateur et Président du Jury ; et Mme Leleu, examinatrice et excusée de sa regrettable absence le jour de la soutenance, mais qui a eu la gentillesse de lire le manuscrit et de me poser indirectement quelques questions.

Je remercie également Yves Confesson, Directeur du groupement Cemagref de Lyon, et Michel Lang, Chef de l'Unité de Recherche Hydrologie-Hydraulique, pour m'avoir accueilli dans leurs locaux. Michel a également été mon directeur de thèse au Cemagref : merci pour ses conseils réguliers, son suivi pointu tout au long de ces travaux de thèse, et son humour, même si le ton peut être moqueur parfois (par exemple lorsqu'il m'imite quand je ne sais pas répondre à une question...).

Comment aurais-je pu mener à bien et à terme ma thèse sans le suivi quotidien et la générosité d'Eric Sauquet ? J'ai eu la chance et l'honneur d'avoir été encadré comme il se doit (et même parfois plus !) : je tiens donc à lui exprimer toute ma gratitude. Sans Eric, le travail n'aurait sans aucun doute pas été aussi abouti. Sa compétence, sa rigueur et son sérieux vont vraiment me manquer (et me manquent déjà !). En parallèle, je salue son ouverture d'esprit et son humour développé, quoique parfois limite ! Encore merci à lui pour m'avoir permis de me relaxer à certains moments difficiles, et aussi pour m'avoir laissé gagner de nombreux matchs de badminton...

M. le professeur Charles Obled a également dirigé ma thèse. Malgré la distance, qui nous a séparés et qui ne nous a pas permis de nous réunir aussi souvent que nous l'aurions voulu, nous avons pu échanger sur des sujets aussi physiques qu'intéressants. Je regrette de ne pas être passé par l'ENSHMG de Grenoble, rien que pour avoir manqué ses cours (on m'en a dit tant de bien !). Le domaine de la météorologie semble nous passionner l'un comme l'autre,

que cela puisse continuer !

Le travail mené ici n'aurait pu être rigoureux sans Guillaume Bontron, qui a encadré et suivi ces travaux avec une grande attention. Je tiens à le remercier pour ses conseils très bénéfiques et ses commentaires particulièrement intéressants, qui m'ont permis d'avancer, tant dans le raisonnement scientifique même que dans la façon de présenter ce raisonnement. J'en profite aussi pour le remercier du sacrifice de quelques soirées de décembre 2009, durant lesquelles il a su relire et corriger avec une attention particulière les chapitres de mon manuscrit. Merci également à Muriel Haond, membre du comité de pilotage, pour son suivi tout au long de la thèse, et pour les corrections qu'elle a eu la gentillesse d'apporter à mon mémoire.

Je tiens à remercier sincèrement Hélène et Anne, les deux fées du secrétariat et de la gestion de l'unité HH. Elles ont eu la gentillesse et la patience de m'accorder régulièrement de leur temps pour toutes les questions administratives. Je tiens encore à m'excuser auprès d'Hélène pour le parapluie que je lui avais emprunté et que j'avais malheureusement oublié... dans un métro lyonnais ! Pour l'anecdote, je dois aussi remercier Michel pour avoir récupéré une réplique du dit parapluie ! Bien sûr, comment pourrais-je oublier notre chère - et tellement efficace - gestionnaire des ressources humaines du Cemagref de Lyon, Adeline, qui a su répondre à mes multiples interrogations par mail (au fait Adeline, j'ai encore une question concernant le CET... !?). Qu'elle en soit ici remerciée !

La technologie actuelle ne nous permet plus de travailler comme autrefois, sans machine ni ordinateur... Et bien sûr, qui dit « informatique » dit « problèmes » ou encore « bugs »... Je tiens donc tout d'abord à remercier Olivier pour le temps et la patience qu'il m'a accordé pour l'installation de la fameuse bibliothèque Netcdf (!), mais aussi Antoine et Yves pour leur aide et leur réactivité dès lors qu'un problème informatique intervenait. Merci à eux !

La présente thèse a été rédigée sous Openoffice, et si la qualité de la présentation du document est à la hauteur de mes exigences, c'est sans aucun doute grâce à l'aide précieuse et aux conseils de Jean-Baptiste. Merci à lui pour le temps qu'il a eu la gentillesse de m'accorder. La recherche bibliographique n'a pu être aboutie sans l'aide d'Aline et d'Anne-Laure, nos documentalistes vertueuses. Je tiens donc à les remercier.

Je souhaite remercier les Services de Prévision des Crues Rhône amont-Saône (Pierre-Yves Valantin et Pierre-Marie Béchon) et Seine amont-Marne amont (Yan Lacaze, Carine Chaléon, Olivier Piotte, Jean-Pedro Silva), avec qui j'ai pu collaborer à plusieurs reprises. Merci à eux pour le temps consacré à la recherche et à la transmission de données, et à la réflexion que nous avons partagée sur l'utilisation des prévisions probabilistes.

Quand je pense « prévisions probabilistes », je pense « Maria-Helena » ! Je tiens à remercier Maria-Helena pour avoir adapté le modèle hydrologique GR3P du Cemagref à la prévision probabiliste. Son aide m'a été précieuse lors de l'utilisation du modèle. Merci également à elle pour la relecture attentive d'une partie du mémoire.

Jamais je n'aurais eu d'aussi nombreux résultats en prévision hydrologique sans l'aide de Julien, stagiaire que j'ai eu la chance d'encadrer durant six mois. Merci à lui pour l'intérêt

porté à la météo et à l'hydrologie, et pour son esprit d'initiative.

Enfin, je remercie les personnes qui ont contribué à ces travaux : Marion Pithon de Météo-France, qui a eu la grande patience et l'amabilité de répondre à mes (trop !) nombreuses interrogations sur la base MARS du centre européen ; Etienne Leblois pour son aide lors de l'utilisation des réanalyses ERA-40 et pour l'extraction des radiosondages, à laquelle avait largement participé Angélique Godart, que je remercie également ; Benjamin et Jean-Philippe pour leur aide et leurs conseils en Fortran et Matlab.

La solitude n'étant pas ma tasse de thé, j'ai eu la chance d'être accompagné d'autres brillants thésards et post-doctorants dans le bureau du fond : Mathieu, Aurélie, Benjamin, Alex, Yann, Audrey ma p'tite biquette, et la petite dernière, Claire, passionnée de clémentines (ou de mandarines, je ne sais plus !!).

Merci également à Christine pour les gourmandises régulières qu'elle donnaient à ses bureaux voisins, à Jérôme et Guillaume pour leur humour taquin sur la météo (oui oui, j'appréciais à la longue !).

Une thèse sans sport, c'est comme un orage sans tonnerre... ça ne peut exister ! Le badminton fut sans aucun doute l'activité sportive que j'ai le plus pratiquée : je tiens donc à remercier tous mes partenaires et adversaires d'Escapade, en espérant que je n'oublie personne : Eric, Elodie, Anne-Laure, Magali, Christophe, Bout'bout', Stéphane, Alex, Benoît, Sandhya, Yann, Audrey, Jean-Philippe, Léa, Julien, Lise, Elise, Charly, Olivier, Marie, et Damien l'excellent joueur qui a su me faire courir au filet jusqu'à me léser le ménisque ! Merci aussi à Raphael pour l'initiation au squash !

Comme tout thésard du Cemagref de Lyon, je me suis diverti dans les bars lyonnais (glup's !), mais aussi en randonnant dans les Alpes... alors merci à Mathieu, Sandra, Clotaire, Jean-Marie, Léa, Stéphane, Eric, Julien, Yann, Audrey and Co pour tous les bons souvenirs !

En revanche, outre les activités sportives, j'ai dû évacuer le stress par l'intermédiaire de cette maudite cigarette dont je me suis à présent débarrassé ! Les pauses cigarettes furent souvent l'occasion d'échanger des pensées philosophiques, alors merci à Clotaire (mon sponsor officiel !!), Yann, Audrey, Benjamin, Céline, Romain, Philippe... pour leur compagnie récurrente, et bleu-bleu alias Jean-Pierre, qui m'a bien souvent fait rire !

Mon père me l'a bien fait remarquer le jour de la soutenance, donc je n'oserai pas oublier de remercier ici... mes parents ! Malgré la distance, ils m'ont toujours soutenu et surtout encouragé tout au long de la thèse (j'entends encore mon père dire « sois bon ! » au téléphone). Merci à eux pour tout, pour m'avoir permis d'accéder aux études supérieures et ainsi d'achever ce long parcours universitaire. Je ne pourrais oublier de remercier Matthieu, qui m'a accompagné durant ces trois ans et qui a su me soutenir lors de ces périodes obscures que l'on s'empresse d'oublier !

Table des matières

Glossaire.....	14
Introduction générale.....	17
Première partie : Contexte et outils.....	21
1 La prévision des précipitations adaptée par analogie.....	23
1.1 Principe de l'analogie en météorologie.....	23
1.1.1 Une impression de « déjà vu ».....	23
1.1.2 Procédure d'identification d'une analogue.....	25
1.2 Historique.....	26
1.2.1 De l'approche statistique vers l'adaptation statistique.....	26
1.2.2 Méthodes développées en France.....	27
1.2.3 Applications externes à la France.....	29
1.3 Description de la méthode de référence.....	30
1.3.1 Sélection de situations analogues à la situation du jour C.....	31
1.3.2 Emission de la prévision de pluie probabiliste.....	36
1.4 Principales limitations.....	38
1.4.1 Longueur des archives.....	38
1.4.2 Homogénéité des archives.....	39
1.4.3 Robustesse et fiabilité de la prévision des variables d'analogie.....	39
1.4.4 Résolution.....	40
1.5 Conclusion et objectif de la thèse.....	41
2 Secteurs d'étude et données utilisées.....	43
2.1 Zones d'étude.....	43
2.1.1 Le bassin versant de la Saône à Lyon.....	43
2.1.2 Le bassin versant de la Seine à Paris.....	45
2.2 Données disponibles.....	46
2.2.1 Archives de réanalyses météorologiques.....	47
2.2.2 Archives pluviométriques.....	53
2.2.3 Archive hydrométrique.....	56
3 Outils d'évaluation et contextes d'élaboration d'une prévision.....	59

3.1	Evaluation globale.....	59
3.1.1	Choix du score CRPS.....	60
3.1.2	Compétence CRPSS du CRPS.....	62
3.1.3	Gain G de performance.....	63
3.2	Evaluation en fonction d'un seuil.....	64
3.2.1	Probabilité de détection (POD).....	64
3.2.2	Taux de fausses détections (POFD).....	65
3.2.3	Taux de fausses alertes (FAR).....	65
3.2.4	Score d'Hansen-Kuipers (TSS).....	65
3.2.5	Diagramme ROC et score AROC.....	66
3.3	Identification du biais de la prévision.....	67
3.3.1	Diagramme de Talagrand.....	67
3.3.2	Diagramme de fiabilité.....	69
3.4	Contextes d'élaboration d'une prévision adaptée par analogie.....	70
4	Analyse exploratoire des bases de données.....	71
4.1	Comparaison et critique des archives pluviométriques.....	71
4.1.1	Objectif et méthodologie.....	71
4.1.2	Résultats.....	72
4.2	Analyse des régimes pluviométriques et hydrologiques.....	76
4.2.1	Régimes pluviométriques.....	76
4.2.2	Régimes hydrologiques.....	78
4.3	Comparaison et critique des archives de réanalyses.....	83
4.3.1	Comparaison des variables d'analogie de la méthode de référence entre les réanalyses ERA-40 et NCEP/NCAR.....	83
4.3.2	Etude des anomalies d'humidité dans les réanalyses ERA-40.....	88
4.4	Sensibilité de la méthode de prévision par analogie à l'archive de réanalyses.....	92
4.4.1	Archives NCEP/NCAR et ERA-40 à 2,5°.....	92
4.4.2	Archive ERA-40 aux résolutions 2,5° et 1,125°.....	95
4.5	Conclusion.....	98
	Seconde partie : Développement de la méthode de prévision par analogie.....	101
5	Analyse préliminaire des variables d'analogie potentielles.....	107
5.1	Définitions.....	108
5.1.1	Variables brutes.....	108
5.1.2	Variables élaborées.....	112
5.2	Analyse des liaisons entre les variables.....	114
5.2.1	Analyse en composantes principales.....	114
5.2.2	Analyse de corrélation et des similitudes entre champs.....	118
5.3	Tests préliminaires sur des variables d'analogie potentielles.....	120
5.4	Conclusion.....	121
6	Prise en compte de la saisonnalité.....	123

6.1 Pourquoi introduire la température ?.....	123
6.1.1 Inconvénient de la sélection des situations candidates.....	123
6.1.2 Principe et test préliminaire.....	126
6.2 Optimisation des paramètres.....	127
6.2.1 Choix d'un niveau de pression et d'une heure de mesure.....	129
6.2.2 Choix de plusieurs heures de mesure et niveaux de pression.....	132
6.2.3 Choix du nombre de situations candidates à retenir.....	135
6.2.4 Choix du domaine d'analogie.....	138
6.2.5 Choix du critère d'analogie.....	140
6.3 Validation et diagnostic.....	141
6.3.1 Adaptation au bassin de la Seine.....	142
6.3.2 Validation sur les deux bassins tests.....	146
6.3.3 Comparaison à la sélection par saisonnalité.....	147
6.4 Conclusion.....	152
7 Prise en compte des mouvements verticaux.....	155
7.1 Un niveau d'analogie supplémentaire.....	156
7.1.1 Objectif et choix de la variable.....	156
7.1.2 Un niveau d'analogie intermédiaire.....	158
7.2 Optimisation des paramètres.....	159
7.2.1 Choix du critère d'analogie.....	160
7.2.2 Nombres d'analogues à sélectionner.....	162
7.2.3 Heures de mesure et domaine d'analogie.....	164
7.3 Validation et diagnostic.....	168
7.3.1 Adaptation au bassin de la Seine.....	169
7.3.2 Validation sur les deux bassins tests.....	171
7.3.3 Diagnostic approfondi des résultats.....	173
7.4 Conclusion.....	176
8 Optimisation des paramètres de référence.....	181
8.1 Domaine d'analogie du niveau 1.....	182
8.2 Des heures de mesure supplémentaires.....	188
8.2.1 Variables du niveau 1.....	188
8.2.2 Variable du niveau 3.....	189
8.3 Des niveaux de pression supplémentaires.....	190
8.3.1 Constat préliminaire.....	190
8.3.2 Application à l'humidité relative.....	191
8.4 Synthèse, validation et diagnostic.....	197
8.4.1 Application au bassin de la Seine.....	198
8.4.2 Validation sur les deux bassins tests.....	199
8.4.3 Diagnostic des résultats.....	200
8.4.4 Application à la méthode A1.....	201
8.5 Conclusion.....	202

Troisième partie : De la prévision des pluies à la prévision des débits.....	205
9 Outils et méthodologie.....	209
9.1 Modèle hydrologique.....	209
9.1.1 Description.....	209
9.1.2 Procédure de calage du modèle.....	211
9.1.3 Evaluation en conditions parfaitement connues.....	213
9.2 Génération des prévisions de débit.....	217
9.3 Prévisions naïves.....	218
9.3.1 Persistance.....	219
9.3.2 Climatologie.....	219
10 Dans un contexte de prévision parfaite.....	221
10.1 Prévision des précipitations.....	221
10.1.1 Ré-optimisation des nombres d'analogues.....	221
10.1.2 Etude statistique et critique des prévisions de précipitations.....	224
10.1.3 Discussion et conclusion.....	234
10.2 Prévision des débits.....	235
10.2.1 Pertinence des scénarios obtenus par tirage aléatoire.....	236
10.2.2 Evaluation des prévisions de débit.....	237
10.2.3 Conclusion.....	245
11 Dans un contexte de prévision opérationnelle.....	247
11.1 Choix préliminaires.....	248
11.1.1 Période d'étude et modèle de prévision.....	248
11.1.2 Disponibilité des variables d'analogie.....	248
11.2 Prévision des précipitations.....	249
11.2.1 Evaluation globale.....	250
11.2.2 Evaluation en fonction d'un seuil.....	255
11.3 Prévision des débits.....	258
11.3.1 Evaluation globale.....	258
11.3.2 Etude des performances en fonction du seuil de débit.....	263
11.4 Etude de cas.....	267
11.4.1 Remarques préliminaires.....	267
11.4.2 Exemple de défaut d'alerte.....	268
11.4.3 Exemple de fausse alerte.....	273
11.4.4 Exemples de bonne alerte.....	275
11.4.5 Comparaison à une prévision déterministe expertisée.....	280
11.5 Conclusions.....	282
Conclusions & perspectives.....	285
Références bibliographiques.....	293
Annexe 1 : Acte de colloque (2009), Colloque Isère 1859.....	299

Annexe 2 : Article scientifique (2009), Atmospheric Science Letters.....	317
Annexe 3 : Article scientifique (2009), La Houille Blanche.....	325
Annexe 4 : Acte de colloque (2008), International conference Hydropredict'.....	333

Glossaire

ACP	Analyse en Composantes Principales
AROC	Aire sous la courbe ROC
C4	Critère de persistance
CDAS	Climate Data Assimilation System
CEPMMT	Centre Européen de Prévision Météorologique à Moyen Terme
CNR	Compagnie Nationale du Rhône
CRPS	Continuous Ranked Probability Score
CRPSS	Continuous Ranked Probability Skill Score
DIV	Divergence Horizontale
DREAL	Direction Régionale de l'Environnement, de l'Aménagement et du Logement
ECA	Eulerian Centripetal Acceleration
ECMWF	European Centre for Medium-Range Weather Forecast
EDF	Electricité de France
ERA ou ERA-40	45-years ECMWF Re-Analysis
FAR	False Alarm Rate
GR3P	Génie Rural à 3 paramètres adapté à la Prévision
HYDRO	Banque Nationale de Données pour l'Hydrométrie et l'Hydrologie
LTHE	Laboratoire d'étude des Transferts en Hydrologie et Environnement
MAE	Mean Absolute Error
MARS	Meteorological Archive and Retrieval System
MCG	Modèles de Circulation Générale
MF	Météo-France
MNP	Modèles Numériques de Prévision du temps
NCEP/NCAR	National Center for Environmental Prediction / National Center for Atmospheric Research
NNR	NCEP/NCAR Reanalysis
PJXA	Pluie Journalière Maximale Annuelle
PNN	Pluie Non Nulle
POD	Probability Of Detection

POFD	Probability Of False Detection
PVU	Potential Vorticity Unit
R	Coefficient de corrélation
RAINFAST	Système de génération et de visualisation de prévisions de précipitations adaptées par analogie
RH	Humidité relative
RMSE	Critère ou score Root Mean Squarred Error
ROC	Relative Operating Characteristic
SAFRAN	Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige
SCHAPI	Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations
SPC	Service de Prévision des Crues
T	Température
TCW	Total column of water
TPE	Température Potentielle Equivalente
TSS	True Skill Statistic ou score de Hansen-Kuipers
TU	Temps Universel
TWS	Critère de Teweles-Wobus
U, V	Vitesse des composantes zonale et méridienne du vent horizontal
VT	Vent horizontal
W	Vitesse verticale
Z	Altitude géopotentielle
Z-2PVU	Altitude géopotentielle du tourbillon potentiel égal à 2PVU

Introduction générale

Pleue en Berry, *pleuve* en Picard, *pluja* en catalan, *chuva* en portugais... Du latin *pluvia*, la *pluie*, et de manière générale la variable « précipitations », est sans doute l'un des phénomènes météorologiques les plus difficiles à prévoir. Pourtant, les enjeux liés à la prévision des précipitations sont importants. En effet, connaître les pluies futures est utile voire nécessaire pour différents acteurs de l'eau¹. Notamment, les producteurs d'hydro-électricité français, tels que Electricité de France (EDF) ou la Compagnie Nationale du Rhône (CNR), ont besoin de prévisions précises et fiables pour anticiper les évolutions de débit au droit de leurs aménagements et assurer la sûreté hydraulique, ainsi que pour optimiser la production de l'énergie et la gestion de la ressource disponible. Le Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations (SCHAPI) et les Services de Prévision des Crues (SPC) souhaitent anticiper les risques de crue, et donc connaître les précipitations attendues pour des échéances de plusieurs jours (Lacaz et al., 2008).

Plus particulièrement, une des volontés du SCHAPI est d'améliorer l'efficacité de l'annonce de crues, pour une meilleure anticipation des inondations et à terme une réduction des pertes humaines et économiques. Le retour d'expérience sur la gestion des événements catastrophiques d'inondation ces dernières années a montré qu'une marge importante de progrès existait dans le domaine de la prévision des crues. Sur les bassins versants de grande taille, le délai de propagation amont-aval permet aux SPC de faire actuellement une prévision des crues de un à trois jours à l'avance. Or, une prévision de précipitations jusqu'à une échéance de sept jours leur permettrait de mettre en sécurité la population et de protéger les biens au plus tôt. Ces besoins font écho à la demande sociale, notamment à Paris, là où les enjeux sont importants : voie routière sur berge, musées au bord de la Seine, Assemblée Nationale, etc. D'un point de vue économique, on estime qu'une crue centennale de la Seine telle que celle qui s'est produite en 1910 occasionnerait aujourd'hui, en Ile-de-France, des dégâts de l'ordre de 12 milliards d'euros (HYDRATEC, 1998).

Pour les raisons évoquées précédemment, on attend des météorologues qu'ils élaborent des prévisions de précipitations les plus fiables possibles. Malheureusement, la complexité des

¹ Les exemples donnés ensuite ne se veulent pas exhaustifs, puisque de nombreux autres acteurs de l'eau sont intéressés par la prévision des précipitations (agriculteurs pour l'irrigation, gestionnaires des réseaux d'assainissement, etc.).

processus responsables de la formation des précipitations et leur connaissance partielle ne permettent pas d'établir des équations robustes les régissant et encore moins de les résoudre de manière analytique. Les Modèles Numériques de Prévision du temps (MNP) les remplacent alors par des paramétrisations, mais force est de constater que la pluie est une variable difficile à prévoir, notamment pour des échéances de l'ordre de plusieurs jours. L'échelle de ces phénomènes est notamment trop petite pour qu'ils soient pris en compte par les MNP (processus sous-maille). Toutefois la recherche continue sur ce type de modèle. La prévision des précipitations qu'ils proposent s'améliore continuellement, notamment grâce à l'augmentation de leur résolution. Cet outil constitue un support incontournable pour un prévisionniste.

Cependant, la qualité des prévisions d'autres variables météorologiques par les MNP se dégrade moins vite avec l'échéance que celle des prévisions de précipitations. Ces variables, généralement mieux prévues pour des échéances éloignées, permettent de caractériser une situation météorologique à venir. En identifiant dans le passé des journées aux conditions synoptiques similaires et en leur associant la pluviométrie observée lors de ces journées, il est possible d'estimer les risques d'événements extrêmes ou d'intensité plus faible. Il s'agit là d'un principe bien connu dans le domaine de la météorologie : celui de la recherches d'analogues. Pour aller plus loin, les risques peuvent être mesurés sous forme probabiliste, grâce à un traitement statistique des cumuls de pluie des journées analogues. Le prévisionniste peut ainsi disposer d'une approche complémentaire de celle de la modélisation numérique, reposant sur une adaptation statistique des produits des modèles numériques.

Par ailleurs, il faut savoir que les MNP proposent de leur côté depuis plusieurs années une prévision probabiliste des précipitations : la prévision d'ensemble. Celle-ci permet de mieux intégrer le caractère chaotique du système atmosphérique qui amplifie rapidement les incertitudes sur l'état initial de l'atmosphère, en fournissant des scénarios équiprobables par perturbation des conditions initiales du modèle numérique ou par perturbation de paramètres du modèle. De nos jours, des prévisions d'ensemble sont produites régulièrement par des centres météorologiques, tels que le centre américain NCEP (Toth and Kalnay, 1993, 1997), ou le centre européen ECMWF (Molteni et al., 1996 ; Houtemaker et al., 1996). En France, Météo-France a développé un système de prévision d'ensemble PEARP basé sur le modèle numérique ARPEGE. Ces prévisions servent notamment à alimenter, pour l'instant à titre expérimental, une chaîne de prévision hydrologique (SAFRAN-ISBA-MODCOU) afin de prévoir de manière probabiliste les débits (Rousset-Regimbeau et al., 2007 ; Thirel et al., 2008 ; Martin et al., 2009). L'incertitude proposée par la prévision d'ensemble est donc une incertitude sur l'évolution de l'atmosphère, alors que celle obtenue par analogie est une incertitude sur la génération des précipitations à situation atmosphérique donnée. Par conséquent, ces deux approches peuvent être utilisées de façon complémentaires, elles ne s'excluent pas.

C'est l'approche par analogie que nous étudierons dans cette thèse. Nos travaux

s'inscrivent évidemment dans la continuité de travaux antérieurs. Cette méthode, initiée en France par Duband (1970), a été mise à jour en 2004 par Bontron, qui a privilégié comme secteur d'application des bassins à crues rapides du Sud-Est de la France. Afin de répondre aux attentes des SPC, nous proposons d'adapter cette méthode, que nous qualifierons de référence, à des bassins à crues lentes, tels que ceux de la Saône en amont de Lyon et de la Seine en amont de Paris, pour lesquels le contexte météorologique à l'origine des crues (succession de perturbations frontales) diffère de celui associé aux petits bassins montagneux (convection et soulèvements orographiques prédominants dans la genèse des crues). L'objectif final est d'établir s'il est possible, grâce à une prévision probabiliste des précipitations adaptée par analogie, d'émettre des prévisions de débits correctes sur ce type de bassin à des horizons supérieurs à ceux que connaissent actuellement les SPC.

Le présent mémoire est divisé en trois parties :

La première partie se propose d'introduire le contexte et les données nécessaires à l'application d'une méthode d'analogie. Le chapitre 1 vise à décrire le principe d'analogie et à dresser un état de l'art de la technique de prévision des précipitations basée sur ce principe. Les zones d'étude et les données utilisées sont présentées dans le chapitre 2, puis le chapitre 3 décrit les critères mathématiques qui nous permettront d'évaluer les performances d'une méthode de prévision probabiliste. Les contextes d'élaboration d'une prévision de pluie (prévision *parfaite* et prévision *opérationnelle*) y sont également mentionnés. Le chapitre 4 s'intéresse à l'analyse exploratoire des différentes bases de données dont nous disposons pour nos travaux. Il permet notamment de faire des choix cruciaux dont les résultats suivants dépendront.

La seconde partie est entièrement dédiée au développement de l'algorithme de recherche d'analogues. Le bassin de la Saône constituera un bassin test pour la recherche du meilleur algorithme, tandis que celui de la Seine permettra d'adapter les différents algorithmes développés en vue de les valider. Dans un premier temps, le chapitre 5 se propose de présenter les variables descriptives d'une situation météorologique que nous considérons comme les plus informatives vis à vis des précipitations, et d'étudier les liaisons entre elles afin de ne retenir que les variables susceptibles d'être introduites dans l'algorithme. Dans les chapitres 6 et 7, nous proposons d'introduire deux variables pertinentes : la température, qui permet notamment de relaxer une hypothèse forte ; et la vitesse verticale, qui vise à mieux caractériser les mouvements atmosphériques verticaux, et par conséquent les fronts, générateurs de l'essentiel des précipitations sur nos secteurs d'intérêt. Enfin, nous envisageons dans le chapitre 8 de ré-optimiser certains paramètres de la méthode de référence.

Dans la troisième partie, nous proposons d'appliquer les algorithmes développés dans la seconde partie dans l'objectif de prévoir les débits. Le chapitre 9 est dédié à la présentation des outils et à la description de la méthodologie adoptée pour la génération de prévisions probabilistes de débit. Le chapitre 10 est consacré à l'évaluation des prévisions de pluie et de

débit dans un contexte de prévision *parfaite* des précipitations. Dans ce contexte, les situations cibles ne sont pas prévues par un modèle numérique, mais sont en quelque sorte observées. Enfin, l'étude des performances de prévision de pluie et de débit dans un contexte de prévision *opérationnelle* est présentée au chapitre 11, avec des situations cibles prévues par un modèle numérique. De cette manière, il est possible d'identifier les différentes sources à l'origine des incertitudes de la prévision hydrologique. A titre d'illustration, quelques cas de crue sont également analysés à la fin de ce chapitre.

Enfin nous résumerons en fin de mémoire les principales conclusions du travail effectué, avec une série de perspectives et de développements possibles.

En annexe sont reportés les articles scientifiques qui ont été produits au cours de ces travaux de thèse.

Première partie

Contexte et outils

1 La prévision des précipitations adaptée par analogie

L'approche reposant sur le principe d'analogie est au cœur des travaux de thèse. Ce n'est qu'une méthode possible parmi celles aptes à émettre des prévisions quantitatives probabilistes de précipitations.

Ce chapitre propose d'introduire le sujet en dressant un état de l'art des approches qui reposent sur le principe de l'analogie en météorologie. Dans un premier temps, les principaux aspects de ce type d'approche seront présentés. Nous dresserons ensuite l'historique de ce principe appliqué à la prévision des précipitations, avant de décrire en détails la dernière version de la méthode d'analogie, validée en France et qui formera le point de départ de nos travaux. Enfin, nous évoquerons les principales limites d'une telle méthode de prévision qu'un utilisateur potentiel se doit de connaître.

1.1 Principe de l'analogie en météorologie

1.1.1 Une impression de « déjà vu »

Le moyen le plus simple de décrire le principe d'analogie est de faire un parallèle avec le comportement d'un prévisionniste en météorologie. En effet, grâce à son expérience et à sa mémoire, il associe instinctivement une situation météorologique prévue à des situations passées sensiblement similaires. Il ressent ainsi une impression de « déjà vu ». Il est alors à même d'évaluer le risque qu'un événement (tel qu'un épisode pluvieux conséquent, un épisode orageux, une tempête, une chute de neige exceptionnelle, etc.) se produise au cours de cette journée, ce après un travail de mémoire sur les événements passés. C'est généralement en raison du caractère exceptionnel du phénomène météorologique qu'il se souvient de situations passées. Sa mémoire, si bonne soit elle, lui fait évidemment défaut si l'événement ne revêt pas de caractère exceptionnel. En revanche, c'est son expérience qui peut lui permettre d'estimer le risque qu'il se produise.

Cette démarche, devenue finalement habituelle pour le prévisionniste expérimenté, repose

sur l'hypothèse que deux situations météorologiques similaires doivent engendrer des effets locaux du même ordre (Lorenz, 1956, 1969). Remarquons au passage que l'hypothèse réciproque est bien moins fondée : Tartaglione et al. (2009) ont en effet montré que si l'on observe un cumul pluviométrique identique pour deux situations, cela n'implique pas nécessairement que les champs de pression associés à ces situations se ressemblent.

Le principe d'analogie ne s'applique pas que dans le domaine de la météorologie. Ainsi par exemple, Bois et Obled (1976) ou encore Obled and Good (1980) ont utilisé ce principe pour la prévision des avalanches. Ils ont recherché des situations nivo-météorologiques analogues pour déterminer le risque de déclenchement d'avalanches. Plus récemment, Kalkstein et al. (2008) ont appliqué le principe d'analogie pour estimer la mortalité qui serait attendue dans certaines villes américaines si une vague de chaleur extrême comme celle observée à Paris en 2003 se produisait. Le principe peut par exemple être utilisé pour étudier les effets du changement climatique prévu par des modèles sur certaines zones urbaines (Kopf et al., 2008). L'objectif est de rechercher une ville B dont le climat actuel correspond statistiquement au climat simulé d'une ville A, ce qui permet d'apprécier les effets du changement climatique sur les zones urbaines, en remplaçant le changement du climat (dans le temps) par le changement de la localisation d'une ville (dans l'espace). On peut même utiliser le principe de l'analogie pour déterminer les effets du changement climatique prévu sur l'enneigement, à l'image de l'application de Martin (1995) aux Alpes françaises ; ou bien sur les caractéristiques de la saison touristique du ski (Dawson et al., 2009), telles que la longueur de la saison de ski, la quantité de neige artificielle produite ou l'énergie utilisée par les canons à neige.

Ces exemples montrent bien la diversité d'application et le fort intérêt d'un principe tel que celui de l'analogie. Dans notre cas, nous l'utiliserons pour la prévision des précipitations. Nous supposons donc que deux situations météorologiques analogues, au sens tel que nous le définirons, génèrent des quantités de pluie équivalentes sur la zone géographique qui nous intéresse. Cette technique permet également de reconstituer des événements pluviométriques passés, sur lesquels on dispose de peu d'information, à partir d'observations plus récentes (cf. Ben Daoud et al. (2009b) en Annexe 1).

Nous connaissons à présent le principe d'analogie, mais nous ne savons pas comment l'appliquer concrètement à la prévision des précipitations. Les questions que nous sommes en droit de nous poser sont les suivantes : Quels paramètres descriptifs d'une situation météorologique allons-nous choisir ? Quel est le critère qui va nous permettre d'affirmer qu'une situation est plus analogue à la situation prévue qu'une autre ? Sur quel domaine spatial doit-on rechercher l'analogie ? Pour y répondre, nous allons maintenant décrire la procédure générique permettant d'identifier une situation analogue.

1.1.2 Procédure d'identification d'une analogue

La recherche d'une situation analogue repose sur un choix de plusieurs paramètres. Tout d'abord, il faut préalablement définir ce que l'on veut prévoir : dans notre cas, il s'agit du cumul de précipitations. Il est également nécessaire de définir la durée sur laquelle les précipitations sont cumulées. Pour que la prévision soit utile aux systèmes d'alerte de crue, la durée d'intégration ne doit bien sûr pas dépasser le temps de réaction du bassin versant. Pour émettre une prévision de précipitations, il faut en outre disposer d'une archive contenant les cumuls pluviométriques observés dans le passé, ce qui peut être limitant pour émettre des prévisions sur des pas de temps fins.

Ensuite, nous devons choisir les variables qui décrivent l'état de l'atmosphère et qui permettent de caractériser une situation de la meilleure façon : les **variables d'analogie**. Ces variables ne sont évidemment pas choisies au hasard. Elles doivent nécessairement être liées à la variable que l'on veut prévoir. Il faut également disposer d'une archive météorologique suffisamment longue, de laquelle les champs des variables d'analogie sont extraits.

Le domaine spatial optimal sur lequel on recherche l'analogie doit être défini : il s'agit du **domaine d'analogie**. La localisation et la taille de ce domaine dépendent d'une part du lieu géographique - et donc du climat - pour lequel on désire émettre une prévision, et d'autre part des échelles d'espace et de temps auxquelles la variable d'analogie correspondent.

Enfin, il faut choisir le critère numérique adéquat qui nous permet d'identifier une situation analogue : le **critère d'analogie**, grâce auquel nous pourrions classer et comparer les situations passées selon leur degré de ressemblance avec la situation prévue, et identifier la meilleure analogue.

Cependant, l'atmosphère est un système très complexe dont l'état exact :

- ne peut être décrit de façon exhaustive. L'analogie est donc recherchée sur un nombre restreint de variables, si bien qu'elle ne présage pas de l'analogie sur les autres variables non prises en compte ;
- s'il peut être décrit de façon exhaustive, ne peut être identique entre deux instants sur une durée raisonnable. Par exemple, Van Den Dool (1994) a estimé la longueur d'archive nécessaire pour trouver deux situations analogues dont l'écart de champ de géopotential¹ à 500 hPa sur l'Hémisphère Nord serait de l'ordre grandeur de l'erreur de mesure du géopotential. Les calculs indiquent une longueur d'archive égale à 10^{30} années ! Ce résultat, déjà pressenti par Lorenz (1969), rejoint celui de Ruosteenoja (1988), qui avait trouvé une longueur de 10^{29} années. Mais en revanche, sur un domaine géographique plus restreint, on peut raisonnablement trouver de bonnes analogues avec des archives de longueur « courte » (10 à 100 ans).

Par conséquent, nous ne chercherons pas à identifier la *meilleure* analogue, mais un ensemble de situations analogues, le degré d'analogie étant alors différent d'une situation à

1 Le lecteur est invité à se reporter au chapitre 5 (5.1.1.1) pour la définition de cette variable météorologique.

une autre. La taille de cet ensemble est définie par le **nombre d'analogues** à sélectionner.

Ce paramètre doit également être choisi astucieusement. En effet, le nombre d'analogues retenues doit être suffisamment grand pour exprimer l'incertitude qui reste sur la variable prévue du fait des caractéristiques météorologiques non prises en compte par les variables d'analogie. Mais la longueur de l'archive n'étant pas infinie, si nous retenons d'emblée un nombre trop important d'analogues, la qualité de l'analogie sera moins bonne vis-à-vis des variables d'analogie. De cette manière, nous diluerons l'information pluviométrique qu'elles peuvent apporter et nous nous rapprocherons de la climatologie. Il est donc nécessaire de trouver un compromis entre un nombre suffisant de situations analogues, afin d'exprimer l'incertitude sur les variables non prises en compte par l'analogie, et un nombre pas trop élevé afin de conserver des situations analogues suffisamment ressemblantes. Toutefois dans certaines études, seules quelques analogues ont été retenues (Roebbert and Bosart, 1998), voire uniquement la meilleure analogue (Wetterhall et al., 2005), mais en toute logique les résultats ne se sont pas avérés concluants.

Par la suite, outre les paramètres que nous venons de définir, nous parlerons de situations cibles et de situations candidates. Une situation *cible* est une situation pour laquelle on cherche à émettre une prévision, tandis qu'une situation *candidate* est une situation historique observée sur laquelle on recherche une analogie avec la situation cible. Nous nommerons ainsi *période candidate* la période dans laquelle les situations analogues seront finalement recherchées.

Pour la prévision des précipitations, il est donc nécessaire de disposer de deux archives :

- une archive de lames d'eau observées sur la *période candidate* ;
- une archive des variables météorologiques qui décrivent les situations passées *candidates*.

1.2 Historique

1.2.1 De l'approche statistique vers l'adaptation statistique

Le concept d'analogie existe depuis longtemps en météorologie. Il a d'abord été utilisé pour la prévision des précipitations avec une approche statistique simple. Une situation présente observée était supposée évoluer à peu près de la même façon qu'une situation passée observée analogue. Etant donné le nombre très important de degrés de liberté de l'évolution de l'état de l'atmosphère, l'écart entre les situations, pourtant négligeable initialement, devient très important avec l'évolution temporelle (Lorenz, 1963, 1969). En effet, la similitude entre deux situations n'est pas durable, même si le domaine d'analogie est de faible étendue : tôt ou tard, les situations deviendront trop éloignées pour qu'elles soient considérées comme analogues.

Bien qu'elle puisse encore être utilisée actuellement, à l'image des travaux récents de

Bannayan and Hoogenboom (2008), cette méthode statistique a rapidement été abandonnée au profit de la modélisation numérique de prévision (MNP). Les MNP ont d'abord permis une meilleure prévision à court et moyen terme. Grâce aux Modèles de Circulation Générale (MCG), la prévision à long terme, dite prévision saisonnière, a ensuite été rendue possible et s'est améliorée avec le progrès de la modélisation numérique (Déqué, 2003; Leroy et Céron, 2007).

Depuis la seconde moitié du XX^{ème} siècle, le concept d'analogie a plutôt été appliqué par adaptation statistique des sorties de MNP. Les modèles actuels fournissent une évolution temporelle des champs synoptiques des jours à venir à partir desquels on peut alors faire l'analogie avec des situations passées. Les situations météorologiques retenues comme analogues à chaque échéance de prévision sont alors choisies indépendamment, contrairement à la méthode purement statistique, qui prenait d'emblée la séquence passée comme prévision.

De nos jours, la recherche d'analogie est souvent utilisée pour réévaluer les valeurs de variables moins bien prévues par les MNP, notamment pour la prévision à long terme. Cette méthode permet d'effectuer ce que l'on appelle une descente d'échelle (de l'anglais *downscaling*), c'est-à-dire qu'elle permet de déduire les variables d'échelle locale (telles que la température, les précipitations, l'insolation, le risque d'avalanches, etc.) à partir de variables définies à grande échelle (telles que les champs de pression, de vent, etc.). Ce type d'approche s'est rapidement répandu au cours des années 1990 (Wilby and Wigley, 1997).

1.2.2 Méthodes développées en France

Historiquement, Duband (1970) a commencé à travailler sur la prévision des précipitations journalières par analogie. Ses travaux ont conduit EDF à implanter un système de prévision opérationnel au début des années 1970 (Duband, 1974). Ce système utilisait les champs de pression observés le jour J et une extrapolation temporelle était effectuée à partir de ces champs pour les deux jours suivants. Il s'agissait là d'une approche statistique simple.

Avec le progrès des MNP, la méthode a évolué dans les années 1980 (Duband, 1980) en devenant une méthode d'adaptation statistique des MNP. Elle s'est développée ensuite jusque dans les années 1990, avec l'amélioration des prévisions de variables synoptiques. L'échéance de prévision a alors été repoussée à 4 jours (jusqu'à l'échéance $J+3$).

Ensuite des améliorations ont été apportées par Guilbaud (1997), avec l'introduction d'un nouveau critère d'analogie. Grâce aux performances croissantes des MNP, l'échéance de prévision a été étendue à cinq jours ($J+4$). Enfin, Bontron (2004) a amélioré la méthode de Guilbaud en hiérarchisant la sélection des situations analogues en deux niveaux et en utilisant un nouveau type d'archive de variables d'analogie (Obled et al., 2002) : les réanalyses météorologiques, que nous définirons précisément dans le chapitre suivant.

Optimisée sur des bassins à crues rapides du Sud-Est de la France, cette méthode permet d'émettre tous les jours une prévision quantitative des précipitations de manière probabiliste

pour les jours à venir. Elle a été adaptée aux sous bassins du Rhône et fonctionne de manière opérationnelle à la Compagnie Nationale du Rhône au sein du système OPALÉ (Outils de Prévision par Analogie des Lames d'Eau). Ce système fournit des prévisions de pluie probabilistes et déterministes au pas de temps de six heures sur un semis de bassins continu (Bompart et al., 2009). Par ailleurs, la méthode a été implantée au Laboratoire des Transferts en Hydrologie et Environnement (LTHE) de Grenoble, en tant que système de démonstration (système RAINFAST). Le laboratoire, qui gère l'observatoire OHMCV (Observatoire Hydro-météorologique Méditerranéen Cévennes-Vivarais), met à disposition les prévisions quotidiennes de précipitations journalières émises sur plusieurs sous-bassins du Sud-Est de la France, notamment sur la région des Cévennes¹. Dans ce contexte, et plus récemment, la méthode d'analogie développée par Bontron (2004) au pas de temps journalier a été adaptée par Marty (2010) à des pas de temps infrajournaliers sur des bassins à crue rapide de la région Cévennes-Vivarais (Marty et al., 2008, 2009).

Le Tableau 1.1 résume l'ensemble des paramètres qui caractérise les versions successives de la méthode en mode de prévision opérationnelle jusqu'à celle de Guilbaud. Ces paramètres sont l'archive météorologique et les variables d'analogie² qu'elle contient, le domaine d'analogie, les , les critères d'analogie, le nombre d'analogues retenues et la variable prévue, souvent exprimée sous la forme d'une ou plusieurs valeurs de quantiles au non-dépassement.

Tableau 1.1 : Récapitulatif des développements de la méthode initiée par Duband en 1970. Les variables Z700 et Z1000 représentent les géopotentiels à 700 et à 1000 hPa.

	Méthodes développées (auteur, année)		
	Duband, 1970	Duband, 1980	Guilbaud, 1997
Archive de variables d'analogie	25 radiosondages, période 1953-1973	37 radiosondages, période 1953-1980	37 radiosondages interpolés sur une grille régulière, période 1953-1985
Variables d'analogie	Z700 à 00 h TU, pression au sol (Pmsl) à 06 h TU	Z700 à 00 h TU, Z1000 à 00 h TU	Z700 à 00 h et à 24 h TU, Z1000 à 00 h et à 24 h TU
Domaine d'analogie	Europe de l'Ouest et proche Atlantique	Europe de l'Ouest et proche Atlantique	Sud de la France (-6,2°E/12,9°E ; 38,0°N/50,3°N)
Critères d'analogie	Analyse en composantes principales (ACP) sur Z700 et Pmsl	ACP sur Z700 et critère de corrélation sur Z700, Z1000 et Z700-Z1000	Critère de Teweles- Wobus (Teweles J. et Wobus H., 1954) sur Z700, Z1000 à 00 h et 24 h
Nombre d'analogues retenues	25	25	50
Variable prévue	P50% et intervalle de confiance P10%-P90%	P20%, P60% et P90%	P20%, P60% et P90%

La méthode développée par Bontron (2004) sera décrite en détails un peu plus loin dans ce chapitre (section 1.3).

1 Site internet : http://www.lthe.hmg.inpg.fr/OHM-CV/P752_analogues.php?echeance=0

2 Nous rappelons au lecteur que le géopotential Z sera défini dans le chapitre 5 (5.1.1.1). Le géopotential ZX définit le géopotential au niveau de pression X hPa.

1.2.3 Applications externes à la France

Parallèlement, le principe d'analogie a été appliqué dans d'autres pays pour la prévision des précipitations. Altava-Ortiz et al. (2006) ont utilisé une méthode d'analogie pour l'étude d'un événement intense en Catalogne. Ils ont notamment comparé les prévisions par analogie obtenues avec celles d'un MNP déterministe (MM5) et ont montré l'intérêt d'utiliser les deux. Ils ont en effet remarqué que les analogues permettent de bien discriminer les situations pluvieuses des situations non pluvieuses et produisent des maximas plus proches des observations si on les compare à ceux fournis par le MNP, tandis que le MNP reproduit une évolution temporelle correcte des précipitations attendues.

La performance d'une méthode d'analogie peut donc être comparée à celle d'un MNP en matière de prévision de pluie. Diomede et al. (2008) ont également comparé, pour un bassin versant italien, les prévisions de pluie issues de méthodes d'analogie avec celles fournies par un MNP. Les méthodes d'analogie sont basées sur les deux approches : prévision statistique simple et adaptation statistique des sorties d'un MNP à aire limitée. Elles ont été optimisées en testant plusieurs variables et critères d'analogie. Les auteurs ont suggéré que les prévisions par analogie doivent être considérées comme un complément aux prévisions déterministes fournies par un MNP. Dans le cadre de l'expérience MAP (Mesoscale Alpine Programme), Djerboua and Obled (2003) ont également comparé les prévisions issues d'une technique d'analogie avec celles fournies par le modèle ARPEGE de Météo-France. Leur conclusion rejoint les précédentes, à savoir qu'une prévision adaptée par analogie doit être considérée comme un complément aux prévisions générées par un MNP.

Il est par ailleurs courant de comparer plusieurs méthodes de sélection d'analogues. Pour la prévision des précipitations automnales en Catalogne, Gibergans-Baguena and Llasat (2007) ont ainsi testé trois méthodes : la première est à peu près identique à celle développée par Guilbaud (1997), la seconde n'utilise que des variables thermodynamiques, et la troisième combine les deux précédentes méthodes. Les auteurs ont montré l'intérêt de cette dernière, ce qui rejoint la conclusion de Bontron (2004) concernant l'analogie supplémentaire sur l'humidité. Notons qu'une étude antérieure (Llasat et al., 1997), effectuée sur quelques bassins du Sud de la France, a également montré l'intérêt d'ajouter des informations thermodynamiques pour améliorer la recherche d'analogues.

Le principe d'analogie pour la prévision des précipitations a également été appliqué sur des bassins versants allemands. Bliefertnicht and Bardossy (2008) se sont intéressés aux événements intenses uniquement. Ils ont effectué une série de tests avec plusieurs critères d'analogie : la distance euclidienne, le coefficient de corrélation et un critère qui combine les deux. Les auteurs ont montré que le critère combiné permettait d'obtenir les meilleures performances de prévision.

Aux Etats-Unis, Roebbert and Bosart (1998) ont étudié la sensibilité des précipitations aux détails de la circulation atmosphérique sur un épisode pluvieux de deux jours (14 au 16 février 1992). Ils n'ont sélectionné des situations analogues que sur les jours de février d'une

période de 25 ans (1965-1989). L'analogie a été recherchée sur les 48 h, en ne retenant que les trois situations les plus analogues. Logiquement de cette manière, les précipitations observées lors de ces trois groupes de deux journées sont bien différentes de celles observées au cours de l'épisode étudié. Une étude complémentaire (Roebbert and Reuter, 2002) a révélé que c'est en raison d'éléments non pris en compte par la méthode d'analogie (humidité, dynamique atmosphérique) que les pluies reconstituées sur les journées du 14 au 16 février 1992 n'ont pas été correctes.

Comme nous l'avons indiqué précédemment (section 1.2.1), le principe d'analogie est également appliqué pour la prévision saisonnière et pour l'étude de l'impact du changement climatique. Nous n'allons pas lister les travaux effectués dans ce contexte puisque les références sont très nombreuses. Nous noterons que ce sont Barnett and Preisendorfer (1978) qui ont notamment montré l'intérêt d'une telle technique pour la prévision saisonnière. Ils ont ainsi adopté une approche statistique simple pour la prévision des anomalies de température de surface. De nos jours, le principe d'analogie est plus souvent utilisé en tant que technique de descente d'échelle. Cette méthode simple à mettre en oeuvre est souvent comparée à d'autres méthodes plus complexes (von Storch, 1995 ; Zorita et al., 1995 ; Zorita and von Storch, 1999 ; Wetterhall et al., 2006 ; Wetterhall et al., 2007).

Enfin, le principe d'analogie peut également être utilisé pour corriger les prévisions produites par un MNP. Cette approche, qui a été adoptée par Hamill et al., 2006, consiste à rechercher des prévisions anciennes similaires aux prévisions actuelles pour une échéance donnée, et ensuite à extraire la variable à prévoir qui a été observée dans le passé à cette échéance. Puisque la structure des MNP change régulièrement, l'approche de Hamill consiste à utiliser des « re-prévisions » générées par un MNP identique sur une longue période. Cela permet de faire de meilleures prévisions que celles fournies par un modèle ensembliste de plus grande résolution. De cette manière, le biais systématique du MNP est corrigé. Cette approche s'est révélée fructueuse pour la correction des prévisions de précipitations (Hamill and Whitaker, 2006), mais aussi pour la prévision des températures et des géopotentiels (Hamill and Whitaker, 2007).

1.3 Description de la méthode de référence

Nos travaux s'inscrivent donc dans la suite de ceux menés en France depuis plusieurs dizaines d'années. La méthode développée par Bontron (2004) formera le point de départ de nos travaux. Nous l'appellerons par la suite la **méthode de référence** ou méthode **A0**. Considérons une journée cible C pour laquelle nous désirons émettre une prévision de précipitations journalières. La variable que nous cherchons à prévoir est représentée par le cumul de pluie de 6 h Temps Universel (TU) le jour C à 6 h TU le jour $C+1$. Les variables d'analogie sont définies toutes les 6 h sur une grille régulière de $2,5^\circ \times 2,5^\circ$.

Le système de prévision permet :

- i) de sélectionner un certain nombre de situations analogues $\{J_{ana}\}$ à la situation météorologique du jour C ;
- ii) d'établir une prévision de pluie sous une forme probabiliste pour le jour C à partir des pluies observées lors des journées analogues pré-sélectionnées : on donne ainsi la distribution des pluies conditionnée par la situation du jour C .

Bontron avait travaillé sur des regroupements de postes pluviométriques EDF. Pour chacun d'entre eux, les domaines d'analogie ont été optimisés. Nous allons détailler la procédure de sélection en reprenant les paramètres qu'il avait déterminés pour le regroupement pluviométrique R5 considéré dans sa thèse (Figure 1.1). Celui-ci correspond à un regroupement de postes pluviométriques situés approximativement dans le bassin versant de la Saône, que nous considérerons comme un bassin test dans le cadre de la thèse.

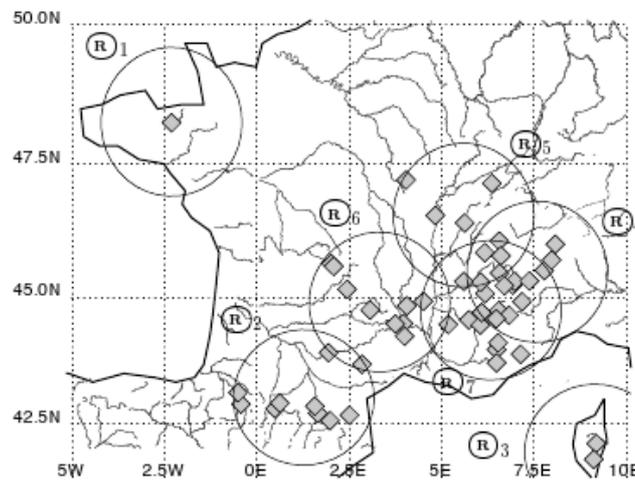


Figure 1.1 : Regroupements pluviométriques utilisés par Bontron (2004) dans ses travaux de thèse (extrait de Bontron (2004)).

1.3.1 Sélection de situations analogues à la situation du jour C

La sélection de situations analogues s'effectue en deux étapes. Une première sélection, que nous appellerons **niveau 1**, permet de constituer un échantillon de situations similaires à la situation cible en considérant la circulation générale à échelle synoptique. Une seconde sélection, que nous appellerons **niveau 2**, permet de ne retenir, parmi les situations de l'échantillon formé au niveau 1, que les situations analogues à la situation du jour C en terme d'hygrométrie à une échelle cette fois très locale.

A chaque niveau de recherche d'analogie correspondent un domaine d'analogie, des variables d'analogie, un critère d'analogie et un nombre de situations analogues retenues. Nous allons définir ces paramètres pour chaque niveau d'analogie, valables pour le regroupement R5 utilisé par Bontron (2004).

Remarque préliminaire sur la période candidate

La méthode a été optimisée en tenant compte des effets saisonniers. On suppose en ce sens que deux situations ayant lieu à une saison différente, mais similaires à échelle synoptique, ne produisent pas forcément les mêmes effets à l'échelle locale (Lorenz, 1969). Afin de respecter cette saisonnalité, les dates des situations candidates au niveau 1 extraites de chaque année candidate sont incluses dans une fenêtre temporelle glissante de quatre mois centrée sur le jour C . Pour l'illustrer, la Figure 1.2 indique la période dans laquelle sont extraites les dates des situations candidates pour la sélection d'analogues à la situation cible du 13/11/2008.

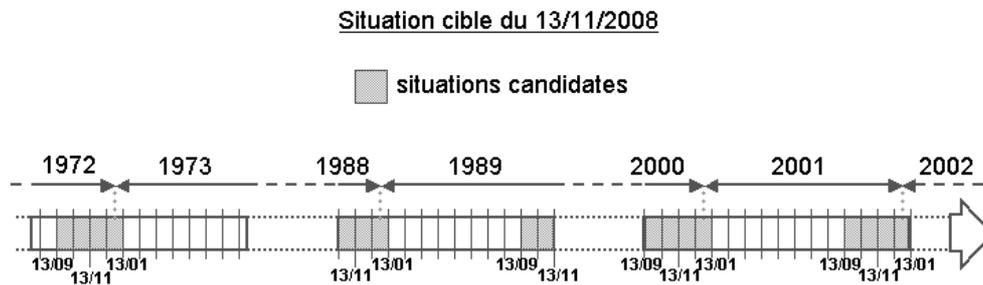


Figure 1.2 : Période dans laquelle sont extraites les dates des situations candidates pour la situation de la journée cible du 13/11/2008.

1.3.1.1 Niveau 1 de la méthode de référence

L'objectif de ce premier niveau de sélection est de rechercher l'analogie sur la circulation générale à échelle synoptique. Ainsi, les variables d'analogie retenues pour décrire l'état de l'atmosphère dans l'espace (niveaux de pression) et dans le temps (heures de mesure) sont le géopotential à 1000 hPa ($Z1000$), mesuré à 12 h TU le jour J , et le géopotential à 500 hPa ($Z500$), mesuré à 00 h TU le jour $J+1$. Par convention, et dans la suite de ce mémoire, les heures de mesure des variables d'analogie seront exprimées relativement à l'instant 00h TU du jour J (Figure 1.3).

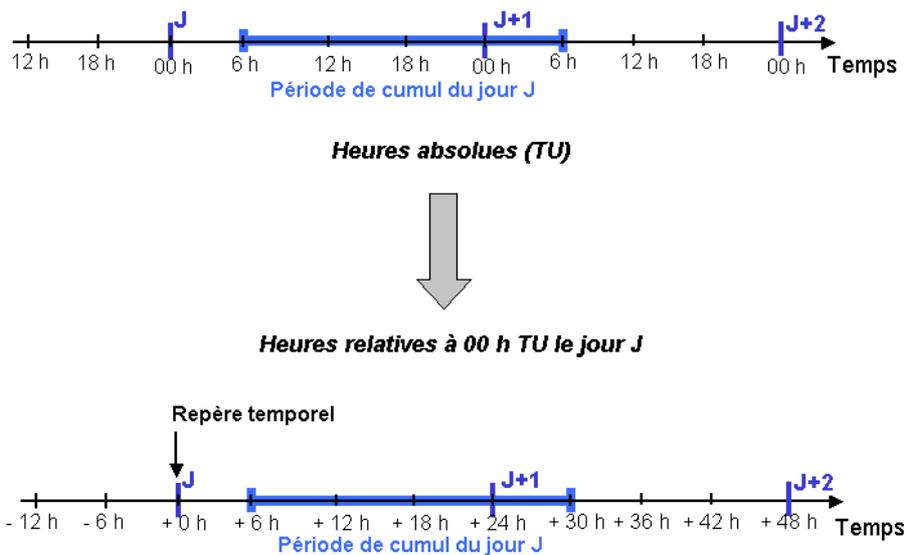


Figure 1.3 : Schématisation de la transformation des heures de mesure absolues en heures de mesure relatives à l'instant absolu 00h TU d'un jour J .

Elles seront identiques pour les journées cibles et candidates. Considérons une journée cible C et une journée analogue A . De cette manière, si l'on recherche l'analogie sur une variable à +30h, alors la variable mesurée à 6 h TU le jour $A+1$ est comparée à la variable mesurée à 6 h TU le jour $C+1$. Dans le cas présent, les variables d'analogie utilisées au niveau 1 sont donc $Z1000$ à +12h et $Z500$ à +24h.

Le domaine d'analogie D1 appliqué au niveau 1, indiqué par la Figure 1.4, tient compte de la position géographique du bassin (ici du regroupement pluviométrique R5), mais aussi des influences météorologiques principales. Le domaine est rectangulaire, centré en (5°E ; 45°N), et s'étire parallèlement au flux directeur principal des masses d'air (régime d'ouest prédominant pour cette région du globe). Il couvre ainsi une grande partie de l'Europe occidentale, permettant de rechercher l'analogie sur les champs de géopotential à échelle synoptique.

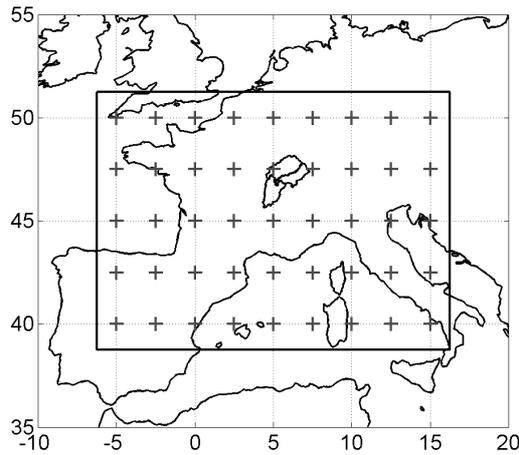


Figure 1.4 : Domaine d'analogie D1 du niveau 1 (-5°E/15°E ; 40°N/50°N) appliqué dans la méthode de référence A0.

Le critère d'analogie appliqué au niveau 1 ne tient pas compte de la valeur absolue du géopotentiel en chaque point, mais de l'analogie générale de forme et donc de la circulation générale sur le domaine considéré. Il s'agit du critère de Teweles and Wobus (1954), que l'on nommera *TWS*. Ce critère, défini par l'équation 1.1, fut initialement introduit pour évaluer la qualité des prévisions de champ de géopotentiel. Calculé pour chaque variable d'analogie du niveau 1, le critère *TWS* compare les formes des deux champs de géopotentiels sur le domaine D1 en se basant les gradients de géopotentiel dans les directions zonale et méridienne. Si sa valeur est égale à 200, les champs comparés ont une forme totalement opposée ; si elle est égale à 0, les champs comparés ont une forme identique.

$$TWS = 100 \cdot \frac{\left(\sum_{i=1}^{I-1} \sum_{j=1}^J |\Delta_{i,j}^{i,cible} - \Delta_{i,j}^{i,candidate}| + \sum_{i=1}^I \sum_{j=1}^{J-1} |\Delta_{i,j}^{j,cible} - \Delta_{i,j}^{j,candidate}| \right)}{\left(\sum_{i=1}^{I-1} \sum_{j=1}^J \max(|\Delta_{i,j}^{i,cible}|; |\Delta_{i,j}^{i,candidate}|) + \sum_{i=1}^I \sum_{j=1}^{J-1} \max(|\Delta_{i,j}^{j,cible}|; |\Delta_{i,j}^{j,candidate}|) \right)} \quad (1.1)$$

Où : I et J sont respectivement les nombres de points dans les directions zonale et méridienne ;
 $\Delta_{i,j}^{i,cible} = V_{i+1,j}^{cible} - V_{i,j}^{cible}$ (resp. $\Delta_{i,j}^{i,candidate}$) est le gradient dans la direction zonale du champ cible (resp. candidat) de la variable V ;
 $\Delta_{i,j}^{j,cible} = V_{i,j+1}^{cible} - V_{i,j}^{cible}$ (resp. $\Delta_{i,j}^{j,candidate}$) est le gradient dans la direction méridienne du champ cible (resp. candidat) de la variable V .

Pour chaque situation candidate, une valeur du critère *TWS* est calculée à partir du champ de géopotentiel Z1000 à +12h ($TWS_{Z1000,+12h}$) et une autre valeur est calculée à partir du champ de Z500 à +24h ($TWS_{Z500,+24h}$). On obtient alors une valeur unique du critère *TWS* pour chaque situation candidate, en calculant la moyenne arithmétique des deux scores :

$$TWS_{\text{niveau1}} = \frac{(TWS_{Z1000,+12h} + TWS_{Z500,+24h})}{2} \quad (1.2)$$

Enfin, les situations candidates sont classées par valeur croissante de TWS . A l'issue de ce premier niveau, les $N1=70$ premières situations les plus similaires en termes de champs de géopotentiels sont alors retenues. Ce nombre $N1$ a été optimisé par Bontron en fonction de la longueur de l'archive de situations candidates utilisée.

1.3.1.2 Niveau 2 de la méthode de référence

Cette seconde sélection permet de ne retenir que les situations les plus analogues en termes d'humidité à une échelle plus locale.

Choisies de manière heuristique par Bontron, les variables d'analogie de ce second niveau sont la quantité d'eau précipitable (TCW , de l'anglais Total Column of Water) et l'humidité relative au niveau de pression 850 hPa ($RH850$). Une variable fictive mais qui rend bien compte de la situation hygrométrique a été créée : il s'agit de $TCW\%$, correspondant au produit de TCW par $RH850$. Ces variables seront précisément définies au chapitre 5 (sections 5.1.1.3 et 5.1.2.3). De cette manière, la variable $TCW\%$ traduit à la fois la quantité d'eau disponible et la proximité à la saturation. Elle est mesurée à +12h (12 h TU le jour J) et à +24h (00 h TU le jour $J+1$).

Le domaine d'analogie D2 du niveau 2 se révèle bien plus restreint que le domaine D1 du niveau 1. Il est réduit à une fenêtre de quatre points de grille centrée approximativement sur le bassin (Figure 1.5).

Le critère d'analogie appliqué au second niveau n'est pas un critère de forme mais de distance : il s'agit du critère $RMSE$ (Root Mean Squarred Error). Il permet ainsi de retenir les situations dont la quantité d'humidité présente dans la troposphère est équivalente à celle de la situation cible. Le critère $RMSE$, défini par l'équation 1.3, mesure donc la distance moyenne entre les valeurs de $TCW\%$ des points du domaine D2. Pour une situation analogue, la valeur du critère est obtenue en sommant les valeurs des critères $RMSE$ calculées pour la variable $TCW\%$ à +12h et à +24h.

$$RMSE = \sqrt{\frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J (TCW\%_{i,j}^{cible} - TCW\%_{i,j}^{candidate})^2} \quad (1.3)$$

Où : $\left\{ \begin{array}{l} I \text{ est le nombre de points dans la direction zonale ;} \\ J \text{ est le nombre de points dans la direction médienne.} \end{array} \right.$

Plus la valeur du critère est faible, plus les valeurs de $TCW\%$ des situations cible et candidate sont proches, *i.e* plus la situation candidate à ce niveau 2 est analogue à la situation cible en terme d'humidité sur la zone proche du bassin.

Finalement, les 70 situations analogues sélectionnées au niveau 1 sont classées par valeur

croissante du critère *RMSE*. Seules les $N2=30$ premières situations sont alors retenues comme analogues.

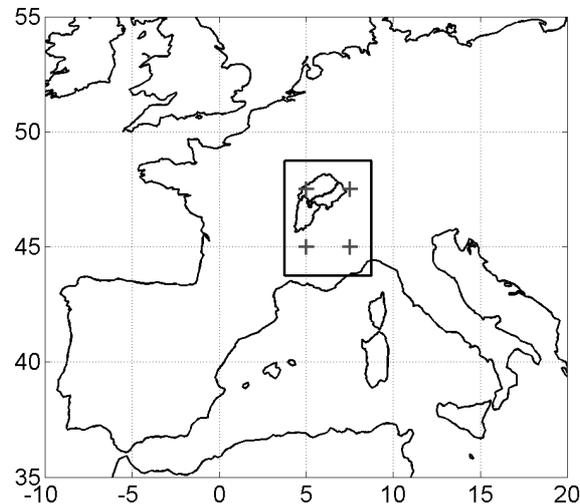


Figure 1.5 : Domaine d'analogie D2 du niveau 2 ($5^{\circ}\text{E}/7,5^{\circ}\text{E}$; $45^{\circ}\text{N}/47,5^{\circ}\text{N}$) appliqué dans la méthode de référence A0.

Notons que les nombres $N1$ et $N2$ d'analogues retenues aux niveaux 1 et 2 ont été optimisés par Bontron sur une période candidate de 44 ans (1953-1996). Nous verrons que pour notre travail, ces nombres seront ré-optimisés car la durée de la période candidate sur laquelle nous allons travailler est différente. Cette ré-optimisation sera effectuée au chapitre 4 (section 4.4.1.1).

1.3.2 Emission de la prévision de pluie probabiliste

Une fois les 30 situations analogues sélectionnées pour la journée C , on extrait les cumuls pluviométriques mesurés au cours de chaque journée. Classés par ordre croissant, ces cumuls permettent alors de construire la distribution empirique des pluies pouvant être observées ce jour C , conditionnées par l'analogie recherchée. La fonction de répartition prévue constitue ainsi une prévision de précipitations sous une forme probabiliste. La Figure 1.6 présente un exemple de prévision obtenue pour le bassin de la Saône.

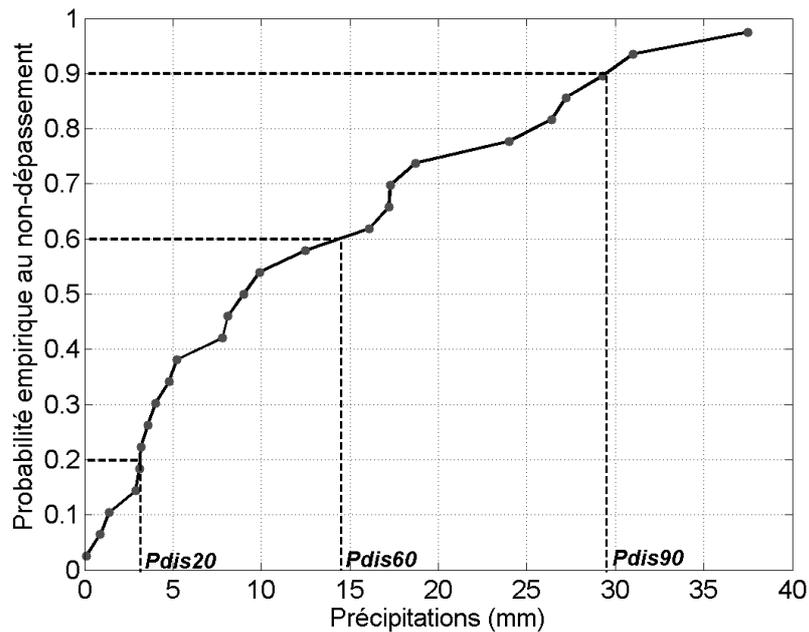


Figure 1.6 : Exemple de prévision émise, caractérisée par la fonction de répartition des pluies observées lors des dates des situations analogues.

De ce graphique peuvent être tirées plusieurs informations. Les éléments les plus couramment étudiés sont trois valeurs de quantile au non-dépassement, leur choix reposant sur l'expérience des prévisionnistes utilisant cette méthode. L'une des valeurs est une indication sur la probabilité de pluie : si la valeur associée dépasse 0 mm, alors on peut en déduire que la probabilité qu'il pleuve est importante. Les deux autres valeurs de quantile forment un intervalle de confiance. La première idée serait de retenir le quantile d'ordre 50 % ($Pdis50$) pour la valeur indiquant la probabilité de pluie, et les quantiles d'ordre 10 % ($Pdis10$) et 90 % ($Pdis90$) pour l'intervalle de confiance. Mais il se trouve que le cumul pluviométrique est borné en zéro, tandis qu'il ne l'est pas vers les grandes valeurs. C'est donc une variable dissymétrique non gaussienne, ce qui implique que la distribution est souvent biaisée vers les petites valeurs. Par conséquent, les quantiles retenus pour l'intervalle de confiance sont ceux d'ordre 20 % ($Pdis20$), et 90 % ($Pdis90$), et celui retenu pour indiquer la probabilité d'occurrence de pluie est le quantile d'ordre 60 % ($Pdis60$).

Enfin, pour se donner une idée de la fiabilité de la prévision, il est également possible de visualiser la distribution des pluies observées lors des situations les plus analogues à la situation cible (Figure 1.7). Si cette distribution se superpose à la fonction de répartition (Figure 1.7 a), les situations retenues comme analogues sont alors réparties de manière homogène, les meilleures analogues étant représentatives de l'ensemble des $N2$ situations de l'échantillon. Dans ce cas, la ressemblance entre les situations retenues et la situation cible est suffisamment forte : on peut ainsi estimer que la fiabilité de la prévision est correcte. Dans le cas contraire (Figure 1.7 b), si les cumuls associés aux dix meilleures analogues s'écartent fortement de la distribution de l'ensemble des situations, on peut considérer que la fiabilité est moins bonne.

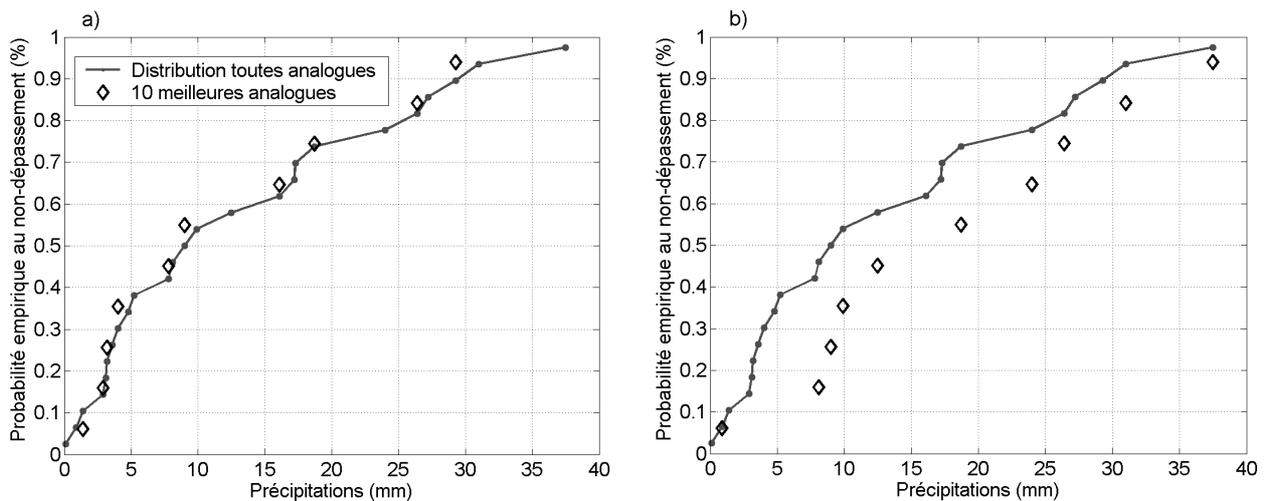


Figure 1.7 : Exemple de distribution empirique de l'ensemble des analogues, accompagnée de celle des dix meilleures analogues. a) Cas où la fiabilité est très correcte ; b) cas où la fiabilité est moins bonne.

Il est également important de savoir que l'on suppose l'équiprobabilité des analogues sélectionnées lors du calcul des quantiles d'une distribution de pluies. Ainsi la pluie observée lors de la situation la plus analogue a un poids identique à celle observée lors de la situation la moins analogue. Cette hypothèse a déjà été relaxée (Guilbaud, 1997), mais les résultats ne se sont pas révélés concluants, peut-être parce que l'hypothèse est globalement vérifiée sur les cas les plus courants, mais pas sur les cas les plus intéressants (*i.e.* les événements pluvieux les plus forts).

1.4 Principales limitations

A l'instar de tout système de prévision, il existe des limitations de performance inhérentes à l'adaptation statistique par analogie. Pour l'utilisateur ou le développeur d'un tel système, il est important de connaître les facteurs limitants et d'en tenir compte lors de l'utilisation et de l'interprétation des résultats.

1.4.1 Longueur des archives

La limitation principale provient de la longueur des archives utilisées. Les situations analogues sont recherchées sur la période commune aux périodes couvertes par les archives de lames d'eau et de variables météorologiques. Si cette période n'est pas assez longue, il sera alors plus difficile de trouver des situations suffisamment analogues, notamment pour les situations météorologiques les moins souvent rencontrées. Bien que les premières mesures météorologiques aient été faites il y a plusieurs centaines d'années, l'archivage régulier des données n'est pas si ancien. On ne dispose donc pas d'observations suffisantes sur une durée très longue (supérieure à 100 ans). D'autre part, la longueur de l'archive pluviométrique limite la performance de prévision des événements extrêmes, ce qui induit nécessairement un biais dans la prévision. En effet, plus l'événement pluvieux est important, plus la probabilité

de sélectionner des situations analogues ayant généré un événement de la même ampleur est faible. Si nous disposons d'une archive de 30 ans et que nous désirons anticiper un événement pluvieux décennal, nous ne pourrions sélectionner en théorie, *dans le meilleur des cas*, que quelques situations (au nombre de trois en théorie) dont la période de retour du cumul observé est de dix ans. La prévision d'un tel événement est donc automatiquement biaisée, puisque si nous retenons 30 analogues, toutes les valeurs des quantiles inférieurs au quantile d'ordre 90 % sont plus faibles que la valeur du cumul décennal. Il est tout de même possible de remédier à ce problème en employant une technique de correction de biais.

Pour que ce facteur limitant ait un impact le plus faible possible sur les performances de prévision, il est nécessaire, compte tenu de la longueur de la période candidate, de rechercher l'analogie sur un domaine spatial assez restreint. Ainsi si l'on dispose de plusieurs dizaines d'années candidates, une fenêtre géographique restreinte, de l'ordre de 20°x20°, semble suffisante pour trouver quelques situations suffisamment analogues (Bontron, 2004). En revanche, plus le domaine est étendu, moins on aura de chance d'en sélectionner (Van Den Dool, 1994).

1.4.2 Homogénéité des archives

Les archives nécessaires au fonctionnement d'une technique d'adaptation par analogie nécessitent d'être homogènes. De leur qualité dépend la qualité de la prévision. En effet, les changements intervenant dans les données assimilées (suppression de capteurs de mesure, déplacement de capteurs, remplacement de capteurs anciens par des appareils plus performants, ajout de nouveaux capteurs) peuvent notamment rendre les données difficilement comparables entre elles. C'est ainsi par exemple que plusieurs stations de radiosondage ont été supprimées au cours des années 1980, notamment dans l'Atlantique Nord.

Cette remarque est également valable pour l'archive pluviométrique. La suppression de postes pluviométriques peut entraîner une hétérogénéité dans les données. Avant de choisir nos archives, il est donc préférable de vérifier leur qualité, comme nous allons le faire dans le chapitre 4.

1.4.3 Robustesse et fiabilité de la prévision des variables d'analogie

Une méthode de prévision adaptée par analogie destinée à fonctionner de manière opérationnelle. Les champs des variables d'analogie étant prévus par un MNP, il est nécessaire de travailler avec des variables d'analogie robustes, dont la prévision doit dépendre le moins possible du MNP.

Or les structures des MNP opérationnels sont fréquemment modifiées, notamment en raison de la constante augmentation de la puissance de calcul. Ainsi par exemple, le modèle ARPEGE de Météo-France a été modifié tous les six mois entre 1992 et 1999 (Pailleux et al., 2000).

Si les variables d'analogie sont fortement dépendantes du MNP utilisé :

- les situations analogues sélectionnées seront en partie conditionnées par la capacité du modèle à prévoir ces variables ;
- l'utilisation d'une archive de variables météorologiques observées, produite par un MNP différent de celui qui a généré les champs de variables d'analogie prévus, pourra poser problème ;
- les changements fréquents du MNP peuvent avoir un impact sur la prévision des variables retenues, ce qui peut rendre délicate l'évaluation de la prévision sur une longue période dans un contexte opérationnel.

Il s'avère notamment que dans la méthode de référence, les variables du niveau 1 (*Z1000*, *Z500*) sont généralement mieux prévues par un MNP que les variables du niveau 2 (*TCW* et *RH850*). En effet, les premières sont calculées dans la partie « dynamique » du MNP, tandis que les secondes sont calculées dans la partie « physique », partie qui s'avère beaucoup plus dépendante du MNP (schémas numériques, paramétrisations, etc.). Afin de remédier au problème lié à la dégradation avec l'échéance de la qualité de prévision des variables, les deux niveaux de sélection ne sont appliqués que jusqu'à l'échéance $J+2$ dans le système RAINFAST implanté au LTHE. A partir de l'échéance $J+3$, la prévision n'est établie qu'avec la première sélection. En outre, on considère dans le système RAINFAST un nombre d'analogues variable, en fonction de l'échéance. En effet, il semble utile de retenir plus de situations analogues, lorsque l'échéance augmente, pour que les incertitudes soient représentatives de la dégradation de la fiabilité de la prévision (Thévenot, 2004 ; Hamill et al., 2006).

Il est donc nécessaire d'utiliser des variables peu dépendantes du MNP. Néanmoins, dans une perspective de prévision limitée à quelques jours d'échéance, il est possible d'exploiter d'autres variables explicatives des précipitations que les géopotentiels ou les variables d'humidité, tant que la fiabilité de la prévision reste satisfaisante.

1.4.4 Résolution

La performance d'une prévision par analogie est également limitée par la résolution des champs de variables d'analogie. Les phénomènes se produisant à une échelle inférieure à la résolution des champs ne peuvent être expliqués par des variables atmosphériques. Ainsi, la convection isolée à l'échelle du kilomètre ne peut être anticipée par une technique d'analogie puisqu'aucune archive ne contient les champs des variables à une résolution équivalente.

C'est la raison pour laquelle les précipitations convectives d'été ne sont généralement pas prévues correctement par une méthode d'analogie (Bliefernicht and Bardossy, 2008). Il en est de même pour la prévision de précipitations fournie par un MNP, même si les progrès récents ont permis d'augmenter sensiblement la résolution des modèles et donc leur capacité à simuler des phénomènes de plus petite échelle.

Par conséquent, les variables qui interviennent dans la recherche d'analogie doivent être décrites à une résolution au moins égale à celle des phénomènes que l'on veut prévoir.

1.5 Conclusion et objectif de la thèse

Lorsqu'un prévisionniste analyse une situation météorologique qu'il a déjà rencontrée auparavant, il se souvient souvent des effets que ces situations similaires ont eu en terme de temps sensible. Il peut alors estimer pour la situation prévue le risque qu'un événement se produise. Ainsi, il applique intuitivement le principe d'analogie lorsqu'il élabore une prévision. Après avoir présenté ce principe, nous avons expliqué dans ce chapitre comment des situations météorologiques analogues sont identifiées.

Historiquement, la technique d'analogie a d'abord été utilisée en tant que prévision statistique simple : une situation présente observée était supposée évoluer à peu près de la même façon qu'une situation passée observée considérée comme analogue. Grâce au progrès des MNP, le principe d'analogie a ensuite été appliqué en tant que technique d'adaptation statistique des sorties des MNP. Les modèles actuels fournissent une évolution temporelle des champs synoptiques des jours à venir à partir desquels on peut alors rechercher l'analogie avec des situations passées.

L'utilisation d'une méthode de prévision des précipitations adaptée par analogie, initiée en France dans les années 1970, s'est progressivement répandue à la fin du XXème siècle et continue à être largement utilisée par la communauté des hydrométéorologues. En France en particulier, les producteurs d'hydroélectricité notamment appliquent cette méthode de manière opérationnelle, en complément des techniques de modélisation numérique habituellement employées.

La dernière version de la méthode d'analogie mise au point par Bontron (2004) a été détaillée dans ce chapitre. Elle est constituée de deux niveaux de sélection d'analogues :

- le niveau 1 permet de rechercher l'analogie sur la circulation générale à échelle synoptique à l'aide de champs de géopotentiel et d'un critère d'analogie de forme (*TWS*) ;
- le niveau 2 permet de sélectionner, parmi les situations analogues retenues au niveau 1, que celles dont la situation hygrométrique est proche de celle prévue pour la journée cible à prévoir, à l'aide d'une variable d'humidité élaborée et d'un critère d'analogie de distance (*RMSE*).

Cette dernière version de la méthode, qui a été adaptée à des bassins à crues rapides du Sud-Est de la France, formera ainsi le point de départ de nos travaux. L'approche par analogie, bien qu'elle soit limitée par certains facteurs, notamment par la longueur et l'homogénéité des archives, a fait ses preuves ces dernières années, notamment grâce à l'amélioration de la prévision issue des MNP.

Nous proposons dans ces travaux de thèse d'améliorer cette technique par l'introduction de nouvelles variables explicatives des précipitations, et de l'adapter aux bassins tests à crues lentes de la Saône et de la Seine dans la perspective de montrer qu'il est possible, sur ce type de bassin, de repousser l'échéance de prévision des crues, actuellement limitée à quelques

jours.

2 Secteurs d'étude et données utilisées

Ce chapitre est dédié à la présentation des secteurs d'étude et des données de base exploitées. La première partie présente les deux secteurs choisis comme domaine d'application, correspondant à deux grands bassins versants français à enjeux. Nous poursuivrons par la description des archives requises pour appliquer une méthode de prévision des pluies par analogie, dont les caractéristiques (longueur, homogénéité, etc.) doivent être les mieux adaptées à ce type d'approche. Outre ces archives météorologiques, des mesures historiques de débit au droit des stations situées aux exutoires des bassins tests sont indispensables pour établir des prévisions de débit à partir d'un modèle hydrologique alimenté par des scénarios de précipitations générés par une méthode d'analogie (pour le calage du modèle et la vérification des prévisions). Les archives de variables d'analogie, les archives pluviométriques et l'archive hydrométrique seront décrites dans la seconde partie.

2.1 Zones d'étude

Deux bassins versants ont été retenus dans le cadre de cette thèse : le bassin de la Saône à Lyon et celui de la Seine à Paris. Sur ces deux bassins, nous avons bénéficié d'une collaboration avec les services opérationnels effectuant des prévisions de débit pour des besoins internes ou pour l'annonce de crue. Sur le bassin du Rhône et en particulier en aval de la Saône, la CNR gère un parc de centrales hydroélectriques et effectue une prévision des débits tout au long de l'année. Le SPC Rhône amont-Saône est en charge, sur le bassin de la Saône, de la gestion d'un réseau hydrométrique pour le suivi hydrologique et l'alimentation du dispositif Vigicrues¹. Sur le bassin de la Seine, deux SPC se partagent la surveillance de l'évolution des crues : le SPC Seine moyenne-Yonne-Loing et le SPC Seine amont-Marne amont.

2.1.1 Le bassin versant de la Saône à Lyon

La rivière Saône prend sa source à Vioménil dans les Vosges, dans le versant sud des monts Faucilles, à environ 400 m d'altitude. La Saône parcourt 482 km avant de se jeter dans le

¹ L'adresse Internet est : <http://www.vigicrues.ecologie.gouv.fr>

Rhône. A sa confluence avec le Rhône à La Mulatière (Lyon), la Saône draine une importante superficie, de l'ordre de 30 000 km², répartie sur cinq régions (Lorraine, Champagne-Ardennes, Franche-Comté, Bourgogne et Rhône-Alpes). L'influence climatique est diverse : en amont du bassin, elle est plutôt continentale ; tandis qu'en aval, depuis Auxonne (Côte d'Or) jusqu'à l'exutoire du bassin, l'influence est plutôt méridionale. Le dénivelé n'est que de 245 m, et il existe une nappe puissante qui joue le rôle de régulateur. Ces deux éléments expliquent en grande partie la lenteur de l'écoulement de la Saône. Cela n'empêche pas que le débit de la rivière puisse être important, comme nous le verrons plus tard dans le chapitre 4 (section 4.2.2).

Le bassin versant a été divisé en trois sous-bassins aux caractéristiques différentes (Figure 2.1) : le Doubs, la Petite Saône et la Saône Intermédiaire.

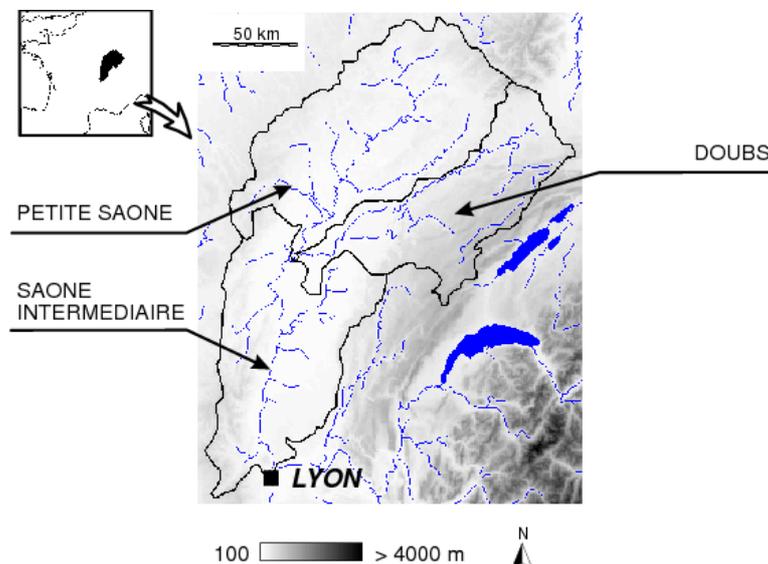


Figure 2.1 : Localisation géographique des trois sous-bassins versants de la Saône, avec indication du réseau hydrographique et de la topographie.

Le bassin du Doubs est le plus montagneux, tandis que les bassins de la Petite Saône et de la Saône Intermédiaire sont vallonnés, avec des altitudes maximales inférieures à celle atteinte sur le bassin du Doubs. L'altitude maximale dépasse néanmoins 1000 m sur le Doubs ainsi que sur la Petite Saône, en raison des massifs montagneux des Vosges (Petite Saône) et du Jura (Doubs). Les caractéristiques géographiques des sous-bassins versants sont reportés dans le Tableau 2.1. Les données hypsométriques ont été obtenues par l'intermédiaire d'un modèle numérique de terrain à maille kilométrique - source Global 30 Arc Second Elevation Data Set, U.S. Geological Survey (Sauquet, 2006).

En outre, il est bon de savoir qu'il n'y a pas d'ouvrages hydrauliques importants sur la Saône. Il s'agit principalement d'ouvrages au fil de l'eau avec de faibles capacités de stockage. Les débits mesurés aux exutoires des bassins seront donc considérés comme naturels par la suite.

Tableau 2.1 : Caractéristiques géographiques des sous-bassins versants de la Saône.

Sous-bassin versant	Superficie (km ²)	Altitude (m)		
		Moyenne	Minimale	Maximale
Doubs	7670	550	170	1400
Petite Saône	12030	320	170	1150
Saône Intermédiaire	9860	290	160	900

2.1.2 Le bassin versant de la Seine à Paris

Le fleuve Seine prend sa source sur le plateau de Langres dans le département de la Côte d'Or, à une altitude de 471 mètres. Il arrose successivement la Champagne, reçoit l'Aube (rive droite) et l'Yonne (rive gauche) à Montereau, baigne Melun et traverse l'Île-de-France, grossit du Loing, de l'Essonne, de l'Orge (rive gauche) et de l'Yerres (rive droite), de la Marne et de l'Oise (rive droite), de part et d'autre de Paris. Le bassin de la Seine en amont de Paris s'étend sur 43 800 km², et le territoire est couvert par trois régions (Bourgogne, Champagne-Ardenne et Ile-de-France). Les bassins de la Seine et de la Saône sont donc mitoyens. L'influence climatique est semi-océanique, avec la pré-dominance d'un régime d'ouest, qui apporte régulièrement de l'humidité en provenance de l'Océan Atlantique. Son relief, d'origine sédimentaire, est peu accidenté avec des altitudes généralement inférieures à 300 m, dépassant rarement 500 m, sauf dans le Morvan où elles culminent à 900 m.

Le bassin versant de la Seine en amont de Paris a été divisé en cinq sous-bassins (Figure 2.2) : le Loing à Episy (Loing-E), la Marne à La Ferté-sous-Jouarre (Marne-F), la Seine à Bazoches-lès-Bray (Seine-B), l'Yonne à Courlon-sur-Yonne (Yonne-C) et la Seine Intermédiaire (Seine-Int), délimité par Paris et les quatre exutoires des autres sous-bassins.

Les bassins les plus vallonnés sont situés à l'est (Yonne-C, Marne-F et Seine-B), tandis qu'en aval le relief est peu marqué, avec une altitude moyenne de l'ordre de 150 à 200 m. Seul le bassin de l'Yonne-C se démarque des autres par une topographie plus marquée, notamment vers le Morvan. Ainsi d'après le modèle de terrain, près de 10% de la surface de ce bassin se situe à plus de 450 m d'altitude. On peut donc s'attendre à des réactions hydrologiques plus fortes en amont du bassin. En revanche en aval, le faible relief explique probablement la lente propagation des eaux dans le réseau hydrographique. L'ensemble des caractéristiques géographiques des sous-bassins versants est reporté dans le Tableau 2.2.

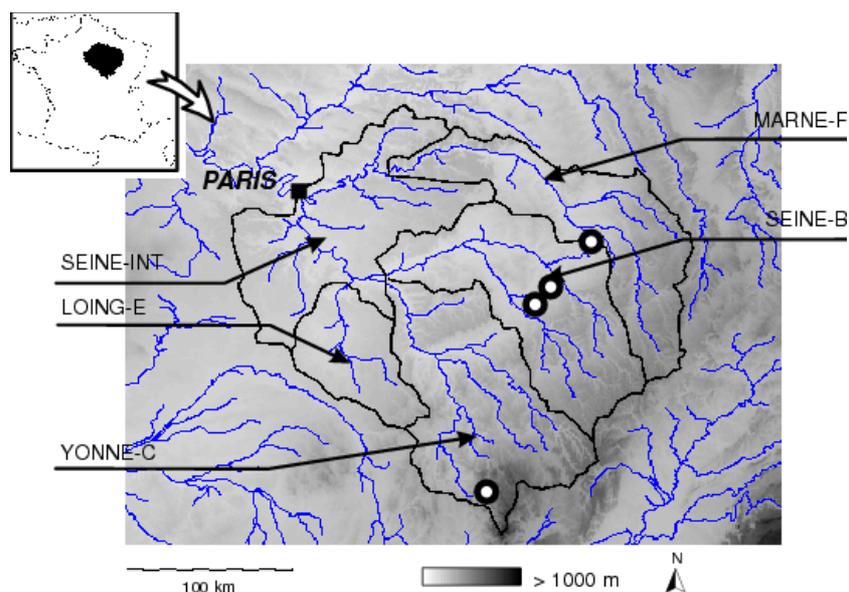


Figure 2.2 : Localisation géographique des cinq sous-bassins versants de la Seine, avec indication du réseau hydrographique, de la topographie et des quatre lacs réservoirs (cercles noirs et blancs).

Tableau 2.2 : Caractéristiques géographiques des sous-bassins versants de la Seine.

Sous-bassin versant	Superficie (km ²)	Altitude (m)		
		Moyenne	Minimale	Maximale
Loing-E	3900	148	52	362
Marne-F	8820	234	54	500
Seine-B	10100	206	51	527
Yonne-C	10700	262	53	882
Seine-Int	10280	122	31	242

Contrairement au bassin de la Saône, il existe quatre lacs-réservoirs sur le bassin de la Seine (Figure 2.2), dont les principaux objectifs sont de réduire les risques dus aux inondations (réduction des débits de crue en aval des aménagements) et d'apporter un soutien des étiages en été et en automne¹. Les lacs-réservoirs se situent sur la Marne, l'Aube, la Seine et l'Yonne.

1 Toutes les informations sur les lacs-réservoirs de la Seine sont disponibles sur le site : <http://www.iibrbs.fr>

2.2 Données disponibles

Nous avons vu, au chapitre 1, qu'il est nécessaire de disposer de longues archives météorologiques pour appliquer une technique d'adaptation par analogie : une archive pluviométrique et une archive de variables d'analogie. De plus, notre objectif final est d'utiliser des prévisions de pluie pour émettre des prévisions de débit. Pour cela, nous utiliserons un modèle hydrologique qu'il faudra préalablement caler. Nous devons donc disposer d'une archive de débits observés au droit des stations situées à l'exutoire de chaque sous-bassin versant, qui servira également à évaluer la prévision hydrologique. Nous allons à présent décrire ces bases de données.

2.2.1 Archives de réanalyses météorologiques

Les développements récents des méthodes de prévision par analogie ont bénéficié de la disponibilité de nouvelles archives : les réanalyses météorologiques. A l'origine, l'élaboration de ce type de données fut motivée par la prise de conscience du changement climatique. Les climatologues ont exprimé le besoin de données fiables et continues dans le temps afin d'étudier ce changement. Une méthode a dû être mise au point pour satisfaire ces scientifiques. La première étape consiste à rassembler toutes les données disponibles sur le globe depuis le début des relevés (généralement dans les années 1940). La seconde étape consiste à construire un système d'assimilation de données, couplé à un MNP de référence, capable de reproduire de façon homogène les champs de variables météorologiques sur une grille régulière et à un pas de temps régulier. Ainsi, les données sont interpolées en respectant une cohérence spatiale mais aussi une cohérence temporelle, assurées par un MNP. De cette manière, toutes les données d'observation disponibles durant la période de réanalyses sont utilisées.

C'est en 1991 que la première série de réanalyses fut initiée par les centres américains NCEP (National Center for Environmental Prediction) et NCAR (National Center for Atmospheric Research). En 1993, le centre européen CEPMMT (Centre Européen de Prévisions Météorologiques à Moyen Terme) a également créé sa propre archive de réanalyses.

Les réanalyses météorologiques ont plusieurs avantages à être utilisées : longueur d'archive, grande variété de variables, données disponibles sur une grille régulière à un pas de temps régulier, etc. Le seul inconvénient est leur dépendance aux systèmes d'observation. Bien que les données assimilées soient les plus nombreuses possibles, ces mesures ne sont pas parfaitement homogènes dans le temps en raison de la modification des réseaux de mesure (suppression de stations de radiosondages, stations météorologiques automatiques ou manuelles), mais aussi de l'apport de nouvelles mesures qui n'étaient pas disponibles en début de période de réanalyse, comme les données satellitaires. Cet inconvénient ne remet pas pour autant en cause leur usage pour la prévision adaptée par analogie. Les avantages à les utiliser sont bien plus nombreux que les inconvénients !

Dans le cadre de notre travail, nous avons rapatrié des variables extraites des réanalyses américaines (NCEP/NCAR) et européennes (ERA-40). Certaines de ces variables sont définies dans la seconde partie de ce mémoire (section 5.1). Nous disposons ainsi de deux bases de données différentes, que nous allons maintenant présenter.

2.2.1.1 Les réanalyses NCEP/NCAR

Historique

Le projet réalisé en 1991 par les centres NCEP et NCAR couvrait initialement la période 1957-1996. Depuis la fin des années 1990, il s'est étendu sur la période 1948 à l'année en cours. Le système d'assimilation, qualifié de CDAS (Climate Data Assimilation System), réanalyse presque en temps réel les données météorologiques. Ce système est resté identique sur l'ensemble de la période pour éviter de bruyier les éventuels signes de changements climatiques, déjà affectés par les changements dans les systèmes d'observation. Une nouvelle archive américaine a été créée en 1998 pour subvenir aux besoins d'une communauté de scientifiques dans le cadre du projet AMIP-II (Atmospheric Model Intercomparison Project). Cette archive, nommée NCEP-DOE AMIP II, ne couvre que les années 1979 à nos jours. Elle a été produite par un système d'assimilation amélioré, avec quelques données supplémentaires, permettant d'éviter certaines erreurs identifiées dans la première archive (Kanamitsu et al., 2002). Cependant la durée couverte par cette archive n'est pas suffisante pour notre travail. Nous retiendrons donc la première archive (1948-2008), que nous nommerons par la suite l'archive NCEP/NCAR.

Description

Kistler et al. (2001) ont détaillé le contenu de l'archive NCEP/NCAR. Les données comprennent non seulement des données réanalysées sur une grille régulière, mais aussi des prévisions de huit jours établies et archivées tous les cinq jours ainsi que toutes les données d'observation atmosphériques codées au format conventionnel international BUFR (Binary Universal Format Representation).

Les réanalyses sont produites par le modèle météorologique opérationnel du NCEP tel qu'il était en 1995 (Kalnay et al., 1996). C'est un modèle spectral global avec 28 niveaux verticaux et une troncature T62 (représentation sphérique harmonique de 62 ondes), ce qui correspond à une résolution horizontale d'environ 210 km. Le schéma d'analyse est un schéma 3DVAR dans l'espace spectral¹. Les données assimilées sont nombreuses : elles proviennent de sondes atmosphériques, de satellites divers, de mesures aéroportées, d'observations sur les océans et d'observations terrestres. Les systèmes d'observation ont cependant évolué au cours du temps. Nous pouvons classer cette évolution en trois phases (Kistler et al., 2001) : de 1948 à 1957, période correspondant au début des observations en haute troposphère; de 1958 à

¹ La représentation spectrale des champs est une méthode de numérisation couramment utilisée en modélisation météorologique. Elle est réservée à la discrétisation des dimensions spatiales, en particulier sur l'horizontale.

1978, correspondant au réseau moderne de sondages; et à partir de 1979, année durant laquelle les premières données satellitaires ont été utilisées.

Les données sont réanalysées, sur l'ensemble du globe, sur une grille de $2,5^\circ \times 2,5^\circ$ toutes les 6 h (03h, 09h, 15h et 21h TU de 1948 à mai 1957 puis 00h, 06 h, 12 h et 18 h TU à partir de juin 1957). Cette différence provient du fait qu'avant le 1er juin 1957, les observations en haute troposphère n'étaient faites que 3 h après les heures synoptiques conventionnelles.

Les variables disponibles sont nombreuses et ont été classées en trois types (A, B et C), selon leur degré de dépendance aux observations ou au modèle (Bontron, 2004). Ces variables sont définies sur des niveaux de pression, des niveaux isentropiques (température potentielle constante), en surface (niveau de la mer) et au niveau de la tropopause. On peut les rapatrier directement à partir du serveur FTP¹ de la NOAA (National Oceanic and Atmospheric Administration), ou commander les CD-Rom annuels conçus par le centre NCAR.

Données rapatriées

Les données dont nous disposons pour notre travail proviennent de CD-Roms mais aussi du serveur FTP. Nous avons choisi le format netCDF (network Common Data Form) pour constituer notre archive. Ce format, privilégié par la communauté des climatologues, permet de compresser les données, d'utiliser un fichier compatible à tous les systèmes d'exploitation (quel que soit le codage des nombres), et permet à un utilisateur d'identifier clairement le contenu du fichier et son organisation grâce à l'en-tête CDL (Common Data form Language).

Notre archive NCEP/NCAR est constituée des variables d'analogie utilisées dans la méthode de référence A0, mais aussi d'autres variables susceptibles d'être exploitées par la suite (Tableau 2.3).

Tableau 2.3 : Variables définies à résolution $2,5^\circ$ rapatriées pour la construction de l'archive NCEP/NCAR.

Variable [unité]	Niveaux de pression [hPa]					
	1000	925	850	700	500	300
Géopotentiel [m]	X				X	
Humidité relative [%]			X			
Vitesse du vent zonal [m/s]			X		X	
Vitesse du vent méridien [m/s]			X		X	
Température de l'air [K]			X		X	
Quantité d'eau précipitable [mm]	Intégrée sur la colonne troposphérique					

Toutes ces variables sont définies aux niveaux de pression qui nous intéressent, sauf la quantité d'eau précipitable qui correspond à une intégration sur la colonne d'air troposphérique. Ces variables ont été extraites :

- sur un vaste domaine ($-60^\circ\text{E}/60^\circ\text{E}$; $15^\circ\text{N}/75^\circ\text{N}$), qui recouvre une grande partie de l'Océan Atlantique, l'Europe et le nord de l'Afrique (Figure 2.3) ;

¹ L'adresse du serveur FTP est : <ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis/>.

- sur toute la période disponible (1948-2006) et à tous les instants de mesure (00h, 06h, 12h et 18h TU).

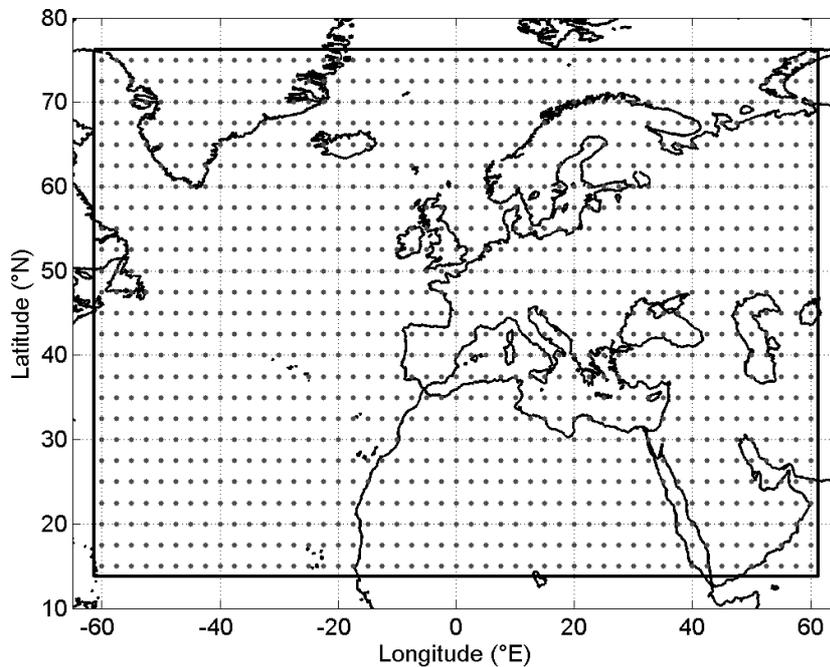


Figure 2.3 : Domaine d'extraction des variables d'analogie de l'archive NCEP/NCAR à résolution 2,5° (60°E/60°E ; 15°N/75°N) et grille correspondante. Les points correspondent aux centres des mailles carrées.

2.2.1.2 Les réanalyses ERA-40

Historique

En 1993, le CEPMMT a construit l'archive de réanalyses ERA-15 qui couvre la période 1979-1993 (Gibson et al., 1997). Dès 1998, le projet a été développé par la création de l'archive ERA-40, qui couvre la période 1957-2002 (Uppala et al., 2005). La construction de cette archive avait pour but de produire et d'inciter à utiliser une gamme complète d'analyses à l'échelle globale, qui décrivent l'état de l'atmosphère, du sol terrestre et de l'océan, sur une période de 45 ans (septembre 1957 à août 2002). Le principal objectif était de promouvoir la recherche européenne et internationale en effectuant le plus possible d'observations, d'analyses et de rapports d'études.

Description

Comme celles de l'archive américaine, les données de l'archive ERA-40 regroupent des analyses mais aussi des prévisions (Kallberg et al., 2004). Le modèle atmosphérique utilisé est le modèle IFS CY23r4, qui fait partie du système de prévision du CEPMMT. C'est un modèle spectral global avec 60 niveaux verticaux, de troncature T159 (représentation sphérique harmonique de 159 ondes), ce qui correspond à une résolution horizontale d'environ 125 km

sur la grille gaussienne. D'après Uppala et al. (2005), le schéma du système d'assimilation est un schéma 3DVAR amélioré, le schéma 4DVAR étant à l'origine trop coûteux en temps de calcul. Comme pour l'archive NCEP/NCAR, les données d'observation proviennent de pays et d'organismes du monde entier. Il s'agit d'observations au sol, de radiosondages, de profileurs, de mesures aéroportées et de données satellitaires.

Les données sont réanalysées, sur l'ensemble du globe, sur une grille gaussienne réduite de résolution comprise entre $1,121^\circ \times 1,121^\circ$ à l'équateur et $1,125^\circ \times 1,125^\circ$ dans les tropiques. Les analyses sont faites toutes les 6 h (00 h, 06 h, 12 h et 18 h TU), tandis que des prévisions à court et moyen terme sont faites à des heures variables.

Les variables disponibles sont également très nombreuses. Néanmoins, seules certaines d'entre elles sont mises à disposition d'organismes de recherche, et ce uniquement à la résolution de $2,5^\circ \times 2,5^\circ$. Il est important de savoir que ces données à $2,5^\circ$ proviennent de celles générées à $1,125^\circ$, qui ont été dégradées. Par ailleurs, l'ensemble des données à haute résolution est stocké dans le système d'archive MARS (Meteorological Archive and Retrieval System) du centre européen, système qui n'est accessible qu'à certains utilisateurs, généralement autorisés par les centres de prévision des différents pays de l'Union Européenne.

Les données nécessaires à notre étude ont été rapatriées dans un premier temps via le site ECMWF Data Services¹, à résolution $2,5^\circ$. Dans un second temps, nous avons eu accès aux données à résolution $1,125^\circ$ via la base MARS. De la même manière que pour l'archive NCEP/NCAR, l'archive ERA-40 a été construite au format netCDF. Le domaine d'extraction des données à $2,5^\circ$ est identique au domaine de la Figure 2.3, tandis que le domaine sur lequel nous avons extrait les données à $1,125^\circ$, représenté sur la Figure 2.4, est moins étendu en raison du volume très conséquent que ces données peuvent représenter. Les données ont été extraites sur la période 1957-2002 aux instants de mesure 00h, 06h, 12h et 18h TU.

1 L'adresse du site Internet est : http://data.ecmwf.int/data/d/era40_daily/.

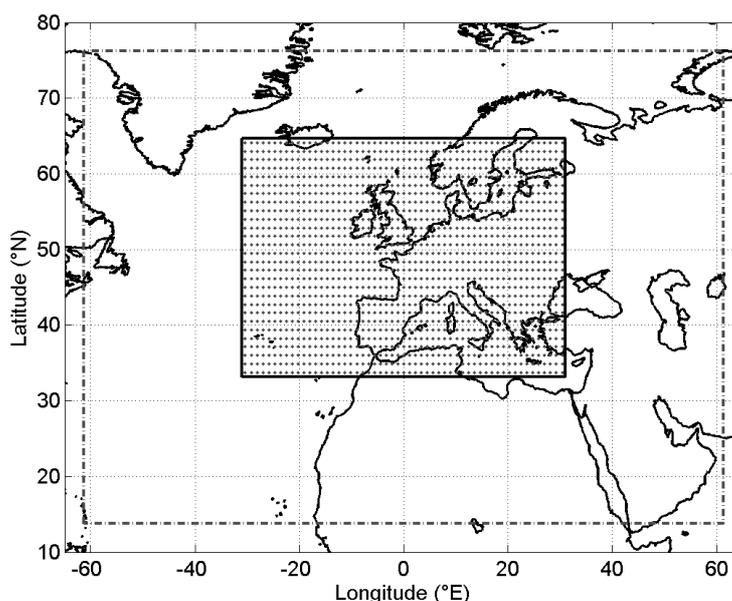


Figure 2.4 : Domaine d'extraction des variables d'analogie de l'archive ERA-40 à résolution 1,125° (-30,375°E/30,375°E ; 33,75°N/64,125°N) et grille correspondante. Les points correspondent aux centres des mailles carrées. Le domaine d'extraction des données à résolution 2,5° est indiqué en traits discontinus.

Les mêmes variables d'analogie que celles de l'archive NCEP/NCAR ont été rapatriées, mais notre archive ERA-40 est plus riche (cf. Tableau 2.4).

Tableau 2.4 : Variables définies aux résolutions 2,5° et 1,125° rapatriées pour la construction de l'archive ERA-40. Les variables représentées en gras sont communes aux archives ERA-40 et NCEP/NCAR.

Variable [unité]	Niveaux de pression [hPa]								
	1000	925	850	775	700	600	500	400	300
Géopotential [m]	X	X	X	X	X	X	X	X	X
Humidité relative [%]		X	X	X	X				
Vitesse du vent zonal [m/s]			X	X	X	X	X	X	
Vitesse du vent méridien [m/s]			X	X	X	X	X	X	
Température de l'air [K]		X	X	X	X	X	X		
Tourbillon potentiel [K.m ² /(kg.s)]						X	X	X	X
Vitesse verticale [Pa/s]		X	X	X	X	X	X		
Divergence horizontale [/s]	X	X	X		X				
Géopotential [m]	Au niveau du tourbillon potentiel égal à 2PVU								
Précipitations totales [mm]	Estimées au sol par le modèle								
Quantité d'eau précipitable [mm]	Intégrée sur la colonne troposphérique								

Nous disposons des mêmes variables sur un nombre de niveaux de pression élargi, et de nouvelles variables qui seront définies au chapitre 5 : tourbillon potentiel, vitesse verticale, divergence horizontale et géopotential au niveau de tourbillon potentiel égal à 2PVU et précipitations totales (donnée sol).

2.2.1.3 Nouvelle génération de réanalyses

La puissance du calcul informatique s'est considérablement développée ces dernières années, ce qui a rendu possible l'utilisation d'un modèle 4DVAR (Uppala, 2006 ; Simmons et al., 2006). Ce modèle a été testé et validé en opérationnel depuis 1997.

Une nouvelle génération de réanalyses météorologiques, nommées ERA-Interim, a ainsi vu le jour en 2009 après plusieurs années de travail sur le projet (Berrisford et al., 2009). Les objectifs de ce projet étaient d'améliorer certains aspects des réanalyses ERA-40, comme la représentation du cycle de l'eau, la qualité de la circulation stratosphérique, et le traitement des biais et changements du système d'observation.

La période d'analyse débute en 1989, et se poursuit jusqu'à nos jours. Le système d'assimilation fonctionne donc en temps réel (CDAS). Le modèle atmosphérique et le système de réanalyse dérive du système de prévision intégré (IFS) du CEPMMT. Le modèle a une troncature T255 (représentation sphérique harmonique de 255 ondes) et discrétise l'atmosphère sur 60 niveaux verticaux, le dernier niveau de pression étant 0,1 hPa. De nouveaux schémas numériques, paramétrisations et corrections de biais ont été inclus dans ce système nouvelle génération.

ERA-Interim est, comme son nom l'indique, une archive temporaire. En effet, le CEPMMT a prévu de créer l'archive ERA-extended, dont les premiers travaux devraient débuter en 2010. Cette archive devrait alors couvrir la période 1940 à nos jours. Cette version tiendra compte de nouvelles données d'observation et de nouvelles méthodes d'assimilation et de modélisation.

Jusqu'en fin de thèse, l'archive ERA-Interim n'était pas encore disponible et n'a donc pas pu être exploitée. Néanmoins, la période actuellement couverte serait trop courte pour que cette base de données puisse être utilisée dans le cadre du développement d'une méthode d'analogie.

2.2.2 Archives pluviométriques

Nous disposons de deux archives contenant des relevés historiques de lames d'eau : l'archive construite par Météo-France (MF) pour les besoins de la CNR, baptisée « MF/CNR » ; et la réanalyse atmosphérique SAFRAN. La première ne contient que des observations sur le bassin de la Saône, tandis que la seconde en contient sur les deux bassins tests retenus dans le cadre de la thèse.

2.2.2.1 L'archive MF/CNR

Il est par nature difficile de se procurer des valeurs de lame d'eau observée moyenne sur les bassins versants : la donnée de pluie en chaque point du secteur géographique est, en effet, inaccessible. Le moyen de s'approcher des lames d'eau est de reconstituer des champs pour chaque pas de temps grâce à des méthodes d'interpolation alimentées par des observations ponctuelles avant de les agréger sur le secteur d'intérêt.

De cette manière, une archive pluviométrique a été élaborée par Météo-France pour les besoins de la CNR. Cette archive contient des lames d'eau au pas de temps de 6h sur les 46 sous-bassins du bassin du Rhône. Elle a été construite par une méthode d'interpolation de données d'observation (Pelletier, 2004).

Les données utilisées proviennent de trois réseaux de mesures de Météo-France :

- i) le réseau de stations humaines (R1), qui a peu évolué en dix ans et qui représente un à deux points de mesure par département ;
- ii) le réseau de stations automatiques (R2), qui a fortement évolué en 10 ans et qui représente actuellement moins de 10 points de mesure par département ;
- iii) le réseau d'observateurs (R3), qui représente 30 à 70 points de mesure par département.

Les réseaux R1 et R2 fournissent des lames d'eau horaires alors que le réseau R3 ne fournit que des lames d'eau journalières, que les observateurs transmettent en temps différé. Pour reconstituer les lames d'eau au pas de temps de 6 h, Météo-France a utilisé les données des trois réseaux.

Cette reconstitution est effectuée à partir de lames d'eau « brutes » au pas de temps de 6 h (données de R1 et R2) et au pas de temps de 24 h (données de R1, R2 et R3). Une méthode de krigeage a été utilisée pour interpoler les données sur une grille régulière de maille 5 km. Afin de mieux représenter les bassins versants du Rhône de petite taille, des fonctions splines ont été exploitées pour obtenir les données sur une grille de maille 2,5 km. Les lames d'eau ont ensuite été ré-aggrégées par sous-bassin. En raison de la faible densité du réseau du réseau R2 au début de la période, une correction des lames d'eau 6 h est apparue nécessaire. Les lames d'eau ont ainsi été corrigées lorsqu'elles répondent à deux critères :

- elles sont supérieures à 5 mm ont d'abord été sélectionnées ;
- l'écart relatif entre les lames d'eau 24 h calculées à partir de quatre lames d'eau 6 h (basées sur les seuls réseaux R1 et R2) et les lames d'eau 24 h calculées directement à partir des trois réseaux (R1, R2 et R3) est supérieur à 10 %.

La correction a été effectuée en considérant la répartition temporelle obtenue par tranches de six heures et en appliquant cette répartition aux lames d'eau 24 h, qui sont *a priori* plus représentatives du cumul journalier d'un point de vue quantitatif. Les corrections ont tout de même été de moins en moins fréquentes avec le temps grâce à la densification rapide du réseau R2.

Dans le cadre de nos travaux, nous avons ainsi à disposition, uniquement pour chaque sous-bassin de la Saône, une archive de lames d'eau moyennes, au pas de temps de 6 h, sur la période 01/01/1992-31/12/2003.

2.2.2.2 L'archive SAFRAN

Le Centre d'Etude de la Neige de Météo France a développé, au cours des années 1990, le

système SAFRAN (Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige) sur les principaux massifs alpins et pyrénéens (Durand et al., 1993, Martin, 1995, Durand et al., 1999) pour répondre en premier lieu au besoin de la prévision opérationnelle des avalanches. SAFRAN assimile les données provenant de divers réseaux de mesures (SYNOP, NIVO-METEO), les ébauches du modèle ARPEGE et la climatologie des précipitations. Il fournit les principales données météorologiques horaires affectant le manteau neigeux dont les précipitations, selon l'altitude et l'exposition, pour chaque massif montagneux considéré comme climatologiquement homogène. La répartition horaire des cumuls journaliers de précipitations est effectuée suivant l'évolution temporelle de l'humidité spécifique du modèle ARPEGE de Météo-France (Déqué et al., 1994). Désormais nommé SAFRAN/Avalanches (SAFRAN/A), ce système est intégré dans la chaîne SAFRAN - CROCUS - MEPRA, utilisée pour la prévision opérationnelle des avalanches sur les massifs français depuis l'hiver 1992.

Plus récemment, les procédures développées pour SAFRAN/A ont été généralisées au territoire français (SAFRAN/F) (Quintana-Segui et al., 2008). Les données ne sont plus fournies par massif/altitude/exposition, mais sur une grille régulière de résolution 8 x 8 km² et permettent de reconstituer les forçages météorologiques depuis 1958. Elles servent entre autre à alimenter le modèle hydrogéologique MODCOU couplé au modèle de transfert sol - végétation - atmosphère ISBA (Habets et al., 2008). Les procédures au cœur de SAFRAN ont été réappliquées *a posteriori* sur de longues périodes, de 1958 à 2008 pour la version SAFRAN/F (Vidal et al., 2009) et de 1958 à 2002 pour SAFRAN/A (Durand et al., 2009). Le lecteur est invité à consulter les publications de Quintana-Segui et al. (2008) et de Vidal et al. (2009) pour une description détaillée du système d'analyse SAFRAN.

Pour notre travail, qui ne concerne pas les Alpes, nous avons utilisé la base de données SAFRAN/F, que nous appellerons l'archive SAFRAN par la suite, pour construire des lames d'eau de bassin historiques sur les bassins tests. En début de thèse, nous ne disposions des données SAFRAN que sur la période 01/08/1972-31/07/2006. Nous avons ensuite eu accès à des données supplémentaires, la période finalement couverte à la fin de la thèse s'étendant du 01/08/1970 au 31/07/2008.

Les lames d'eau fournies par SAFRAN sur la France sont disponibles au pas de temps horaire sur des mailles carrées de huit kilomètres de côté. Le contour des bassins versants obtenu à partir du modèle numérique de terrain kilométrique (Sauquet, 2006) a été plaqué sur la grille de restitution de SAFRAN. Les précipitations horaires de bassin ont ensuite été estimées par intégration spatiale sur le secteur drainé, puis finalement cumulées de façon à produire trois archives de lames d'eau aux pas de temps de 6 h, 12 h et 24 h. Nous disposons ainsi de lames d'eau de bassin sur chaque sous-bassin de la Saône et de la Seine.

2.2.2.3 Transformation des précipitations

Les sous-bassins que nous considérons ne sont pas tous soumis aux mêmes régimes

pluviométriques, comme nous le verrons dans le chapitre 4 (section 4.2.1). Un cumul identique n'est donc pas forcément observé avec la même fréquence sur nos différents bassins. En outre, la lame d'eau est une variable dont la distribution est fortement dissymétrique. Bornée en 0, la lame d'eau n'a pas de limite supérieure. Si nous évaluons une méthode de prévision de manière globale, comme nous le ferons dans la phase de développement, alors les écarts sur les fortes valeurs auront beaucoup plus de poids que les autres dans le score global. Afin d'effacer ces effets, nous avons repris la procédure de Bontron (2004) qui consiste à transformer la lame d'eau journalière brute selon l'équation 2.1 :

$$P = \sqrt{\frac{P_{brute}}{PJXA10}} \quad (2.1)$$

où P est la lame d'eau transformée, P_{brute} est la lame d'eau brute et $PJXA10$ est la lame d'eau décennale, c'est-à-dire le quantile de pluie journalière maximale annuelle de période de retour 10 ans, obtenu par ajustement statistique.

Nous utiliserons donc ces lames d'eau transformées lorsque nous développerons la méthode d'analogie dans la seconde partie de ce mémoire.

2.2.3 Archive hydrométrique

Nous appliquerons dans la troisième partie les méthodes d'analogie développées dans la seconde partie pour générer des prévisions probabilistes de débit. Il est donc nécessaire, afin de caler le modèle hydrologique, de disposer d'une archive contenant les mesures historiques de débit moyen journalier au droit des stations situées aux exutoires de chaque sous-bassin : pour la Saône il s'agit de Le Châtelet [Saône], Neublans [Doubs] et Couzon [Saône] ; et pour la Seine il s'agit d'Episy [Loing], La Ferté [Marne], Bazoches [Seine], Courlon [Yonne] et Paris [Seine].

Les chroniques de débits journaliers ont été extraites de la banque de données HYDRO¹ (Banque Nationale de Données pour l'Hydrométrie et l'Hydrologie), pour toutes les stations à l'exception de celle de Couzon. La base de données HYDRO est alimentée par différents organismes (Direction Régionale de l'Environnement, de l'Aménagement et du Logement (DREAL), SPCs, EDF, CNR, etc.) et gérée par le SCHAPI. Dans notre cas, les données sont transmises par la Diren Ile-de-France pour les stations du bassin de la Seine, par la Dreal Bourgogne pour la station de Le Châtelet et par la Dreal Franche-Comté pour celle de Neublans. La chronique de débit de la Saône à Couzon a été transmise par la CNR, gestionnaire de la station.

Les périodes sur lesquelles les mesures de débit sont disponibles sont discontinues pour certaines stations. Nous avons décidé d'extraire toutes les données possibles pour chaque station. Compte tenu des nombreuses lacunes présentes dans les chroniques, nous avons

¹ L'adresse du site Internet est : <http://www.hydro.eaufrance.fr/>.

considéré un mois comme complet s'il comporte au moins 75 % des données et une année hydrologique comme *complète* s'il y a au plus un mois de lacunes. La période couverte (en fin de thèse) et le nombre d'années *complètes* sont reportés dans le Tableau 2.5 pour chaque station. Ainsi, le nombre d'années complètes est très élevé pour la station de Couzon, pour laquelle on dispose de données dès 1920. A l'opposé, les données pour la station de La Ferté ne sont disponibles qu'à partir de 1993. Bien que les premières données datent de 1964, les lacunes sont très nombreuses pour la station de Bazoches¹, avec seulement 21 années complètes (contre 44 s'il n'y avait aucune donnée manquante).

Nous utiliserons dans la section suivante toutes les données disponibles pour analyser le régime hydrologique de chaque sous-bassin. En revanche, la période sur laquelle nous utiliserons un modèle hydrologique dans la partie III s'étire du 01/08/1970 au 30/08/2002. Sur cette période, les données de débit manquantes sont très nombreuses pour les stations de La Ferté et de Bazoches, avec une proportion de lacunes respectivement de 71,3 % et 68,9 % sur cette période restreinte.

Tableau 2.5 : Périodes de disponibilité des données de débit et nombres d'années hydrologiques complètes pour les huit stations hydrométriques de la Saône et de la Seine, dont les codes de station HYDRO sont également indiqués, à l'exception de Couzon pour laquelle nous lui avons associé un code fictif (en italique).

Station [cours d'eau]	Code HYDRO	Période disponible	Nombre d'années complètes
Le Châtelet [Saône]	U1420010	01/01/1965-01/11/2009	44
Neublans [Doubs]	U2722010	24/08/1966-01/11/2009	43
Couzon [Saône]	<i>U3000000</i>	01/01/1920-31/05/2008	87
Episy [Loing]	H3621010	01/09/1964-01/11/2009	45
La Ferté [Marne]	H5321010	01/01/1993-01/11/2009	14
Bazoches [Seine]	H1940010/H1940020	01/09/1964-01/11/2009	21
Paris [Seine]	H5920010	01/01/1974-01/11/2009	35
Courlon [Yonne]	H2721010	01/01/1964-01/11/2009	45

¹ La station de Bazoches a été modifiée au cours du temps. Il s'agissait d'une station à une échelle au barrage de la Grande Bosse jusqu'en 1978 (code HYDRO : H1940010) ; puis une nouvelle station à ultrason a été installée en 1990 à quelques kilomètres du barrage (code HYDRO : H1940020).

3 Outils d'évaluation et contextes d'élaboration d'une prévision

Dans notre travail, nous serons amenés à évaluer un système de prévision afin d'en mesurer les performances. Puisqu'une série de prévisions ne peut se juger de manière subjective, il est nécessaire de s'appuyer sur des scores pour évaluer les performances de manière objective et systématique. Nous utiliserons par la suite plusieurs scores d'évaluation d'une prévision probabiliste que nous allons définir dans ce chapitre.

La première section s'intéresse aux scores d'évaluation globale d'une prévision. Ces derniers seront notamment utilisés dans la seconde partie lorsque nous développerons la méthode de sélection d'analogues. La seconde section s'emploie à présenter des scores d'évaluation en fonction d'un seuil de précipitations ou de débit, qui seront utilisés dans la troisième partie de ce manuscrit. Deux méthodes permettant d'identifier le biais d'une prévision probabiliste seront présentées dans la troisième section. Enfin dans la dernière partie de ce chapitre, nous définirons les contextes dans lesquels les prévisions de précipitations seront évaluées : le contexte de prévision parfaite et celui de prévision opérationnelle.

3.1 Evaluation globale

L'évaluation globale permet de connaître la performance d'une méthode de prévision pour tous les types d'événement pluvieux. Ceci nous permet d'avoir une vision globale du gain de qualité et d'utilité apporté par une nouvelle méthode. Nous avons notamment retenu un unique score pour l'évaluation globale, ce qui lève certaines ambiguïtés liées à l'usage de plusieurs scores de vérification. En effet, lorsque l'évaluation repose sur plusieurs scores, il est possible que tous ne mènent pas à la même conclusion.

De cette manière, nous utiliserons un seul score pour déterminer le gain d'utilité d'une méthode de prévision par rapport à une autre. Le lecteur pourra se référer aux travaux de Bontron (2004) ou à l'ouvrage de Jolliffe and Stephenson (2003) pour une revue des différents scores d'évaluation.

3.1.1 Choix du score CRPS

Nous avons retenu le score utilisé par Bontron (2004) pour évaluer une prévision probabiliste : le *CRPS* (Continuous Ranked Probability Score). Ce score dérive du *RPS* (Ranked Probability Score), utilisé par Guilbaud (1997), qui est un score de vérification d'une prévision par catégories ordonnées. Pour la prévision d'une variable continue, comme c'est le cas pour les précipitations, le score *CRPS* (Brown, 1974 ; Matheson and Winkler, 1976) est défini par l'équation 3.1 :

$$CRPS = \int_{-\infty}^{+\infty} [F(x) - H_{x_{obs}^0}(x)]^2 \cdot dx \quad (3.1)$$

Où: x_{obs}^0 est la valeur observée ;
 F est la fonction de répartition prévue de la variable x ;
 $H_{x_{obs}^0}$ est la fonction indicatrice, telle que $H_{x_{obs}^0}(x) = \begin{cases} 0 & \text{si } x < x_{obs}^0 \\ 1 & \text{si } x \geq x_{obs}^0 \end{cases}$.

Ce score rend compte de l'utilité d'une prévision pour un ensemble d'utilisateurs, c'est-à-dire qu'il mesure le coût effectif de la prévision lorsque la valeur observée s'écarte de la valeur que l'ensemble d'utilisateurs avait retenue (cf. Bontron, 2004). Il présente l'avantage, pour une prévision probabiliste, de tenir compte de la finesse (dispersion) et de la justesse (décalage entre prévision et observation) de la prévision (Hersbach, 2000) (équation 3.2). Ces deux parties, définies par les équations 3.3 et 3.4, permettent de quantifier l'utilité d'une prévision.

$$CRPS = Finesse + Justesse \quad (3.2)$$

$$Finesse = \int_{-\infty}^{+\infty} [F(x) - H_{Q_{50\%}}(x)]^2 \cdot dx \quad (3.3)$$

$$Justesse = \int_{x_{obs}^0}^{Q_{50\%}} (1 - 2 \cdot F(x)) \cdot dx \quad (3.4)$$

Où: $Q_{50\%}$ est le quantile au non-dépassement d'ordre 50 % de la fonction de répartition F ;
 $H_{Q_{50\%}}$ est la fonction indicatrice, telle que $H_{Q_{50\%}}(x) = \begin{cases} 0 & \text{si } x < Q_{50\%} \\ 1 & \text{si } x \geq Q_{50\%} \end{cases}$.

La Figure 3.1 permet d'interpréter graphiquement le score *CRPS* et sa décomposition en finesse et en justesse. Sur la Figure 3.1 a), l'aire hachurée correspond à la valeur du *CRPS* pour une prévision. Sur la Figure 3.1 b) (resp. Figure 3.1 c)), l'aire hachurée correspond à la valeur de la finesse (resp. justesse) du *CRPS*. Ces aires doivent être les plus faibles possibles,

le CRPS parfait étant égal à 0.

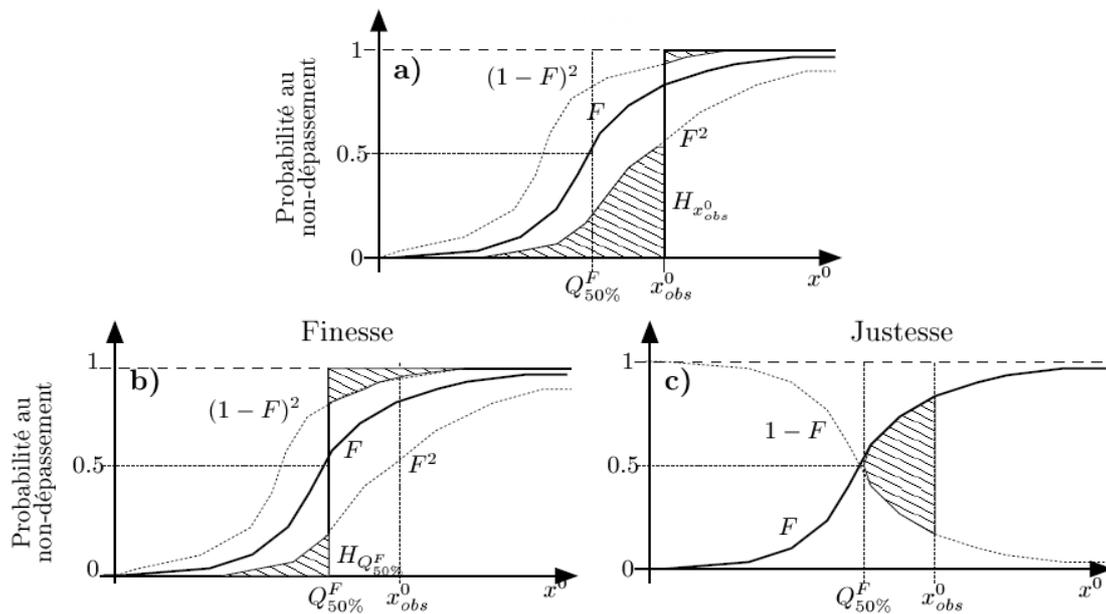


Figure 3.1 : Interprétation graphique a) du CRPS et de sa décomposition en b) finesse et c) justesse. Les aires hachurées doivent être les plus faibles possibles (d'après Bontron (2004)).

La Figure 3.2 présente quatre exemples types de distribution prévue. Une prévision probabiliste peut être a) parfaitement juste (centrée en la valeur finalement observée), mais peu fine (la plage de pluie pouvant être observée est grande) ; b) fine (peu d'incertitude sur la valeur prévue), mais elle peut être peu juste (valeur médiane éloignée de la valeur finalement observée) ; c) peu fine et peu juste à la fois ; d) très fine et très juste. Une prévision probabiliste ne présentant pas, par nature, une finesse et une justesse parfaites, il est donc nécessaire de trouver un équilibre entre les deux.

Notons que dans le cadre de cette thèse, les distributions prévues se seront pas ajustées à une loi statistique. Le score CRPS et ses composantes seront donc évalués directement sur les distributions empiriques des précipitations.

Par ailleurs, pour évaluer la performance d'un système de prévision, nous considérons une série de prévisions. Le score global correspond alors à la moyenne des scores obtenus pour chaque prévision. Lorsque nous développerons la méthode d'analogie, nous évaluerons donc la valeur moyenne du CRPS sur la période de prévision étudiée, et ce pour chaque sous-bassin.

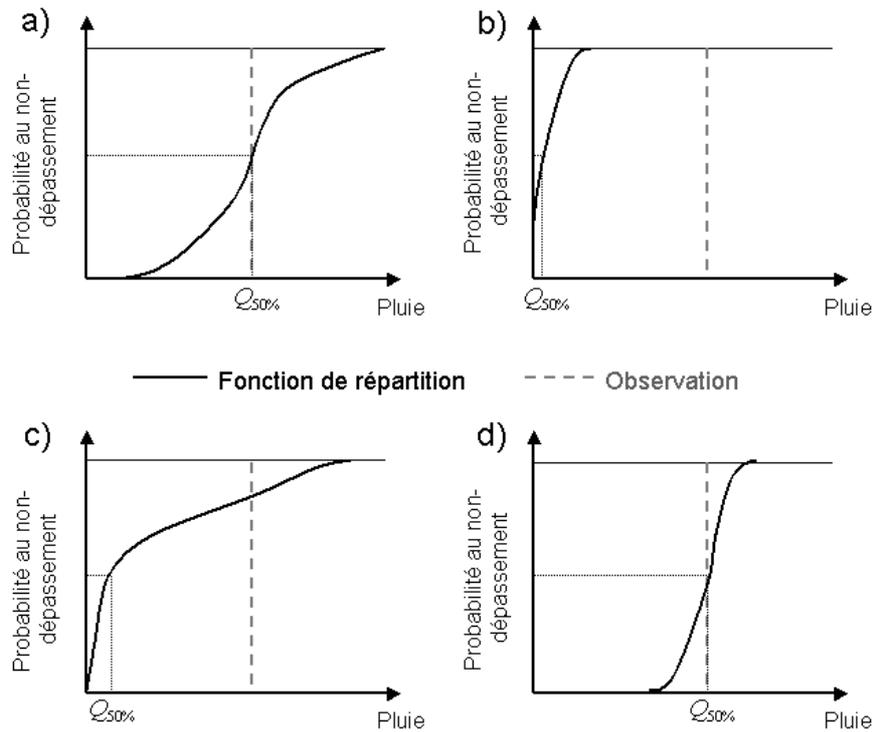


Figure 3.2 : Exemples de fonction de répartition prévue. a) Prévision parfaitement juste mais peu fine ; b) prévision peu juste mais très fine ; c) prévision peu juste et peu fine ; d) prévision très juste et très fine.

3.1.2 Compétence CRPSS du CRPS

Il est fréquent d'exprimer un score en compétence par rapport à une prévision de référence. Cela permet de comparer la performance d'une méthode, mesurée par le score, à la performance d'une prévision de référence généralement naïve. Cela donne ainsi plus de signification aux valeurs numériques. La compétence du CRPS d'une méthode M de prévision est mesurée par le score CRPSS (Continuous Ranked Probability Skill Score, Hamill and Whitaker, 2007), défini par l'équation 3.5 :

$$CRPSS_M = \frac{CRPS_M - CRPS_{ref}}{CRPS_{parfait} - CRPS_{ref}} \quad (3.5)$$

Où :

- $CRPS_M$ est la valeur moyenne du CRPS obtenue par la méthode M évaluée ;
- $CRPS_{ref}$ est la valeur moyenne du CRPS obtenue par une prévision de référence ;
- $CRPS_{parfait}$ est la valeur du CRPS obtenue pour une prévision parfaite (égale à 0).

Ce nouveau score CRPSS est égal à 1 pour une prévision parfaite, égal à 0 pour une prévision équivalente à la prévision de référence, et est négatif pour une prévision moins bonne que la prévision de référence. Il est donc sans dimension.

La prévision de référence communément choisie est la prévision climatologique. Dans ce cas, la prévision probabiliste des précipitations correspond à la fonction de répartition de

toutes les précipitations recueillies sur le bassin. D'autres prévisions de référence peuvent également être utilisées, telles que la persistance ou le tirage aléatoire. La persistance prévoit pour les jours suivants un cumul pluviométrique identique à celui mesuré la veille du jour de prévision. Notons que cette référence est souvent utilisée pour la prévision de débit. Le tirage aléatoire peut également être considéré comme prévision de référence (Wetterhall et al., 2005 ; Diomedea et al., 2008). Dans ce cas, la prévision émise pour chaque échéance est constituée de valeurs de cumuls tirés aléatoire dans la distribution climatologique. Dans notre cas, nous avons choisi arbitrairement la climatologie comme prévision de référence. Pour respecter l'effet de la saisonnalité pris en compte dans la méthode de référence A0, nous avons choisi de construire la prévision climatologique, pour un jour cible, en considérant seulement les cumuls pluviométriques observés dans une fenêtre temporelle de plus ou moins deux mois autour du jour cible.

Par ailleurs, sauf avis contraire, les performances seront exprimées pour un ensemble de sous-bassins appartenant à un même bassin. Le score *CRPSS* résultant sera évalué pour l'ensemble des sous-bassins en moyennant les scores *CRPSS* obtenus pour chaque sous-bassin, selon l'équation 3.6 :

$$CRPSS = \frac{1}{N_{sbv}} \sum_{i_{sbv}=1}^{N_{sbv}} CRPSS_{sbv}(i_{sbv}) \quad (3.6)$$

où N_{sbv} est le nombre de sous-bassins versants et $CRPSS_{sbv}(i_{sbv})$ est le score *CRPSS* évalué pour le i_{sbv} ème sous-bassin.

3.1.3 Gain *G* de performance

Lorsque nous développerons la méthode de prévision par analogie, nous construirons des algorithmes intermédiaires dont nous chercherons à évaluer les performances. Afin de comparer la performance obtenue par une méthode *B* à celle obtenue par une méthode *A* antérieure, nous utiliserons le score *G* défini par le gain de performance de la méthode *B* par rapport à la méthode *A* (équation 3.7). Le gain *G* n'est rien d'autre que la compétence de la compétence du *CRPS* d'une méthode *B* par rapport à une méthode *A*. Cela revient finalement à calculer la compétence du *CRPS* d'une méthode *B* par rapport à une prévision de référence définie par la méthode *A*.

$$G(B/A) = \frac{CRPSS_B - CRPSS_A}{CRPSS_{parfait} - CRPSS_A} \quad (3.7)$$

Ainsi dans la seconde partie de ce mémoire, nous utiliserons généralement ce score pour évaluer le gain de performance apporté par une nouvelle méthode par rapport à une version antérieure.

3.2 Evaluation en fonction d'un seuil

Si nous nous plaçons en tant qu'utilisateur d'un système de prévision, nous sommes certes intéressés de savoir que notre système est globalement plus performant qu'un autre, mais nous ne savons pas s'il permet de bien anticiper un type particulier d'événement. Dans ce contexte, un score d'évaluation globale tel que le *CRPS* n'est pas adapté. En revanche, on s'intéresse à des scores spécifiques à un type d'événement. Dans le cas des précipitations, on peut fixer des seuils de cumuls pluviométriques pour lesquels on évalue la capacité de notre système à bien prévoir le dépassement ou le non-dépassement de ces seuils. Pour un seuil S donné, les prévisions générées et les observations associées sont classées dans les événements réalisés ou dans les événements non réalisés. De cette manière, on remplit un tableau de contingence (Tableau 3.1) en comptant le nombre d'événements :

- A) observés et correctement prévus (vraie alerte) ;
- B) non observés mais prévus (fausse alerte) ;
- C) observés mais non prévus (défaut d'alerte) ;
- D) non observés et non prévus.

Tableau 3.1 : Tableau de contingence valable pour un seuil S de précipitations.

	$P_{observée} > S$	$P_{observée} \leq S$
$P_{prévue} > S$	A	B
$P_{prévue} \leq S$	C	D

Une méthode de prévision est d'autant meilleure que les nombres B et C sont faibles. Ainsi, une méthode parfaite ne génère aucune fausse alerte ni aucun défaut d'alerte. Il existe de nombreux scores de vérification qui peuvent être calculés à partir d'un tableau de contingence (Jolliffe and Stephenson, 2003). Nous ne les listerons pas ici, mais nous proposons quelques scores très utiles pour apprécier la qualité d'une méthode de prévision.

3.2.1 Probabilité de détection (*POD*)

Le score *POD* (Probability Of Detection), défini par l'équation 3.8, permet de connaître la probabilité que le seuil S soit dépassé par une lame d'eau prévue, sachant que la lame d'eau observée dépasse effectivement le seuil. Ce score, très sensible à la fréquence climatologique de l'événement, tient notamment compte des vraies alertes mais ignore totalement les fausses alertes. Le score *POD* est généralement analysé conjointement avec le score *FAR* que nous définirons dans cette section. Il est également à la base du diagramme ROC que nous présenterons un peu plus loin. Le score *POD* est compris entre 0 et 1, la valeur du score parfait étant 1.

$$POD = \frac{A}{A+C} \quad (3.8)$$

3.2.2 Taux de fausses détections (*POFD*)

Le score *POFD* (Probability Of False Detection), défini par l'équation 3.9, permet de connaître la probabilité que le seuil *S* soit dépassé par une lame d'eau prévue, sachant que la lame d'eau observée ne dépasse pas le seuil. Ce score est sensible aux fausses alertes mais ignore totalement les alertes manquées. C'est également un composant essentiel du diagramme ROC. Le score *POFD* est compris entre 0 et 1, la valeur du score parfait étant 0.

$$POFD = \frac{B}{B+D} \quad (3.9)$$

3.2.3 Taux de fausses alertes (*FAR*)

Le score *FAR* (False Alarm Rate), défini par l'équation 3.10, permet de connaître la probabilité, lorsque le seuil *S* est dépassé par une lame d'eau prévue, que la lame d'eau observée ne dépasse pas ce seuil. Ce score, très sensible à la fréquence climatologique de l'événement, tient compte des fausses alertes mais ignore les alertes manquées. Le score *FAR* est compris entre 0 et 1, la valeur du score parfait étant 0.

$$FAR = \frac{B}{A+B} \quad (3.10)$$

Il existe cependant des confusions concernant la définition du taux de fausses alertes, comme en a parlé Golding (2000). En météorologie, le taux de fausses alertes s'exprime du point de vue de l'observation, c'est-à-dire qu'il évalue le rapport entre le nombre de fausses alertes (*B*) et le nombre total de fois où le seuil *S* n'a pas été dépassé en réalité (*B+D*). Ce score en météorologie correspond au taux de fausses détections (*POFD*) en médecine, tel que nous l'avons défini précédemment. En médecine en revanche, le taux de fausses alertes s'exprime du point de vue de la prévision, c'est-à-dire qu'il évalue le rapport entre le nombre de fausses alertes (*B*) et le nombre total d'alertes (*A+B*). C'est donc cette définition que nous avons employé ici. Il faut tout de même avouer que la définition utilisée par la médecine a plus de sens : en effet dans le contexte d'évaluation d'un système de prévision, on cherche bien à minimiser le nombre de fausses alertes par rapport au nombre total d'alertes annoncées, *i.e.* à minimiser le taux de fausses alertes.

3.2.4 Score d'Hansen-Kuipers (*TSS*)

Le score *TSS* (True Skill Statistic, ou score de Hansen-Kuipers), défini par l'équation 3.11, tient compte à la fois des vraies alertes (*POD*) et des alertes manquées (*POFD*), et permet

d'évaluer plus globalement une méthode de prévision pour un type d'événement donné. Il utilise tous les éléments d'un tableau de contingence, et tient compte de l'influence de la climatologie de l'événement ciblé sur la réalisation de bonnes prévisions du simple fait du hasard. Ce score, peu sensible à la fréquence climatologique de l'événement, ne tient pas compte du biais éventuel de la prévision. Il est donc moins adapté aux événements rares. Le score *TSS* est compris entre -1 et 1, une valeur nulle indiquant aucune compétence, la valeur du score parfait étant égale à 1.

$$TSS = \frac{A}{A+C} - \frac{B}{B+D} = POD - POFD \quad (3.11)$$

3.2.5 Diagramme ROC et score AROC

Le diagramme ROC (Relative Operating Characteristic) est utilisé en médecine mais aussi depuis quelques années en météorologie (Stanski et al., 1989 ; Golding, 2000). Ce diagramme a l'avantage de présenter plusieurs caractéristiques utiles pour la comparaison entre deux méthodes de prévision. Ce diagramme trace, pour un seuil *S* fixé, la probabilité de détection (*POD*) en fonction du taux de fausses détections (*POFD*). Les scores *POD* et *POFD* sont donc évalués pour différentes fréquences au non-dépassement. Pour évaluer une méthode d'analogie, il est également possible de calculer ces scores pour chacune des valeurs de la distribution empirique. Ainsi, si l'on retient finalement *N* analogues, le diagramme sera constitué de *N* points.

Le diagramme ROC n'est pas sensible au biais de la prévision, en revanche il permet de mesurer l'utilité d'une méthode de prévision. Si la courbe est une droite qui se confond avec la première bissectrice, alors la prévision évaluée est équivalente à la climatologie ; tandis que la prévision est parfaite si la courbe se superpose à l'angle supérieur gauche (Figure 3.3).

Dans ce cas, pour toutes les fréquences étudiées, la probabilité de détection est de 1 et le taux de fausses détections est de 0. En outre, les diagonales tangentes à la courbe ROC en chaque point représentent les isovaleurs de *TSS*.

Il est généralement utile de condenser l'information contenue dans un diagramme ROC par l'aire évaluée sous la courbe ROC (partie hachurée de la Figure 3.3), que nous nommerons le score *AROC* par la suite. Ce score est compris entre 0 et 1, une valeur de 1 correspondant à une prévision parfaite et une valeur de 0,5 indiquant aucune compétence par rapport à la climatologie.

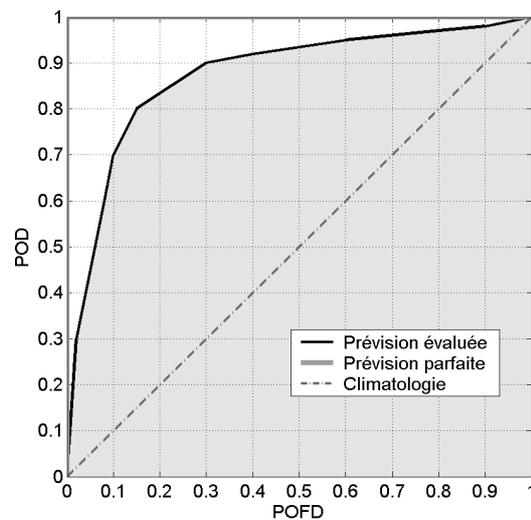


Figure 3.3 : Exemple de diagramme ROC pour une prévision évaluée (courbe pleine noire). La partie grisée représente l'aire sous la courbe ROC.

3.3 Identification du biais de la prévision

Nous avons retenu deux éléments permettant d'identifier et d'analyser le biais que peut contenir une prévision probabiliste : le diagramme de Talagrand et le diagramme de fiabilité.

3.3.1 Diagramme de Talagrand

Le diagramme de Talagrand (Anderson, 1996 ; Talagrand et al., 1997), également appelé histogramme des rangs, permet d'évaluer la fiabilité d'une méthode de prévision probabiliste. En d'autres termes, il permet de vérifier si la pluie observée est en moyenne incluse dans l'ensemble des scénarios, et par conséquent de détecter les défauts systématiques du système de prévision. Il est plus largement utilisé dans le cadre de la vérification des prévisions d'ensemble (Hamill, 2001). L'inconvénient majeur est qu'il ne tient pas compte des pluies nulles.

Pour un bassin versant, le diagramme est construit de la manière suivante. Pour chaque jour J de prévision de la période étudiée, au cours duquel la pluie observée est non nulle, les cumuls de précipitation observés lors des situations analogues ainsi que le cumul observé le jour J sont classés par ordre croissant. On note ensuite le numéro du rang du cumul observé. On obtient ainsi, sur la période, une fréquence pour chaque rang. Le diagramme est finalement construit en traçant, pour chaque rang, la fréquence associée.

La Figure 3.4 illustre quatre exemples types de diagramme, obtenu pour une prévision probabiliste constituée de 10 membres. L'interprétation des diagrammes est la suivante :

- a) forme plate : la prévision n'est pas nécessairement correcte, mais les scénarios probabilistes représentent parfaitement la distribution empirique des observations ;
- b) forme en « U » : la prévision probabiliste est trop fine et n'est pas assez juste, les

observations étant bien souvent situées en dehors des scénarios extrêmes ;

- c) forme en dôme (ou « n ») : la prévision probabiliste n'est pas assez fine, la plupart des observations étant comprises entre les scénarios extrêmes, et bien souvent proches du scénario médian ;
- d) forme asymétrique : la prévision probabiliste est biaisée, ce qui se traduit par une sous-estimation ou une sur-estimation quasi-systématique des précipitations observées.

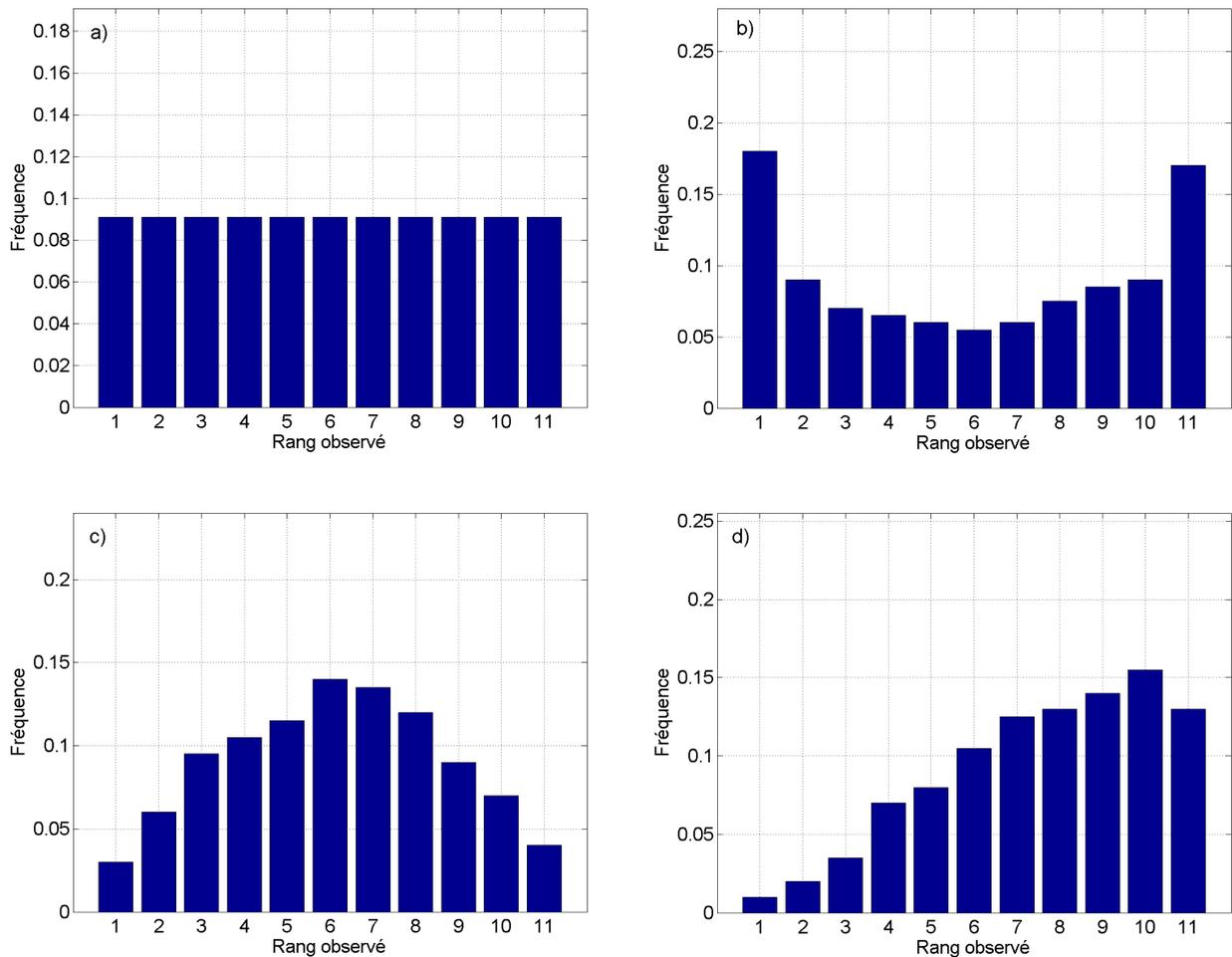


Figure 3.4 : Exemples de diagramme de Talagrand.

Par conséquent, nous pouvons imaginer que le diagramme d'un système de prévision parfait serait représenté par un histogramme concentré sur le rang médian.

3.3.2 Diagramme de fiabilité

Ce diagramme permet, pour un événement donné, de vérifier si la probabilité prévue que cet événement se produise correspond à sa fréquence observée (Wilks, 1995 ; Hamill, 1997). Cela consiste finalement à tracer des fréquences relatives observées de l'événement F_{obs} en fonction des probabilités prévues P_{emp} . On associe généralement au diagramme

l'histogramme des fréquences d'utilisation des probabilités prévues.

Il se construit de la manière suivante. Considérons une série de N prévisions probabilistes, dont les distributions sont constituées de P points, et intéressons-nous aux événements pluvieux dont la lame d'eau observée est non nulle. Il est alors possible de calculer la fréquence observée F_{obs} de ces événements pour chacune des probabilités prévues P_{emp} :

$$P_{emp}(i) = \frac{i}{P} \quad \text{et} \quad F_{obs}(i) = \frac{N_{obs}(i)}{N_{emp}(i)} \quad (3.12)$$

où $N_{obs}(i)$ et $N_{emp}(i)$ sont évalués sur la série de prévisions pour le i ème rang de l'ensemble des P points. Pour chaque jour J de la série de N prévisions, on procède comme suit :

- iv) si la pluie observée le jour J est nulle, et que 10 des P valeurs de la distribution sont non nulles, alors pour $P_{emp}(10) = \frac{10}{P}$, $N_{obs}(10)$ reste identique et $N_{emp}(10) = N_{emp}(10) + 1$;
- v) si la pluie observée le jour J est non nulle, et que 5 des P valeurs de la distribution sont non nulles, alors pour $P_{emp}(5) = \frac{5}{P}$, $N_{obs}(5) = N_{obs}(5) + 1$ et $N_{emp}(5) = N_{emp}(5) + 1$.

Finalement sur la série de N prévisions, pour chacun des P rangs, la fréquence relative observée F_{obs} est déduite et la probabilité prévue P_{emp} lui est associée. Ceci permet alors de tracer les fréquences observées en fonction des probabilités prévues (cf. Rousset-Regimbeau (2007) pour une description détaillée du principe de construction du diagramme).

L'interprétation du diagramme est la suivante (Figure 3.5). La fiabilité (de l'anglais *reliability*) de la prévision, équivalente à la notion de justesse introduite précédemment, est indiquée par la proximité de la courbe à la diagonale. L'écart à la diagonale représente ainsi le *biais conditionnel* de la prévision. Dans le cas d'un système de prévision probabiliste, si la courbe est en-dessous (resp. au-dessus) de la diagonale, alors la prévision tend à sur-estimer (resp. sous-estimer) les précipitations. Dans l'exemple de la Figure 3.5, la prévision sous-estime les précipitations observées pour les faibles probabilités (inférieures à 0,25), tandis qu'elle les sur-estime pour les probabilités moyennes à fortes (supérieures à 0,5). En outre, plus la courbe est plate, moins la prévision a de résolution (équivalente à la notion de finesse). Ainsi la prévision climatologique, qui ne permet pas de discriminer les situations pour lesquelles l'événement a eu lieu de celles pour lesquelles il n'a pas eu lieu, n'a aucune résolution. Ce diagramme est par conséquent conditionné par les prévisions, et traduit ainsi le vrai sens de la prévision. Il est complémentaire du diagramme ROC, dans la mesure où celui-ci est conditionné par les observations.

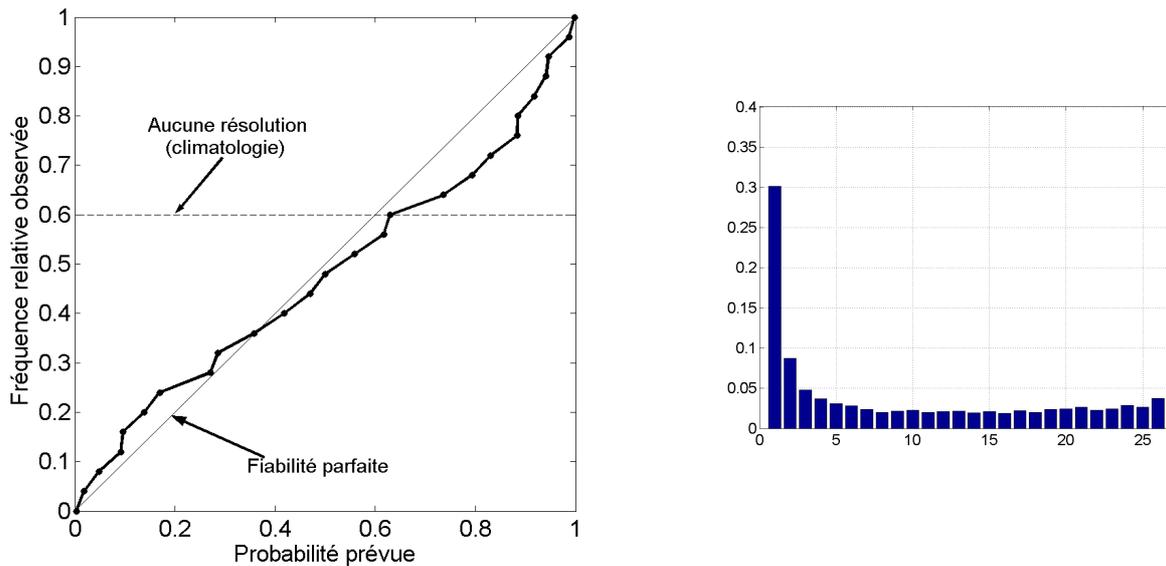


Figure 3.5 : Exemple de diagramme de fiabilité accompagné de l'histogramme des fréquences observées, pour une prévision probabiliste formée de 25 individus.

3.4 Contextes d'élaboration d'une prévision adaptée par analogie

Les scores présentés précédemment seront évalués dans deux contextes distincts : le premier correspond à un contexte de simulation (*prévision parfaite*), tandis que le second correspond à un contexte de prévision réelle (*prévision opérationnelle*).

Le développement et l'optimisation d'une méthode ne doivent pas être influencés par les biais des MNP en prévision ; les choix doivent être argumentés à partir de la meilleure connaissance des situations météorologiques dont nous disposons. C'est pourquoi nous nous placerons dans un premier temps dans un contexte de *prévision parfaite* (de l'anglais *perfect prognosis*), afin que nos conclusions ne soient pas influencées par les erreurs de prévision des variables d'analogie. Ainsi les situations cibles pour lesquelles une prévision est émise ne sont pas des situations prévues, mais sont des situations directement extraites des réanalyses.

Une fois l'algorithme développé, nous pourrions alors évaluer ses performances en conditions réelles, c'est-à-dire dans un contexte de *prévision opérationnelle*. Dans ce cas, les situations cibles pour lesquelles on souhaite émettre une prévision de précipitations ne sont pas observées ou réanalysées, mais directement produites par un MNP.

4 Analyse exploratoire des bases de données

Nous avons présenté dans le chapitre précédent les différents jeux de données dont nous disposons dans le cadre de la thèse. Cependant, avant d'entamer les travaux de développement de la méthode d'analogie, convient de se poser des questions sur ces données : les différentes archives sont-elles cohérentes entre elles ? Lesquelles allons nous utiliser ?

Les archives pluviométriques seront considérées dans un premier temps, ce qui nous conduira ensuite à analyser les régimes pluviométriques et hydrologiques auxquels sont soumis les bassins tests.

Dans un second temps, notre attention se portera sur les archives météorologiques, en focalisant plus particulièrement sur les valeurs brutes des variables d'analogie utilisées dans la méthode de référence. La mise en évidence d'écarts nous amènera à étudier la sensibilité de la méthode au choix de l'archive de réanalyses dans la dernière section.

Cette analyse exploratoire nous permettra ainsi de déterminer les archives qui seront utilisées dans la suite de ce mémoire, et d'en connaître la qualité.

4.1 Comparaison et critique des archives pluviométriques

4.1.1 Objectif et méthodologie

Dans le cadre du développement de la méthode de prévision par analogie, nous disposons des deux archives pluviométriques, SAFRAN et MF/CNR, présentées dans le chapitre 2. Cependant, ces deux archives ne couvrent pas la même période : l'archive MF/CNR ne couvre que la période 1992-2003 (12 ans), tandis que l'archive SAFRAN couvre la période 1972-2005 (33 ans). Il serait donc plus judicieux d'utiliser SAFRAN puisque la période de situations candidates, dans laquelle les situations analogues sont sélectionnées, est bien plus longue que celle couverte par l'archive MF/CNR. De plus, puisque notre objectif est également d'appliquer une méthode d'analogie au bassin de la Seine, il est nécessaire de choisir une

archive contenant les mesures historiques de précipitations sur ce bassin test. Dans notre cas, seule l'archive SAFRAN en contient, ce qui plaide pour son utilisation.

Néanmoins, les réanalyses pluviométriques SAFRAN n'ont pas directement été produites à partir de données de précipitations observées seules (cf. section 2.2.2.2). Les lames d'eau journalières sont effectivement issues d'observations seules et d'une ébauche constante constituée d'un gradient avec l'altitude), mais ensuite c'est un modèle météorologique qui permet de conserver une cohérence spatiale et temporelle lors de la décomposition horaire des données, selon le principe de la réanalyse. Avant d'utiliser l'archive SAFRAN, nous nous sommes donc interrogés sur la qualité de ses données de précipitations. Or celles de l'archive MF/CNR proviennent uniquement de données observées sur un réseau dense. Malgré les corrections effectuées sur ses lames d'eau (cf. section 2.2.2.1), l'archive MF/CNR sera considérée dans cette section comme une « référence » : nous proposons de comparer les deux bases de données par l'analyse de leur cohérence (pour le bassin de la Saône), et ainsi de critiquer les précipitations extraites de l'archive SAFRAN.

Les données de l'archive MF/CNR ont été validées par Pelletier (2004). La cohérence des lames d'eau a été contrôlée au pas de temps annuel, pour chaque sous-bassin de la Saône, à l'aide de données de mesures pluviométriques de quelques stations du réseau de Météo-France, prises comme référence. Cette validation a été effectuée avec les lames d'eau 6 h corrigées et avec les lames d'eau 24 h. D'autre part, les données de l'archive SAFRAN ont également été validées par plusieurs auteurs. Quintana-Segui et al. (2008) ont évalué la qualité d'un panel de variables météorologiques (température, humidité, vitesse du vent, radiation incidente et précipitations) fournies par SAFRAN sur la France par comparaison avec des observations provenant des réseaux de mesure de Météo France. Vidal et al. (2009) ont également effectué une validation des réanalyses SAFRAN à l'aide de données dépendantes et indépendantes. Ils ont montré que les différences entre les données analysées et les données observées étaient faibles et relativement constantes sur la période.

Nous proposons à présent d'étudier la cohérence entre les deux archives pluviométriques pour différents pas de temps d'intégration (6 h, 12 h et 24 h), sur la période commune aux archives (01/01/1992-31/12/2003). Pour chaque pas de temps, les échantillons de pluies de bassin SAFRAN ont été comparés avec les lames d'eau MF/CNR correspondantes au travers d'une analyse de corrélation.

4.1.2 Résultats

Dans un premier temps, nous avons considéré toutes les observations disponibles sur la période. Nous avons calculé le coefficient de détermination R^2 et tracé le nuage de points, la droite de régression linéaire et la première bissectrice. La Figure 4.1 illustre les corrélations entre les lames d'eau MF/CNR et SAFRAN pour le bassin du Doubs, aux pas de temps de 6 h, 12 h et 24 h. Les coefficients de détermination correspondant à chaque sous-bassin de la Saône et pour les trois pas de temps sont rassemblés dans le Tableau 4.1. Ils ont été calculés

pour toutes les observations (TP), mais aussi pour certaines observations seulement (PNN), pour lesquelles au moins une des deux valeurs de lame d'eau est non nulle (SAFRAN>0 ou MF/CNR>0). Cet échantillon PNN permet de ne pas tenir compte des lames d'eau nulles, qui *a priori* ont un poids important dans la corrélation.

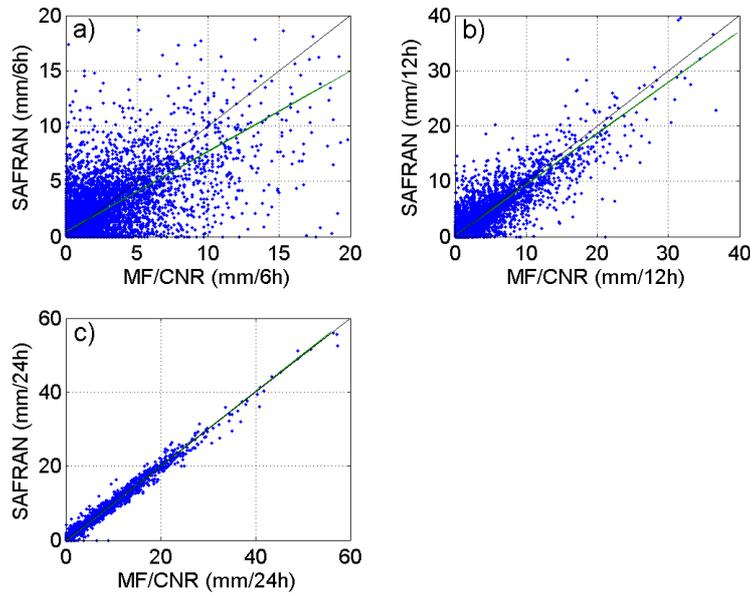


Figure 4.1: Corrélation entre les lames d'eau MF/CNR et SAFRAN aux pas de temps a) 6 h, b) 12 h et c) 24 h, sur la période du 01/01/1992 au 31/12/2003, pour le bassin du Doubs.

Tableau 4.1 : Coefficients de détermination (R^2) calculés entre les lames d'eau SAFRAN et MF/CNR, pour toutes les pluies (TP) et lorsqu'au moins l'une des deux lames d'eau est non nulle (PNN), aux pas de temps 6 h, 12 h et 24 h, sur la période du 01/01/1992 au 31/12/2003, pour les trois sous-bassins de la Saône.

Pas de temps	Echantillon	Sous-bassin		
		Doubs	Petite Saône	Saône Intermédiaire
6 h	TP	0,52	0,52	0,46
	PNN	0,44	0,44	0,39
12 h	TP	0,81	0,81	0,78
	PNN	0,77	0,78	0,75
24 h	TP	0,99	0,98	0,97
	PNN	0,98	0,98	0,97

Cette comparaison met en évidence plusieurs éléments. Si l'on s'intéresse au pas de temps journalier, la cohérence est forte avec des valeurs de R^2 proches de 1. La Figure 4.1 c) montre un nuage de points presque rectiligne, la droite de régression linéaire s'ajustant à la première bissectrice. En revanche aux pas de temps infra-journaliers, une certaine incohérence apparaît. Ainsi au pas de temps de 12 h, la cohérence est encore raisonnable mais moins consistante qu'au pas de temps de 24 h, avec des valeurs de R^2 de l'ordre de 0,8. Au pas de temps de 6 h, les valeurs de R^2 sont proches de 0,5, ce qui indique une faible relation entre les

deux sources de données SAFRAN/F et MF/CNR. Pour le bassin de la Saône Intermédiaire (cf. Tableau 4.1), la cohérence est plus faible que pour les autres bassins, avec un coefficient R^2 inférieur à 0,5. Pour ces pas de temps infra-journaliers, les nuages de points sont plus dispersés et la droite de régression linéaire s'écarte de la première bissectrice avec la diminution du pas de temps, ce qui traduit un biais marqué (Figure 4.1 b) et c)). Si nous nous intéressons aux lames d'eau non nulles (PNN), cette hiérarchie est conservée. Conformément à nos attentes, la cohérence s'avère plus faible que celle observée pour l'ensemble des observations (TP).

Les différences entre pas de temps considérés peuvent s'expliquer par le fait que le système SAFRAN analyse la quantité de précipitation en assimilant des données observées dont la plupart provient d'un réseau de mesures essentiellement journalières. Une désagrégation temporelle est ensuite effectuée, principalement fondée sur les analyses du MNP utilisé, et non pas sur des données observées, contrairement aux données de l'archive MF/CNR.

Nous avons également examiné la cohérence par gamme d'intensité de précipitations. Pour ce faire, seules les lames d'eau MF/CNR supérieures à un seuil ont été prises en compte. Les seuils étudiés sont les quantiles de pluie $PNN50$, $PNN70$ et $PNN90$, le quantile $PNNX$ correspondant à la lame d'eau qui n'est pas dépassée dans X % des lames d'eau non nulles. Le Tableau 4.2 représente, pour tous les bassins et pour tous les pas de temps, les coefficients R^2 en fonction du quantile au non-dépassement.

Tableau 4.2 : Coefficients de détermination (R^2) calculés entre les lames d'eau SAFRAN et MF/CNR, en considérant toutes les lames MF/CNR supérieures aux quantiles 50 %, 70 % et 90 % de pluie non nulle, aux pas de temps de 6 h, 12 h et 24 h.

Pas de temps	Quantile seuil	Sous-bassin		
		Doubs	Petite Saône	Saône Intermédiaire
6 h	$PNN50$	0,32	0,34	0,32
	$PNN70$	0,26	0,29	0,28
	$PNN90$	0,18	0,18	0,21
12 h	$PNN50$	0,69	0,7	0,69
	$PNN70$	0,64	0,66	0,64
	$PNN90$	0,56	0,58	0,57
24 h	$PNN50$	0,97	0,97	0,96
	$PNN70$	0,97	0,97	0,95
	$PNN90$	0,96	0,96	0,93

Le comportement observé précédemment selon le pas de temps est inchangé. Plus les événements pluvieux sont rares, *i.e.* plus les lames d'eau considérées sont élevées, moins la cohérence est avérée entre les archives SAFRAN et MF/CNR. Au pas de temps de 6 h, le coefficient R^2 ne vaut plus que 0,15 à 0,2 pour l'échantillon des événements rares. Les lames d'eau 24 h restent assez cohérentes, et ce quel que soit l'événement pluvieux, avec un

coefficient R^2 supérieur à 0,9.

La corrélation entre les lames d'eau SAFRAN et MF/CNR a également été étudiée selon les saisons. Les coefficients de détermination correspondant à la saison hivernale, de décembre à février, et à la saison estivale, de juin à août, sont reportés dans le Tableau 4.3. Les résultats laissent apparaître une cohérence différente selon la saison considérée. Ainsi en hiver, le lien semble plus fort que toutes saisons confondues aux pas de temps infra-journaliers. En été, en revanche, la corrélation est moindre avec des coefficients R^2 nettement plus faibles, notamment aux pas de temps inférieurs à 24 h. Cette différence s'expliquerait notamment par les systèmes précipitants d'origine convective, plus courts et très localisés en période estivale, et donc modérément interceptés par les réseaux horaires et même journaliers.

Tableau 4.3 : Coefficients de détermination (R^2) calculés entre les lames d'eau SAFRAN et MF/CNR, en considérant les relevés en hiver (décembre à février) et en été (juin à août), aux pas de temps de 6 h, 12 h et 24 h.

Pas de temps	Saison	Sous-bassin		
		Doubs	Petite Saône	Saône Intermédiaire
6 h	Hiver	0,67	0,62	0,52
	Eté	0,37	0,37	0,31
12 h	Hiver	0,81	0,78	0,7
	Eté	0,65	0,63	0,59
24 h	Hiver	0,92	0,91	0,88
	Eté	0,83	0,81	0,8

En conclusion, la cohérence entre l'archive MF/CNR et l'archive SAFRAN est solide au pas de temps journalier, tandis qu'elle est fragile voire insuffisante aux pas de temps infrajournaliers. Les résultats obtenus dans cette section permettent donc de valider les précipitations fournies dans les réanalyses SAFRAN au pas de temps journalier, mais pas aux pas de temps infrajournaliers.

Quoi qu'il en soit, les sous-bassins de la Saône sur lesquels la méthode de prévision sera développée ont un temps de réaction assez lent, de l'ordre de la journée voire de quelques jours, comme nous le verrons dans la section suivante. Il n'est donc pas nécessaire de connaître les pluies futures sur des durées infrajournalières : une prévision de pluie au pas de temps de 24 h devrait *a priori* être suffisante. Le travail d'optimisation de la méthode se fera par conséquent avec l'archive de cumuls pluviométriques fournis par le système SAFRAN, qui couvre une période suffisamment longue pour permettre un travail de ce type.

4.2 Analyse des régimes pluviométriques et hydrologiques

L'objectif de cette section est d'identifier les régimes pluviométriques et hydrologiques auxquels sont soumis les bassins tests. Ces analyses nous permettront plus tard, dans la troisième partie, d'interpréter au mieux les résultats et de situer les épisodes pluvieux et les

événements de crue étudiés par rapport à des événements majeurs.

4.2.1 Régimes pluviométriques

Les données pluviométriques utilisées pour cette analyse proviennent de l'archive SAFRAN. Nous avons étudié la climatologie de chaque bassin à l'aide des pluies brutes journalières de bassin extraites sur la période de disponibilité des données en fin de thèse (01/08/1970-30/07/2008), soit 13879 jours.

Pour chaque sous-bassin de la Saône et de la Seine, nous avons évalué la fréquence de jours sans pluie ($F0$), les cumuls de pluie dont la période de retour est de 5 ans ($PJXA5$) et 10 ans ($PJXA10$), ainsi que le cumul maximal ($Pmax$) observé sur la période (Tableau 4.4). Les valeurs extrêmes $PJXA5$ et $PJXA10$ ont été obtenues par un ajustement à la loi de Gumbel sur les maxima annuels. L'ensemble des ajustements effectués est représenté sur la Figure 4.3.

Tableau 4.4 : Caractéristiques de la pluviométrie de chaque sous-bassin de la Saône et de la Seine ($F0$: fréquence de jours sans pluie ; $PJXA5$ et $PJXA10$: cumuls de pluie de période de retour 5 et 10 ans ; $Pmax$: cumul maximal enregistré sur la période 01/08/1970-30/07/2008).

Bassin versant	$F0$ (%)	$PJXA5$ (mm)	$PJXA10$ (mm)	$Pmax$ (mm)
Doubs	20,5	46,3	51,5	58,1
Petite Saône	20	35,9	39,7	58,5
Saône Int.	23	45,5	51,8	71,1
Loing-E	30	29,5	33,3	51,4
Marne-F	19,8	32	35,3	43,9
Seine-B	22,2	29,2	32	35,6
Seine-Int	19,8	28,1	31,9	47,9
Yonne-C	21,7	31	34,4	39,4

Le bassin le moins arrosé est le bassin du Loing-E avec près d'un jour sur trois sans pluie observée. De manière générale, la fréquence des journées sans pluie est de l'ordre de 20 %, ce qui paraît bien faible. En effet, le système SAFRAN a des difficultés à discriminer les jours sans pluie puisque l'analyse est faite par zone. Ainsi, il suffit qu'un poste pluviométrique affiche une valeur non nulle pour que l'ensemble de la zone soit affectée d'une lame d'eau non nulle. Dans le cas présent, cet effet est sans doute accentué par l'aggrégation spatiale sur les sous-bassins.

Par ailleurs, selon le Tableau 4.4, le bassin de la Saône connaît des pluies fortes plus importantes que celles que connaît le bassin de la Seine. Les cumuls quinquennaux et décennaux sont en effet bien plus élevés sur les sous-bassin de la Saône, notamment sur le Doubs et la Saône Intermédiaire. Ces résultats montrent l'influence de l'orographie sur les précipitations : le bassin de la Saône est effectivement plus vallonné avec des altitudes moyennes plus élevées que sur le bassin de la Seine (cf. 2.1). Pour le cas particulier de la Saône Intermédiaire, signalons que les fortes valeurs de cumul quinquennal et décennal sont en partie liées à l'extension vers ce bassin des épisodes pluvieux en flux de Sud, qui se produisent typiquement près de la Méditerranée et dans le sud du Massif-Central. Par

ailleurs, le cumul maximal observé sur la période varie du simple (35,6 mm sur la Seine-B) au double (71,1 mm sur la Saône Intermédiaire).

Les régimes pluviométriques peuvent également être caractérisés aux échelles mensuelle et annuelle. La Figure 4.2 montre pour chaque sous-bassin a) l'évolution du cumul mensuel moyen et b) le cumul annuel moyen, évalué entre le 1er septembre et le 31 août de chaque année.

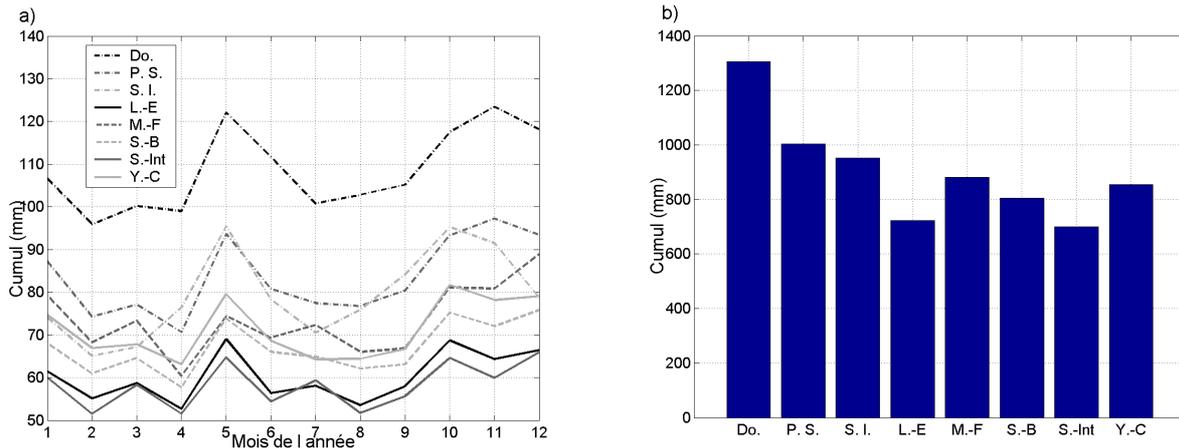


Figure 4.2 : Caractéristiques de la pluviométrie aux échelles mensuelle et annuelle sur les huit sous-bassins de la Saône et de la Seine (Do : Doubs ; P. S. : Petite Saône ; S. I. : Saône Intermédiaire ; L.-E : Loing-E ; M.-F : Marne-F ; S.-B : Seine-B ; S.-Int : Seine-Int ; Y.-C : Yonne-C), obtenues sur la période 01/08/1970-30/07/2008. a) Evolution du cumul mensuel moyen ; b) Cumul annuel moyen.

Les périodes de l'année au cours desquelles les précipitations sont les plus importantes sont le mois de mai ainsi que les mois d'octobre à décembre. Le bassin du Doubs, pour lequel l'orographie est la plus importante, est le plus arrosé avec un cumul annuel moyen supérieur à 1300 mm. De manière générale, ce sont les sous-bassins de la Saône qui connaissent les précipitations les plus abondantes. Les sous-bassins de la Seine recueillent chacun à peu près la même quantité de précipitations, les moins arrosés étant ceux du Loing-E et de la Seine-Int.

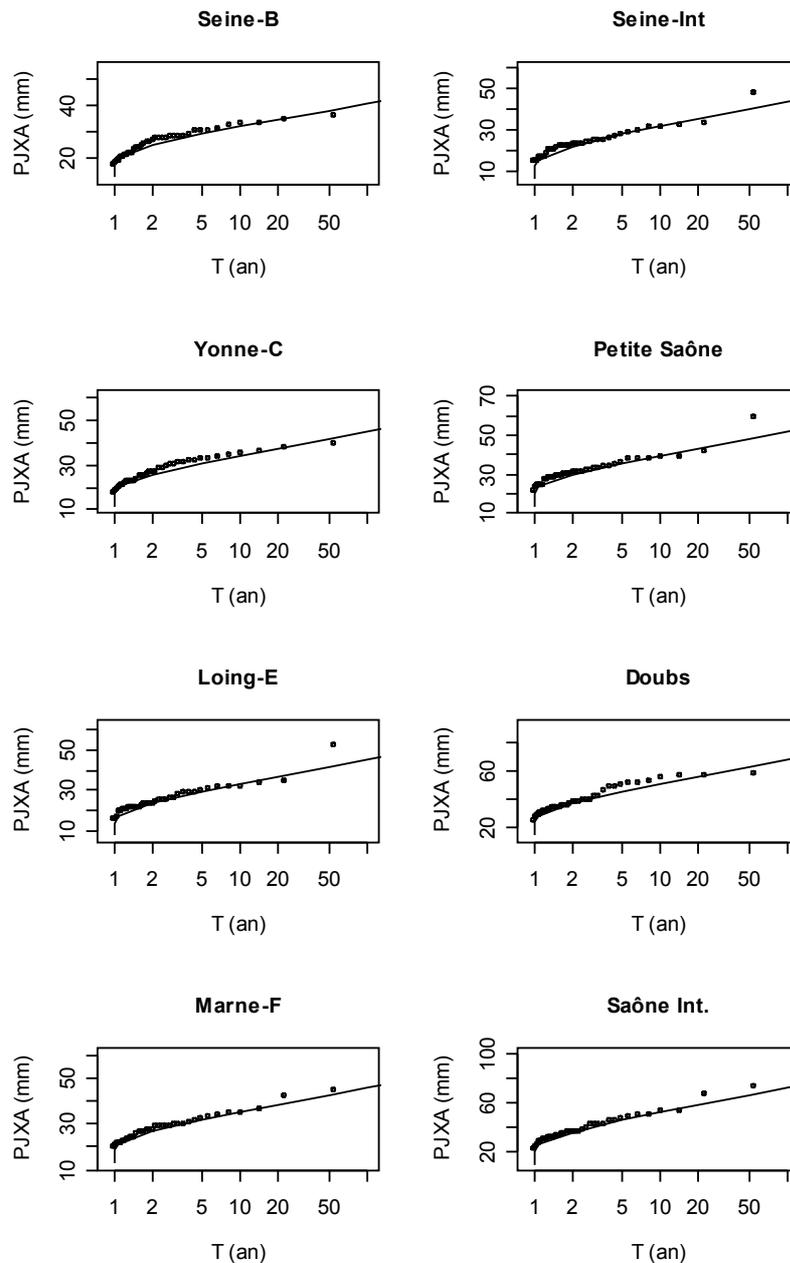


Figure 4.3 : Ajustements à la loi de Gumbel sur les cumuls journaliers maximums annuels de la période 01/08/1970-30/07/2008, pour les sous-bassins de la Saône et de la Seine.

4.2.2 Régimes hydrologiques

Les régimes hydrologiques auxquels sont soumis les bassins tests ont été caractérisés à l'aide de l'ensemble des données de débit disponibles (cf. 2.2.3). Afin d'avoir une vue d'ensemble de ces régimes, nous avons tracé sur la Figure 4.4 l'évolution des débits mensuels moyens observés aux stations hydrométriques des différents sous-bassins de la Saône et de la Seine. Les débits ont été exprimés en débit spécifique ($l/s/km^2$) pour qu'ils soient comparables entre les bassins.

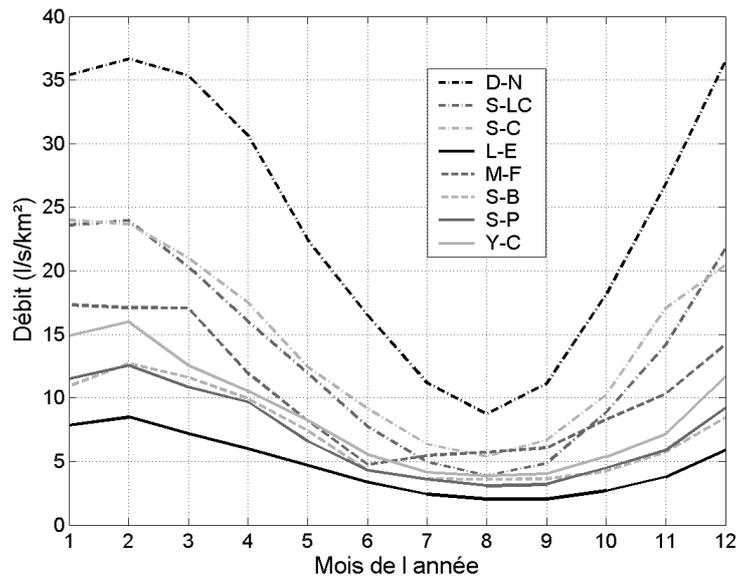


Figure 4.4 : Evolution des débits mensuels moyens, exprimés en $l/s/km^2$, évalués sur les périodes de disponibilité des données, pour les sous-bassins de la Saône (D-N : Doubs à Neublans ; S-LC : Saône à Le Chatelet ; S-C : Saône à Couzon) et de la Seine (L-E : Loing à Episy ; M-F : Marne à La Ferté ; S-B : Seine à Bazoches ; S-P : Seine à Paris ; Y-C : Yonne à Courlon).

Les débits spécifiques les plus forts sont observés pour le Doubs à Neublans, tandis que la station hydrométrique du Loing à Episy enregistre les débits les plus faibles. Nous retrouvons ici des débits en moyenne plus importants sur le bassin de la Saône que sur le bassin de la Seine, du fait d'une pluviométrie globalement plus importante (la Figure 4.4 est à comparer avec la Figure 4.2 a)). Les forts débits s'observent typiquement entre les mois de décembre et de mars, tandis que la phase d'étiage est maximale au mois d'août. Ils révèlent un régime hydrologique de type pluvial au sens de Pardé (1933). Nous ne décelons pas à l'échelle mensuelle d'influence particulière de la fonte de la neige. Le manteau neigeux n'est pas suffisamment étendu et régulier d'une année sur l'autre pour induire un comportement pluvio-nival.

Nous avons caractérisé plus particulièrement les régimes de crues de chaque bassin. Afin de caractériser la dynamique de crue, nous avons identifié pour chaque station différents hydrogrammes selon le débit maximal atteint au cours de chaque année hydrologique (01/09-31/08). Nous les avons ensuite normés par le débit maximal observé de chaque événement. Les hydrogrammes obtenus pour chaque station sont indiqués sur la Figure 4.5 (courbes grisées). L'hydrogramme médian est représenté en noir.

Nous pouvons déduire de chaque hydrogramme médian la durée caractéristique de crue. Celle-ci est exprimée par la médiane des durées de dépassement de la moitié du débit de pointe de chaque crue. Les durées caractéristiques et les nombres d'épisodes considérés sont reportés dans le Tableau 4.5.

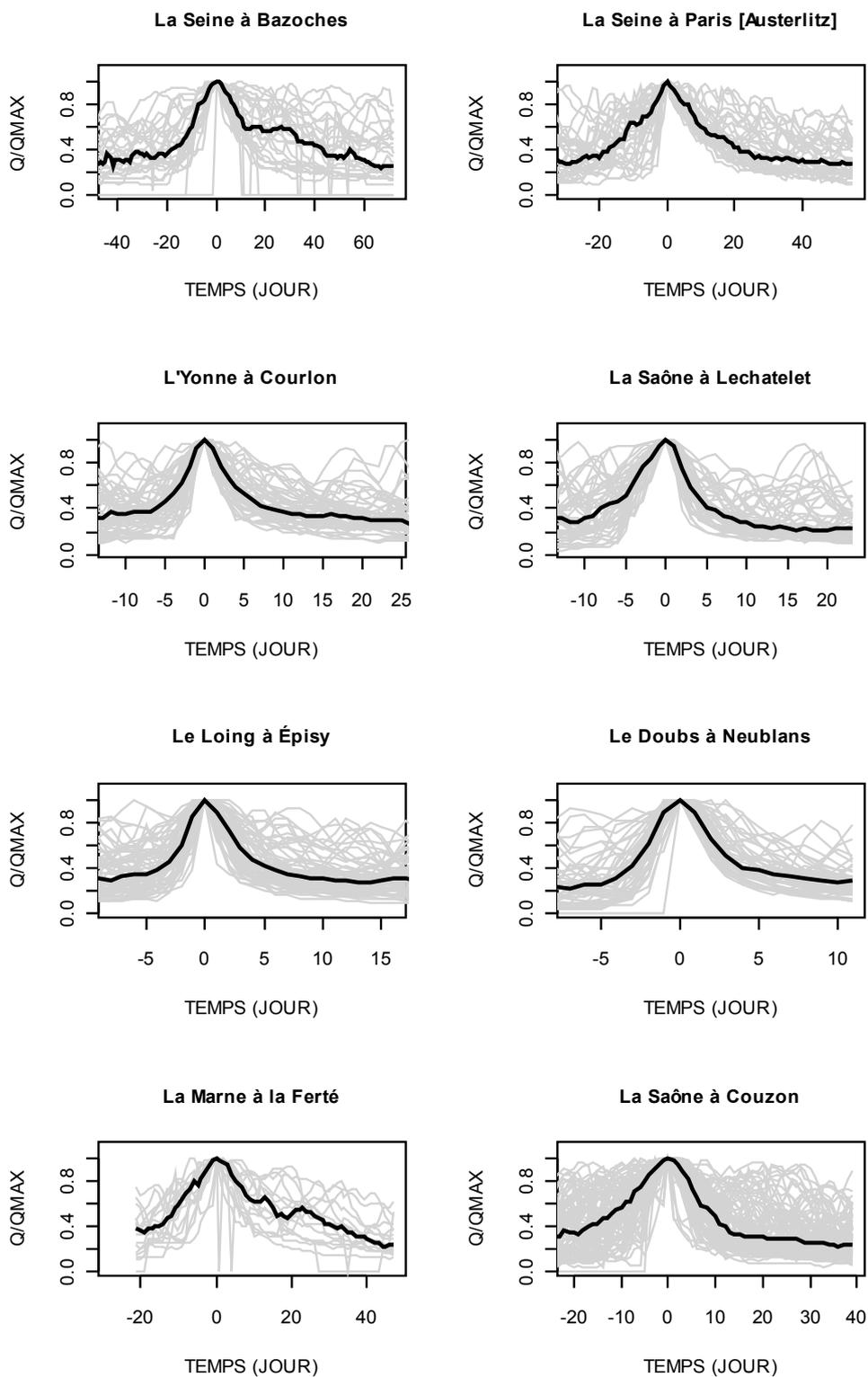


Figure 4.5 : Hydrogrammes de crue normés sélectionnés sur la période de disponibilité des données (courbes grises). La courbe noire représente l'hydrogramme médian.

Tableau 4.5 : Durées caractéristiques des hydrogrammes de crue médians et nombres d'épisodes considérés.

Station [cours d'eau]	Durée caractéristique (jour)	Nombre d'épisodes
Le Châtelet [Saône]	17	45
Neublans [Doubs]	7	44
Couzon [Saône]	32	89
Episy [Loing]	17	61
La Ferté [Marne]	35	17
Bazoches [Seine]	60	28
Paris [Seine]	38	36
Courlon [Yonne]	17	52

Le nombre d'épisodes est assez faible pour les stations de la Seine à Bazoches et de la Marne à La Ferté en raison des nombreuses lacunes présentes dans les données. L'examen du Tableau 4.5 indique que la dynamique de crue la plus rapide est observée pour le Doubs à Neublans, avec une durée caractéristique de sept jours seulement. En revanche la dynamique est bien plus lente pour la Saône à Couzon avec une durée de 32 jours, ce qui justifie le fait que ce bassin soit considéré comme un *bassin à crue lente*. Les sous-bassins de la Seine sont soumis à des dynamiques variées. Ainsi la dynamique est moins lente sur l'Yonne à Courlon et le Loing à Episy avec une durée caractéristique de 17 jours. La dynamique de la Seine à Paris est beaucoup plus lente, tandis que celle de la Seine à Bazoches est extrêmement lente. Ainsi en moyenne pour cette station, le débit est supérieur à la moitié du débit de pointe de crue durant 60 jours. Ces nuances sont partiellement dues à l'existence de stocks tampons (nappe d'accompagnement, lac-réservoir) qui ralentissent les écoulements de manière naturelle ou artificielle.

Outre la dynamique des crues, il est également intéressant d'analyser leur intensité. Nous avons estimé les quantiles de débit de crue associés à différentes périodes de retour (Tableau 4.6). Pour cela, nous avons d'abord formé, pour chaque station, un échantillon représentatif des débits de crue. Nous avons privilégié la méthode d'échantillonnage par valeur supérieure à un seuil (sup-seuil). Cette méthode est jugée plus fiable que la méthode des maxima annuels avec des valeurs d'échantillon plus homogènes, même si l'extraction des valeurs selon l'approche sup-seuil nécessite une analyse complémentaire pour éliminer les valeurs liées et une étude de sensibilité sur le choix du seuil (Lang et al., 2006). Nous avons choisi un nombre moyen d'événements de deux par an. Nous avons ensuite testé deux lois recommandées par la théorie des valeurs extrêmes : la loi exponentielle (EXPO) et la loi de Pareto Généralisée (GPD). Nous avons privilégié la loi exponentielle ; cependant cette dernière s'est révélée inadaptée pour certains échantillons (les quantiles semblent être bornés supérieurement). Les lois d'ajustement sont reportées dans le Tableau 4.6 pour chaque station. La Figure 4.6 présente, à titre indicatif, les échantillons de valeurs de débit supérieures au seuil prédéterminé ainsi que les ajustements aux lois EXPO (noir) et GPD (rouge).

Tableau 4.6 : Quantiles de débit (m^3/s) associés à différentes périodes de retour T (an), obtenus par l'ajustement des débits supérieurs à un seuil à une loi (EXPO ou GDP selon le sous-bassin).

Station [cours d'eau]	Bazoches [Seine]	Courlon [Yonne]	Épisy [Loing]	La Ferté [Marne]	Paris [Seine]	Le Chatelet [Saône]	Neublans [Doubs]	Couzon [Saône]
Loi ajustée	GPD	GPD	EXPO	GPD	GPD	EXPO	EXPO	EXPO
T (an)	1	147	378	72	282	889	720	936
	2	205	484	108	342	1139	890	1105
	5	271	600	156	386	1401	1114	1329
	10	314	672	193	403	1557	1283	1498
	20	352	733	229	413	1685	1452	1668

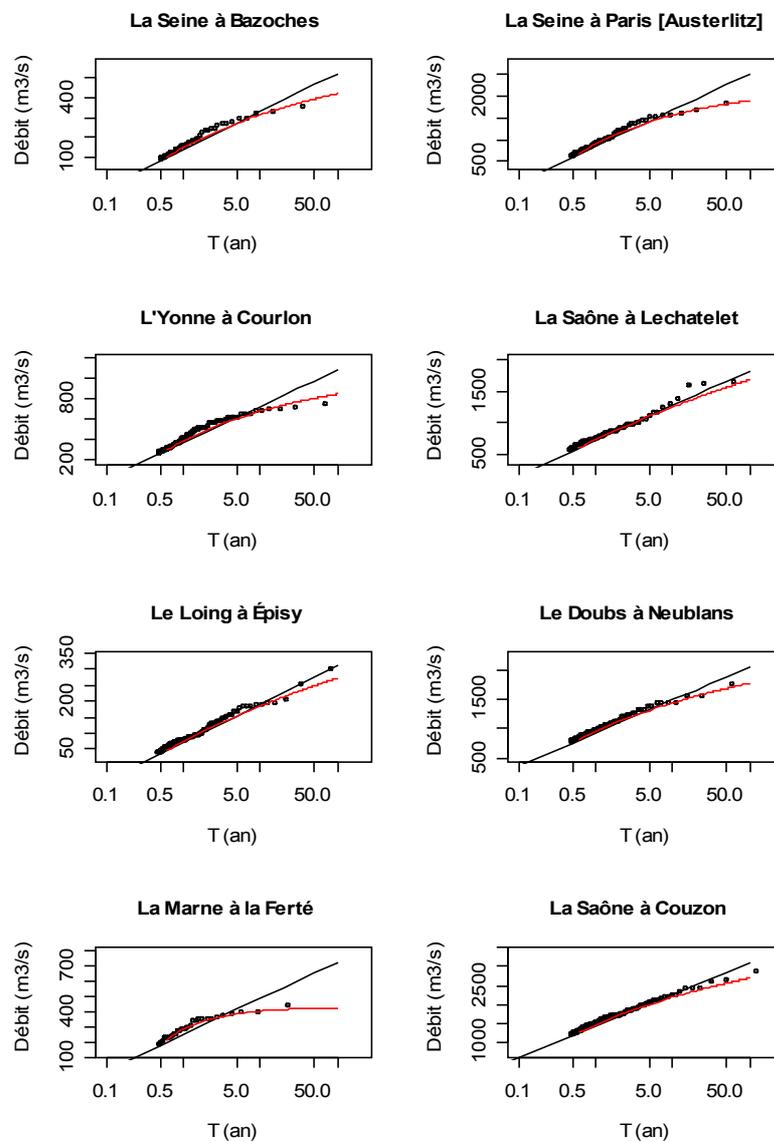


Figure 4.6 : Echantillon de valeurs de débit sup-seuil (points) et ajustements aux lois EXPO (noir) et GPD (rouge) obtenus pour chaque station hydrométrique des bassins de la Saône et de la Seine.

4.3 Comparaison et critique des archives de réanalyses

Nous avons à disposition deux archives de réanalyses (ERA-40 et NCEP/NCAR) contenant les variables d'analogie. Or pour développer une méthode de manière cohérente, il est nécessaire d'utiliser une même archive. On ne peut pas changer d'archive au cours du développement : il faut donc faire un choix entre l'une ou l'autre. Puisque le choix n'est pas évident à première vue, nous avons (1) comparé les valeurs des variables d'analogie utilisées dans la méthode de référence A0 (Bontron, 2004) extraites des deux archives, et (2) étudié la sensibilité de la méthode A0 au choix de l'archive de réanalyses.

4.3.1 Comparaison des variables d'analogie de la méthode de référence entre les réanalyses ERA-40 et NCEP/NCAR

Les deux archives météorologiques ERA-40 (ERA) et NCEP/NCAR (NNR) ont déjà fait l'objet de comparaisons dans des cadres souvent différents, principalement pour des études climatologiques. Les variables et les zones d'étude diffèrent donc fortement d'une étude à une autre.

Gleisner et al. (2005) ont examiné la relation entre l'activité solaire et le climat en comparant les deux réanalyses, à partir de l'épaisseur de la couche 1000 hPa - 500 hPa et de celle de la couche 500 hPa - 200 hPa sur le globe et sur la période 1958-2001. Des anomalies apparaissent dans l'archive ERA, contrairement à l'archive NNR, notamment au milieu des années 1970 et en 1986.

La température moyenne de surface a été comparée par Simmons et al. (2004) sur le globe et sur la période 1958-2001. La température et son évolution dans l'archive ERA sont plus proches des observations que celles de l'archive NNR en raison de l'intégration de nouvelles données sol.

L'homogénéité de l'archive ERA a également été étudiée par Sterl (2004). Les valeurs mensuelles de pression au niveau de la mer et du géopotential à 500 hPa des archives ERA et NNR ont été comparées à l'aide de trois coupes longitudinales. Des changements sont détectés en 1967, 1969, 1973 (autour de 40°N), 1976 (autour de 40°N et 60°S), 1979 (probablement en raison de l'assimilation de données satellitaires supplémentaires) et en 1993 dans le Pôle Sud. La plupart de ces changements sont observés pour chaque variable et pour chaque coupe longitudinale. D'après l'auteur, ces changements sont probablement liés à des modifications de la dynamique et de la physique du modèle ainsi que des systèmes d'assimilation.

Dans le cadre d'une comparaison de méthodes de reproduction des modes saisonniers de circulation atmosphérique, Vrac et al. (2007) ont utilisé les deux archives de réanalyses. Les géopotentiels à 500 hPa, 700 hPa et 850 hPa ainsi que la pression au niveau de la mer ont été étudiés sur la période 1980-1999 sur le domaine de l'Amérique du nord (26°N-65°N/110°W-67°W). Leur étude fut originale dans la mesure où ils ont utilisé l'archive ERA avec la résolution initiale de 1,125° et avec la résolution de 2,5°. Il apparaît clairement que la

résolution explique en partie les différences observées, mais la paramétrisation physique de chaque modèle jouerait un rôle non négligeable dans les niveaux inférieurs. Des erreurs systématiques sont même produites dans certaines régions du globe, en raison des techniques de modélisation qui ne sont pas encore tout à fait maîtrisées (Jung, 2005).

D'autres variables ont été étudiées par Ito and Sasai (2006) sur la période 1982-2001 et sur l'ensemble du globe. Ils ont notamment comparé la température de l'air à 2m et les précipitations des archives ERA et NNR avec celles de bases de données directement issues d'observations (du Climate Research Unit (CRU) notamment). Leurs résultats indiquent que les températures moyennes sont peu éloignées mais que, sur la base d'une appréciation visuelle, les fluctuations annuelles de cette variable sont mieux représentées dans l'archive ERA. Les précipitations de l'archive NNR sont 10 % supérieures à celles de l'archive ERA, mais la variabilité interannuelle est bien représentée dans les deux réanalyses. En revanche la tendance à long terme diverge : alors que la tendance est à la hausse dans l'archive du CRU et dans l'archive NNR, les précipitations tendent à diminuer dans l'archive ERA.

Les études précédemment citées, souvent réalisées à l'échelle climatologique sur l'ensemble du globe, ne nous permettent pas cependant de connaître avec exactitude les différences de valeurs des variables d'analogie utilisées dans la méthode de référence. Nous allons donc comparer les valeurs des quatre variables qui nous intéressent, à savoir celles des géopotentiels $Z500$ et $Z1000$, d'humidité $RH850$ et de quantité d'eau précipitable TCW . Dans le cadre de la comparaison, nous considérons la période commune aux deux archives (01/09/1957-31/08/2002 toutes les 6h, soit 65744 dates), et le domaine d'extraction initial (cf. 2.2.1.1, Figure 2.3).

A l'aide du critère de forme TWS

Pour les géopotentiels, nous avons mesuré la similitude des champs de $Z500$ et $Z1000$ entre les deux archives par l'intermédiaire du critère TWS défini dans le chapitre 1 (cf. 1.3.1.1). La valeur moyenne du critère TWS pour les variables $Z500$ et $Z1000$ s'élève respectivement à 13,9 et 18,6, ce qui est assez faible. La cohérence entre les deux archives est donc consistante en termes de forme des champs de géopotentiels. Mais cette similitude est-elle constante au cours du temps ? Pour le vérifier, la Figure 4.7 représente, pour la variable $Z500$, l'évolution du critère TWS annuel sur le domaine d'extraction et sur le domaine d'analogie D1 du niveau 1 de la méthode de référence (cf. 1.3.1.1, Figure 1.4).

Une baisse significative de la valeur du TWS annuel s'observe au fil du temps, et notamment avant le milieu de la période. Un test de Pettitt (Pettitt, 1979) a été effectué et confirme une rupture dans les années 1970. Cette rupture pourrait être due à l'assimilation de nouvelles données en 1973 et 1979 pour la construction de l'archive ERA-40 (Figure 4.8). Ce résultat garantit ainsi une plus forte cohérence des archives pour la seconde partie de la période, en particulier celle couverte par l'archive pluviométrique SAFRAN. D'autre part, les critères TWS obtenus sur les deux domaines ont un comportement similaire dans le temps.

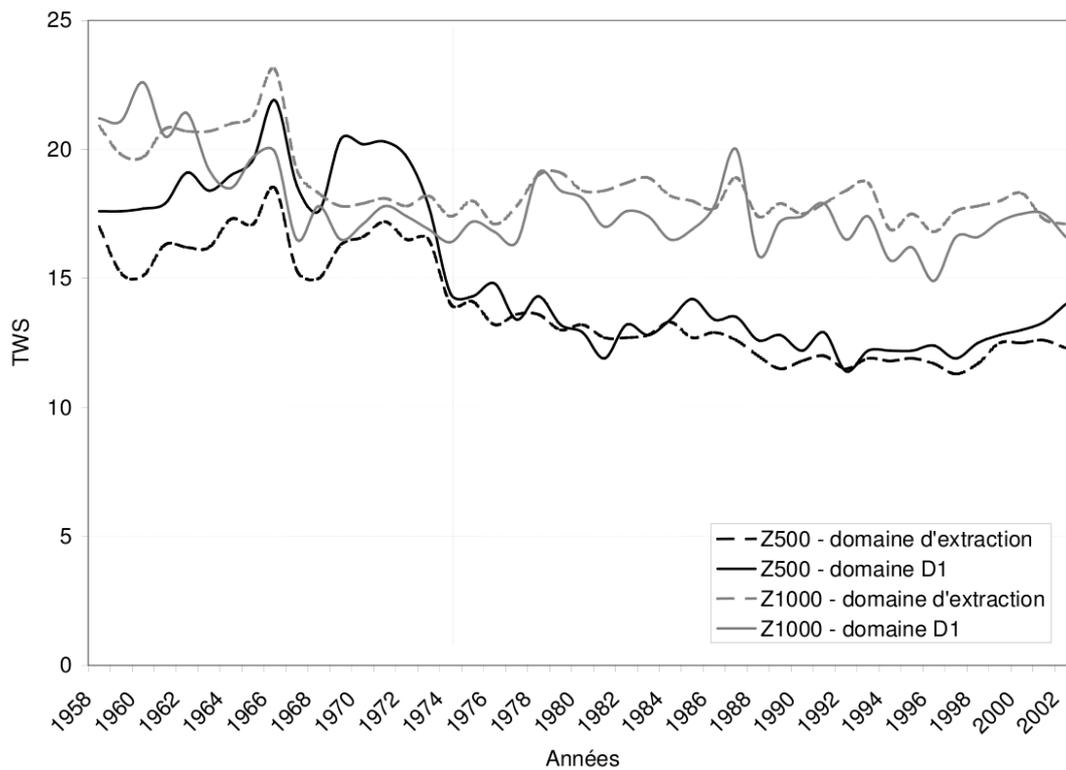


Figure 4.7 : Evolution de la valeur moyenne annuelle du critère *TWS* annuel de la variable Z500 sur le domaine d'extraction et sur le domaine D1, calculé entre les champs extraits des archives ERA et NNR.

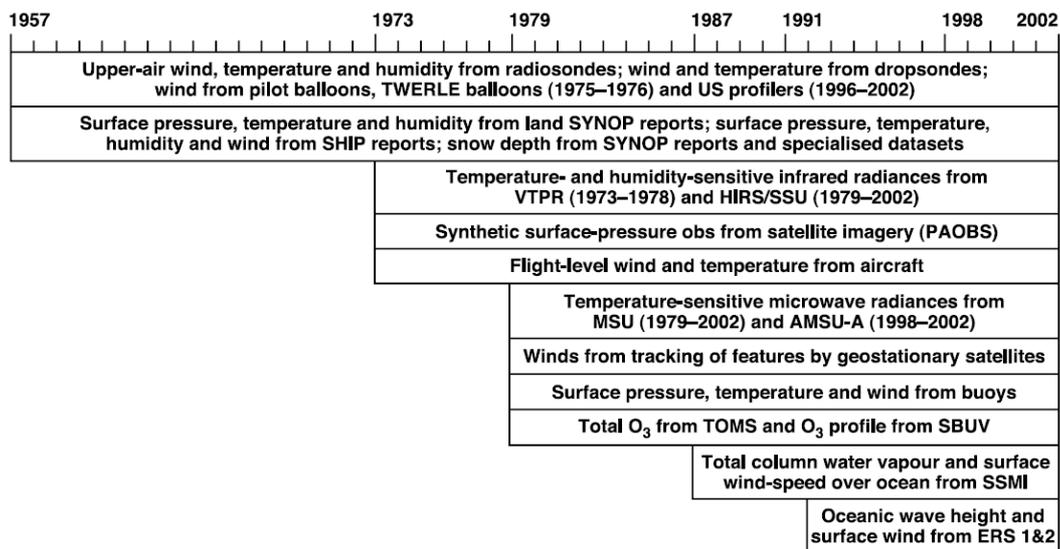


Figure 4.8 : Chronologie des types d'observations assimilées dans l'archive ERA de 1957 à 2002 (d'après Uppala et al., 2005).

Cependant, la valeur calculée sur le domaine D1 est très souvent supérieure à celle calculée sur le domaine d'extraction, même si l'écart diminue également à partir de la fin des années 1970. Cette différence peut s'expliquer par le fait que les grands centres d'action à l'échelle synoptique et les flux environnants sont mieux représentés dans le domaine

d'extraction que ceux de plus basse échelle dans le domaine D1. En effet, la variabilité des gradients de géopotential devient plus importante lorsque la taille du domaine (*i.e.* l'échelle) diminue. En revanche, aucun changement significatif n'a été perçu au cours du temps pour la variable *Z1000*, même si la valeur moyenne annuelle du critère *TWS* est un peu plus élevée en début de période (décennie 1960-1970).

A l'aide de l'écart moyen *MD*

Pour chacune des quatre variables (*Z500*, *Z1000*, *RH850* et *TCW*), nous avons simplement calculé, en chaque point de grille du domaine, l'écart *MD* de valeur entre les archives ERA et NNR moyenné sur la période, selon l'équation 4.1 :

$$MD(i, j) = \frac{1}{nT} \sum_{t=1}^{nT} (V_{ERA}(t) - V_{NNR}(t)) \quad (4.1)$$

où $MD(i, j)$ est l'écart moyen au point de grille (i, j) , nT est le nombre de pas de temps (65744 dans le cas présent), $V_{ERA}(t)$ est la valeur de la variable V extraite de l'archive ERA (resp. NNR) à la date t .

Les écarts *MD* obtenus pour chacune des quatre variables sont cartographiés sur la Figure 4.9. Les différences de valeur entre les deux archives sont dans l'ensemble assez faibles pour les variables *Z500*, *Z1000* et *TCW* (Figure 4.9 (a), (b) et (d)).

Sur le globe, la valeur de la variable *Z500* est généralement comprise entre 5000 et 6000 m, ce qui correspond à une étendue moyenne de 1000 m. Sur le domaine d'extraction, l'écart moyen global entre les valeurs de *Z500* des archives ERA et NNR est d'environ 0,9 m, ce qui est négligeable devant l'étendue de 1000 m. Cependant, des différences assez marquées s'observent dans la région des Balkans.

Le résultat peut être interprété de la même façon pour la variable *Z1000*. L'écart *MD* global s'élève à -0,5 m, ce qui est négligeable devant l'étendue de valeurs de cette variable (de l'ordre de 800 m). Nous notons cependant de forts écarts dans la région du Groënland, avec un écart moyen minimal de -47,1 m, ce qui traduit un géopotential en moyenne bien plus bas dans l'archive ERA que dans l'archive NNR. Le nombre de dépressions détectées dans les régions orographiques comme le Groënland a été étudié et s'avère en effet plus important dans l'archive ERA (Hanson et al., 2004 ; Trigo, 2005). Cette différence serait due à la résolution plus grande du modèle utilisé pour l'archive ERA, mais aussi aux méthodes d'extrapolation. Les écarts sont également négatifs dans d'autres régions comme dans le bassin méditerranéen. D'après Trigo (2005), la résolution initiale de l'analyse (1,125°) serait à l'origine du plus grand nombre de petites dépressions dans ce secteur.

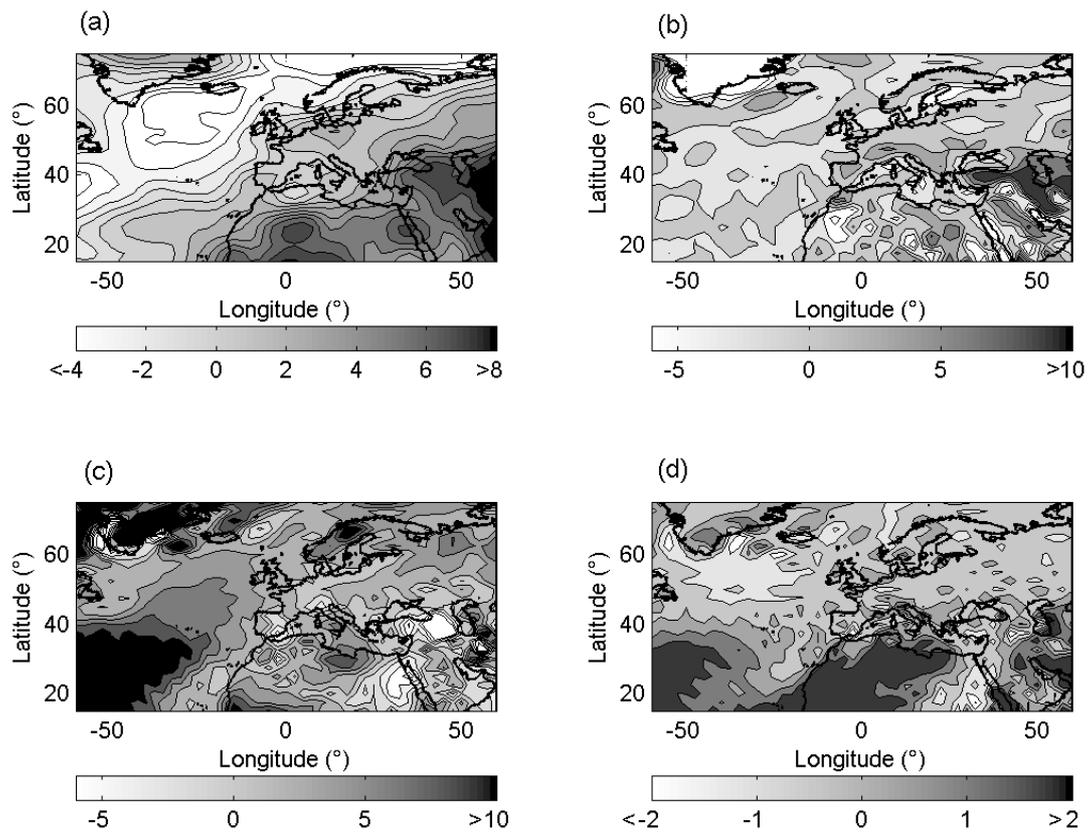


Figure 4.9 : Ecart moyen *MD* entre les archives ERA et NNR pour les variables (a) Z500 (m), (b) Z1000 (m), (c) RH850 (%) et (d) TCW (mm) ; obtenu sur la période 01/09/1957-31/08/2002. Les intervalles entre les contours sont respectivement de 1 m, 2 m, 2 % et 0,5 mm.

Les valeurs de la variable *TCW* extraites des deux archives sont globalement similaires. L'écart *MD* global s'élève ainsi à 0,5 mm. De même, cette valeur apparaît négligeable devant l'intervalle de valeurs que cette variable occupe (0 à 50 mm). Nous remarquons sur la Figure 4.9 (d) que l'écart est plus faible dans les hautes latitudes alors qu'il est plus conséquent dans les basses latitudes. C'est en effet dans ces régions que l'air troposphérique est le plus humide et les quantités d'eau précipitable sont les plus élevées. Dans le bassin méditerranéen, la quantité d'eau précipitable est globalement plus importante dans l'archive ERA, ce qui est en accord avec ce que Trigo (2005) suggère dans ses conclusions.

En revanche, la Figure 4.9 (c) indique des différences marquées entre les deux archives sur la variable *RH850*. Des écarts de l'ordre de +5 à +20 % sont observés sur l'océan Atlantique, et même jusqu'à +49 % au Groënland, ce qui traduit une humidité relative dans l'archive ERA largement supérieure à celle de l'archive NNR. Sur le domaine de l'Europe occidentale, les écarts sont plus faibles mais apparemment non négligeables. L'écart *MD* global s'élève ainsi à +4,6 % (8,3 % en valeur absolue), ce qui traduit un certain biais dans l'une des deux archives.

L'importance de ce biais peut être mesurée en analysant la chronique des écarts en un point qui nous intéresse, par exemple près du bassin de la Saône. La Figure 4.10 montre une chronique au point de coordonnées (5°E ; 45°N). La variation de l'écart est importante tout au long de la période. Compris en moyenne entre -20 % et +20 %, ces écarts sont bien plus

élevés en valeur absolue dans les années 1960 et au début des années 1970. Le 9 février 1967, la valeur extraite de l'archive ERA est supérieure de 65,6 % à celle extraite de l'archive NNR. Par ailleurs le 14 janvier 1965, la valeur de l'archive ERA est inférieure de 55,2 % à celle de l'archive NNR. Ces écarts sont bien trop significatifs pour être négligés, d'autant plus que le point étudié appartient à une zone continentale riche en observations (radiosondages notamment). Nous soupçonnons donc l'existence de valeurs aberrantes dans l'une ou l'autre archive de réanalyses.

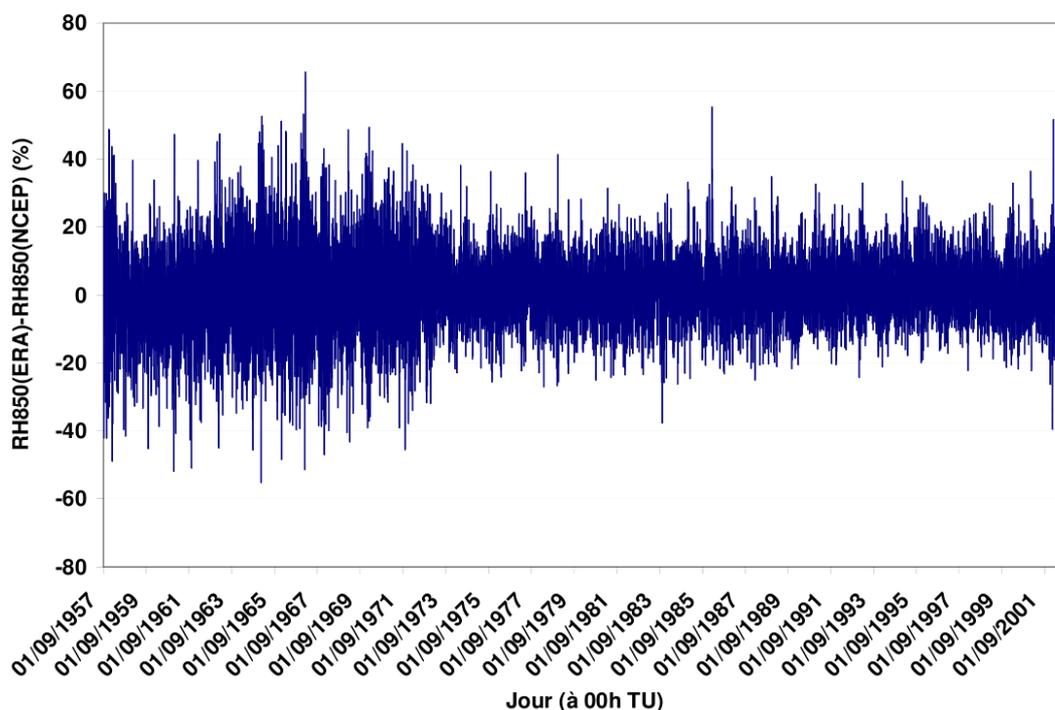


Figure 4.10 : Chronique de l'écart de RH850 entre les archives ERA et NNR sur la période 01/09/1957-31/08/2002, au point de coordonnées (5°E ; 45°N).

4.3.2 Etude des anomalies d'humidité dans les réanalyses ERA-40

Nous nous sommes rapidement aperçus, par une analyse détaillée, que des valeurs totalement aberrantes de la variable *RH850* figuraient dans l'archive ERA à résolution 2,5°. En effet, cette variable prend à certains endroits et à certains moments des valeurs aberrantes, *i.e.* inférieures à 0 % (anomalies négatives) ou des valeurs supérieures à 100 % (anomalies positives). Or en réalité, cette variable est bornée, comprise entre 0 % et 100 %, comme c'est le cas dans l'archive NNR. L'humidité relative peut certes dans certains cas dépasser légèrement 100 %, lorsque l'air est sursaturé en vapeur d'eau, mais cela ne peut expliquer les écarts constatés. Par ailleurs, nous avons également vérifié que ces anomalies étaient présentes dans l'archive ERA à résolution 1,125°.

Cette observation nous amène à nous poser plusieurs questions : ces valeurs aberrantes

sont-elles réparties de manière homogène dans l'espace et dans le temps ? Sont-elles vraiment irréalistes ? Sont-elles fréquentes ? Peut-on les corriger ? Pour y répondre, nous avons étudié plus précisément ces anomalies dans l'archive ERA à résolution 1,125°. Le lecteur est invité à consulter l'article (Ben Daoud et al., 2009a) reporté en annexe 2 pour visualiser l'ensemble des résultats.

Les anomalies positives et négatives ont été caractérisées, en chaque point du domaine d'extraction à 1,125° (cf. 2.2.1.2, Figure 2.4), par leur valeur extrême et leur proportion évaluées sur la période 1958-2001. Les valeurs extrêmes, qui s'élèvent à -16,2 % et +121,8 % en deux points distincts de l'Océan Atlantique nord, sont très anormales. Les proportions d'anomalies négatives sont relativement faibles, mais les proportions d'anomalies positives sont significatives. Par conséquent, nous avons envisagé de corriger les valeurs aberrantes. D'une part, nous les avons tronquées en fixant les valeurs négatives à 0 % et les valeurs supérieures à 100 % à 100 %, mais cela n'a pas eu d'impact sur les performances de la méthode de référence A0. D'autre part, nous avons essayé de corriger les anomalies en utilisant des données observées. Les valeurs de RH850 extraites de l'archive ERA ont été comparées à celles fournies par huit stations de radiosondage situées dans l'est de la France et au voisinage. Les valeurs de chaque station ont été comparées à celles extraites de l'archive ERA au point de grille le plus proche de la station. Les résultats montrent globalement une assez bonne cohérence entre les deux bases de données, mais la correction éventuelle des anomalies à partir des données des radiosondages semble à écarter.

Enfin, nous avons voulu vérifier s'il y avait une saisonnalité des valeurs aberrantes. Pour cela, les anomalies détectées sur le domaine d'extraction à 1,125° ont été comptabilisées pour chaque mois de l'année (Figure 4.11). On remarque effectivement une saisonnalité, puisqu'il y a bien plus d'anomalies en hiver et en automne et bien moins en été, ce qui n'est guère encourageant pour notre travail puisque les événements pluvieux que nous désirons anticiper se produisent en automne, hiver et au printemps. Nous avons également voulu savoir s'il n'y avait une tendance à l'échelle annuelle.

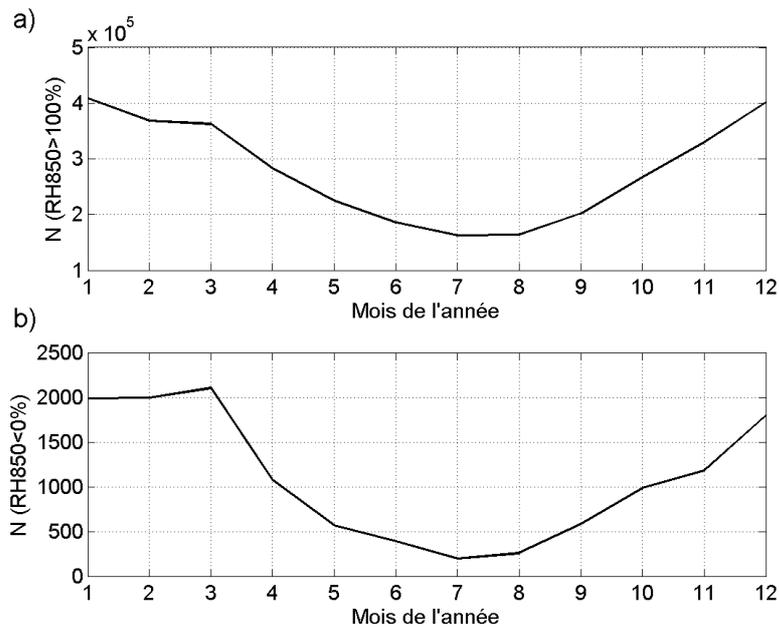


Figure 4.11 : Evolution du nombre total d'anomalies, détectées sur la période 1957-2002 sur le domaine d'extraction à 1,125°, en fonction du mois de l'année. a) Anomalies positives (>100 %) ; b) anomalies négatives (<0 %).

De même que précédemment, les anomalies ont été comptabilisées pour chaque année (Figure 4.12). Le nombre d'anomalies positives est assez variable, presque stable de 1957 à 1972, puis en forte hausse entre 1973 et 1978 où il a presque doublé. Ensuite, ce nombre chute de près de moitié pour se stabiliser malgré quelques irrégularités.

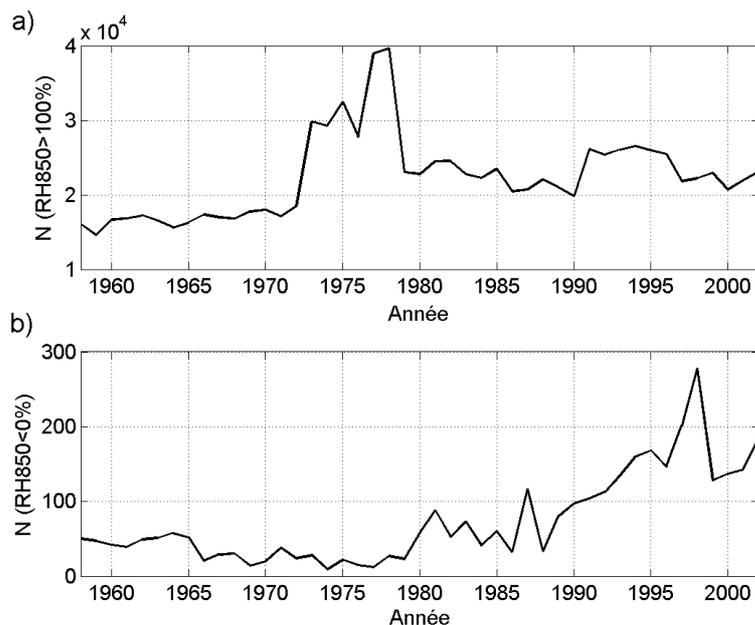


Figure 4.12 : Evolution du nombre total d'anomalies, comptabilisé sur la période 1957-2002 sur le domaine d'extraction à 1,125°, en fonction de l'année. a) Anomalies positives (>100 %) ; b) anomalies négatives (<0 %).

Il est intéressant de croiser ces résultats avec la chronologie des types d'observations assimilées dans l'archive ERA (cf. Figure 4.8). Cette période correspond exactement à celle de l'assimilation de nouvelles données de température et d'humidité fournies par le satellite VTPR. Elle furent remplacées en 1979 par d'autres données issues du satellite HIRS/SSU (Uppala et al., 2005). En revanche, le nombre d'anomalies négatives, stable entre les années 1957 et 1980, tend à augmenter au fil des années (nombre multiplié par 4 entre le début et la fin de la période). Cette tendance n'a pas pu être expliquée.

Nous nous sommes également assuré de la cohérence entre les archives ERA à résolution $2,5^\circ$ et $1,125^\circ$. Pour cela, nous avons simplement comparé les valeurs d'humidité $RH850$ aux trois points communs aux deux grilles d'extraction, de coordonnées $(-22,5^\circ E ; 45^\circ N)$, $(0^\circ E ; 45^\circ N)$, et $(22,5^\circ E ; 45^\circ N)$. Le résultat obtenu pour chacun des trois points est équivalent. Le nuage des points associés aux valeurs de $RH850$ extraites de l'archive ERA à $1,125^\circ$ et à $2,5^\circ$ est tracé sur la Figure 4.13, pour le point de coordonnées $(-22,5^\circ E ; 45^\circ N)$. Les données sont cohérentes puisque le nuage de point s'ajuste bien avec la première bissectrice, même si certains points sont très éloignés. Le coefficient de détermination R^2 est d'environ 0,9, ce qui justifie une bonne corrélation entre les données.

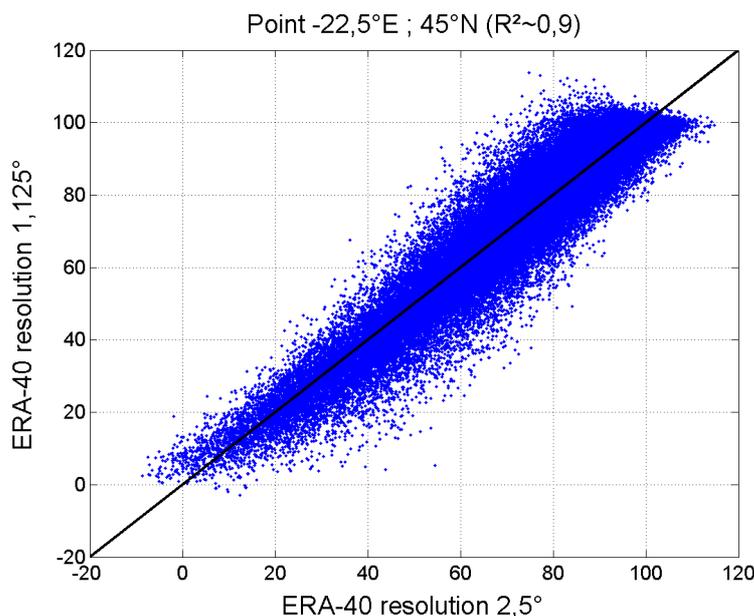


Figure 4.13 : Comparaison entre les valeurs de $RH850$ extraites de l'archive ERA à $1,125^\circ$ (ordonnée) et de l'archive ERA à $2,5^\circ$ (abscisse), sur la période 1957-2002, au point de coordonnées $(-22,5^\circ E ; 45^\circ N)$.

En conclusion, des valeurs aberrantes existent dans les réanalyses ERA, que ce soit à résolution $2,5^\circ$ ou à résolution $1,125^\circ$, parfois en proportion significative. Cependant aucune alternative n'existe à l'heure actuelle pour éviter de manipuler ces données. Nous disposons donc des deux archives de réanalyses NNR et ERA contenant les variables d'analogie utilisées dans la méthode de référence A0, mais aussi d'autres variables que nous exploiterons dans la seconde partie de ce mémoire. Nous avons vu qu'il existait des différences entre les deux

archives, mais il faut faire un choix pour la suite de notre travail : la méthode d'analogie de référence est-elle sensible à l'archive de variables d'analogie ? L'utilisation des valeurs aberrantes d'humidité présentes dans l'archive ERA peut-elle avoir un impact sur la qualité d'une prévision adaptée par analogie ?

4.4 Sensibilité de la méthode de prévision par analogie à l'archive de réanalyses

Nous allons maintenant étudier la sensibilité de la méthode de référence A0 au choix de l'archive de variables d'analogie. Pour cela, nous travaillons dans un contexte de prévision parfaite, *i.e.* les situations cibles sont des situations non prévues, dont les champs météorologiques sont extraits des réanalyses. Les périodes cibles et candidates sont identiques et s'étendent du 01/08/1972 au 30/08/2002. Afin que la meilleure analogue à une situation cible ne soit pas la situation cible elle-même, nous nous interdisons de rechercher des situations analogues dans la même année que la situation cible.

4.4.1 Archives NCEP/NCAR et ERA-40 à 2,5°

Nous proposons ici d'évaluer les performances de prévision obtenues lorsque les archives de réanalyses NNR et ERA à résolution 2,5° sont utilisées. L'analyse de sensibilité a été menée à l'aide de deux scores d'évaluation : le *CRPSS* et le diagramme ROC (et le score *AROC* associé). En outre, les performances n'ont été évaluées que pour le bassin de la Saône. L'archive pluviométrique à présent utilisée est l'archive SAFRAN, et nous rappelons que dans ce contexte nous utilisons les précipitations transformées (cf. 2.2.2.3).

Les résultats présentés dans cette section ont fait l'objet d'un article (Ben Daoud et al., 2009c), que le lecteur trouvera en annexe 3 ; et d'une conférence (Ben Daoud et al., 2008b), dont l'article extrait des actes est reporté en annexe 4.

4.4.1.1 A l'aide du *CRPSS*

Nous avons appliqué la méthode de référence A0, dans un contexte de prévision parfaite, au bassin versant de la Saône. Il est nécessaire d'adapter la méthode à ce bassin en optimisant les domaines d'analogie et les nombres d'analogues à sélectionner.

Les domaines d'analogie aux niveaux 1 et 2 ont été optimisés pour ce bassin, et se révèlent finalement identiques à ceux optimisés par Bontron (2004) pour le regroupement pluviométrique R5 considéré dans ses travaux de thèse (cf. 1.3.1, Figures 1.4 et 1.5). Les nombres d'analogues *N1* et *N2* à retenir à l'issue des deux niveaux de sélection dépendent de la longueur des archives utilisées. Dans notre cas, ces nombres sont à ré-optimiser puisque nos archives ne sont pas identiques à celles utilisées par Bontron (2004).

Dans un premier temps, nous avons donc étudié la sensibilité de la méthode au choix de l'archive au travers de l'optimisation des nombres d'analogues *N1* et *N2*. Les performances

sont exprimées en termes de *CRPSS* pour l'ensemble du bassin de la Saône. Nous avons optimisé les nombres d'analogues en calculant le score *CRPSS* associé à chaque couple ($N1$, $N2$), pour un nombre $N1$ variant de 10 à 100 par pas de 10, et pour un nombre $N2$ variant de 5 à $N1$ par pas de 5, et ce en utilisant les archives ERA et NNR. Le score *CRPSS* maximal a été atteint pour un nombre $N1$ de 70 avec l'archive ERA, et de 80 avec l'archive NNR. Au niveau 2, le nombre d'analogues $N2$ optimal est de 25 avec les deux archives (Figure 4.14).

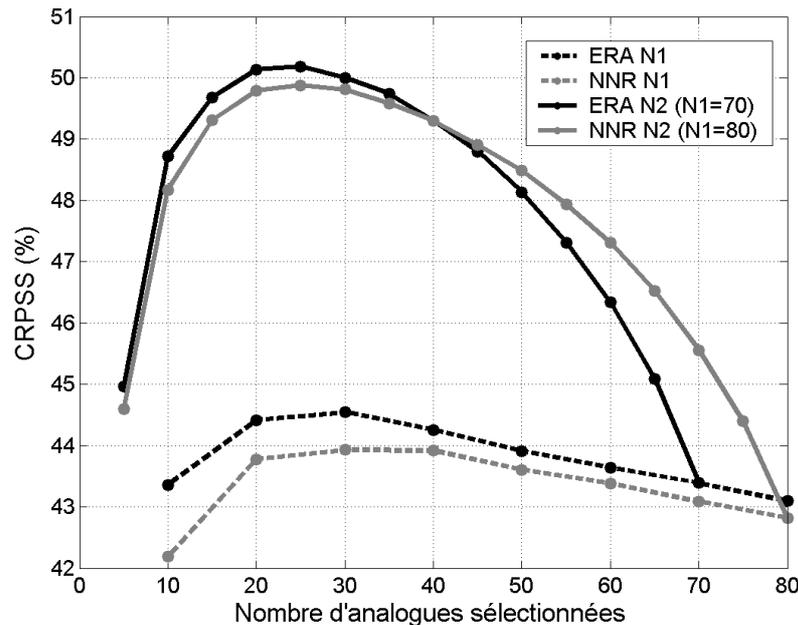


Figure 4.14 : Evolution du score *CRPSS* en fonction du nombre d'analogues sélectionnées au niveau 1 ($N1$) et au niveau 2 ($N2$), lorsque les archives de réanalyses ERA et NNR sont utilisées.

Bien que l'archive ERA contienne des valeurs irréalistes d'humidité, les performances de prévision semblent équivalentes avec les deux archives. Contrairement à ce que l'on aurait pu craindre, les performances au niveau 2 sont même légèrement supérieures avec l'archive ERA. Ce faible gain de performance avec ERA est également obtenu à l'issue du premier niveau d'analogie (courbes en trait discontinu). Par ailleurs, l'examen de la Figure 4.14 indique que le niveau 2 d'analogie est bien justifié du fait d'un gain de performances non négligeable apporté par son application (de l'ordre de 5 points de *CRPSS*).

4.4.1.2 A l'aide du diagramme ROC

L'étude de sensibilité a été complétée par l'analyse des diagrammes ROC, présentés au chapitre 3 (cf. 3.2.5). Un diagramme a été construit pour différents seuils de pluie : le seuil 0 mm (*i.e.* le seuil pluie/non pluie), et les seuils correspondant aux quantiles de pluie non nulle d'ordre 50, 70, 90 et 99 % ($PNN50$, $PNN70$, $PNN90$ et $PNN99$). Cependant les courbes obtenues lorsque les archives ERA et NNR sont utilisées ne présentent pas de différences significatives à l'échelle du diagramme. Pour distinguer les différences de performance, nous allons travailler avec le score *AROC* (cf. 3.2.5).

Supposons qu'en prévision opérationnelle, un utilisateur dispose des sorties d'un MNP différent du MNP utilisé pour les réanalyses, mais de bonne qualité pour les variables d'analogie. L'utilisateur pourrait-il rechercher dans ces réanalyses des situations analogues à la situation prévue sans que les performances ne soient dégradées ? Cette question a toute son importance : outre la qualité et la longueur limitées de l'archive de réanalyses, la qualité des variables d'analogie prévues reste l'une des limitations de la méthode en mode opérationnel. Bien que nous soyons placés dans un contexte de prévision parfaite, nous nous sommes également intéressés à l'impact sur les performances de l'utilisation d'une archive, dans laquelle on extrait les situations cibles, différente de celle dans laquelle on extrait les situations analogues. Nous avons ainsi envisagé les quatre cas suivants (Tableau 4.7) : situations cibles extraites d'ERA et analogues extraites d'ERA (*EE*), situations cibles extraites de NNR et analogues extraites de NNR (*NN*), situations cibles extraites d'ERA et analogues extraites de NNR (*EN*), situations cibles extraites de NNR et analogues extraites d'ERA (*NE*).

Tableau 4.7 : Valeurs du score *AROC* obtenues pour les différents seuils de pluie, lorsque les archives ERA (E) et NNR (N) sont utilisées. Le cas XY correspond aux situations cibles extraites de X et aux situations analogues extraites de Y. Le nombre de valeurs de chaque échantillon est également donné.

Cas	Valeur seuil				
	<i>0 mm</i>	<i>PNN50</i>	<i>PNN70</i>	<i>PNN90</i>	<i>PNN99</i>
<i>EE</i>	0,932	0,936	0,930	0,922	0,895
<i>NN</i>	0,930	0,935	0,930	0,922	0,888
<i>EN</i>	0,929	0,933	0,927	0,922	0,891
<i>NE</i>	0,927	0,932	0,927	0,923	0,897
$\Delta(\text{max-min})$	0,005	0,004	0,003	0,001	0,009

Concernant les cas *EE* et *NN*, aucune différence notable n'est observée entre les scores. Cependant, les performances sont légèrement supérieures lorsque l'archive ERA est utilisée, notamment pour les pluies extrêmes (supérieures au quantile *PNN99*). L'avantage de ce score est aussi de pouvoir discriminer la prévision entre deux événements : on remarque, lorsque les deux archives sont utilisées, que le score *AROC* maximal est atteint pour le seuil *PNN50*, c'est-à-dire pour les situations dont les cumuls pluviométriques sont compris entre les quantiles *PNN50* et *PNN70* (i.e. pour les journées faiblement pluvieuses).

En outre, les résultats montrent qu'en général les performances, bien qu'elles soient équivalentes, diminuent légèrement si deux archives différentes sont utilisées pour les situations cibles et candidates. Un gain de performance est cependant obtenu pour les pluies les plus importantes. Ce résultat doit être relativisé compte tenu du faible nombre de valeurs de l'échantillon (environ 110 valeurs, contre environ 7000 valeurs pour le seuil *0 mm*). Ainsi si l'on dispose des sorties de plusieurs modèles, cette perspective en mode opérationnel s'avèrerait intéressante dans certaines conditions, par exemple si les sorties de l'un d'entre eux ne sont pas disponibles. L'utilisateur qui se trouverait dans un tel cas devra tout de même effectuer une série de tests via les sorties archivées des MNP opérationnels.

4.4.2 Archive ERA-40 aux résolutions 2,5° et 1,125°

Malgré la présence dans l'archive ERA de valeurs aberrantes d'humidité relative à 850 hPa, nous venons de voir que les performances de prévision sont légèrement plus élevées lorsque les variables d'analogies, décrivant les situations cibles et candidates, sont extraites de l'archive ERA plutôt que de l'archive NNR.

Cette archive serait donc à privilégier, d'autant plus que les champs sont disponibles à une plus haute résolution (1,125°). Une des perspectives évoquée par Bontron (2004) était l'utilisation de l'archive ERA à 1,125° au second niveau d'analogie, étant donnée l'augmentation des performances avec la résolution des champs des variables d'analogie (cf. pages 228-229 de son mémoire de thèse). En revanche au premier niveau de sélection, l'augmentation de la résolution (*i.e.* la diminution de la taille du maillage) entraînerait selon lui peu de gain de performance sur l'analogie à échelle synoptique.

Nous proposons ici d'étudier la sensibilité de la méthode A0 à la résolution des champs météorologiques extraits des réanalyses ERA et ainsi de valider l'hypothèse émise par Bontron (2004). L'objectif est d'appliquer la méthode de référence au bassin de la Saône en utilisant les variables *RH850* et *TCW* non plus définies à résolution 2,5°, mais définies à résolution 1,125°. Bien que les valeurs de ces variables soient issues des mêmes calculs, les champs extraits des réanalyses peuvent être différents selon la résolution, comme nous l'avons remarqué précédemment pour la variable *RH850* (Figure 4.13). Des incohérences implicites pourraient ainsi se produire entre les situations analogues retenues au niveau 1 (caractérisées à 2,5°) et celles finalement retenues au niveau 2. C'est la raison pour laquelle nous allons dans un premier temps utiliser les variables d'analogie du niveau 1 définies à 1,125°, puis nous comparerons les performances obtenues lorsque les variables sont définies à 2,5° ou à 1,125° au niveau 1 et/ou au niveau 2.

Dans cette section, les nombres d'analogues retenues sont ceux que nous avons optimisés dans la section précédente pour l'archive ERA ($N1=70$, $N2=25$), et les performances sont à présent exprimées en termes de *CRPSS*.

4.4.2.1 Au cours de la première sélection

La résolution des champs des variables d'analogie n'étant plus identique, il est nécessaire d'optimiser le domaine d'analogie du niveau 1. Le domaine d'analogie associé à la première sélection a ainsi été optimisé à 1,125° avec un nombre $N1$ égal à 70, et ce pour l'ensemble des sous-bassins de la Saône. Le domaine optimal à 1,125° est représenté sur la Figure 4.15. A titre de comparaison, le domaine appliqué à 2,5° est également tracé. Le domaine optimisé à 1,125° est moins étendu que le domaine appliqué à 2,5°, même s'il reste suffisamment grand pour tenir compte de la circulation atmosphérique à échelle synoptique.

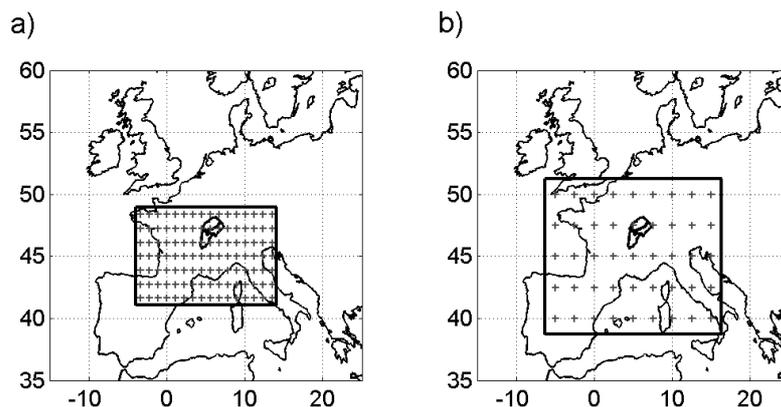


Figure 4.15 : Domaine d'analogie du niveau 1 optimisé pour l'ensemble des sous-bassins de la Saône et grille correspondante. La résolution des champs des variables est de a) 1,125° ; b) 2,5°. Les croix correspondent aux centres des mailles carrées.

Les performances obtenues lorsque les champs des géopotentiels sont décrits aux résolutions 1,125° et 2,5° sont indiquées dans le Tableau 4.8, pour chaque sous-bassin et pour l'ensemble des sous-bassins. Un gain de performance faible mais non négligeable est ainsi obtenu lorsque l'analogie est faite à 1,125° par rapport à l'analogie à 2,5°. Le gain G de performance moyenne, exprimé entre l'utilisation des champs à 1,125° par rapport à celle des champs à 2,5°, s'élève à 1,57 % pour l'ensemble des sous-bassins.

Tableau 4.8 : Performances (CRPSS (%)) obtenues à l'issue de la première sélection lorsque l'analogie est faite à 1,125° et à 2,5°, et gains G de performance correspondants.

	Résolution 1,125°	Résolution 2,5°	G (1,125°/2,5°) (%)
Doubs	45,01	44,05	1,72
Petite Saône	44,86	43,92	1,68
Saône Intermédiaire	42,98	42,19	1,37
Ensemble des sous-bassins	44,28	43,39	1,57

Ce gain de performance modéré peut s'expliquer par l'utilisation de données initialement fournies à 1,125°. Bontron (2004) avait étudié l'évolution des performances en fonction de la taille de la maille élémentaire à partir des données fournies à résolution 2,5° (réanalyses NNR). Il avait appliqué un domaine d'analogie de taille identique en considérant successivement l'ensemble des points de grille (2,5°x2,5°), un point sur deux (5°x5°), un point sur trois (7,5°x7,5°) et enfin un point sur six (15°x15°). Il en avait finalement déduit qu'il était peu probable que l'emploi de données définies à résolution 1° génère une hausse vraiment sensible des performances. Or dans notre cas, les réanalyses ERA ont été générées par un modèle de troncature supérieure à celle du modèle utilisé pour l'archive NNR (cf. 2.2.1). Les données de cette dernière ne sont donc pas dégradées par rapport à celles de l'archive ERA, comme c'est le cas entre les données de l'archive ERA à 2,5° et à 1,125°. Nous aurions pu nous attendre à des conclusions équivalentes à celles de Bontron (2004) si nous avions

effectué la même opération à partir des données décrites à $1,125^\circ$.

4.4.2.2 Au cours de la seconde sélection

A l'issue de la première sélection, le gain de performance se révèle faible mais non négligeable, ce qui laisse supposer un gain de performance encore plus significatif à l'issue de la seconde sélection lorsque l'analogie est recherchée à résolution $1,125^\circ$.

Le domaine d'analogie du niveau 2 a été optimisé pour différentes résolutions des variables d'analogie, selon les deux cas suivants :

- niveau 1 à $2,5^\circ$ et niveau 2 à $1,125^\circ$;
- niveau 1 à $1,125^\circ$ et niveau 2 à $1,125^\circ$.

Le domaine optimal est finalement identique dans les deux cas. La résolution des champs de géopotential n'a donc pas d'impact significatif sur la fenêtre spatiale sur laquelle l'analogie sur l'humidité est recherchée. Les domaines obtenus lorsque les variables du second niveau sont définies à $1,125^\circ$ et lorsque les variables des deux niveaux sont définies à $2,5^\circ$ sont représentés sur la Figure 4.16. A $1,125^\circ$, le domaine apparaît un peu plus petit qu'à $2,5^\circ$, avec 3×3 points de grille (soit un domaine carré de $2,25^\circ$ de côté).

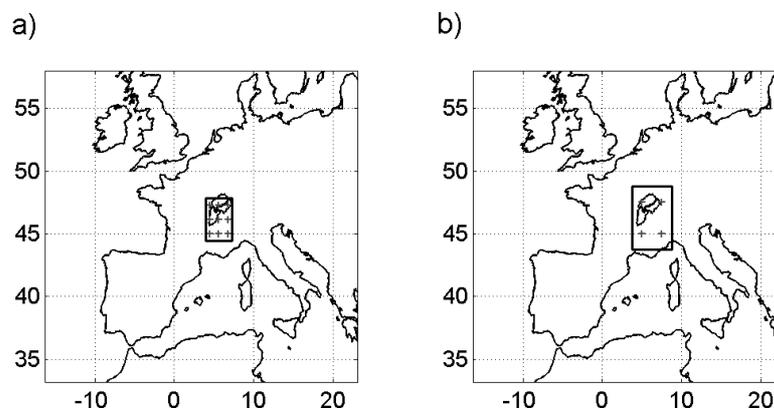


Figure 4.16 : Domaine d'analogie du niveau 2 optimisé pour l'ensemble des sous-bassins de la Saône et grille correspondante. La résolution des champs des variables est de a) $1,125^\circ$; b) $2,5^\circ$. Les croix correspondent aux centres des mailles carrées.

Les performances (*CRPSS*) obtenues lorsque les variables des deux niveaux sont définies à $2,5^\circ$ ($2,5^\circ/2,5^\circ$), ainsi que celles obtenues lorsque les variables du niveau 1 sont définies à $2,5^\circ$ ou à $1,125^\circ$ et les variables du niveau 2 à $1,125^\circ$ ($2,5^\circ/1,125^\circ$ et $1,125^\circ/1,125^\circ$), sont reportées dans le Tableau 4.9. Le gain final est décevant, avec une valeur de 1,65 % pour l'ensemble des sous-bassins lorsque toutes les variables sont définies à $1,125^\circ$ au lieu de $2,5^\circ$. Le gain *G* obtenu sur le bassin de la Saône Intermédiaire (2,15 %) est tout de même significatif. Le Tableau 4.9 révèle également qu'il peut exister des incohérences entre les dates des analogues retenues au niveau 1 avec la résolution $2,5^\circ$ et les dates sélectionnées au

1 Nous aurions alors considéré successivement des domaines élémentaires de $1,125^\circ \times 1,125^\circ$, $2,25^\circ \times 2,25^\circ$, $4,5^\circ \times 4,5^\circ$, etc.

niveau 2 à 1,125°. En effet, le score *CRPSS* évalué pour le bassin de la Petite Saône est plus faible lorsque le niveau 2 est appliqué à 1,125° (cas 2,5°/1,125° dans le Tableau 4.9) que lorsque les deux niveaux sont appliqués à 2,5° (cas 2,5°/2,5°).

Tableau 4.9 : Performances (*CRPSS* (%)) obtenues à l'issue de la seconde sélection sur le domaine optimal, lorsque l'analogie est faite à 2,5° aux deux niveaux (2,5°/2,5°), à 2,5° au premier et à 1,125° au second niveau (2,5°/1,125°), et à 1,125° aux deux niveaux (1,125°/1,125°).

	2,5°/2,5°	2,5°/1,125°	1,125°/1,125°	G max (%)
Doubs	51,38	51,61	52,13	1,54
Petite Saône	50,82	50,78	51,45	1,28
Saône Intermédiaire	48,35	49,12	49,46	2,15
Ensemble des sous-bassins	50,19	50,5	51,01	1,65

Finalement, contrairement à ce que nous pouvions nous attendre, la recherche de l'analogie sur les champs définis à résolution 1,125° au lieu de 2,5° n'entraîne pas de gain substantiel de performance.

4.5 Conclusion

L'objectif de ce chapitre était d'analyser, de critiquer et de comparer les différentes bases de données recueillies pour le développement de la méthode d'analogie : les archives pluviométriques d'une part, et les archives de réanalyses météorologiques d'autre part.

Nous avons évalué la cohérence entre les deux archives MF/CNR et SAFRAN contenant les mesures historiques de précipitations sur le bassin de la Saône, et ce sur la période 01/01/1992-31/12/2003 commune aux périodes couvertes par ces deux bases de données. L'analyse de la corrélation des lames d'eau à différents pas de temps (6 h, 12 h et 24 h) a révélé une cohérence assez faible aux pas de temps infrajournaliers, mais plutôt satisfaisante au pas de temps de 24 h. Toutefois l'incohérence au pas de temps de 6 h ne devrait pas troubler la suite de notre travail. En effet, les bassins versants de la Saône et de la Seine n'ont pas un temps de réaction aux précipitations aussi rapide que celui des bassins du Sud-Est de la France, sur lesquels Bontron (2004) avait optimisé et validé la méthode de référence A0. La prévision des précipitations au pas de temps journalier sur ces bassins tests devrait donc être suffisante pour améliorer la prévision des crues.

Les archives ERA et NNR contenant les champs des variables d'analogie ont également été comparées (Ben Daoud et al., 2008a). Nous avons analysé les différences entre les valeurs des variables d'analogie de la méthode de référence A0. Les champs des variables *Z500*, *Z1000* et *TCW* sont relativement similaires, même si des différences plus importantes sont observées sur les massifs montagneux et dans le bassin méditerranéen. En revanche, nous avons observé d'importants écarts de la variable *RH850* entre les archives. Après une étude plus

approfondie, des valeurs aberrantes d'humidité ont été détectées dans l'archive ERA¹, que ce soit à résolution 2,5° ou à résolution 1,125° : certaines valeurs sont inférieures à 0 % et d'autres sont supérieures à 100 %. Les valeurs irréalistes d'humidité détectées à 1,125° ont ensuite été caractérisées. Les anomalies négatives d'humidité sont finalement assez peu nombreuses, en revanche les anomalies positives représentent une proportion non négligeable. Nous avons cherché à corriger ces valeurs, mais aucune solution simple et objective ne s'est révélée efficace pour la suite de notre travail. Cette analyse est détaillée dans Ben Daoud et al. (2009a) (cf. annexe 2).

Enfin, nous avons étudié la sensibilité de la méthode de référence au choix de l'archive de réanalyses, d'abord en évaluant les performances obtenues avec les deux archives ERA et NNR à résolution 2,5°. Nous en avons profité pour ré-optimiser les nombres $N1$ et $N2$ d'analogues à sélectionner, qui s'élèvent à 70 et 25 lorsque l'archive ERA est utilisée, et 80 et 25 lorsque l'archive NNR est utilisée. Cette analyse de sensibilité a finalement montré qu'en dépit des valeurs aberrantes contenues dans l'archive ERA, l'utilisation de cette dernière conduisait à de meilleures performances que celle de l'archive NNR, que ce soit au premier ou au second niveau de sélection. Toutefois nous ne pouvons quantifier l'impact de la présence d'anomalies d'humidité sur les performances de prévision. Les résultats de cette étude ont été présentés à une conférence et dans un article (Ben Daoud et al., 2008b ; Ben Daoud et al., 2009c) (cf. annexes 3 et 4).

Ensuite nous avons étudié la sensibilité de la méthode de référence à la résolution des champs des variables d'analogie extraits de l'archive ERA. Selon Bontron (2004), l'usage d'une archive possédant une résolution de 1°x1° pour les variables du second niveau pourrait améliorer la qualité des prévisions. Cependant, nous avons constaté que le gain de performance s'avère peu significatif lorsque les variables des deux niveaux de sélection sont définies à résolution 1,125°. Ce résultat pourrait s'expliquer par le fait que les données des réanalyses ERA ont été générées à résolution 1,125°. Ainsi une partie du gain apporté par cette meilleure résolution est déjà obtenue lorsque les champs sont définis à 2,5°. C'est notamment pour cette raison que les performances sont plus élevées lorsque l'archive ERA est utilisée au lieu de l'archive NNR. En fin de compte, la résolution 2,5° des champs météorologiques contenus dans l'archive ERA semble suffisante pour rechercher des situations analogues. Ainsi nous utiliserons des variables d'analogie définies à 2,5° pour développer l'algorithme de prévision, quitte à employer la résolution 1,125° pour certaines variables lorsque la phase de développement sera achevée.

A l'issue de cette analyse exploratoire des bases de données, nous avons décidé de retenir les archives suivantes pour le développement de la méthode d'analogie :

- archive pluviométrique : SAFRAN, puisque sa longueur (environ 34 ans en début de thèse) est bien supérieure à celle de l'archive MF/CNR (12 ans), et les cumuls

¹ Pour information, le problème a été posé au CEPMMT, et cela proviendrait du système d'analyse (communication P. Bechtold). La nouvelle génération de réanalyses (ERA-Interim et ERA_extended, cf. 2.2.1.3) ne devrait plus contenir d'anomalies d'humidité.

journaliers semblent cohérents avec ceux de l'archive MF/CNR ;

- archive de variables d'analogie : ERA, étant donné que :
 - cette archive est plus riche ;
 - les performances de la méthode A0 obtenues avec cette archive sont un peu plus élevées qu'avec l'archive NNR malgré les valeurs aberrantes d'humidité ;
 - et les champs sont disponibles à plus haute résolution (1,125°).

Finalement, la période sur laquelle l'algorithme sera développé s'étend du 01/08/1972 au 30/08/2002, les variables d'analogie potentielles pouvant être utilisées à quatre instants différents (00h, 06h, 12h et 18h TU).

Seconde partie

Développement de la méthode de prévision par analogie

Introduction à la partie II

L'objectif principal de cette partie est d'améliorer les performances de prévision de l'algorithme A0 de sélection d'analogues, construit par Bontron (2004) et qui nous servira de point de départ. Bien que cette méthode soit déjà performante sur les grands bassins français, nous tenterons de l'améliorer en tenant compte de la climatologie de ces bassins, notamment en caractérisant les fronts, qui sont responsables de la plupart des événements pluvieux sur nos secteurs d'intérêt. Pour y répondre, deux approches complémentaires sont possibles :

- i) Le développement. Il s'agit d'introduire de nouvelles variables ou de nouveaux critères d'analogie. Cela peut notamment aboutir à la création d'une nouvelle étape de sélection d'analogues.
- ii) L'optimisation. Cette voie d'amélioration suppose les familles auxquelles les variables d'analogie appartiennent et les critères d'analogie déjà fixés. Le travail consiste alors à optimiser les paramètres : domaine d'analogie (taille, forme, localisation géographique), heures et niveaux de mesure des variables d'analogie, nombres d'analogues à sélectionner (variabilité des nombres par exemple).

Dans le cadre de cette thèse, nous avons décidé d'adopter la première approche en nous inspirant dans un premier temps du comportement du prévisionniste en météorologie. Lorsqu'il débute son activité, son premier réflexe est d'analyser la situation synoptique, c'est-à-dire de localiser les centres d'action qui seront *a priori* responsables de phénomènes tels qu'un front pluvieux, des cellules orageuses, des chutes de neige, etc. En outre, le prévisionniste s'intéresse à d'autres éléments secondaires susceptibles de favoriser le développement de ces phénomènes : convergence des vents, contraste de masses d'air, anomalies et advection de tourbillon¹, énergie convective, etc. Nous nous efforcerons ainsi, dans le cadre de ce travail, à rendre objective l'analyse du prévisionniste, qui par nature est subjective. Cette tâche peut paraître intuitive pour un expert, mais elle n'en est pas moins difficile à formaliser tant la subjectivité de l'analyse est forte. La première étape de notre travail sera d'identifier les variables descriptives d'une situation météorologique. Selon notre expertise, ces variables jouent un rôle dans la formation et dans l'évolution des systèmes précipitants. Nous effectuerons ensuite une série de tests afin de rendre la méthode de prévision plus utile, et donc plus performante. Dans un second temps, nous adopterons la

¹ Le lecteur curieux de connaître la définition de cette variable se reportera au chapitre 5 (section 5.1.1.2).

seconde approche pour les étapes de sélection de la méthode de référence.

A chaque étape du développement, la performance sera mesurée en termes de gain de performance G d'une méthode A par rapport à une méthode B, tel que nous l'avons défini dans la partie I (section 3.1.3). Le bassin de la Saône, présenté dans la partie I (section 2.1.1), servira de bassin test pour mener les travaux de développement. Le bassin de la Seine quant à lui servira de bassin test pour « valider » les méthodes développées. Sauf avis contraire, le gain G indiqué au cours des développements sera évalué pour l'ensemble des sous-bassins.

Dans le cadre du développement de l'algorithme de sélection, nous nous affranchirons des erreurs de prévision des variables d'analogie que peut faire un MNP. Nous nous placerons ainsi dans un contexte de prévision parfaite, comme nous l'avons fait lors de l'étude de sensibilité au choix de l'archive de réanalyses. Toutes les situations cibles ne sont donc prévues par un MNP, mais directement tirées des réanalyses ERA.

Choix d'une période de calibration et d'une période de validation

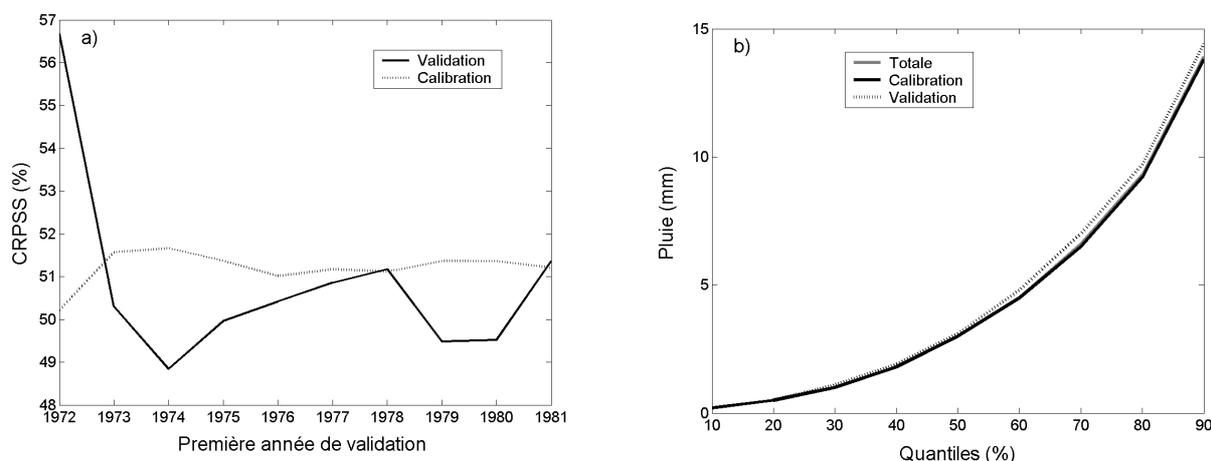
Au cours de chaque phase du développement de la méthode, nous chercherons à valider les résultats que nous aurons obtenus sur un échantillon de situations cibles indépendant de celui utilisé pour le développement. Il nous faut donc choisir une période cible pour la calibration de l'algorithme, et une autre pour la validation. La période commune aux périodes couvertes par les archives ERA et SAFRAN est d'environ 30 ans (01/08/1972-30/08/2002). Nous avons choisi une durée de 25 ans pour calibrer l'algorithme et une durée de 5 ans pour le valider.

La démarche intuitive consiste à conserver les cinq dernières années pour la validation de l'algorithme. Mais dans ce cas, si un changement était détecté dans nos données (dû à un changement dans les appareils de mesure, à la procédure pour la construction des réanalyses ou même à un changement climatique), un biais pourrait apparaître dans les résultats. De plus, l'algorithme étant validé sur un échantillon de situations cibles relativement restreint, la climatologie de ces situations pourrait avoir une incidence sur les résultats si elle diffère de celle des situations de la période de calibration.

Afin de nous affranchir de ces problèmes éventuels, nous avons décidé de choisir astucieusement les deux sous-périodes. Nous avons décidé de tirer une année tous les cinq ans pour rendre la validation homogène. Notre période couvrant environ 30 ans, dix choix sont possibles : les périodes de validation dont la première et la dernière année sont 1972 et 1992, 1973 et 1993, 1974 et 1994, ... , et 1981 et 2001. Les dix couples possibles de périodes de validation et de calibration ont ainsi été considérés. Pour chacun d'entre eux nous avons évalué, pour le bassin de la Saône, les performances ($CRPSS$) moyennes obtenues sur la période de calibration et sur la période de validation (figure suivante a)).

L'homogénéité serait parfaite si les deux courbes se superposaient. Le score $CRPSS$ calculé sur les différentes périodes de calibration varie peu, contrairement à celui calculé sur les périodes de validation. Ce résultat s'explique par la faible longueur de la période de

validation, rendant ainsi le résultat fortement dépendant des années choisies. La différence la plus faible de scores *CRPSS* entre les deux sous-périodes est obtenue pour la période de validation dont la première année est 1978. De plus, nous nous sommes assurés que la climatologie des pluies est la même au cours des périodes de calibration et de validation. Pour ce faire, nous avons calculé des statistiques pour chacun des bassins et au cours des périodes de validation (débutant par l'année 1978) et de calibration associée, et au cours de la période totale. Ces statistiques sont la proportion de jours secs (cf. tableau suivant) et les quantiles de pluie non nulle *PNN10* à *PNN90*. Pour ces derniers, les résultats sont similaires pour les trois sous-bassins. L'illustration représentée par la figure b) est valable ici pour le bassin du Doubs.



a) Evolution des performances (*CRPSS*) moyennes obtenues sur les périodes de calibration et de validation, en fonction de la première année de la période validation. b) Comparaison des quantiles au non-dépassement de pluie non nulle d'ordre 10 à 90% calculés sur les périodes de validation et de calibration ainsi que sur la période totale, pour le bassin du Doubs, lorsque la période de validation débute en 1978.

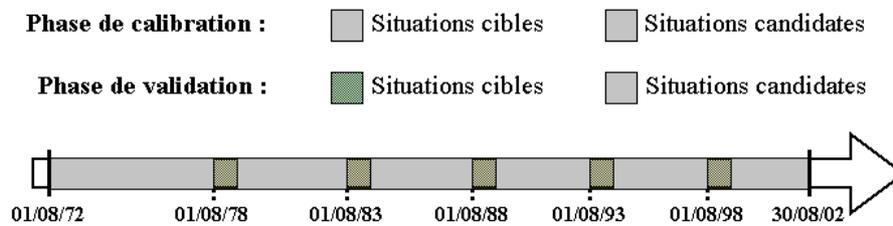
Proportion (%) de jours secs calculée pour chaque sous-bassin et sur les périodes de validation (1978, 1983, 1988, 1993, 1998), de calibration (1972-2002 sans les années de la période de validation), et sur la période totale.

Sous-bassin	Période cible		
	Validation	Calibration	Totale
Doubs	33,3	32,8	32,9
Petite Saône	35,3	35,8	35,7
Saône Int.	38,3	36,5	36,8

Les différences obtenues entre la période de validation débutant en 1978 et la période de calibration correspondante ne sont pas significatives, avec une proportion de jour secs qui ne diffère pas de plus de 1,5 % entre les deux sous-périodes et la période totale. Les quantiles au-non dépassement sont également semblables (figure b), ce qui laisse transparaître une certaine homogénéité dans l'archive pluviométrique.

Pour les travaux de développement que nous entreprendrons dans cette partie, nous allons donc travailler sur la période de calibration 1972-2002 dont les années 1978, 1983, 1988,

1993 et 1998 sont exclues. En phase de calibration, les situations cibles et candidates sont extraites de la période de calibration ; tandis qu'en phase de validation, les situations cibles sont extraites de la période de validation et les situations candidates de la période de calibration. Cette procédure est décrite par la figure suivante :



Périodes dans lesquelles sont extraites les situations cibles et candidates en phase de calibration et en phase de validation.

5 Analyse préliminaire des variables d'analogie potentielles

Les variables météorologiques produites dans les réanalyses sont nombreuses. Elles ne sont pas toutes indispensables pour décrire les phénomènes physiques responsables de la formation des précipitations. Un tri préliminaire appuyé par notre expertise et notre expérience de prévisionniste, nous a permis de retenir les variables qui nous semblent les plus informatives pour décrire une situation météorologique. Certaines sont liées à d'autres par des relations physiques et mathématiques. Or notre objectif est de développer la méthode de prévision par analogie en introduisant de nouvelles variables. Si ces dernières sont liées à une autre variable déjà utilisée dans notre méthode de référence, leur apport ne sera probablement pas substantiel. Nous chercherons par conséquent à les écarter des autres variables d'analogie potentielles.

L'objectif de ce chapitre est donc d'analyser les liaisons entre les variables d'analogie potentielles. Dans un premier temps, nous présenterons ces variables en donnant leur définition et éventuellement leur mode de calcul afin de permettre au lecteur de comprendre leur rôle et leur intérêt en prévision météorologique. Elles peuvent être séparées en deux groupes : les variables brutes, directement extraites des réanalyses ; et les variables élaborées qui, comme leur nom l'indique, sont construites à partir d'une ou plusieurs variables brutes. Les liaisons entre les variables seront ensuite analysées. Cette analyse nous permettra finalement d'écarter les variables les plus liées aux variables de la méthode de référence A0 (géopotentiels à 500 et 1000 hPa, humidité relative à 850 hPa et quantité d'eau précipitable). Nous avons choisi des variables d'analogie potentielles sans tenir compte de leur prévisibilité opérationnelle. En effet, l'intérêt d'utiliser des variables peu robustes est assez limité. Dans un futur relativement proche, nous pouvons espérer une amélioration de la prévision de certaines variables. Pour le moment, nous nous en affranchissons et nous retiendrons les variables qui semblent informatives et descriptives d'une situation météorologique.

5.1 Définitions

5.1.1 Variables brutes

Les variables brutes sont des variables directement produites par le MNP dans les réanalyses. Nous en avons choisi sept, dont l'abréviation et la nomenclature usuelle sont reportées dans le Tableau 5.1. Celles dotées d'un astérisque sont définies sur des niveaux de pression.

Tableau 5.1 : Liste des variables brutes analysées : abréviation et nomenclature usuelle correspondante. Les variables dotées d'un astérisque sont des variables mesurées à des niveaux de pression.

Abréviation	Nomenclature usuelle
Z^*	Altitude géopotentielle d'un niveau de pression
$Z-2PVU$	Altitude géopotentielle au niveau de tourbillon potentiel 2PVU
RH^*	Humidité relative
T^*	Température
W^*	Vitesse verticale
DIV^*	Divergence horizontale
VT^*	Vitesse du vent horizontal

Nous supposons que le lecteur connaît les définitions de la température T et du vent VT . Nous allons à présent donner les définitions des variables brutes restantes afin de comprendre leur rôle et leur intérêt pratique en prévision météorologique. Pour des définitions plus précises, le lecteur pourra par exemple se référer à l'ouvrage de Malardel (2005).

5.1.1.1 Altitude géopotentielle d'un niveau de pression (Z)

La constante de gravité, spécifique à la Terre, varie avec l'altitude, à l'instar de la pression atmosphérique. Toute particule présente dans l'air possède ainsi une énergie potentielle. Le géopotential $\phi(h)$ d'une parcelle d'air de masse unitaire, située à l'altitude h , correspond à l'énergie potentielle gagnée par cette parcelle entre le niveau de la mer (altitude 0 m) et l'altitude h . La définition mathématique est indiquée par l'équation 5.1 :

$$\phi(h) = \int g(\varphi, z) \cdot dz \quad (5.1)$$

où g est l'accélération de la pesanteur, φ est la latitude et z est l'altitude.

On en déduit alors l'altitude géopotentielle Z définie par l'équation 5.2 :

$$Z = \frac{\phi}{g_0} \quad (5.2)$$

où g_0 est l'accélération de la pesanteur au niveau de la mer. On peut donc définir l'altitude géopotentielle ZP d'une parcelle d'air située à un niveau de pression P comme étant l'altitude à laquelle la pression atmosphérique est égale à P . Son unité est le « mètre géopotential » (m_{gp}). Sous nos latitudes, on peut indiquer les valeurs moyennes des altitudes géopotentielles. Aux niveaux de pression 1000, 850 et 500 hPa, ces valeurs sont respectivement de l'ordre de 0 m, 1500 m et 5500 m. Notons également que, par abus de langage, on parle souvent de géopotential pour désigner l'altitude géopotentielle, terme que nous utiliserons dans ce manuscrit pour simplifier l'écriture.

Intérêt pratique

Le champ du géopotential Z d'une surface isobare donnée est couramment visualisé en prévision météorologique. Il permet notamment de situer ses anomalies, c'est-à-dire les différents centres d'action¹, et par conséquent d'identifier les grandes lignes de la circulation atmosphérique. Le niveau de pression du géopotential le plus souvent analysé est le niveau 500 hPa (niveau standard). On considère également qu'à ce niveau de pression, les lignes d'égal géopotential, encore appelées isohypses, représentent assez bien la direction du vent réel, alors assimilé au vent géostrophique. L'analyse du champ de géopotential à 500 hPa ($Z500$) est généralement couplée à l'analyse d'un champ de pression ou de géopotential proche du sol, tel que le champ de pression au niveau de la mer. Cela permet d'associer les centres d'action en altitude à ceux proches de la surface.

5.1.1.2 Altitude géopotentielle au niveau de tourbillon potentiel 2PVU (Z-2PVU)

Nous venons de définir l'altitude géopotentielle d'un niveau de pression. De manière générale, on peut associer l'altitude géopotentielle à une variable météorologique quelconque, à partir du moment où la valeur de cette variable est fixée. Ainsi on peut définir l'altitude géopotentielle d'une valeur du tourbillon potentiel (TP). Nous n'entrerons pas dans le calcul du tourbillon TP puisque cette variable fait intervenir de nombreuses notions de dynamique et de thermodynamique atmosphérique. Néanmoins, il est bon de savoir que ce paramètre représente, sous l'hypothèse adiabatique², l'effet de la rotation de la Terre sur la circulation de l'air atmosphérique. L'unité du système international est le $rad.m^2.K.s^{-1}.kg^{-1}$. Pour simplifier, on appelle PVU (Potential Vorticity Unit) cette unité.

Ainsi, à l'instar de la pression atmosphérique, si l'on définit une surface d'égal TP (par exemple 1,5 ou 2 PVU), on peut caractériser le champ de TP égal à 2 PVU par le géopotential de la surface 2 PVU ($Z-2PVU$). Sous nos latitudes, l'ordre de grandeur de la variable $Z-2PVU$ est d'environ 10 000 m, soit le niveau moyen de la tropopause, limite entre les deux premières couches de l'atmosphère terrestre.

1 Un centre d'action désigne soit un anticyclone ou une dorsale anticyclonique, qui représente une zone de hautes altitudes géopotentielles (anomalie positive) ; soit une dépression ou un talweg, qui représente une zone de basses altitudes géopotentielles (anomalie négative).

2 Un processus est dit adiabatique lorsque l'on considère que les particules d'air n'échangent pas de chaleur avec leur environnement au cours de leur mouvement.

Intérêt pratique

Le champ du géopotential $Z-2PVU$ permet notamment de localiser les zones d'anomalie négative, *i.e.* les zones de faibles valeurs de tourbillon TP . En d'autres termes, il permet de localiser les zones dans lesquelles la tropopause est plus basse que dans leur environnement. L'abaissement du niveau de la tropopause est dû à l'intrusion d'air stratosphérique dans la partie haute de la troposphère. L'intérêt principal de cette variable réside dans l'identification d'une zone, située en aval d'une anomalie de tropopause, où la cyclogénèse et/ou la frontogénèse est susceptible de se produire en raison du forçage induit par la présence de l'anomalie. Ce forçage se traduit par des mouvements verticaux ascendants : par conséquent, un front pluvieux peut se former ou évoluer à l'avant d'une anomalie. Le géopotential $Z-2PVU$ est principalement utilisé pour la prévision des précipitations convectives. Il n'est cependant ni nécessaire ni suffisant qu'une anomalie de tropopause marquée soit identifiée pour que des précipitations se produisent.

5.1.1.3 Humidité relative (RH)

L'air atmosphérique contient généralement de l'eau sous forme vapeur. La quantité de vapeur d'eau dans l'air peut être calculée de différentes façons. L'une des définitions les plus courantes est l'humidité relative (RH). Cette variable mesure le rapport entre le contenu en vapeur d'eau de l'air et la capacité maximale de l'air à en contenir dans des conditions de température et de pression donnée. Elle n'a donc pas d'unité, mais est généralement indiquée en %. La formule de la variable RH est donnée par l'équation 5.3 :

$$RH = 100 \cdot \frac{e}{e_{sat}(T)} \quad (5.3)$$

où e est la tension (ou pression) partielle de vapeur d'eau, $e_{sat}(T)$ est la tension de vapeur saturante à la température T .

Intérêt pratique

La variable RH indique quel est le taux d'humidité de l'air. A un niveau de pression donné par exemple, elle nous renseigne notamment sur la possibilité de formation de nuages. Si l'humidité est de 100 %, alors on considère que l'air est saturé, et donc qu'il y a probablement présence de nuages. Cette variable est typiquement utilisée, parmi d'autres, pour anticiper la formation de brouillard. Ainsi, si l'air est saturé près du sol, la probabilité d'observer du brouillard est élevée. Dans le cas des précipitations, la variable RH doit être légèrement supérieure à 100 % pour que les gouttelettes d'eau nuageuse puissent grossir et ainsi précipiter. On considère alors que l'air est sur-saturé.

5.1.1.4 Vitesse verticale (W)

L'air est un fluide en mouvement dans l'espace tridimensionnel. Le mouvement de l'air a donc plusieurs composantes : le mouvement horizontal caractérisé par la vitesse et la direction du vent (VT) ; et le mouvement vertical, caractérisé par sa vitesse verticale (W). En moyenne, la vitesse verticale est négligeable devant la vitesse du vent. Néanmoins cette composante du mouvement joue un rôle fondamental dans la formation et l'évolution de certains phénomènes météorologiques. La vitesse verticale peut être déduite, à un instant donné, par le calcul du gradient de pression atmosphérique d'une particule d'air par unité de temps. Son unité est donc le $Pa.s^{-1}$. Compte tenu de la diminution de pression avec l'altitude, la vitesse d'une particule d'air en mouvement ascendant (resp. subsident) est négative (resp. positive).

Intérêt pratique

Comme nous l'avons mentionné en introduction à ce mémoire, les mouvements verticaux dans l'atmosphère sont à l'origine de la formation des nuages et donc des précipitations. Par conséquent, la variable W permet de localiser les zones d'instabilité, *i.e.* les zones d'ascendance, responsables de la formation des fronts. A l'inverse, une zone de vitesse verticale positive induit un état de stabilité de l'atmosphère. Par ailleurs, les zones d'ascendance et de subsidence sont couplées, indissociables. C'est pour cette raison que, fort heureusement, notre ciel est généralement alterné de nuages et de ciel clair. La variable W est notamment utilisée en météorologie pour la prévision des orages. C'est néanmoins une variable fortement dépendante du MNP utilisé en raison de sa forte relation avec d'autres paramètres sensibles aux schémas décrits dans le MNP.

5.1.1.5 Divergence horizontale (DIV)

Le mouvement horizontal d'une particule d'air est caractérisé par la vitesse VT et la direction (U, V) du vent horizontal. Le vecteur vent \vec{V} est défini par ses composantes méridienne U et zonale V (équation 5.4).

$$\vec{V} = U \cdot \vec{u} + V \cdot \vec{v} \quad (5.4)$$

La divergence DIV du vent horizontal en un point de l'espace est un réel qui caractérise la structure spatiale du champ de vent dans l'environnement de ce point. Elle est définie, dans un repère cartésien (x, y), par la relation mathématique :

$$DIV(\vec{V}) = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} \quad (5.5)$$

Son unité est le s^{-1} . La valeur que peut prendre la variable DIV peut être négative (air convergent) ou positive (air divergent).

Intérêt pratique

En météorologie, un anticyclone est caractérisé en son centre par une divergence positive, tandis qu'une dépression est caractérisée en son centre par une divergence négative (convergence). Les éléments significatifs tels que les dorsales ou les talwegs sont caractérisés de la même façon. Ainsi un talweg de surface, au passage duquel une perturbation frontale est généralement observée, est associé à une zone de convergence de l'air (*DIV* négative). Cette variable est également utilisée en prévision convective dans les basses couches comme dans les hautes couches de la troposphère. Un air convergent dans les basses couches peut favoriser, dans certaines conditions de température et d'humidité, la formation de nuages convectifs. Une zone de divergence positive située en amont d'un jet (zone de vent fort) d'altitude peut favoriser, dans certaines conditions, l'organisation de nuages convectifs.

Dans notre cas, cette variable a été choisie notamment en raison de son rôle dans la formation et l'évolution des perturbations frontales. C'est en effet une excellente signature de ces dernières.

5.1.2 Variables élaborées

Les variables élaborées sont définies à partir de variables brutes. Souvent créées pour tenir compte de plusieurs éléments et ainsi décrire un état particulier de l'atmosphère, les variables élaborées sont nombreuses et seules quelques unes suffisent à caractériser cet état. Dans le cas des précipitations, nous en avons retenu trois (Tableau 5.2), que nous allons à présent définir.

Tableau 5.2 : Liste des variables élaborées analysées : abréviation et nomenclature usuelle correspondante. Les variables dotées d'un astérisque sont des variables mesurées à des niveaux de pression.

Abbréviation	Nomenclature usuelle
<i>ECA</i>	Accélération centripète eulérienne
<i>TPE*</i>	Température potentielle équivalente
<i>TCW%</i>	Produit de l'humidité relative au niveau de pression 850 hPa par la quantité d'eau précipitable

5.1.2.1 Accélération centripète eulérienne (*ECA*)

La variable élaborée *ECA* a été définie par Lefevre and Nielsen-Gammon (1995), et récemment utilisée par Dal Piva et al. (2008). L'accélération *ECA* représente l'accélération centripète (composante perpendiculaire à la direction du vent) subie par une parcelle d'air portée par le vent géostrophique¹. Elle peut être interprétée comme étant le transport de tourbillon par le vent géostrophique. Généralement analysée au niveau de pression 500 hPa,

¹ Le vent géostrophique est une partie du vent réel pour laquelle la force d'inertie de Coriolis et la force du gradient horizontal de pression sont en équilibre. La force de frottement lié au sol est alors considérée comme nulle. On assimile le vent réel au vent géostrophique dans les hautes couches de la troposphère où l'influence du sol est négligeable. Ainsi à 500 hPa, il peut être déduit directement des gradients de géopotentiel.

l'accélération *ECA* peut également être utilisée à d'autres niveaux de pression. L'équation 5.6 donne la définition mathématique de cette variable.

$$ECA \equiv \frac{V_g^2}{R_s} \quad (5.6)$$

où V_g est le module du vent géostrophique et R_s le rayon de courbure du vent géostrophique, le vent géostrophique étant calculé au niveau de pression désiré (500 hPa dans notre cas). Son unité est le $m.s^{-2}$. Pour le calcul de l'accélération *ECA* en un point, nous avons utilisé les valeurs de *Z500* des points environnants et la latitude du point.

Intérêt pratique

Le champ d'accélération *ECA* permet d'identifier les talwegs d'altitude mobiles. C'est donc un excellent traceur de ces centres d'action. Une anomalie positive d'*ECA* indique ainsi une accélération du vent géostrophique, c'est-à-dire le fond d'un talweg dynamique. Nous avons initialement choisi cette variable pour son intérêt à détecter ces zones sensibles à l'avant desquelles une dépression ou un talweg de surface est susceptible d'évoluer. Elle peut notamment être utile pour la prévision des précipitations convectives engendrées par des systèmes de méso-échelle.

5.1.2.2 Température potentielle équivalente (TPE)

La température *TPE* correspond à la température d'une particule d'air à laquelle on aurait enlevé toute sa vapeur d'eau, par un processus adiabatique, et que l'on aurait ramené au niveau de pression 1000 hPa. C'est donc une température fictive. L'équation 5.7 donne la formule mathématique que nous avons utilisée, dérivée de l'équation de Rossby (1932). Le rapport de mélange en vapeur d'eau est calculé à l'aide de l'humidité relative et de la température.

$$TPE = \left(T + \frac{L_v}{C_p} \cdot r \right) \cdot \left(\frac{1000}{P} \right)^k \quad (5.7)$$

Où :

- P est la pression atmosphérique ;
- T est la température au niveau de pression P ;
- r est le rapport de mélange en vapeur d'eau ;
- L_v est le coefficient de chaleur latente d'évaporation ;
- C_p est la chaleur spécifique de l'air sec à pression constante ;
- k est le rapport de la constante des gaz parfaits sur la chaleur spécifique de l'air sec à pression constante.

De cette manière, la température *TPE* est calculée avec des variables brutes extraites des réanalyses (température et humidité relative).

Intérêt pratique

En prévision météorologique, la température *TPE* permet de caractériser la masse d'air dans les basses couches de la troposphère. Une zone où l'air est plus sec et plus froid est associée à un champ de faible *TPE* (à l'arrière d'un front froid par exemple) ; tandis qu'une zone où l'air est plus humide et plus chaud est associée à un champ de *TPE* élevée (à l'avant d'un front froid par exemple, ou au coeur du front occlus d'une dépression). Cette variable est un très bon traceur des masses d'air, mais est aussi un témoin d'une masse d'air instable. Une autre variable aurait également pu être utilisée comme traceur de masse d'air : la température pseudo-potentielle du thermomètre mouillé θ'_w . Les deux variables sont étroitement liées (Bindon, 1940). Par simplicité, nous avons choisi la température *TPE* puisque celle-ci fait intervenir moins de paramètres que la température θ'_w .

5.1.2.3 Produit de l'humidité par la quantité d'eau précipitable (TCW%)

Nous avons défini précédemment l'humidité relative *RH*. La variable élaborée *TCW%* est une variable qui résulte des recherches menées par Bontron (2004), et est utilisée dans la méthode de référence A0 (cf. 1.3.1.2). Il s'agit du produit de l'humidité relative *RH* par la quantité d'eau précipitable *TCW*. Cette dernière correspond à la quantité d'eau, présente dans l'air sous toutes ses formes (vapeur, liquide, solide), intégrée sur la colonne troposphérique. Son unité est le mm.

Intérêt pratique

Le produit des deux variables *RH* et *TCW* permet de tenir compte à la fois de l'état de saturation en humidité de l'air au niveau auquel est prise la variable *RH*, et également de la quantité d'eau contenue dans l'air entre le sol et la tropopause. A noter que cette variable n'est pas couramment utilisée et cartographiée en prévision météorologique.

Dans la méthode de référence A0, la variable *RH* était prise, faute de mieux, au niveau 850 hPa. La variable *TCW%* renseigne ainsi sur l'état hygrométrique de l'air dans les basses couches mais également sur le potentiel pluviométrique représenté par la variable *TCW*. Si par exemple l'humidité relative est proche de 100 % au niveau 850 hPa et que la quantité d'eau précipitable est faible, cela traduit, en dépit d'un air presque saturé dans les basses couches, l'absence d'eau en grande quantité dans les couches moyennes et hautes de la troposphère. Le risque de fortes précipitations est alors assez faible.

5.2 Analyse des liaisons entre les variables**5.2.1 Analyse en composantes principales**

Nous venons de définir dix variables météorologiques. Elles sont disponibles pour certaines à différents niveaux de pression. Afin d'évaluer la redondance de l'information apportée par

ces variables, nous avons étudié les liaisons entre elles. Les valeurs des variables ont été extraites des réanalyses ERA, à tous les pas de temps (toutes les 6h), sur la période commune à celle de l'archive SAFRAN en fin de thèse (1970-2002).

Une analyse en composantes principales (ACP) normées a été effectuée sur 34 variables, en quatre points de grille, de coordonnées (5°E ; 45°N), (5°E ; 47,5°N), (7,5°E ; 45°N) et (7,5°E ; 47,5°N), représentatifs de la zone du bassin de la Saône. La liste des variables analysées est reportée dans le Tableau 5.3.

Tableau 5.3 : Liste des 34 variables sur lesquelles l'ACP est effectuée.

Variable	Niveaux de pression (hPa)
<i>Z</i>	1000 ; 500
<i>Z-2PVU</i>	-
<i>RH</i>	925 ; 850 ; 775 ; 700
<i>T</i>	925 ; 850 ; 775 ; 700 ; 600 ; 500
<i>W</i>	925 ; 850 ; 775 ; 700 ; 600 ; 500
<i>DIV</i>	1000 ; 925 ; 850 ; 700
<i>VT</i>	850 ; 700 ; 600 ; 500 ; 400
<i>ECA</i>	500
<i>TPE</i>	925 ; 850 ; 775 ; 700
<i>TCW%</i>	-

La Figure 5.1 représente, pour ces quatre points, la part de variance expliquée par chacun des 34 axes factoriels issus de l'ACP. Les deux premiers axes expliquent à eux seuls 55 à 60 % de la variance. On peut considérer que le troisième axe est également pertinent, avec une part de variance expliquée de l'ordre de 10 à 15 %, ce qui n'est pas négligeable au vu du grand nombre de variables considérées. Nous avons analysé la projection de nos variables sur les trois premiers axes.

La Figure 5.2 montre la projection sur le plan factoriel formé par les axes 1 et 2. Sur ces axes, les regroupements sont similaires pour les quatre points considérés. Nous pouvons identifier trois grands groupes (ou quatre, selon le point de grille considéré) :

- le géopotential à 500 hPa, la variable *Z-2PVU*, les températures et les températures potentielles équivalentes (groupe 1) ;
- les vitesses verticales et les divergences (sauf la divergence à 700 hPa) (groupe 2) ;
- le vent et l'accélération *ECA* (groupe 3), ainsi que les humidités relatives pour les deux points d'extraction (5°E ; 47,5°N) et (7,5°E ; 47,5°N).

Seuls la variable *TCW%*, la divergence à 700 hPa et le géopotential à 1000 hPa semblent isolés.

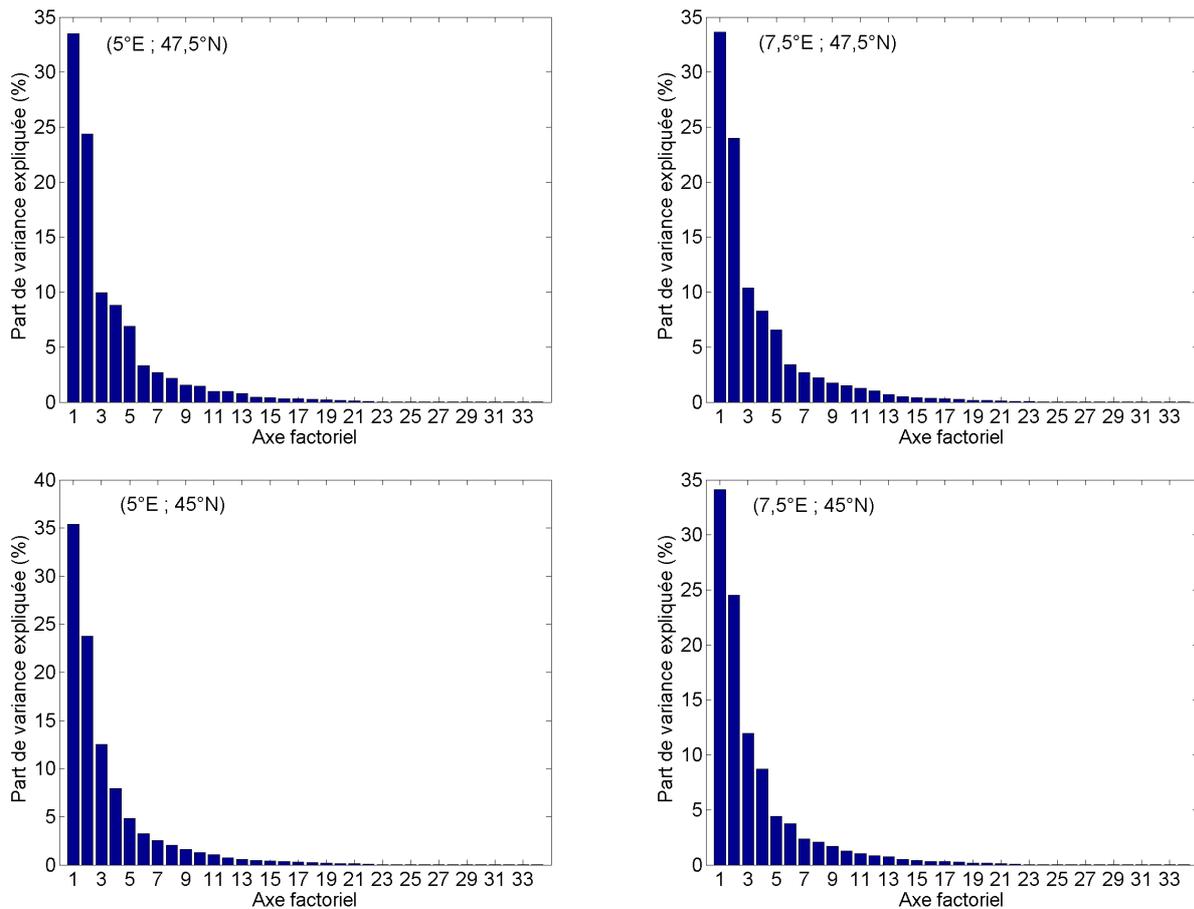


Figure 5.1 : Part de variance expliquée par chacun des 34 axes factoriels issus de l'analyse en composantes principales, évaluée en quatre points de grille.

Dans le groupe 1, les températures et le géopotential à 500 hPa sont fortement liées par l'équation thermodynamique des gaz parfaits, et la température potentielle équivalente est également très liée à la température (par une relation de proportionnalité), tandis que la variable $Z-2PVU$ traduit l'altitude de la tropopause, et celle-ci est en partie liée à l'altitude géopotentielle d'une pression assez basse comme la pression 500 hPa.

Dans le groupe 2, la vitesse verticale est fortement liée à la divergence. En effet, si l'air converge (divergence négative) à un certain niveau (par exemple à 850 hPa), alors l'air est forcé de s'élever, créant une vitesse verticale négative au-dessus (à un niveau de pression inférieur assez proche, par exemple à 700 hPa).

Enfin dans le groupe 3, l'accélération ECA traduit la présence ou non de fort vent d'altitude (500 hPa). Cette variable est donc fortement liée au vent d'altitude. En revanche, l'humidité relative a théoriquement peu de lien avec la vitesse du vent. Elle peut d'ailleurs ne pas être regroupée avec le vent aux points de coordonnées (5°E ; 45°N) et (7,5°E ; 45°N).

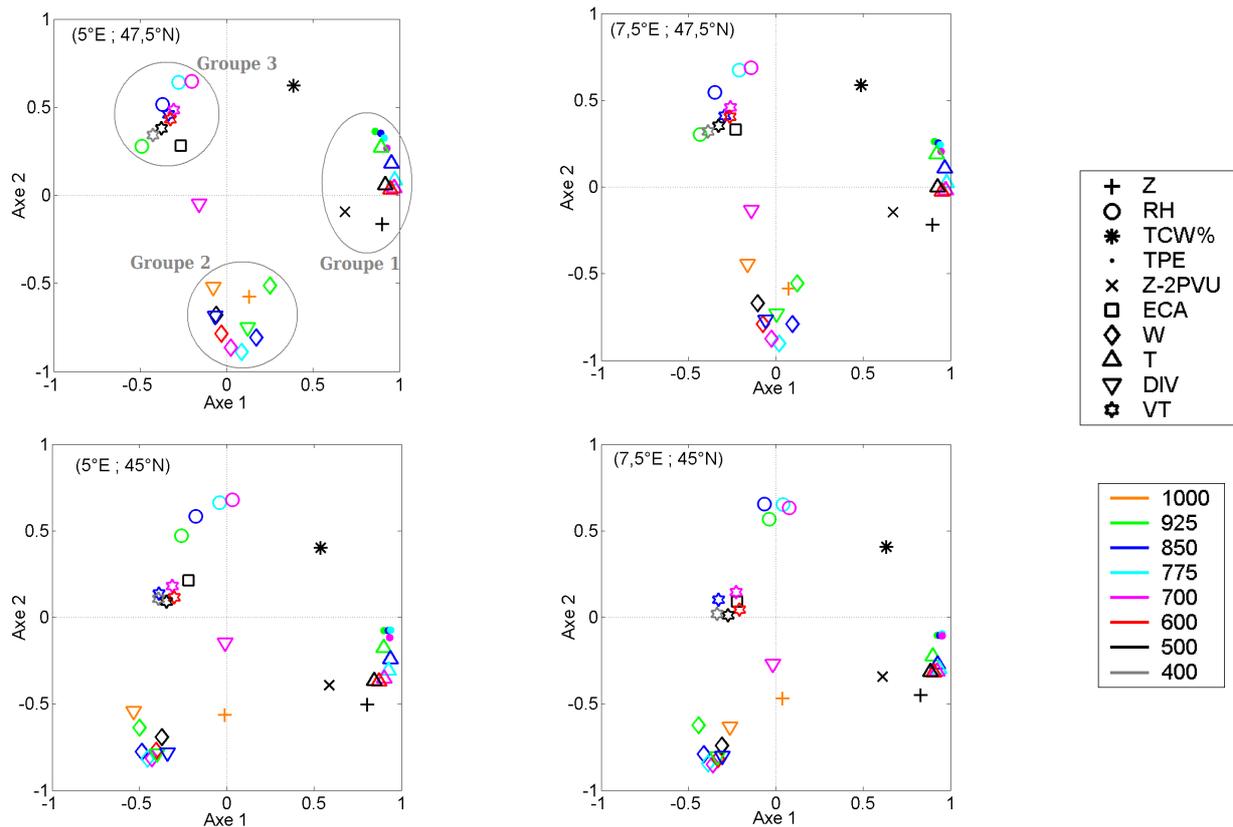


Figure 5.2 : Projection des 34 variables sur le plan factoriel formé par les deux premiers axes, pour les quatre points de grille considérés. Les trois groupes identifiés sont également représentés pour le point (5°E ; 47,5°N).

La représentativité des axes est assez éloquent. Les températures (groupe 1) sont très fortement corrélées avec l'axe 1, avec des coefficients de l'ordre de 0,9. Cet axe est a priori représentatif de l'état thermique de l'atmosphère. Les vitesses verticales (groupe 2) sont fortement anti-corrélées avec l'axe 2, notamment dans les couches moyennes de la troposphère (à 775 hPa et à 700 hPa), avec des coefficients proches de 0,8 à 0,9. Nous pouvons donc supposer que cet axe est représentatif de la stabilité ou de l'instabilité de l'atmosphère, en particulier dans les couches moyennes.

Si nous nous intéressons à la projection des variables sur les axes 2 et 3 (Figure 5.3), les regroupements sont moins évidents. Les variables de même nature mais définies à des niveaux de pression différents sont néanmoins regroupées. Seuls les vents sont assez fortement corrélés avec l'axe 3, notamment à assez haute altitude (à 600 hPa et à 500 hPa), avec des coefficients proches de 0,8. Nous pouvons donc supposer que l'axe 3 est représentatif du flux d'altitude, en d'autres termes de la circulation atmosphérique.

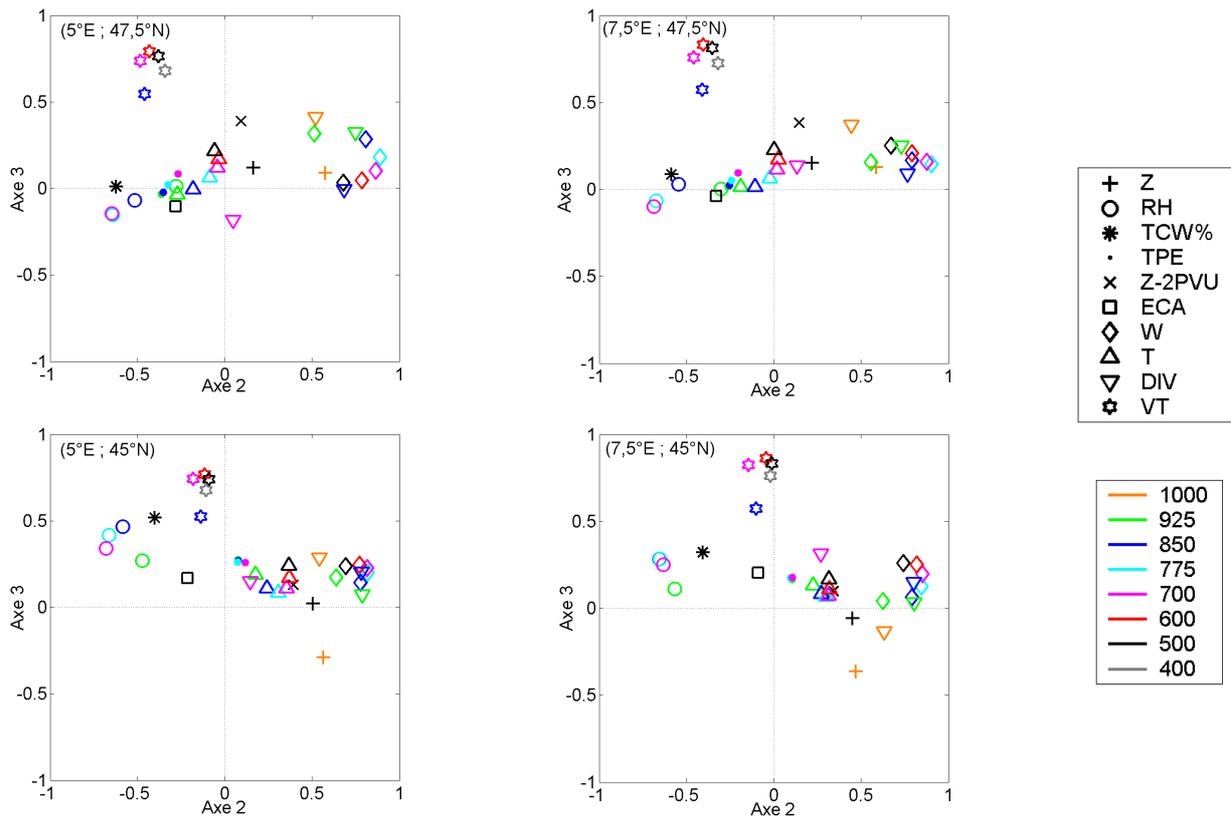


Figure 5.3 : Projection des 34 variables sur le plan factoriel formé par les axes 2 et 3.

5.2.2 Analyse de corrélation et des similitudes entre champs

L'ACP effectuée précédemment permet d'avoir une vision globale des liaisons entre les variables et de déterminer les grandes familles. Elle ne nous permet pas d'apprécier la corrélation entre deux variables, ni de savoir si les champs de deux variables co-fluctuent. Pour compléter l'analyse des liaisons entre les variables, nous proposons ainsi d'étudier la corrélation entre certaines d'entre elles, mais aussi la ressemblance de forme des champs. Notre objectif est notamment de déterminer les variables les plus liées à celles utilisées dans A0, *a priori* redondantes, mais aussi les variables les moins liées, susceptibles d'apporter des informations nouvelles et d'améliorer la prévision.

Dans cette section, nous avons sélectionné les variables utilisées dans la méthode A0 (Z_{1000} , Z_{500} , $TCW\%$), les variables uniques ($Z-2PVU$, ECA), ainsi que les variables restantes prises au niveau de pression auquel ces variables sont usuellement analysées en météorologie (TPE_{850} , W_{700} , T_{850} , T_{500} , DIV_{850} , VT_{500}). Nous avons calculé les coefficients de détermination pour tous les couples de variables, au point le plus proche du bassin de la Saône ($5^{\circ}E$; $47,5^{\circ}N$).

En complément, nous avons calculé la valeur moyenne du critère TWS pour chaque couple de variables afin d'estimer la cohérence entre la forme des champs de ces variables. Nous avons choisi arbitrairement le domaine sur lequel a été calculé le critère. Celui-ci est constitué de 24 points de grille et centré approximativement sur le bassin de la Saône. Les

variables ont été préalablement normalisées afin qu'elles aient la même étendue. Cette normalisation est décrite par l'équation (5.1).

$$V'_{i,j} = \frac{V_{i,j} - \text{Min}(V)}{\text{Max}(V) - \text{Min}(V)} \quad (5.8)$$

Où : $V_{i,j}$ est la composante (i,j) du champ de la variable V ;
 $\text{Min}(V)$ et $\text{Max}(V)$ sont respectivement les valeurs minimale et maximale absolues des champs de la variable V sur la période (1970-2002) ;
 $V'_{i,j}$ est la composante (i,j) de la variable normalisée V' .

Par simplicité, les valeurs du coefficient de détermination sont indiquées dans la partie supérieure droite du Tableau 5.4, et les valeurs du critère TWS sont indiquées dans la partie inférieure gauche.

Tableau 5.4 : **Coefficients de détermination** calculés pour chaque couple de variables (**en haut à droite** de la diagonale grisée) et valeurs moyennes du **critère TWS** obtenues sur les champs de chaque couple de variable (**en bas à gauche**). Les valeurs en gras indiquent une faible liaison entre variables.

	Z500	Z1000	TCW%	TPE850	Z-2PVU	ECA	W700	T850	T500	DIV850	VT500
Z500		0,27	0,07	0,52	0,52	0,06	0,03	0,65	0,78	0,02	0,07
Z1000	80,58		0,09	0,02	0,09	0,03	0,23	0	0,03	0,23	0,02
TCW%	102,60	104,54		0,46	0,02	0	0,09	0,18	0,2	0,03	0,01
TPE850	82,58	113,49	80,21		0,24	0,01	0,04	0,89	0,67	0,04	0,04
Z-2PVU	77,09	93,79	99,88	91,26		0,13	0,01	0,33	0,58	0	0
ECA	105,94	101,58	100,77	101,82	106,30		0,04	0,02	0,05	0,01	0,02
W700	97,24	92,72	109,65	107,66	100,90	100,26		0,02	0	0,75	0,05
T850	70,53	114,12	104,26	59,67	88,14	102,45	103,23		0,71	0,05	0,08
T500	56,80	96,06	99,31	84,19	70,64	106,14	100,20	77,87		0	0,02
DIV850	98,21	93,74	102,95	105,12	100,86	99,61	53,13	105,55	100,18		0,02
VT500	106,16	102,76	95,22	102,74	97,08	99,23	102,24	106,67	98,51	101,28	

Les valeurs du coefficient de détermination montrent que les variables $T850$, $T500$ et $TPE850$ sont assez corrélées avec la variable $Z500$. En revanche, il est intéressant de voir que les températures $T850$ et $T500$ sont moins corrélées avec la variable $TCW\%$ que la température $TPE850$. Puisque l'analogie sur la variable $TCW\%$ est une analogie de proximité, la température pourrait ainsi être introduite avec un critère de distance. La variable $Z-2PVU$ est assez fortement liée au géopotential $Z500$, et risque donc d'être peu informative comme variable complémentaire. A l'inverse, les variables ECA , $W700$, $DIV850$ et $VT500$ sont peu ou très peu corrélées avec les variables de la méthode A0. Leur introduction dans l'algorithme de sélection mérite donc d'être étudiée.

Si on s'intéresse aux valeurs de TWS , on s'aperçoit que les formes des champs des variables étudiées sont globalement peu ressemblantes, avec de nombreuses valeurs comprises entre 90 et 110, ce qui traduit une certaine indépendance des champs. Néanmoins,

les champs de certaines variables sont plus ressemblants que d'autres aux champs des variables dans la méthode de référence A0. Ainsi, les formes des champs des variables thermiques (*TPE850*, *T850* et *T500*) semblent être similaires à celles des champs de hauteurs géopotentielles, avec des valeurs du score *TWS* assez faibles, de l'ordre de 50 à 80. Une analogie supplémentaire sur la forme des champs de température n'apporterait donc pas beaucoup plus d'information.

5.3 Tests préliminaires sur des variables d'analogie potentielles

Les analyses précédentes ont permis de mettre en évidence des liaisons entre certaines variables d'analogie potentielles et les variables déjà utilisées dans la méthode de référence.

Ces variables n'ont pas été écartées d'office pour autant. En effet, nous avons effectué des tests exploratoires avec les variables *Z-2PVU* et *TPE*, dont les résultats ne seront pas développés dans ce mémoire. Les conclusions sont mentionnées ici. Ainsi, nous avons tenté d'introduire :

- la variable *Z-2PVU* entre les niveaux 1 et 2 de la méthode A0, et ce dans le but de rechercher l'analogie sur les anomalies de tourbillon. Plusieurs critères d'analogie ont été testés, mais les résultats ne se sont pas révélés intéressants (perte notable de performance) ;
- la variable *TPE* avant le niveau 2, mais aussi après le niveau 2. L'objectif initial fut de rechercher l'analogie sur la forme des champs de *TPE*, en espérant que les gradients caractérisant les contrastes de masse d'air (et donc les fronts) soient pris en compte. Malheureusement, le gain de performance s'est avéré négligeable.

Par ailleurs, nous avons tenté d'introduire l'accélération *ECA*. En dépit de l'indépendance de cette variable vis-à-vis des variables de l'algorithme A0, son usage n'a guère apporté d'avancée significative. Nous pensons notamment que l'usage de cette variable peut être légitime pour les bassins soumis à des épisodes de précipitations convectives intenses, dont les systèmes de méso-échelle peuvent être favorisés par la présence d'un noyau d'accélération *ECA* élevée. Dans le cas des bassins testés (Saône et Seine), principalement soumis à des précipitations frontales, les résultats ne se sont pas avérés concluants.

Nous avons également introduit la variable *VT* dans le second niveau de sélection de la méthode A0. Nous avons remplacé la variable *TCW%* par le produit du vent *VT* par *TCW%*. Là encore, les performances évaluées pour le bassin de la Saône ont été dégradées par rapport aux performances initiales obtenues par la méthode A0. Néanmoins, l'usage de cette variable peut être justifié pour des bassins montagneux soumis à la convergence d'air humide, à l'instar des bassins à crue rapide situés dans les Cévennes.

Le développement d'une méthode d'analogie semble donc dépendant de la zone d'application. Les résultats obtenus au cours de cette thèse n'auraient probablement pas été les mêmes pour d'autres bassins aux caractéristiques différentes (orographie, climatologie,

régimes de temps, etc.).

5.4 Conclusion

Dans ce chapitre, nous avons effectué une analyse préliminaire des variables d'analogie qui nous semblent les plus informatives pour décrire une situation météorologique. La liste de ces variables n'est évidemment pas exhaustive, mais elle contient un certain nombre de variables pertinentes et souvent utilisées en prévision météorologique.

Après les avoir définies, nous avons analysé les liens possibles entre ces variables afin de faire un premier tri. En effet, il ne serait pas judicieux d'introduire une variable fortement liée à une des variables déjà utilisées dans la méthode d'analogie. Une ACP normée a été effectuée sur 34 variables en quatre points de grille proches du bassin de la Saône. Elle a permis de dégager, grâce à la projection des variables sur le plan factoriel formé par les deux premiers axes, plusieurs groupes de variables *a priori* liées. Ainsi la température T , la température potentielle équivalente TPE , le géopotential Z d'un niveau de pression et le géopotential du tourbillon potentiel égal à 2PVU ont pu être regroupés. Un second groupe indépendant a été mis en évidence. Celui-ci est constitué de la vitesse verticale et de la divergence horizontale. Enfin, nous avons identifié un troisième groupe formé par la vitesse du vent horizontal et l'accélération centripète eulérienne.

La seconde partie de ce chapitre a été consacrée à l'examen de la corrélation entre les variables ainsi que de la ressemblance des champs. Cette analyse a permis de consolider les résultats obtenus par l'ACP. Finalement, les variables les moins liées à celles utilisées dans la méthode de référence A0 sont : ECA , W , DIV et VT .

Enfin dans la troisième partie, nous nous sommes assurés du faible intérêt d'introduire une variable déjà liée avec l'une des variables de la méthode A0. Les tests effectués, d'une part avec le géopotential $Z-2PVU$ et d'autre part avec la température TPE , n'ont pas permis d'obtenir des résultats probants. Par ailleurs, bien que les variables ECA et VT soient peu liées avec celles de la méthode de référence, la tentative de leur introduction n'a pas non plus mené à un gain positif de performance.

Par conséquent, l'ensemble de ces dernières variables ne sera pas considéré dans la suite de nos travaux. En conclusion, les prochains chapitres seront consacrés aux variables d'analogie potentielles W et DIV . Nous débuterons par un chapitre dédié à la température T malgré sa forte liaison avec les géopotentials.

6 Prise en compte de la saisonnalité

Dans la méthode de référence, l'échantillon de situations météorologiques passées, sur lesquelles les analogies sur la circulation et sur l'humidité sont recherchées, est formé en prenant en compte la saison à laquelle appartiennent les dates de ces situations (cf. section 1.3.1). Nous proposons dans ce chapitre de relaxer cette hypothèse en introduisant la notion de masse d'air, caractérisée par la température.

Mais cette variable n'est-elle pas déjà liée à l'une des variables de la méthode de référence ? Nous venons de voir, dans le chapitre 5, que la température est assez fortement corrélée avec le géopotential utilisé au niveau 1. Cependant, la sélection d'analogues du niveau 1 utilise un critère de forme, ce qui nous laisse le loisir de tester un critère de distance. De plus, d'après les résultats de l'étude de corrélation menée au chapitre 5, la température n'apporte pas la même information que la variable *TCW%* du second niveau d'analogie, auquel est appliqué un critère de distance.

Dans un premier temps, nous allons donc justifier l'introduction de la température et préciser le lien avec l'effet saisonnier pris en compte dans la construction de l'échantillon de situations candidates. La stratégie d'optimisation suivie et les résultats seront ensuite présentés. Nous les validerons en adaptant au bassin de la Seine la méthode mise à jour, puis nous achèverons cette étude par un diagnostic afin de mettre en évidence l'apport de cette variable dans la méthode de sélection de référence. Enfin, nous évaluerons les performances sur la période de validation.

6.1 Pourquoi introduire la température ?

6.1.1 Inconvénient de la sélection des situations candidates

Dans la méthode de référence, pour un jour cible *C*, les situations candidates au premier niveau de sélection (analogie sur les champs de géopotential) sont automatiquement pré-sélectionnées dans une fenêtre temporelle de plus ou moins deux mois autour du jour *C* (soit une largeur de fenêtre temporelle de quatre mois). Cette pré-sélection, qui n'est pas

considérée comme un niveau d'analogie, est effectuée pour tenir compte de l'effet de saisonnalité. Certains facteurs comme l'état de la végétation, le rayonnement solaire, l'état du manteau neigeux ou la température de surface de la mer, diffèrent d'une saison à une autre. Dans le cadre de la prévision des pluies par analogie, le respect de la saisonnalité est basé sur l'hypothèse émise par Lorenz (1969) selon laquelle les précipitations qui se produisent pour une situation météorologique donnée sont différentes lorsque cette situation est observée au cours de deux saisons contrastées. Ainsi par exemple, pour la prévision des précipitations automnales, Diomede et al. (2008) n'ont recherché des situations analogues que durant la saison d'automne. Gibergans-Baguena and Llasat (2007) ont appliqué une méthode d'analogie sur la saison pluvieuse en Catalogne, et n'ont recherché des analogues que durant cette même saison (septembre à novembre).

L'optimisation de la largeur de la fenêtre temporelle a déjà fait l'objet de recherches (Wetterhall et al., 2005). Nous avons souhaité également réaliser quelques tests de sensibilité. La méthode de référence A0 a été appliquée au bassin de la Saône en faisant varier la largeur de la fenêtre temporelle entre 1 et 12 mois (Figure 6.1).

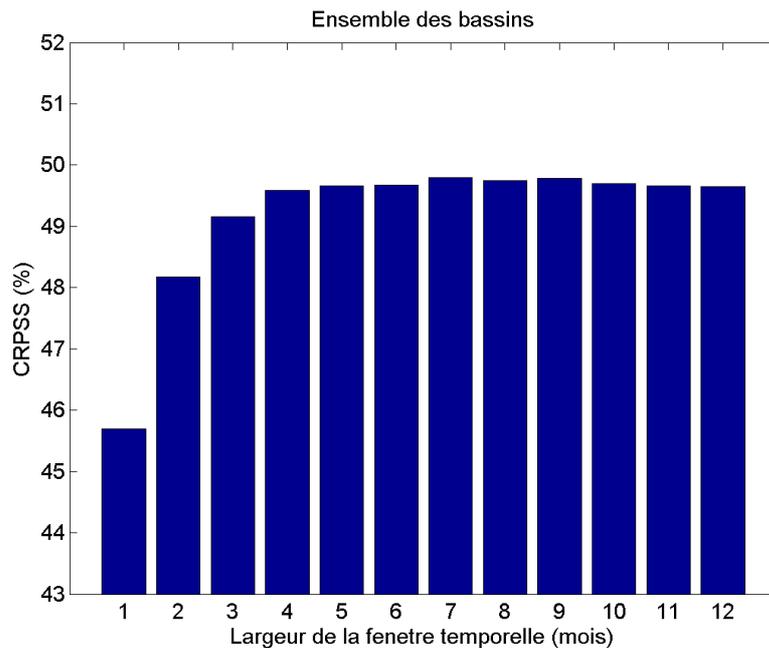


Figure 6.1 : Evolution des performances moyennes (*CRPSS* (%)), évaluées sur le bassin de la Saône, en fonction de la largeur de la fenêtre temporelle dans laquelle les situations candidates sont sélectionnées.

Les performances obtenues sur la période de calibration montrent que la valeur optimale est de sept mois, avec une valeur du score *CRPSS* de 49,8 %¹. Nous nous écartons ainsi de la largeur de fenêtre temporelle initialement choisie (de quatre mois). Or, il se trouve qu'un palier apparaît à partir d'une largeur de quatre mois et aussi que le temps de calcul augmente rapidement avec l'élargissement de la fenêtre. Le choix d'une fenêtre de largeur quatre mois

¹ La largeur de fenêtre optimale dépend bien sûr de la taille de l'archive. Avec une archive de longueur infinie, la taille optimale serait probablement de un jour !

résulte donc d'un compromis entre efficacité et temps de calcul.

La Figure 6.2 illustre la sensibilité de la distribution des précipitations à la largeur de la fenêtre temporelle des situations candidates. Nous avons volontairement choisi deux situations à pluviométrie distincte. La Figure 6.2 a) présente les distributions empiriques des précipitations obtenues, dans un contexte de prévision parfaite, sur le bassin de la Petite Saône le 20/09/1981 (cumul observé de 0 mm), lorsque la largeur de la fenêtre temporelle est de quatre mois (courbe grise discontinue) et cinq mois (courbe noire pleine). Avec un seul mois supplémentaire par année candidate, la distribution apparaît plus juste avec par exemple une valeur médiane de 4,5 mm, au lieu de 9 mm lorsque la fenêtre est large de quatre mois. La Figure 6.2 b) présente les distributions obtenues sur le bassin du Doubs le 20/02/1977 (cumul observé de 29,4 mm), lorsque la largeur de la fenêtre temporelle est de quatre mois (courbe grise discontinue) et sept mois (courbe noire pleine). La distribution obtenue avec la fenêtre de sept mois est un peu moins fine que celle obtenue par la fenêtre de quatre mois. Néanmoins, une fenêtre plus large permet dans ce cas de sélectionner finalement des situations analogues dont le cumul pluviométrique est plus proche de celui observé. L'extension de la fenêtre a notamment permis de considérer cinq situations dont le cumul est supérieur à 20 mm, contre une seule avec la fenêtre de quatre mois.

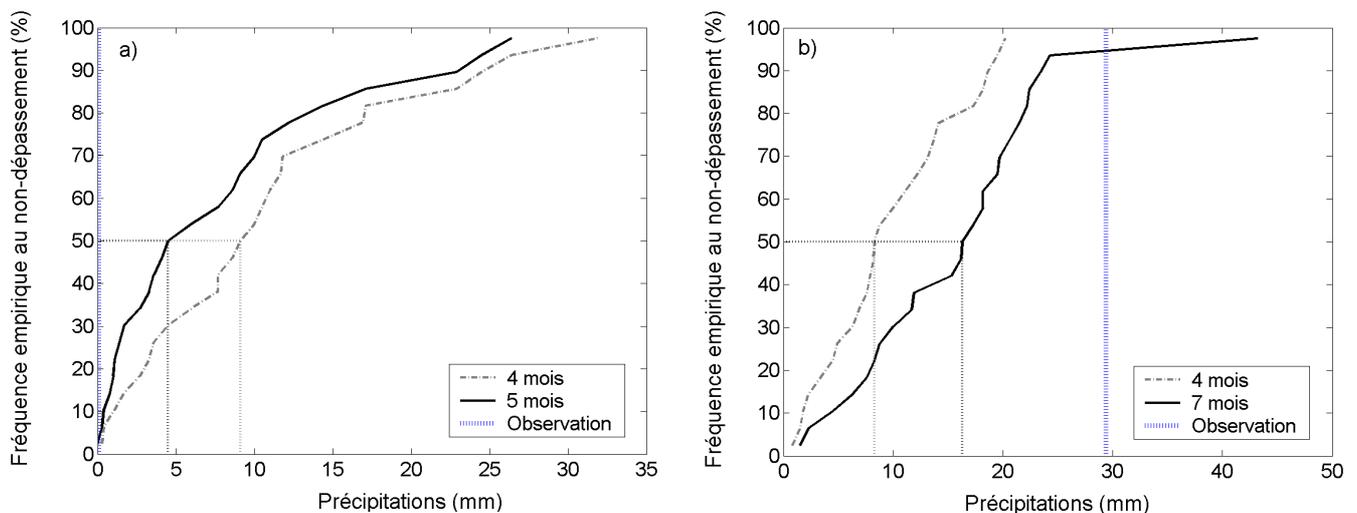


Figure 6.2 : Distributions empiriques des précipitations obtenues dans un contexte de prévision parfaite. a) sur le bassin de la Petite Saône le 20/09/1981 (cumul observé de 0 mm), en considérant une fenêtre temporelle de largeur 4 mois (courbe grise discontinue) et de 5 mois (courbe noire pleine) ; b) sur le bassin du Doubs le 20/02/1977 (cumul observé de 29,4 mm), en considérant une fenêtre temporelle de largeur 4 mois (courbe grise discontinue) et de 7 mois (courbe noire pleine).

En conclusion, si la largeur de cette fenêtre est fixée à quatre mois, des situations candidates potentiellement « bonnes analogues » peuvent être négligées de part et d'autre de la fenêtre. De même, il est possible que des situations potentiellement « mauvaises analogues » soient proposées comme candidates si on considère cette fenêtre fixe. Une fenêtre temporelle fixe, dans laquelle sont extraites les situations candidates au premier niveau de sélection, n'est donc pas adaptée à toutes les situations. Il nous semble intéressant

de sélectionner d'une manière différente les situations candidates, sans restreindre le choix en fonction de la période de l'année dans laquelle elles se situent. L'effet de saisonnalité, qui n'est rien d'autre que la conséquence de la variation du rayonnement solaire au cours de l'année, se retrouve notamment dans les variables liées à la température. Ainsi, l'idée que nous proposons est de remplacer un critère figé basé sur les dates par un critère dynamique basé sur la température. Nous allons à présent décrire le principe de cette pré-sélection et évaluer son intérêt en effectuant un test préliminaire.

6.1.2 Principe et test préliminaire

La température est donc dépendante de la saison. L'importance de cette dépendance varie selon l'altitude à laquelle la température est mesurée dans la troposphère. Nous disposons, dans l'archive ERA rapatriée, de la température mesurée à six niveaux de pression (925, 850, 775, 700, 600 et 500 hPa). La Figure 6.3 montre, pour les niveaux de pression 925, 700 et 500 hPa, l'évolution de la température sur l'année civile.

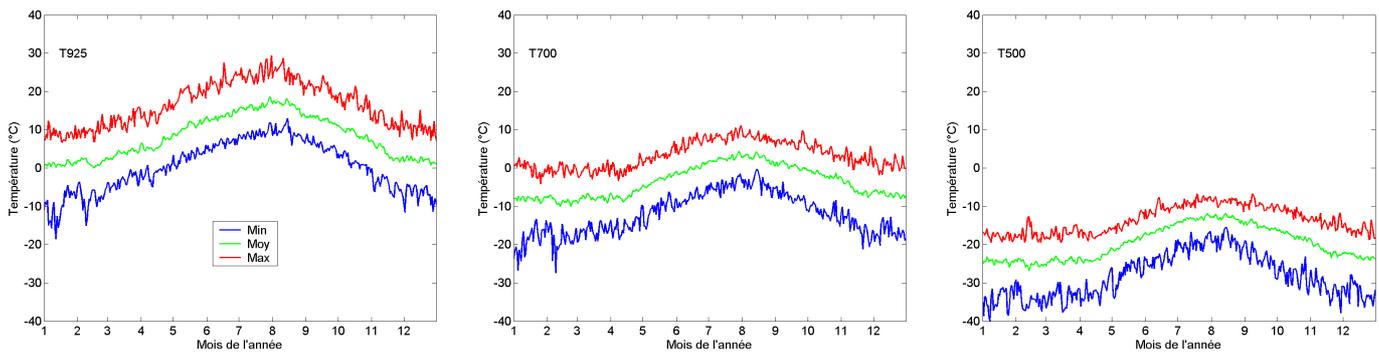


Figure 6.3 : Evolution des températures moyennes journalières minimales (bleu), moyennes (vert) et maximales (rouge), estimées sur la période de calibration. Les niveaux de pression considérés sont a) 925 hPa, b) 700 hPa et c) 500 hPa.

La température moyenne de chaque journée a été estimée en effectuant la moyenne arithmétique des valeurs relevées à +6h, +12h, +18h, et +24h. Pour chaque jour de l'année figurent le minimum absolu, la valeur moyenne et le maximum absolu, obtenus sur la période 1/8/1972 - 31/7/2002, soit 30 années. Ainsi chaque statistique a été évaluée sur 30 valeurs de température.

Ces courbes reflètent bien l'effet de saisonnalité. Cet effet est d'autant plus marqué que l'altitude à laquelle on se situe est basse. Ainsi, la température moyenne à 925 hPa varie fortement d'une saison à une autre, avec une amplitude de 15 à 20°C ; tandis que l'amplitude de la température moyenne à 500 hPa est de l'ordre de 10 à 15°C. Cette différence s'explique principalement par l'influence du sol sur les très basses couches de la troposphère, comme au niveau de pression 925 hPa¹. En revanche, cette influence s'avère nettement plus faible dans les couches moyennes ou hautes de la troposphère comme au niveau de pression 500 hPa. On remarque également que l'évolution de la température, au cours du printemps, est un peu

¹ Le rayonnement diurne absorbé par le sol tend à échauffer les très basses couches de la troposphère, tandis que le rayonnement nocturne (refroidissement par émission de chaleur sensible) tend à les refroidir.

plus rapide dans les très basses couches.

L'effet de la saison peut donc être pris en compte en effectuant une sélection par analogie sur la température. Ce niveau de pré-sélection sera nommé « niveau 0 » par la suite, puisqu'il ne s'agit pas d'un niveau d'analogie permettant de discriminer les situations météorologiques suivant leur pluviosité. Pour évaluer l'intérêt de sélectionner les situations candidates selon la température, nous avons effectué un premier test en utilisant la température mesurée à 850 hPa et à 500 hPa, à +12h et +24h, sur un domaine proche du bassin constitué de quatre points (domaine correspondant à celui du niveau 2 de la méthode A0 pour le cas de la Saône).

Etant donné que l'on cherche à sélectionner les situations candidates dont la température est la plus proche de celle de la situation cible, le critère d'analogie appliqué est le critère de distance *RMSE*. Les performances (*CRPSS*) ont été calculées à l'issue du niveau 1 puis à l'issue du niveau 2 de la méthode de référence A0. Pour une situation cible donnée, contrairement à la sélection initiale, les situations candidates de l'année cible ont été prises en compte, exceptées celles qui entourent le jour cible. Une fenêtre de plus ou moins quatre jours autour du jour cible a été arbitrairement retenue. Afin que les résultats soient comparables, le nombre de situations retenues à l'issue de ce niveau de sélection de situations candidates est identique à celui obtenu par la sélection initiale dans une fenêtre de plus ou moins deux mois, soit environ 3000 situations.

Tableau 6.1 : Performances moyennes (*CRPSS*) obtenues sur le bassin de la Saône à l'issue des niveaux 1 et 2, lorsque les situations candidates sont sélectionnées automatiquement dans une fenêtre temporelle fixe (sans niveau 0) et lorsqu'elles sont sélectionnées selon la température (avec niveau 0).

	Niv. 1 sans niv. 0	Niv. 1 avec niv. 0	Niv. 2 sans niv. 0	Niv. 2 avec niv. 0
Doubs	43,34	43,74	50,91	52,15
Petite Saône	43,26	43,76	50,32	51,29
Saône Intermédiaire	41,18	41,77	47,54	48,93
Ensemble des bassins	42,59	43,09	49,59	50,79

Le gain de *CRPSS* est positif, s'élevant à 0,5 points au niveau 1 et à 1,2 points au niveau 2, ce qui est assez faible mais non négligeable. Or, à l'issue du niveau 0, le nombre de situations candidates est élevé (près de 3000). Une grande partie des situations conservées par la sélection sur la température doivent aussi l'être par la sélection basée sur la saisonnalité, ce qui explique probablement ce faible gain.

Néanmoins, les paramètres, choisis arbitrairement ici, n'ont pas été optimisés. Nous allons donc procéder à une optimisation afin de vérifier si cette pré-sélection est justifiée.

6.2 Optimisation des paramètres

Dans le cadre de l'introduction d'une étape supplémentaire de sélection d'analogues, les paramètres à déterminer sont : le critère d'analogie, le nombre *NO* de situations à

sélectionner, le domaine d'analogie, le (s) heure(s) de mesure et niveau(x) de pression de la température.

Les degrés de liberté étant trop nombreux, une optimisation complète serait très gourmande en temps de calcul, d'autant plus que dans le cas de ce nouveau niveau de sélection, le nombre de situations potentiellement candidates au niveau 1 est très élevé.

Dans cette phase exploratoire, il s'avère donc nécessaire de faire des compromis. Dans un premier temps, nous avons choisi de fixer arbitrairement le critère d'analogie, le nombre $N0$ et le domaine d'analogie.

Critère d'analogie

Quatre critères d'analogie ont été retenus : d'une part les critères de distance $RMSE$ et MAE (équation 6.1), et d'autre part le critère de forme TWS et le coefficient de corrélation R (équation 6.2) liés à la co-fluctuation des champs. Etant donné que l'on cherche à sélectionner les situations candidates dont la température est la plus proche de celle de la situation cible, le critère d'analogie devrait *a priori* mesurer une distance. Le critère $RMSE$, utilisé au cours du niveau 2 de notre méthode de référence, a donc été arbitrairement choisi. Nous nous assurerons, au paragraphe 6.2.5, que ce choix intuitif est justifié.

$$MAE = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J |V_{i,j}^{cible} - V_{i,j}^{candidate}| \quad (6.1)$$

$$R = \frac{\frac{1}{IJ-1} \cdot \sum_{i=1}^I \sum_{j=1}^J [V_{i,j}^{cible} - m(V^{cible})][V_{i,j}^{candidate} - m(V^{candidate})]}{s(V^{cible}) \cdot s(V^{candidate})} \quad (6.2)$$

Où :

- $V_{i,j}^{cible}$ (resp. $V_{i,j}^{candidate}$) est la valeur de la variable V de la journée cible (resp. candidate) au point de grille de coordonnées (i,j) ;
- I et J sont les nombres de points de grille du domaine dans les directions zonale et méridienne ;
- $m(V^{cible})$ (resp. $m(V^{candidate})$) est la moyenne arithmétique des $V_{i,j}^{cible}$ (resp. $V_{i,j}^{candidate}$) sur le domaine d'analogie ;
- $s(V^{cible})$ (resp. $s(V^{candidate})$) est l'écart-type des $V_{i,j}^{cible}$ (resp. $V_{i,j}^{candidate}$) .

Nombre $N0$ de situations à sélectionner

Notre période de calibration est constituée de 25 années. En tenant compte des effets de bord, dus à la non prise en compte des situations candidates de l'année cible, le nombre de situations candidates au niveau 1 de notre méthode de référence varie entre 2904 et 2994 selon la date de la situation cible. Dans un premier temps, afin de nous affranchir de l'optimisation assez lourde de ce nombre $N0$, nous avons choisi de fixer ce nombre à 3000, ce qui est équivalent au nombre de situations candidates retenues initialement par le critère lié à

la date cible.

Domaine d'analogie

La température est un paramètre fortement variable dans le temps et dans l'espace, notamment lorsqu'un changement de masse d'air intervient. En dépit de cette variabilité, nous supposons en première approximation qu'un seul point suffit à décrire la masse d'air à l'échelle du bassin. Ainsi, nous avons choisi le point de grille le plus proche du secteur d'étude pour définir l'analogie. Dans le cas du bassin de la Saône, nous avons considéré le point (5°E ; $47,5^{\circ}\text{N}$). Cette hypothèse permet notamment de réduire encore le temps de calcul de l'optimisation de nos paramètres.

Quelques paramètres étant à présent fixés, nous pouvons optimiser le niveau de pression et l'heure auxquels la température doit être mesurée. Le gain de performance moyenne est à présent exprimé par le score G défini précédemment. Ici nous exprimons donc G par rapport à notre méthode de référence A0.

6.2.1 Choix d'un niveau de pression et d'une heure de mesure

6.2.1.1 Toutes situations confondues

Nous avons à disposition, dans l'archive ERA que nous avons construite, six niveaux de pression (925, 850, 775, 700, 600 et 500 hPa). Pour chaque niveau, dix heures de mesure ont été successivement considérées : de -12h à +42h par pas de 6h (Figure 6.4). Les niveaux 1 et 2 de la méthode de référence A0 ont alors été appliqués sur l'échantillon de 3000 situations candidates les plus proches, d'un point de vue thermique, de la situation cible.

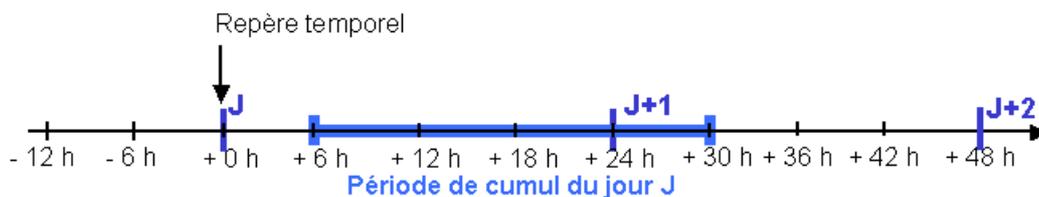


Figure 6.4 : Heures de mesure de la température considérées pour l'optimisation, exprimées relativement à l'instant 00 h TU du jour J (repère temporel).

La Figure 6.5 présente le gain de performance moyenne G par rapport à la méthode de référence A0, en fonction de l'heure de mesure (axe des abscisses) et du niveau de pression (axe des ordonnées). Si l'on devait mesurer la température à un seul niveau de pression et à une seule heure de mesure, il faudrait la mesurer à 925 hPa à +36h (gain G de +2,69 % sur l'ensemble des bassins de la Saône). Contrairement à ce que l'on aurait pu s'attendre, cette heure de mesure optimale ne s'inscrit dans la période du cumul pluviométrique, dont les limites sont indiquées par les lignes épaisses discontinues. De manière générale, la Figure 6.5

révèle que plus la température est mesurée tôt dans la période de cumul, moins la prévision semble bonne, et ce quel que soit le niveau de pression. L'influence de l'heure de mesure est plus marquée lorsque la température est mesurée dans les basses couches (fort gradient de gain de performance), que lorsqu'elle est mesurée en altitude (vers 700 à 500 hPa). Par ailleurs, l'heure optimale associée à chacun des niveaux de pression augmente lorsque la pression augmente (de +18h à 500 hPa à +36h à 925 hPa).

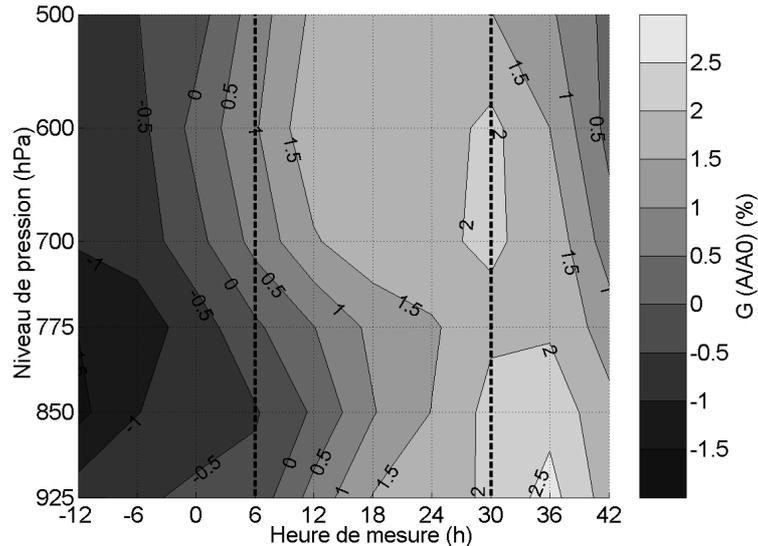


Figure 6.5 : Gain G de performance moyenne, par rapport à la méthode A0 sur le bassin de la Saône, en fonction du niveau et de l'heure de mesure de la température. La période du cumul pluviométrique journalier est limitée par les lignes verticales discontinues.

A première vue, il est assez difficile d'interpréter le comportement du gain de performance avec l'heure de mesure. Le principal résultat indique, pour la séquence temporelle liée au cumul pluviométrique d'une journée, que l'état thermique final importe plus que l'état thermique initial. Si l'on considère une journée anticyclonique, la masse d'air initiale est généralement similaire à la masse d'air finale. En revanche, au cours d'une journée pluvieuse à situation synoptique dépressionnaire, les masses d'air observées sur notre zone d'intérêt sont différentes entre le début et la fin de la période de cumul pluviométrique. En effet, le changement de masse d'air est la conséquence du passage d'une perturbation frontale.

6.2.1.2 Selon la pluviosité des situations

Afin de confirmer notre interprétation, nous avons distingué les situations selon leur pluviosité. Cependant, les épisodes pluvieux qui affectent le bassin de la Saône ne touchent pas nécessairement les trois sous-bassins de manière simultanée. En effet, les précipitations d'origine convective, par exemple, peuvent être locales et ne concerner qu'une partie du bassin de la Saône. Sur l'ensemble des cumuls pluviométriques relevés sur notre période de calage, 26,9% d'entre eux sont nuls sur chacun des trois sous-bassins et 56,5% sont non nuls sur chacun des trois sous-bassins. Par conséquent, 16,6 % des journées de notre archive

correspondent à des situations intermédiaires (journées sans pluie généralisée) .

Afin de discriminer le plus possible les situations cibles non pluvieuses (situations sèches, notées S) des situations cibles pluvieuses (notées P), nous avons considéré l'occurrence ou non de pluie simultanée sur les trois sous-bassins. L'optimisation de l'heure et du niveau de mesure a ainsi été effectuée pour chacun des deux groupes de situations (Figure 6.6). Le gain G de performance moyenne est toujours calculé par rapport à la méthode de référence A0. Pour le calcul du score $CRPS$ de chaque méthode, le score $CRPS$ de la climatologie a donc été évalué pour le groupe S (resp. P), sachant que seules les situations cibles sèches (resp. pluvieuses) ont été considérées, les situations candidates restant les mêmes.

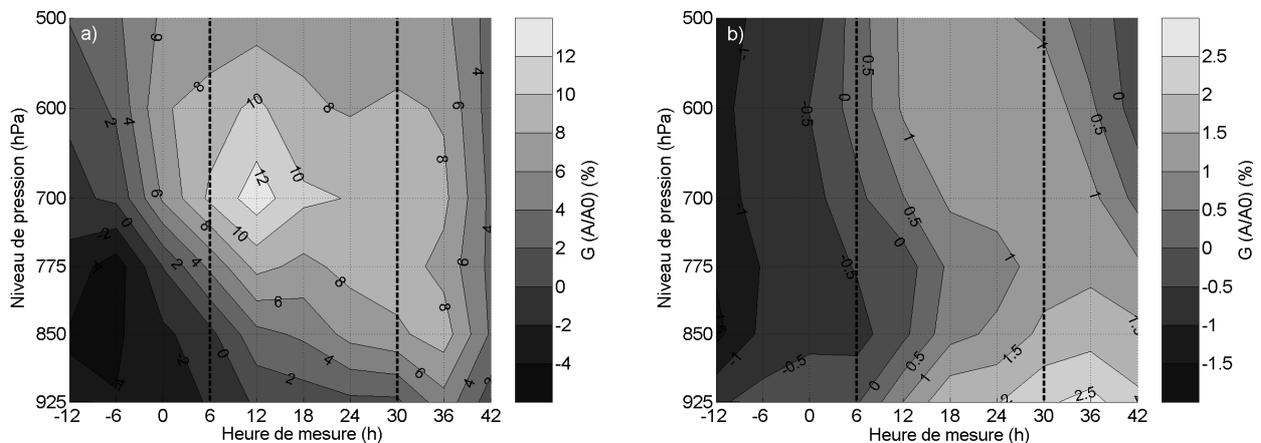


Figure 6.6 : Gain G de performance moyenne, par rapport à la méthode A0 sur le bassin de la Saône, en fonction du niveau et de l'heure de mesure de la température. a) pour le groupe S ; b) pour le groupe P.

La Figure 6.6 révèle des résultats différents selon la pluviosité des situations cibles. Ainsi les journées au cours desquelles il n'a pas plu sur l'ensemble du bassin (Figure 6.6 a)) privilégient un niveau de pression situé dans les couches moyennes de la troposphère (700 hPa). L'heure de mesure optimale est plutôt située en milieu ou en début de période du cumul pluviométrique, même si l'heure reste tardive si on mesure la température dans les basses couches (+36h à 925 et 850 hPa, +30h à 775 hPa). Lorsque l'on considère les situations pluvieuses (Figure 6.6 b)), les résultats sont différents. Les niveaux inférieurs mènent aux gains de performances les plus élevés, et l'heure de mesure optimale se situe à nouveau à +36h, que la température soit mesurée à 925, 850 ou 775 hPa. On peut considérer qu'au passage d'un front, la masse d'air change dans les basses couches, mais ce changement est souvent moins sensible en altitude. C'est peut-être pour cette raison que le niveau optimal se situe dans les très basses couches de la troposphère.

Si nous comparons le graphique de la Figure 6.6 b) avec celui de la Figure 6.5, nous constatons que les résultats sont fortement similaires. Les situations cibles pluvieuses influencent fortement l'optimisation de nos paramètres notamment en raison de leur plus forte proportion, mais cela s'explique aussi par le poids plus important du score $CRPS$, pour ce type de situations, dans le score $CRPS$ moyenné sur l'ensemble des situations. En effet, il

est généralement plus facile de prévoir un cumul de pluie nul, lorsqu'il s'agit de conditions anticycloniques, qu'un cumul de pluie non nul lorsque la situation synoptique est dépressionnaire. Par ailleurs, nous remarquons que le gain G est nettement plus élevé pour les situations sèches (gain d'environ 13 % si la température est mesurée à 700 hPa à +12h) que pour les situations pluvieuses (gain d'environ 2,5 % si la température est mesurée à 925 hPa à +36h). Ce résultat ne signifie pas que si l'on applique ce niveau de pré-sélection sur la température, nous améliorerons plus fortement la prévision des journées du groupe S par rapport à celle des journées du groupe P. En effet, le gain G est un gain relatif de performance, qui se traduit par une valeur élevée si le score $CRPSS$ de référence (ici obtenu par application de la méthode A0) est déjà élevé, ce qui est le cas des situations du groupe S¹.

En conclusion, si la température doit être mesurée à un niveau de pression et à un instant donné, le niveau et l'instant optimaux sont 925 hPa et +36h pour l'ensemble des situations. Néanmoins, en distinguant une situation pluvieuse d'une situation non pluvieuse, nous venons de voir que les niveaux et heures de mesure optimaux sont radicalement différents. Nous pouvons imaginer que ces combinaisons puissent être utilisées simultanément pour que la prévision soit encore meilleure pour l'ensemble des situations. C'est ce que nous allons vérifier en considérant plusieurs niveaux de pression et plusieurs instants de mesure.

6.2.2 Choix de plusieurs heures de mesure et niveaux de pression

Il est possible de combiner plusieurs niveaux de pression et plusieurs heures de mesure de la température. Nous avons choisi d'analyser toutes les combinaisons possibles de deux heures de mesure et d'un niveau de pression, puis certaines combinaisons d'une heure de mesure et de deux niveaux de pression différents. Le critère d'analogie, le nombre de situations candidates retenues et le domaine d'analogie restent identiques à ceux appliqués précédemment.

6.2.2.1 Un niveau de pression, deux heures de mesure

Les six niveaux de pression et les dix heures de mesure retenues précédemment ont été utilisés afin d'examiner toutes les combinaisons possibles d'un niveau de pression et de deux heures de mesure différentes. Par souci de simplicité, seules les cinq meilleures combinaisons sont indiquées dans le Tableau 6.2.

Lorsque la température est mesurée à deux instants différents, le niveau optimal semble se situer dans les couches moyennes ou hautes de la troposphère. Pour les cinq meilleures combinaisons présentées dans ce tableau, les heures de mesure sont bien espacées, l'une se situant en début de période du cumul pluviométrique (généralement +12h), et l'autre se situant en fin de période voire à l'extérieur de la période (+30 ou +36h). Ces résultats expriment l'intérêt de mesurer la température à deux instants bien séparés. Le gain G de performance moyenne par rapport à notre méthode de référence A0 est de +3,21% pour la

¹ Le score $CRPSS$ moyen calculé par la méthode A0 s'élève à 82,53% et celui calculé par cette méthode, dont la température est mesurée à 700 hPa à +12h, s'élève à 84,81%. Le gain G s'élève alors à 13,05%.

combinaison optimale, ce qui représente un gain un peu plus élevé par rapport à l'utilisation d'une seule heure de mesure (gain G de +2,69%).

Tableau 6.2 : Cinq meilleures combinaisons de deux heures de mesure et d'un niveau de pression et gains G de performance moyenne associés, obtenus sur le bassin de la Saône.

	Niveau (hPa)	Heure de mesure (h)									G (A/A0) (%)	
		-12	-6	0	+6	+12	+18	+24	+30	+36		+42
1	600					X			X			3,21
2	600					X				X		3,16
3	700					X				X		3,00
4	500					X			X			2,96
5	700						X			X		2,91

6.2.2.2 Deux niveaux de pression, une heure de mesure chacun

Nous avons ensuite optimisé la combinaison de deux niveaux de pression auxquels la température peut être mesurée, en considérant une heure de mesure pour chaque niveau. Malheureusement, l'optimisation complète (*i.e.* le test de toutes les combinaisons possibles) rendrait le temps de calcul prohibitif. Nous avons donc adopté la stratégie utilisée par Bontron (2004), lorsqu'il avait combiné deux niveaux de pression pour le géopotential au premier niveau de sélection. Cette approche consiste à fixer le niveau de pression et l'heure de mesure de la première température, et à faire varier le niveau et l'heure de mesure de la seconde température. Pour la première température, les dix meilleures combinaisons obtenues lorsque la température est mesurée à un niveau et à une heure ont été retenues. La seconde température a été successivement prise à un niveau de pression parmi les cinq niveaux restant et à une seule heure parmi les dix heures possibles. Le Tableau 6.3 indique les cinq meilleures combinaisons obtenues à l'issue de ce test.

Les résultats révèlent, pour chacune des cinq meilleures combinaisons, qu'un niveau de pression doit être considéré dans les basses couches et un autre niveau dans les couches moyennes ou hautes de la troposphère, même si cette distinction est moins forte pour la cinquième combinaison. En revanche, il est clair que la température doit être mesurée à deux instants bien distincts. En effet, pour ces cinq combinaisons, dont une des heures est fixée à +36h, l'heure de mesure complémentaire est systématiquement +12h. Cette préférence pourrait être spécifique aux situations cibles sèches, puisque pour ces situations, dont les scores *CRPS* ont moins de poids que les autres, nous avons montré que l'heure de mesure optimale était +12h pour les niveaux moyens et élevés (Figure 6.6 a)).

Tableau 6.3 : Cinq meilleures combinaisons de deux niveaux de pression et d'une heure de mesure associée à chaque niveau, et gains G de performance moyenne associés, obtenus sur le bassin de la Saône.

	Niveaux (hPa)	Heure de mesure (h)									$G (A/A0)(\%)$	
		-12	-6	0	+6	+12	+18	+24	+30	+36		+42
1	925									X		3,99
	600					X						
2	925									X		3,82
	700					X						
3	925									X		3,73
	500					X						
4	850									X		3,63
	600					X						
5	775									X		3,61
	600					X						

Les résultats précédents sont résumés par la Figure 6.7, laquelle présente la meilleure combinaison et le gain G de performance associé, obtenu par la sélection des situations candidates selon la température mesurée à un niveau et une heure, un niveau et deux heures, ainsi qu'à deux niveaux et une heure chacun.

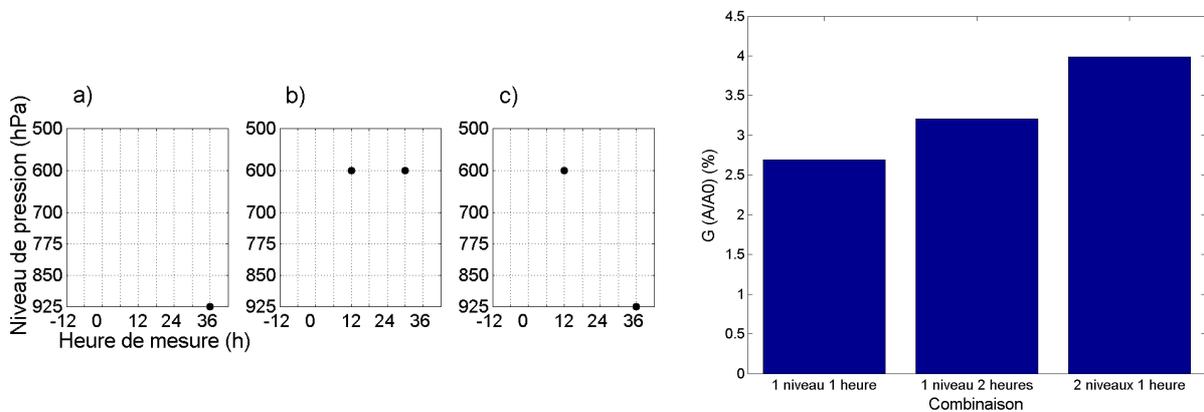


Figure 6.7 : Récapitulatif des meilleures combinaisons (gauche) obtenues. a) un niveau et une heure de mesure ; b) un niveau de pression et deux heures de mesure ; c) deux niveaux de pression et une heure de mesure chacun. Les gains G de performance moyenne associés sont indiqués à droite, pour le bassin de la Saône.

Le gain de performance apporté par cette sélection de situations candidates est manifeste, quel que soit le nombre d'heures de mesure et de niveaux de pression considérés. Il est néanmoins plus élevé en mesurant la température à deux niveaux différents, ainsi qu'à deux instants distants. Le gain ne tend pas à se stabiliser entre les cas b) et c) : on peut donc imaginer qu'en mesurant la température à trois niveaux différents, le gain soit encore plus intéressant. Néanmoins, nous arrêterons notre optimisation ici, même si cette perspective mériterait d'être exploitée.

6.2.3 Choix du nombre de situations candidates à retenir

Le nombre de situations candidates, baptisé NO par la suite, a été arbitrairement fixé à 3000 pour optimiser les niveaux et les heures de mesure. Ces paramètres étant maintenant choisis, nous pouvons à présent optimiser ce nombre NO . Le domaine sur lequel on recherche l'analogie de température reste identique, à savoir qu'il est constitué du point de grille le plus proche du bassin.

Avec une période candidate de 25 années (période de calibration), nous disposons de 9160 situations candidates pour chaque jour cible. Néanmoins nous ne nous autorisons pas à sélectionner les jours qui entourent chaque date cible, puisqu'une forte corrélation existe entre ces journées. Afin de nous affranchir de cette corrélation, nous avons exclu les situations candidates situées dans une fenêtre de plus ou moins quatre jours autour du jour cible, ce qui porte le nombre maximal de situations candidates à 9152 jours.

Afin d'avoir un aperçu de l'évolution des performances en fonction du nombre NO , nous avons d'abord fait varier le nombre de situations entre 500 et 9000 par pas de 500 (Figure 6.8 a)), puis nous avons fait varier, par pas de 100, le nombre NO dans la gamme qui contient le nombre optimal, à savoir l'intervalle [1500 ; 4000] dans le cas de la Saône (Figure 6.8 b)).

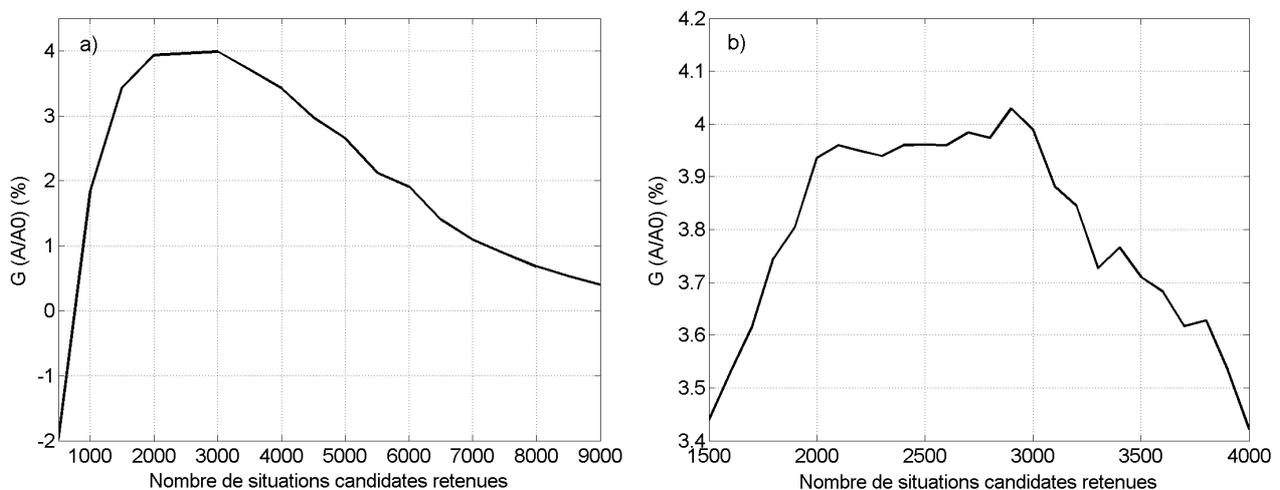


Figure 6.8 : Evolution du gain G de performance moyenne par rapport à la méthode A0 sur le bassin de la Saône, en fonction du nombre NO de situations candidates retenues. a) pour un nombre variant de 500 à 9000 par pas de 500 ; b) pour un nombre variant de 1500 à 4000 par pas de 100.

Les performances atteignent un palier lorsque le nombre NO est de l'ordre de 2000 à 3000, avec un nombre optimal de 2900. Ce dernier est d'ailleurs très proche de celui obtenu lorsque l'on applique une fenêtre temporelle de plus ou moins deux mois autour du jour cible. Néanmoins, le palier qui apparaît entre 2000 et 3000 traduit une robustesse plutôt faible. Certaines situations cibles doivent donc préférer un nombre proche de 2000, tandis que d'autres doivent préférer un nombre de l'ordre de 3000. En effet, la Figure 6.3 indique que la gamme de température observée est sensiblement différente au cours de l'été de celle observée au cours des autres saisons. Ainsi, si l'on fixe un écart maximal de température

entre situation candidate et situation cible, le nombre de situations candidates associé à une journée cible d'été devrait être plus faible que celui associé à une journée d'hiver. Finalement, le nombre NO optimal dépend-il de la température ?

Nous avons donc cherché à optimiser le nombre NO en fonction de la température. Pour cela, nous avons calculé, pour chaque journée cible, la moyenne arithmétique de la température mesurée à 925 hPa à +36h et de la température mesurée à 600 hPa à +12h. Dans un premier temps, la distribution des températures a été divisée en trois échantillons de même taille (classes *INF*, *MOY* et *SUP*). La Figure 6.9 représente la distribution des températures moyennes observées au cours de notre période cible, et les trois classes de température retenues.

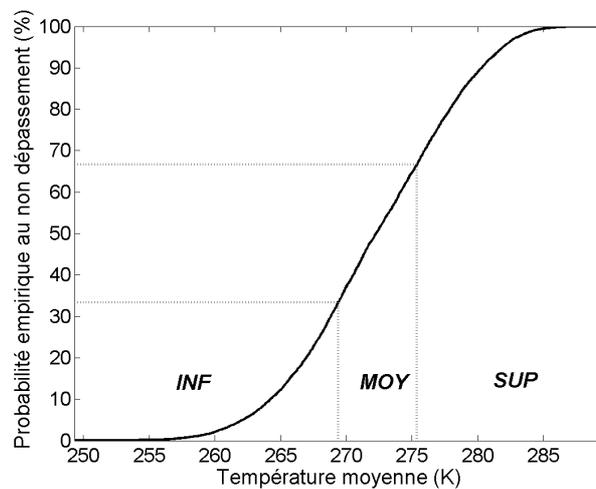


Figure 6.9 : Distribution des températures moyennes, à 925 hPa à +36h et à 600 hPa à +12h, observées sur la période de calibration, au point de grille le plus proche du bassin de la Saône (5°E ; 47,5°N). Les trois classes de température (*INF*, *MOY* et *SUP*) sont délimitées par les lignes pointillées.

La Figure 6.10 présente l'évolution du gain G de performance moyenne, pour un nombre NO variant de 1000 à 4500, pour chacune des trois classes de situations cibles.

Le nombre NO dépend effectivement de la gamme de températures dans laquelle une journée cible se situe. Pour une température moyenne inférieure à 275,38 K (soit environ +2°C), le nombre NO optimal est proche de 3000. Néanmoins un palier apparaît dès 2000 à 2200, avec un gain de performance presque stable jusqu'à un nombre de l'ordre de 4000. En revanche, une tendance assez nette se dessine pour les températures les plus élevées (température moyenne supérieure à 275 K). Le nombre NO optimal est proche de 2000, puis le gain de performance décroît presque linéairement. Ces résultats confirment notre hypothèse initiale, suivant laquelle le nombre de situations candidates à retenir doit être plus petit lorsque la température est élevée.

Il serait donc intéressant de retenir un nombre NO variable, fonction de la température : si la température moyenne dépasse le seuil de 275,38 K, nous pourrions retenir 2000 situations candidates, et dans le cas contraire, le nombre NO resterait fixé à 2900. Cependant, dans un

contexte opérationnel, des incertitudes pèsent sur la prévision de la température moyenne. Par conséquent, retenir un nombre NO sensiblement différent en fonction d'un seuil peut être pénalisant. Mais d'après la Figure 6.10, si nous retenons 2000 situations au lieu de 2900 (ou inversement) pour une température relativement proche du seuil, cela ne devrait pas avoir d'incidence sur les résultats.

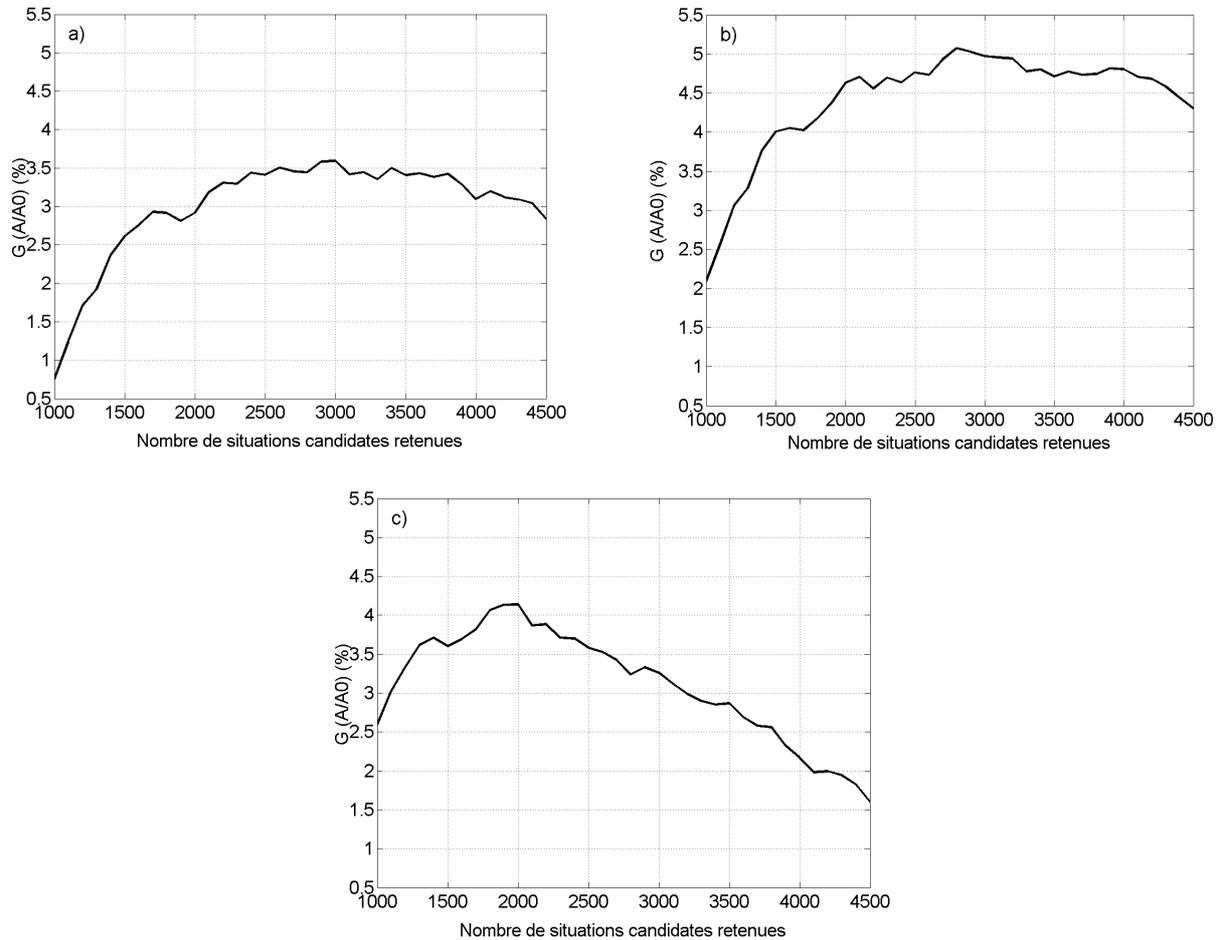


Figure 6.10 : Evolution du gain G de performance moyenne par rapport à la méthode A0, en fonction du nombre NO de situations candidates retenues. Pour les situations cibles dont la température moyenne appartient à la classe a) *INF* ; b) *MOY*; c) *SUP*.

Nous avons donc évalué le gain G en fixant le nombre NO à 2000 pour les situations cibles pour lesquelles la température moyenne dépasse 275,38 K, et à 2900 pour les situations restantes. De plus, le gain G a été évalué en distinguant les situations sèches des situations pluvieuses, la discrimination étant identique à celle appliquée pour l'optimisation du niveau et de l'heure de mesure (groupes S et P). Les gains G de performance par rapport à la méthode A0, exprimés en %, sont indiqués dans le Tableau 6.4.

Tableau 6.4 : Gains de performance G (%) obtenus par rapport à la méthode de référence A0 sur le bassin de la Saône, lorsque le nombre NO est variable ou fixe, et selon le type de situation cible.

	NO variable	NO fixe
Toutes situations confondues	4,29	4,03
Situations sèches	14,98	12,50
Situations pluvieuses	3,24	3,32

Pour l'ensemble des situations cibles, le gain G est légèrement supérieur quand on retient un nombre de situations candidates variable, fonction de la température. En revanche, si on distingue les situations suivant leur pluviosité, le gain, cela bénéficie aux situations sans pluie sur l'ensemble du bassin de la Saône ; tandis que la prévision des journées pluvieuses semble en moyenne légèrement dégradée. Pour cette raison, nous retiendrons donc un nombre NO fixe pour la suite de nos travaux, nombre égal à 2900 dans le cas de la Saône.

6.2.4 Choix du domaine d'analogie

6.2.4.1 Cartographie de la pertinence d'analogie

L'ensemble des paramètres optimisés précédemment a été obtenu en appliquant un domaine d'analogie élémentaire, constitué du point de grille le plus proche du bassin de la Saône. Nous pouvons à présent optimiser ce domaine. Afin d'évaluer la pertinence des points de grille, nous avons cartographié, sur la Figure 6.11, le gain de performance obtenu sur chaque domaine élémentaire au sein d'un plus grand domaine ($-5^{\circ}\text{E}/15^{\circ}\text{E}$; $37,5^{\circ}\text{N}/57,5^{\circ}\text{N}$).

Les paramètres appliqués sont les suivants :

- variables d'analogie : températures mesurées à 925 hPa à +36h et à 600 hPa à +12h ;
- nombre de situations candidates retenues : 2900 ;
- critère d'analogie : RMSE.

Les niveaux 1 et 2 de la méthode de référence A0 ont été appliqués avec les mêmes paramètres que ceux indiqués en début de partie.

Les points de grille les plus pertinents sont concentrés en périphérie immédiate du bassin de la Saône. Par conséquent, l'analogie sur la température doit être recherchée à une échelle spatiale locale. La zone la plus pertinente s'étend plutôt dans la direction zonale que dans la direction méridienne. En effet, le gradient thermique est généralement plus prononcé dans la direction méridienne (gradient latitudinal), si bien que sélectionner des situations candidates dont la température est mesurée à une latitude bien différente tend à dégrader la prévision. Au vu de cette cartographie, le domaine d'analogie optimal est a priori de faible étendue et centré sur le bassin. Le point de grille le plus pertinent est bien le point le plus proche du bassin, avec un gain G de +4,03 % par rapport à la méthode A0. Ce résultat n'est pas vraiment surprenant, puisque ce point a été choisi pour l'optimisation des paramètres précédents.

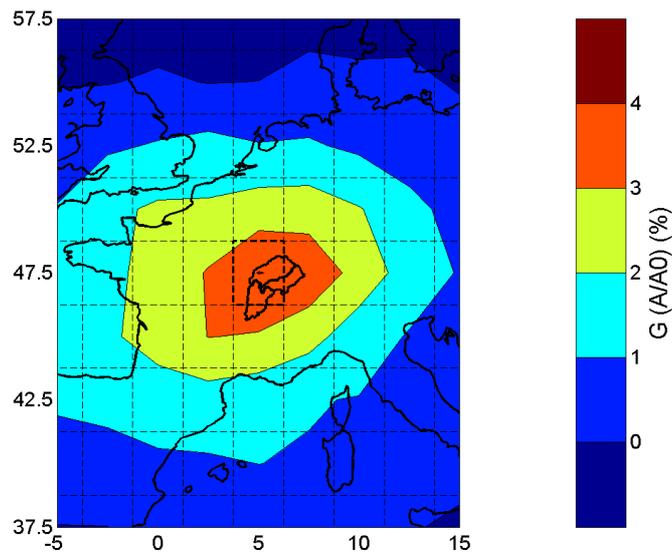


Figure 6.11 : Cartographie de la pertinence des points de grille exprimée en gain G de performance par rapport à la méthode de référence A_0 , sur le bassin de la Saône. Le point le plus proche du bassin, considéré comme domaine élémentaire au cours des optimisations précédentes, est indiqué par la maille au contour plus épais.

Et si nous avons choisi un autre point de grille, les paramètres optimisés précédemment auraient-ils été différents ? Pour répondre à cette question, qui a toute son importance dans le cadre d'une adaptation à un autre bassin, nous avons effectué le même travail d'optimisation sur les niveaux et heures de mesure, en considérant d'autres points proches du bassin de la Saône. Les résultats, non présentés ici, montrent que les paramètres optimaux sont finalement identiques si l'on considère un autre point de grille proche du bassin.

6.2.4.2 Optimisation du domaine

Le point le plus pertinent identifié sur la Figure 6.11 (choisi arbitrairement en début de section) constitue le point de départ de l'optimisation du domaine d'analogie. La méthode d'optimisation est identique à celle appliquée jusqu'à présent (Bontron, 2004).

La Figure 6.12 présente l'évolution des performances en fonction de la longueur diagonale du domaine d'analogie. Il s'avère que le domaine optimal correspond au domaine de départ, c'est-à-dire au point de grille le plus proche du bassin (5°E ; $47,5^{\circ}\text{N}$). Le gain décroît ensuite presque linéairement avec l'agrandissement du domaine. Bien que la figure ne l'indique pas, nous avons également remarqué que les domaines correspondant aux longueurs diagonales croissantes sont systématiquement étendus dans la direction zonale, ce qui corrobore les résultats obtenus par cartographie de la pertinence d'analogie.

Finalement, nous retenons le domaine d'analogie constitué du point de grille le plus proche du bassin.

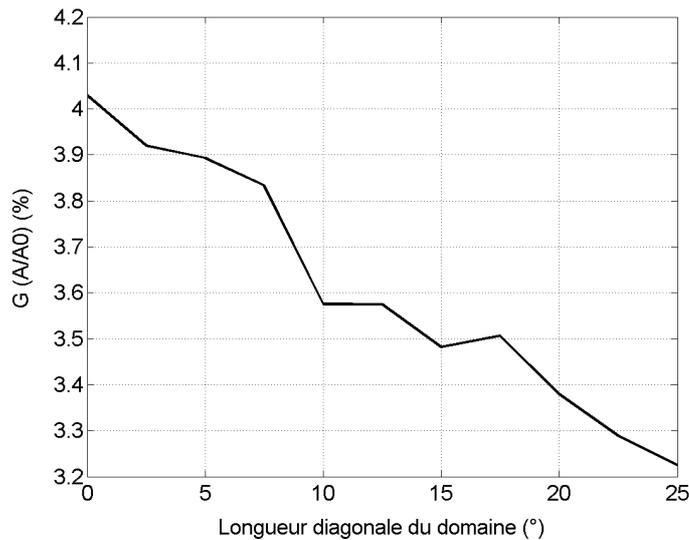


Figure 6.12 : Evolution du gain G de performance, obtenu par rapport à la méthode A0 sur le bassin de la Saône, en fonction de la longueur diagonale du domaine d'analogie.

6.2.5 Choix du critère d'analogie

Les optimisations précédentes ont été entreprises à l'aide d'un critère d'analogie choisi arbitrairement : le critère de distance $RMSE$. Le choix arbitraire d'un critère dépend du type d'analogie : dans le cas de la température, le choix d'un critère de distance nous est apparu intuitif puisque le but de cette pré-sélection n'est pas de retenir des situations dont les champs de température se ressemblent dans leur forme, ou dont les champs sont bien corrélés, mais dont les champs sont plutôt proches en termes de distance (masse d'air similaire).

Dans cette partie, nous avons succinctement étudié l'influence du critère d'analogie sur les performances. Nous avons appliqué les quatre critères cités au début de la section 6.2 : ceux de distance moyenne $RMSE$ et MAE , le critère de corrélation R lié à la co-fluctuation, ainsi que le critère de forme TWS . Par simplicité, nous avons repris certains paramètres optimisés avec le critère $RMSE$ (niveaux et heures de mesure, nombre de situations candidates à retenir).

Pour pouvoir appliquer les critères R et TWS , nous avons dû choisir un domaine d'au moins quatre points de grille. Deux domaines différents ont été testés (Figure 6.13) : un domaine carré de faible étendue, constitué des quatre points de grille les plus proches du bassin de la Saône ; et un domaine plus large, de forme rectangulaire, centré approximativement sur le bassin et constitué de 15 points de grille.

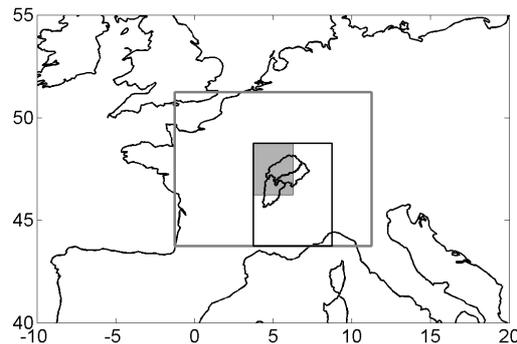


Figure 6.13 : Domaines d'analogie testés pour les différents critères d'analogie (maillage grisé : *RMSE* et *MAE* ; domaines de contour noir et gris : *R* et *TWS*).

Les gains G de performance par rapport à la méthode A0 sont reportés dans le Tableau 6.5. L'utilisation d'un critère de forme tel que le critère *TWS* ou d'un critère lié à la co-fluctuation tel que le critère *R* mène à des pertes de performance notables par rapport à la méthode de référence A0. Ces pertes sont presque équivalentes selon que le domaine d'analogie appliqué soit de faible étendue ou d'étendue plus élevée. En revanche, les critères de distance moyenne tels que le *RMSE* et le *MAE* permettent d'obtenir un gain positif lorsque les situations candidates sont sélectionnées par la température.

Il est évident que, pour les critères *TWS* et *R* notamment, ces valeurs dépendent des paramètres optimisés précédemment. Néanmoins, les pertes de performance sont telles que la ré-optimisation des paramètres (niveaux et heures de mesure, nombre $N0$) ne devrait pas inverser les tendances obtenues ici. Le gain G le plus élevé est obtenu avec le critère *RMSE* : nous le retenons définitivement pour ce niveau préliminaire de sélection des situations candidates selon la température.

Tableau 6.5 : Gain G (%) de performance moyenne, obtenu par rapport à la méthode de référence A0 sur le bassin de la Saône, en fonction du critère d'analogie utilisé (Np = domaine constitué de N points de grille).

Critère	<i>RMSE</i>	<i>MAE</i>	<i>R</i> (4p)	<i>R</i> (12p)	<i>TWS</i> (4p)	<i>TWS</i> (12p)
G (%)	4,03	3,80	-5,29	-4,56	-4,95	-3,95

6.3 Validation et diagnostic

Nous venons d'optimiser un niveau de pré-sélection des situations candidates au premier niveau de sélection d'analogues de la méthode de référence A0. Le gain de performance étant acquis, nous proposons de retenir et de nommer ce nouvel algorithme A1, qui inclut donc trois niveaux de sélection :

- un niveau 0 de pré-sélection préliminaire, basé sur une analogie de l'état thermique qui remplace le critère basé sur la saisonnalité ;

- un niveau 1 de sélection d'analogues en termes de circulation générale ;
- et un niveau 2 de sélection d'analogues en termes d'humidité.

Il est à présent temps de valider ce nouvel algorithme et d'effectuer un diagnostic des performances. Cette section est scindée en trois parties. La première consiste à effectuer les mêmes optimisations de paramètres sur le bassin de la Seine que sur le bassin de la Saône. Seuls les principaux résultats seront présentés. La seconde partie est consacrée à la validation de l'algorithme sur les deux bassins tests. Les performances seront alors évaluées sur la période de validation que nous avons définie précédemment. Enfin, un diagnostic des performances par comparaison à la sélection par saisonnalité sera établi dans la dernière partie de cette section.

6.3.1 Adaptation au bassin de la Seine

Dans le cadre d'une adaptation à un autre bassin versant comme celui de la Seine, nous pourrions choisir de ré-optimiser certains paramètres en conservant, par exemple, les variables d'analogie retenues sur le bassin de la Saône. Mais afin de valider les résultats obtenus sur le bassin de la Saône, nous avons décidé d'appliquer la même démarche au bassin de la Seine. Le seul paramètre que nous avons conservé est le critère d'analogie, qui ne dépend pas du bassin auquel on s'intéresse : le critère *RMSE*.

La méthode de référence A0 a été adaptée au bassin de la Seine par Plasse (2008). Nous appliquons donc les niveaux de référence avec les paramètres optimaux pour ce bassin.

Choix du point de grille

Il est d'abord nécessaire de choisir le point de grille approprié sur lequel les optimisations des niveaux et heures de mesure seront faites. En effet, le bassin de la Seine n'est pas centré à proximité d'une maille de la grille à résolution $2,5^\circ$ (Figure 6.14 a) : on ne peut *a priori* connaître le point le plus pertinent. Nous avons donc considéré la combinaison optimale obtenue sur le bassin de la Saône (925 +36h et 600 +12h) et retenu 2900 situations candidates. La pertinence d'analogie a ensuite été cartographiée sur un domaine de taille réduite, à peu près centré sur le bassin de la Seine et constitué de neuf points de grille (Figure 6.14 b)). Le point de grille le plus pertinent est le point de coordonnées ($2,5^\circ\text{E}$; $47,5^\circ\text{N}$). Nous allons par conséquent optimiser les niveaux et heures de mesure en ce point.

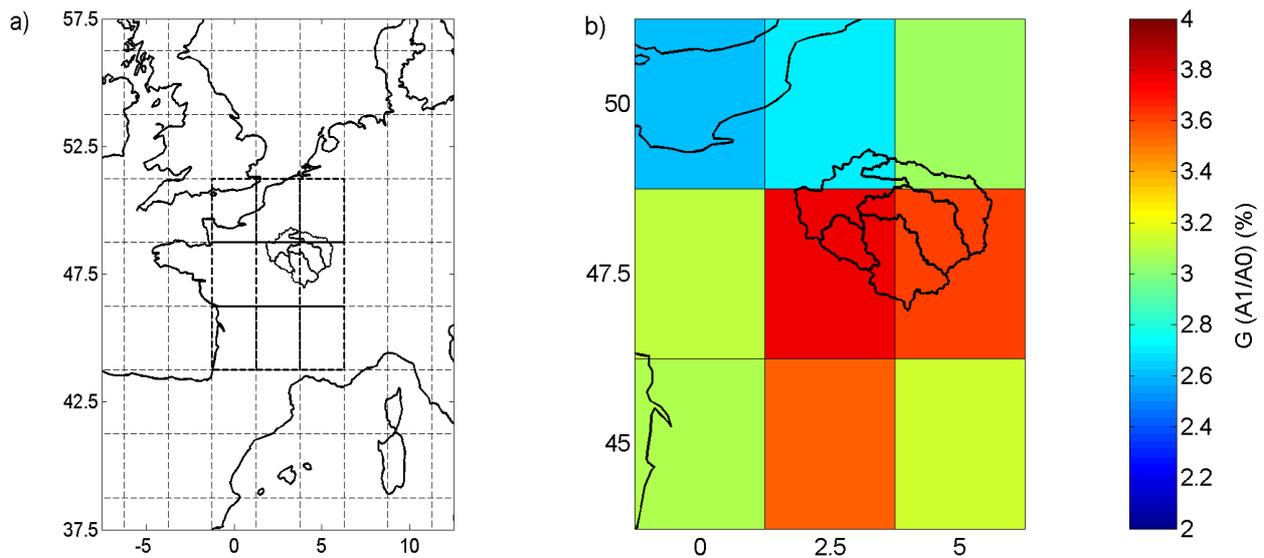


Figure 6.14 : a) Représentation du bassin de la Seine et des mailles de la grille à résolution 2,5° ; b) Cartographie de la pertinence d'analogie, exprimée en gain G de performance par rapport à la méthode de référence A_0 , obtenue sur le bassin de la Seine avec les variables d'analogie retenues sur le bassin de la Saône.

Niveaux et heures de mesure

Nous avons effectué la même optimisation que pour le bassin de la Saône. Dans ce contexte, le nombre de situations candidates retenues a arbitrairement été fixé à 3000. La Figure 6.15 présente l'évolution du gain G en fonction du niveau et de l'heure de mesure auxquels la température est mesurée.

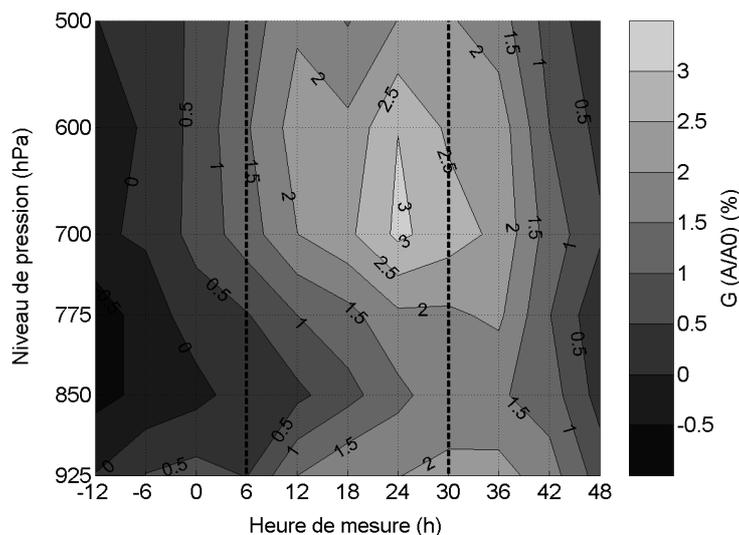


Figure 6.15 : Gain G de performance moyenne, par rapport à la méthode A_0 , en fonction du niveau et de l'heure de mesure de la température, pour le bassin de la Seine. La période du cumul pluviométrique journalier est limitée par les lignes verticales discontinues.

Nous retrouvons à peu près la même répartition du gain G que celle obtenue sur la Saône (Figure 6.5). Néanmoins, le gain G maximal est ici obtenu au niveau de pression 700 hPa à

+24h (contre 925 hPa à +36h pour la Saône), avec un gain G de 3,05 %.

Nous avons également distingué les situations suivant leur pluviosité, en les séparant en trois groupes : le groupe S des situations sèches (pluie journalière nulle simultanément sur les cinq sous-bassins, soit environ 13,8 % de l'échantillon), le groupe P des situations pluvieuses (pluie non nulle simultanément sur les cinq sous-bassins, soit environ 64,3 % de l'échantillon), et les situations restantes. Comme sur le bassin de la Saône, l'optimisation du niveau et de l'heure de mesure a été effectuée pour le groupe S d'une part (Figure 6.16 a)), et pour le groupe P d'autre part (Figure 6.16 b)).

Nous retrouvons approximativement les mêmes résultats que ceux obtenus sur le bassin de la Saône, même si pour le groupe P, le gain maximal est atteint dans les couches moyennes au lieu des couches basses de la troposphère. L'hypothèse suivant laquelle les paramètres optimaux de mesure de la température dépendent du passage ou non d'une perturbation, et donc de l'occurrence d'un changement de masse d'air, reste donc plausible.

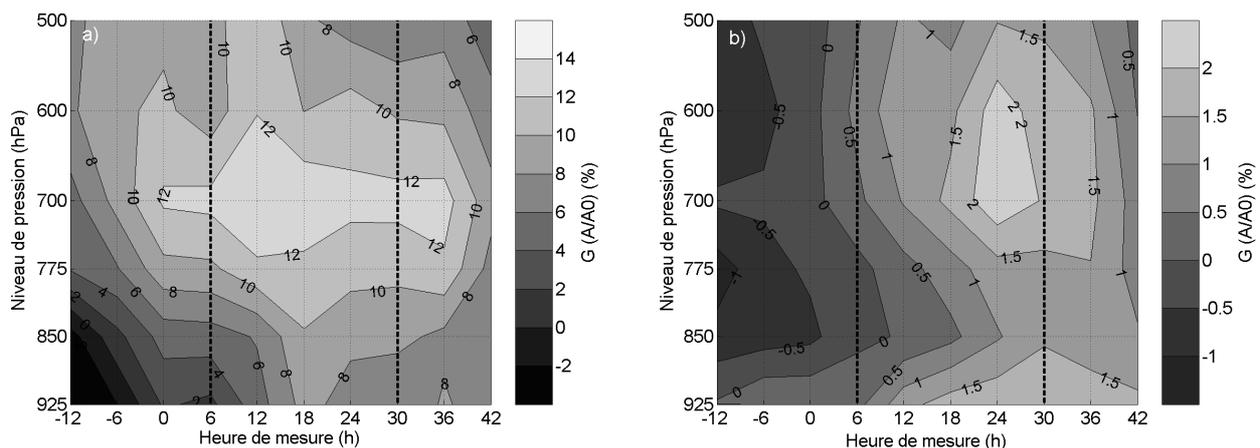


Figure 6.16 : Gain G de performance moyenne, par rapport à la méthode A0, en fonction du niveau et de l'heure de mesure de la température, pour le bassin de la Seine. a) pour le groupe S ; b) pour le groupe P.

Nous avons également optimisé la combinaison de deux niveaux différents avec une heure de mesure chacun. De la même manière que sur la Saône, les dix meilleures combinaisons d'un niveau et d'une heure de mesure ont été retenues ; et pour chacune de ces combinaisons, nous avons fait varier le niveau et l'heure de mesure de la seconde température. Finalement, nous avons retrouvé la même combinaison que pour le bassin de la Saône. Ainsi, pour le bassin de la Seine, la température doit être mesurée à 925 hPa à +36h et à 600 hPa à +12h. Le gain G par rapport à A0 s'élève alors à 3,82 % sur l'ensemble des sous-bassins de la Seine.

Nombre de situations candidates à retenir

Le nombre de situations candidates à retenir a également été optimisé pour le bassin de la Seine (Figure 6.17). Le nombre $N0$ optimal s'élève à 3000 (gain G de 3,82 % par rapport à A0). Il est donc proche du nombre retenu pour le bassin de la Saône.

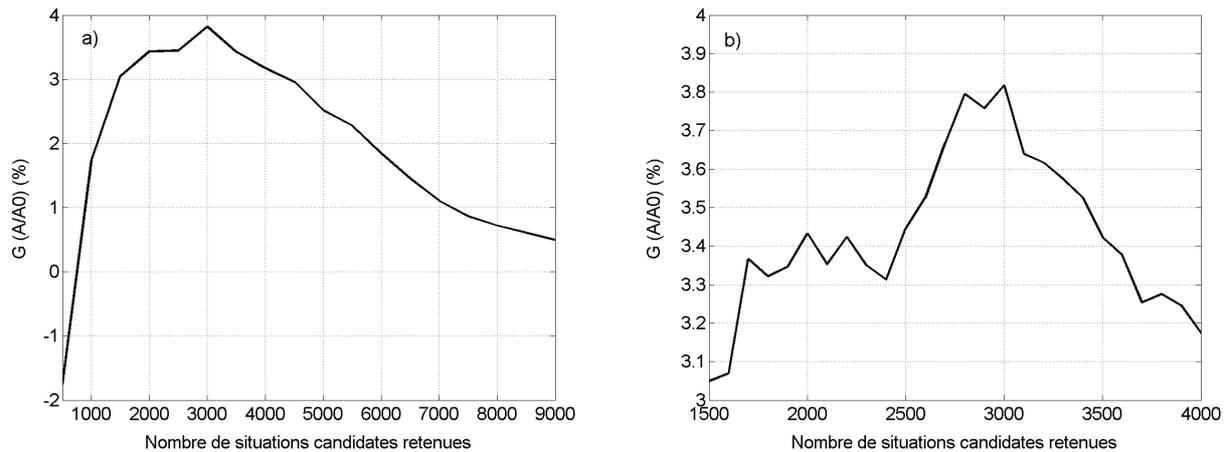


Figure 6.17 : Evolution du gain G de performance moyenne par rapport à la méthode A0, en fonction du nombre N_0 de situations candidates retenues, pour le bassin de la Seine. a) pour un nombre variant de 500 à 9000 par pas de 500 ; b) pour un nombre variant de 1500 à 4000 par pas de 100.

Domaine d'analogie

Nous pouvons à présent optimiser le domaine d'analogie avec les paramètres optimisés précédemment (variables d'analogie et nombre N_0), le domaine de départ étant le domaine élémentaire le plus pertinent utilisé jusqu'ici. La Figure 6.18 présente l'évolution du gain G en fonction de la longueur diagonale du domaine d'analogie. Comme pour la Saône, le domaine optimal correspond au domaine de départ, c'est-à-dire au point de grille le plus pertinent, situé à proximité du bassin ($2,5^\circ\text{E}$; $47,5^\circ\text{N}$).

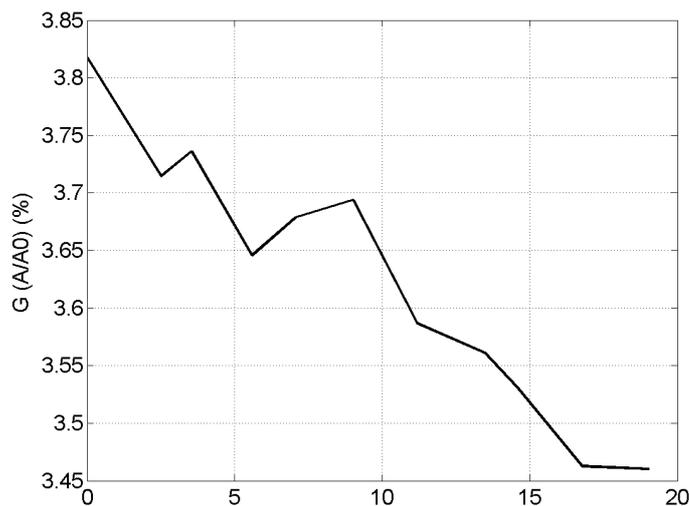


Figure 6.18 : Evolution du gain G de performance, par rapport à la méthode A0, en fonction de la longueur diagonale du domaine d'analogie, pour le bassin de la Seine.

Finalement, les paramètres optimisés sur le bassin de la Seine sont les suivants :

- variables d'analogie : températures mesurées à 925 hPa à +36h et à 600 hPa à +12h ;

- domaine d'analogie : élémentaire, constitué d'un point de grille proche du bassin (2,5°E ; 47,5°N)
- nombre de situations candidates retenues : 3000 ;
- critère d'analogie : *RMSE*.

Le gain G de performance par rapport à notre méthode de référence A0 s'élève à 3,82%. Malgré la proximité géographique des deux bassins tests, nous pouvons considérer que notre méthode A1 est validée sur le bassin de la Seine, le gain positif obtenu sur le bassin de la Saône étant conservé.

6.3.2 Validation sur les deux bassins tests

Nous nous sommes réservés une période cible indépendante de la période sur laquelle la méthode A1 a été développée. Nous proposons à présent de valider l'algorithme par l'évaluation du gain G sur cette période pour les bassins de la Saône et de la Seine. Les gains obtenus sur chaque sous-bassin et sur l'ensemble du bassin sont reportés dans le Tableau 6.6 (Saône) et dans le Tableau 6.7 (Seine). A titre de comparaison, les gains G obtenus en calibration y figurent également.

Tableau 6.6 : Gains G de performance moyenne, obtenue par la méthode A1 par rapport à la méthode A0 sur le bassin de la Saône, évalués sur les deux périodes cibles indépendantes de validation et de calibration.

	Validation	Calibration
Doubs	4,6	4,03
Petite Saône	4,98	3,38
Saône Intermédiaire	3,93	4,64
Ensemble du bassin	4,5	4,03

Tableau 6.7 : Gains G de performance moyenne, obtenue par la méthode A1 par rapport à la méthode A0 sur le bassin de la Seine, évalués sur les deux périodes cibles indépendantes de validation et de calibration.

	Validation	Calibration
Yonne-C	3,67	4,41
Seine-Int	3,29	3,37
Seine-B	3,78	3,95
Marne-F	3,96	3,61
Loing-E	3,11	3,77
Ensemble du bassin	3,55	3,82

Le gain obtenu sur la période de calibration est conservé sur la période de validation, et ce pour l'ensemble des bassins, avec des valeurs largement positives. Nous remarquons cependant que le gain obtenu en calibration peut être supérieur à celui obtenu en validation et inversement. Deux raisons à cela nous paraissent possibles :

- sur chacune des périodes, et malgré le soin apporté à leur construction, certaines caractéristiques non explorées de la climatologie d'un bassin peuvent être légèrement différentes et peuvent influencer le score *CRPSS* ;
- la période cible de la période de calibration est identique à la période candidate. Les climatologies des périodes cible et candidate étant les mêmes, la prévision climatologique est par conséquent avantagée sur la période de calibration. Cette raison expliquerait notamment que le gain en validation puisse être supérieur au gain en calibration, le score *CRPSS* étant plus élevé en validation qu'en calibration.

6.3.3 Comparaison à la sélection par saisonnalité

Maintenant que l'algorithme est validé, un diagnostic est proposé afin de comprendre l'apport de cette pré-sélection par la comparaison des résultats avec ceux obtenus par la pré-sélection initiale par saisonnalité. Ce diagnostic est établi sur la période de calibration.

6.3.3.1 Gain en finesse et en justesse

Nous avons distingué finesse et justesse de la prévision par l'intermédiaire du score *CRPS* (cf. 3.3.3). Il est effectivement intéressant de voir si en moyenne la prévision obtenue par la méthode A1 est plus fine et si elle est plus juste que celle obtenue par la méthode A0. Après décomposition du score *CRPS*, nous pouvons évaluer le gain G de performance moyenne, obtenu par la méthode A1 par rapport à la méthode A0, en termes de finesse et justesse. Les tableaux 6.8 et 6.1 indiquent respectivement les gains obtenus sur le bassin de la Saône et sur celui de la Seine.

Nous remarquons que les gains sont tous positifs, preuve de l'intérêt de cette pré-sélection selon la température. La prévision émise par la méthode A1 est en moyenne bien plus fine que celle émise par A0. En revanche, le gain en justesse est plus faible mais non négligeable. Certains bassins (Saône Intermédiaire, Yonne-C, Loing-E) se distinguent même des autres par un gain en justesse assez élevé par rapport au gain en finesse.

Tableau 6.8 : Gains de performance, obtenus par la méthode A1 par rapport à la méthode A0, sur le bassin de la Saône, par décomposition du gain global en finesse et en justesse.

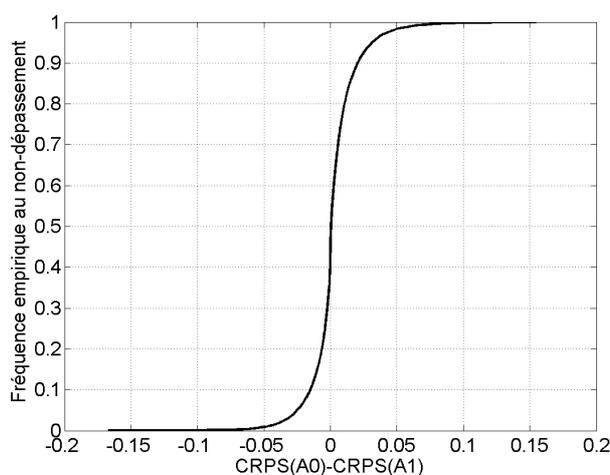
	Gain global	Gain en finesse	Gain en justesse
Doubs	4,04	5,82	2,17
Petite Saône	3,38	4,98	1,79
Saône Intermédiaire	4,64	5,27	4,04
<i>Ensemble des bassins</i>	<i>4,03</i>	<i>5,36</i>	<i>2,7</i>

Tableau 6.9 : Gains de performance, obtenus par la méthode A1 par rapport à la méthode A0, sur le bassin de la Seine, par décomposition du gain global en finesse et en justesse.

	Gain global	Gain en finesse	Gain en justesse
Yonne-C	4,41	4,99	3,91
Seine-Int	3,37	3,95	2,91
Seine-B	3,95	4,75	3,28
Marne-F	3,61	4,38	2,95
Loing-E	3,77	4,03	3,56
<i>Ensemble des bassins</i>	<i>3,82</i>	<i>4,42</i>	<i>3,32</i>

6.3.3.2 Robustesse du gain de performance

Le gain de performance moyenne apporté par ce niveau pré-sélectif n'est pas négligeable, mais est-il systématique ? Pour le vérifier, nous avons simplement évalué, pour chaque situation cible de la période de calibration, l'écart de score *CRPS* entre les méthodes A0 et A1. La distribution empirique des écarts a ensuite été construite. La Figure 6.19 présente cette distribution pour le bassin du Doubs.

Figure 6.19 : Distribution des écarts de score *CRPS* entre les méthodes A0 et A1 calculés sur la période de calibration, pour le bassin du Doubs.

La distribution est presque centrée en 0, ce qui signifie que le gain¹ de performance n'est pas du tout systématique. Il y a presque autant de situations cibles pour lesquelles le gain est positif que de situations pour lesquelles le gain est négatif. Néanmoins, la valeur médiane de la distribution est très légèrement positive. Si nous effectuons un zoom sur les parties extrêmes de la distribution (Figure 6.20), nous remarquons tout de même que les écarts positifs sont un peu plus importants que les écarts négatifs.

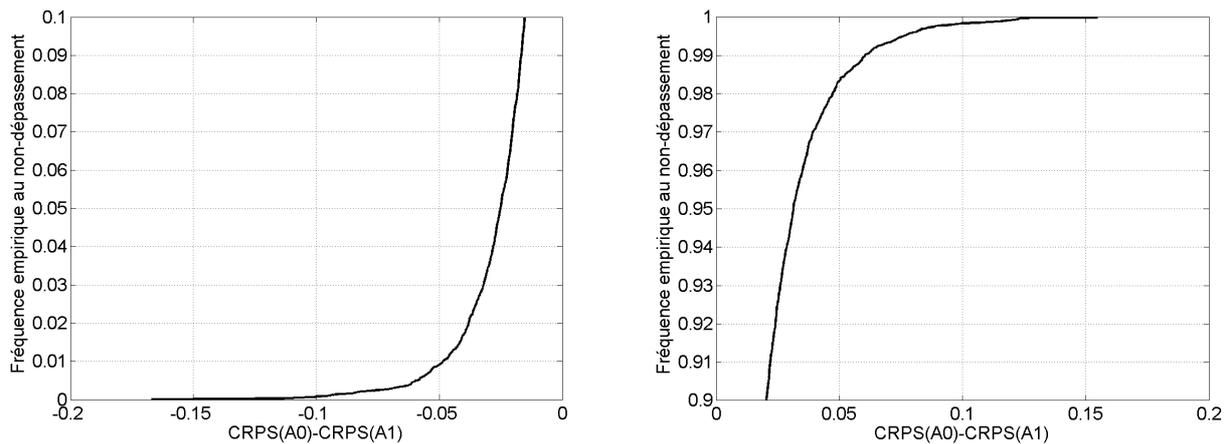


Figure 6.20 : Agrandissements des parties extrêmes de la distribution de la Figure 6.19 : fréquence comprise entre 0 et 0,1 (gauche) et fréquence comprise entre 0,9 et 1 (droite).

Intéressons-nous aux deux situations extrêmes, pour lesquelles l'écart de score *CRPS* est le plus faible et le plus élevé. Il s'agit respectivement du 9 juin 1993 et du 30 avril 1973, journées au cours desquelles les précipitations observées sur le bassin du Doubs sont respectivement de 11,2 mm et 0 mm. Nous avons évalué, pour chacune d'elle, la distance temporelle entre les dates des situations analogues retenues à l'issue du niveau 2 et la date de la situation cible. La Figure 6.21 présente, pour ces deux dates, les distances évaluées pour chacune des analogues et les précipitations correspondantes. Afin de comprendre les écarts de score *CRPS* entre les méthodes A0 et A1, les distances et les précipitations associées obtenues par la méthode A0 sont également indiquées.

¹ Le gain évoqué ici n'est pas assimilable au gain G utilisé jusqu'à présent. Le gain de performance correspond à un écart positif de score *CRPS* entre les méthodes A0 et A1.

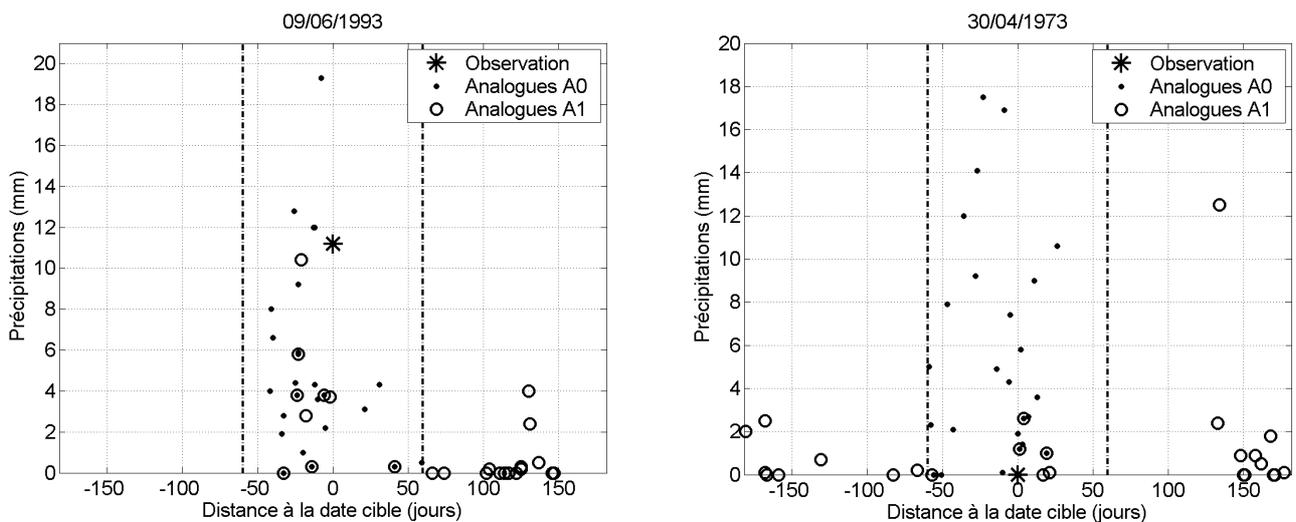


Figure 6.21 : Distance temporelle (en jours) entre chacune des dates analogues et la date cible (09/06/1993 à gauche, 30/04/1973 à droite), et précipitations observées sur le bassin du Doubs lors des dates analogues ; par les méthodes A0 (points) et A1 (cercles).

On comprend mieux pourquoi de tels écarts ont été observés. Ainsi, le nombre de dates analogues communes aux méthodes A0 et A1 n'est que de 6 le 9 juin 1993 et de 3 le 30 avril 1973. Pour le 9 juin 1993, bon nombre des situations analogues finalement sélectionnées au niveau 2, qui se situent à l'extérieur de la fenêtre temporelle de quatre mois (121 jours), présentent des cumuls proches de 0 mm, alors que le cumul observé est de 11,2 mm. En revanche, pour le 30 avril 1973, les nombreuses dates analogues extérieures à la fenêtre de quatre mois présentent des cumuls proches de 0 mm, pour un cumul observé de 0 mm. A l'intérieur de la fenêtre, les précipitations observées lors des dates analogues retenues par la méthode A1 sont inférieures à 3 mm, tandis que celles retenues par A0 sont souvent élevées, avec pas moins de cinq cumuls supérieurs à 10 mm.

Pour le 30 avril 1973, on peut comprendre que des dates analogues situées à près de six mois de la date cible (soit aux environs des mois d'octobre-novembre) aient pu être candidates au niveau 1. En effet, au printemps comme à l'automne, saisons de transition entre les deux saisons contrastées d'hiver et d'été, les températures peuvent être du même ordre dans la basse et la moyenne troposphère. Ainsi par exemple, les températures à 925 hPa à +36h et à 600 hPa à +12h observées le 31 octobre 1975 (jour le plus éloigné, situé à 181 jours de la date cible et classé en 2816ème position parmi les 2900 situations candidates) sont de +9,6°C et -6,3°C ; tandis que celles observées le jour cible sont de +15°C et -7,8°C. C'est notamment la faible différence de T_{600} qui explique la sélection de cette date candidate.

Mais comment expliquer, pour le 9 juin 1993, que des dates analogues situées à quatre voire cinq mois de la date cible aient pu être candidates ? Considérons le jour le plus éloigné, le 3 novembre 1994, qui se situe à 147 jours de la date cible (et classé pourtant en 1269ème position parmi les 2900 dates candidates retenues). Les températures observées à cette date, à 925 hPa à +36h et à 600 hPa à +12h (resp. +15,7°C et -6,6°C), sont assez proches de celles observées le jour cible (+19,1°C et -4,4°C). C'est encore une fois le faible écart de T_{600} qui

explique la sélection de cette date candidate. Cependant la masse d'air était également douce en basses couches. Nous nous sommes intéressés aux champs de géopotentiels $Z1000$ et $Z500$ pour comprendre pourquoi cette situation, pour laquelle le cumul observé est de à 0 mm, a été sélectionnée au niveau 1 parmi les situations analogues à la situation cible du 09/06/1993. Les champs de géopotential $Z500$ à +24h et de $Z1000$ à +12h de la date cible et de cette date analogue sont présentés sur la Figure 6.22.

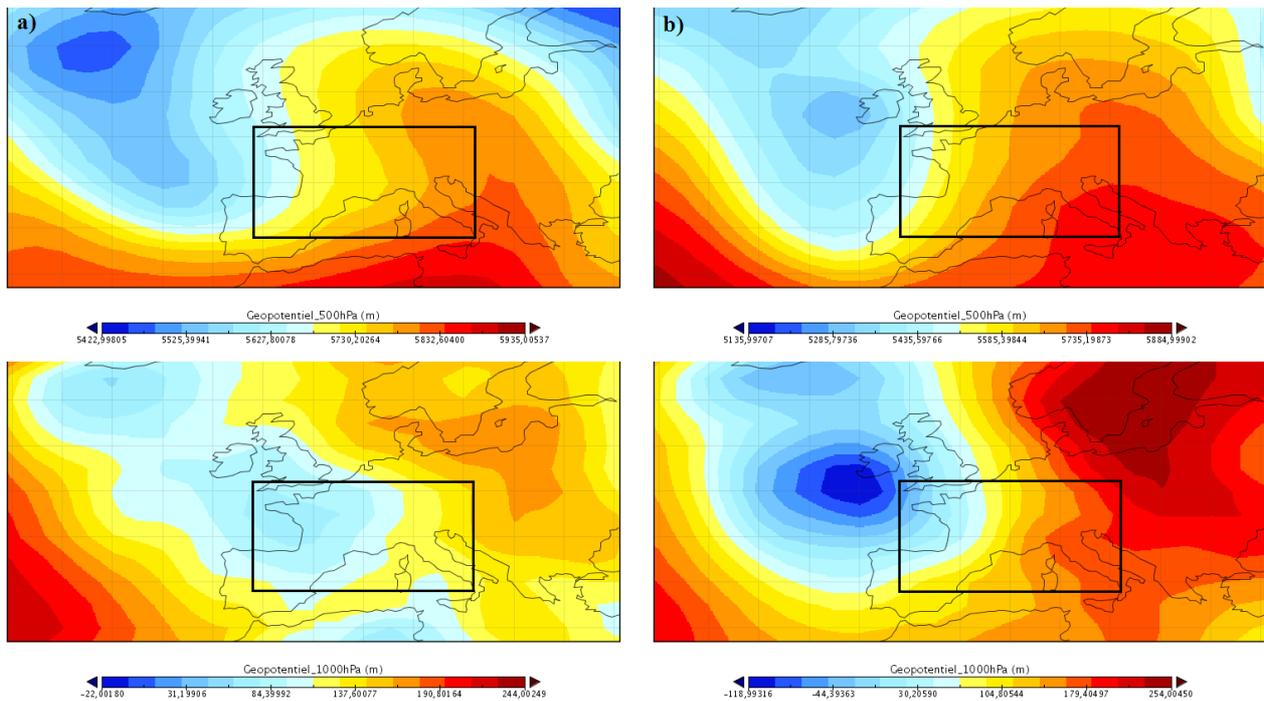


Figure 6.22 : Champs de $Z500$ à +24h (haut) et de $Z1000$ à +12h (bas), extraits des réanalyses pour a) la date cible (09/06/1993) et b) pour la date analogue (03/11/1994) la plus éloignée de la date cible. Le domaine spatial d'analogie du niveau 1 est indiqué par le rectangle noir.

On remarque que l'analogie sur la variable $Z500$ est plutôt bonne sur le domaine spatial du niveau 1, avec une circulation de sud à sud-ouest visible sur les deux champs, même si la situation cible présente un axe de talweg d'altitude prononcé du nord de l'Espagne au sud de la France, axe qui n'apparaît pas pour la situation analogue considérée. En revanche, le champ de géopotential $Z1000$ est différent entre les deux situations. Le champ observé pour la situation cible révèle un minimum dépressionnaire centré au sud de la Bretagne, avec un talweg de surface axé du centre de la dépression au nord-est de la France. La situation analogue ne présente pas de talweg aussi prononcé, et la dépression de surface, plus étendue, est centrée au sud-ouest de l'Irlande. Finalement, les champs, même si leur forme est globalement similaire, en particulier pour la variable $Z500$, ne contiennent pas les mêmes éléments. Ces derniers sont, pour la situation cible, révélateurs d'une perturbation frontale, ce qui n'est pas le cas pour la situation analogue.

6.3.3.3 Distances des dates analogues aux dates cibles

Pour les deux situations extrêmes que nous venons d'analyser, le nombre de dates

communes est très faible, ce qui explique évidemment les valeurs très différentes de CRPS. Finalement, cela nous amène à nous poser une question plus générale : quelle est la proportion de situations analogues, finalement retenues par la méthode A1, situées à l'extérieur de la fenêtre initiale de quatre mois ? Pour y répondre, nous sommes allés plus loin dans l'analyse. Nous avons évalué, pour une distance D comprise entre -184 et 184, la proportion de dates des situations analogues, retenues au niveau 2, situées à la distance D de la date de la situation cible (Figure 6.23).

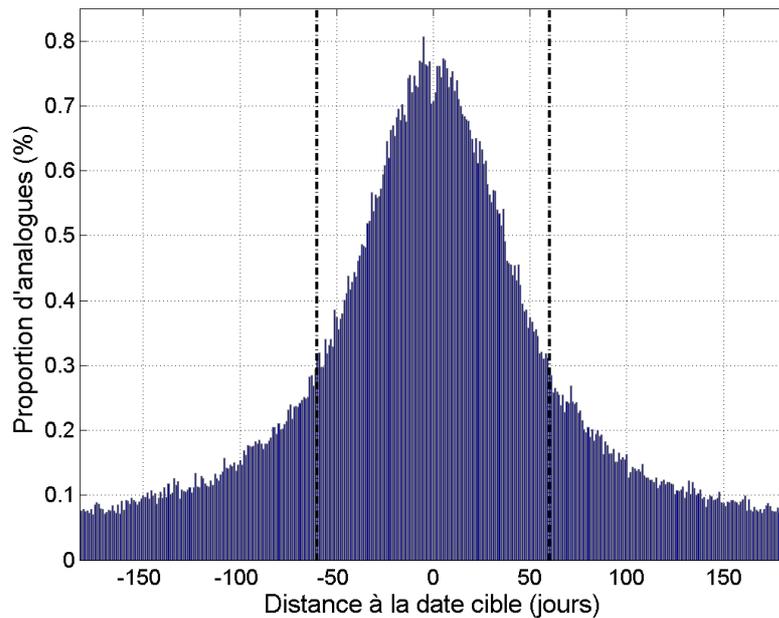


Figure 6.23 : Proportion de dates des situations analogues retenues à l'issue du niveau 2, par application de la méthode A1 au bassin de la Saône, en fonction de leur distance à la date cible. La fenêtre de largeur quatre mois dans laquelle les situations candidates retenues par la méthode A0 est indiquée par les traits verticaux discontinus épais.

La répartition des distances est fortement symétrique, centrée en 0, ce qui signifie que pour une distance absolue fixée, une situation analogue retenue au niveau 2 a autant de chance de se situer avant le jour cible qu'après. Ainsi d'après cette répartition, la probabilité qu'une date analogue se situe à plus de cinq mois du jour cible est d'environ 6%, ce qui est assez faible. En revanche, la probabilité qu'une date analogue se situe dans la fenêtre initiale de quatre mois est de 66,6%, soit deux tiers des situations. En considérant qu'en moyenne, le gain apporté par cette sélection est positif, nous retrouvons également dans ce graphique le résultat obtenu en testant les différentes largeurs de fenêtre (Figure 6.1). En effet, plus la distance est élevée, plus la chance qu'une situation analogue soit retenue est faible.

6.4 Conclusion

La pré-sélection initiale des situations candidates de l'algorithme A0 permet de prendre en compte explicitement les effets saisonniers, et donc implicitement la température. Or, étant

basé sur un critère calendaire figé, elle s'avère trop rigide et semble dans certains cas écarter des situations intéressantes, si bien qu'une situation analogue potentiellement informative du cumul de précipitation prévu, située à l'extérieur de la fenêtre temporelle, ne peut être considérée.

Nous avons donc souhaité prendre en compte la température de la masse d'air explicitement dans la procédure de pré-sélection. Ceci a impliqué l'optimisation de l'ensemble des paramètres : critère d'analogie, nombre de situations à sélectionner, domaine d'analogie, heure(s) de mesure et niveau(x) de pression de la variable d'analogie. Cependant, la combinatoire liée à cette optimisation s'est révélée trop importante pour que tous les paramètres optimaux soient ajustés simultanément. Nous avons donc préalablement choisi le critère d'analogie (*RMSE*), le domaine d'analogie (un seul point de grille proche du bassin) et le nombre de situations candidates à retenir, que nous avons fixé à un nombre équivalent à celui obtenu par l'application d'une fenêtre temporelle de quatre mois centrée sur le jour cible.

La première étape a concerné le choix des niveaux de pression et des heures de mesure auxquels la température est mesurée. Nous avons considéré successivement un niveau et une heure de mesure, puis un niveau et deux heures de mesure, et enfin deux niveaux associé chacun à une heure de mesure. Les résultats ont montré, sur les deux bassins tests, que la température doit être mesurée à deux niveaux de pression différents, l'un situé dans les très basses couches de la troposphère (925 hPa) et associé à une heure de mesure tardive, extérieure à la période du cumul journalier (+36h), et l'autre dans les couches moyennes (600 hPa) et associé à un instant de la journée inclus dans la période du cumul (+12h). Le gain *G* de performance moyenne s'est révélé positif. Afin d'expliquer cette combinaison, nous avons optimisé ces paramètres pour deux groupes de situations distinctes : les unes correspondant à des cumuls pluviométriques nuls sur l'ensemble des bassins, et les autres à des cumuls non nuls sur l'ensemble des bassins. Ainsi, pour les journées pluvieuses, la température devrait être mesurée à 925 hPa à +36h, ce qui pourrait permettre de caractériser un éventuel changement de masse d'air au-dessus du bassin, conséquence du passage d'une perturbation frontale. Pour les journées sèches, en revanche, nous n'avons pas retrouvé ce caractère. La température devrait alors être mesurée à 700 hPa à +12h.

Une fois les variables d'analogie fixées, l'optimisation du nombre *NO* de situations candidates à retenir a constitué la seconde étape. Le nombre optimal s'est ainsi révélé très proche de celui que l'on obtient lorsqu'une fenêtre temporelle fixe de largeur quatre mois est utilisée. Néanmoins un palier est apparu pour ce nombre, si bien que cela nous a amené à vérifier si finalement, le nombre *NO* ne pouvait pas dépendre de la température elle-même. Les résultats n'ont cependant pas été encourageants : le nombre *NO* sera donc identique pour toutes les situations cibles

Enfin, le domaine d'analogie a été optimisé. Sur les deux bassins tests, le domaine optimal n'est finalement constitué que d'un seul point. Ce résultat nous semble logique, puisque nous ne cherchons pas à sélectionner des situations analogues à la situation cible en termes de

champ de température afin d'en déduire les précipitations, mais simplement à pré-sélectionner les situations candidates au premier niveau. Le critère d'analogie *RMSE* utilisé précédemment a été remplacé par d'autres critères afin de vérifier la pertinence de son utilisation. Les résultats ont bien montré qu'un critère de distance est approprié pour cette sélection préliminaire.

Nous avons terminé ce chapitre par une validation suivie d'un diagnostic des résultats. Après avoir effectué les mêmes optimisations sur le bassin de la Seine, le gain de performance obtenu sur ce bassin a également été significatif. L'algorithme nommé A1 a ensuite été validé par l'évaluation des performances sur un échantillon de situations cibles sur laquelle la méthode n'a pas été optimisée. Cela nous a permis d'obtenir un gain de performance du même ordre que celui obtenu sur la période de calibration. La décomposition du score *CRPS* en finesse et en justesse a révélé que cette nouvelle pré-sélection permet d'augmenter de manière significative la finesse de la prévision, mais aussi d'accroître légèrement la justesse. En revanche, nous avons remarqué que le gain n'était pas systématique. Au contraire, il y a presque autant de situations pour lesquelles la prévision est meilleure en termes de *CRPS* que de situations pour lesquelles la prévision est moins bonne qu'avec la sélection initiale des situations candidates dans une fenêtre temporelle fixe. L'étude de la distance entre la date d'une situation cible et celle des situations analogues finalement retenues a montré que près des deux-tiers des dates des situations analogues étaient incluses dans une fenêtre temporelle de quatre mois.

Finalement, nous définissons une nouvelle procédure de sélection A1, qui se déroule en trois étapes (Figure 6.24) : un niveau 0 de pré-sélection, sur toute la période disponible, des situations candidates selon la température ; un niveau 1 de sélection d'analogues en termes de circulation générale ; et un niveau 2 de sélection d'analogues en termes d'humidité.

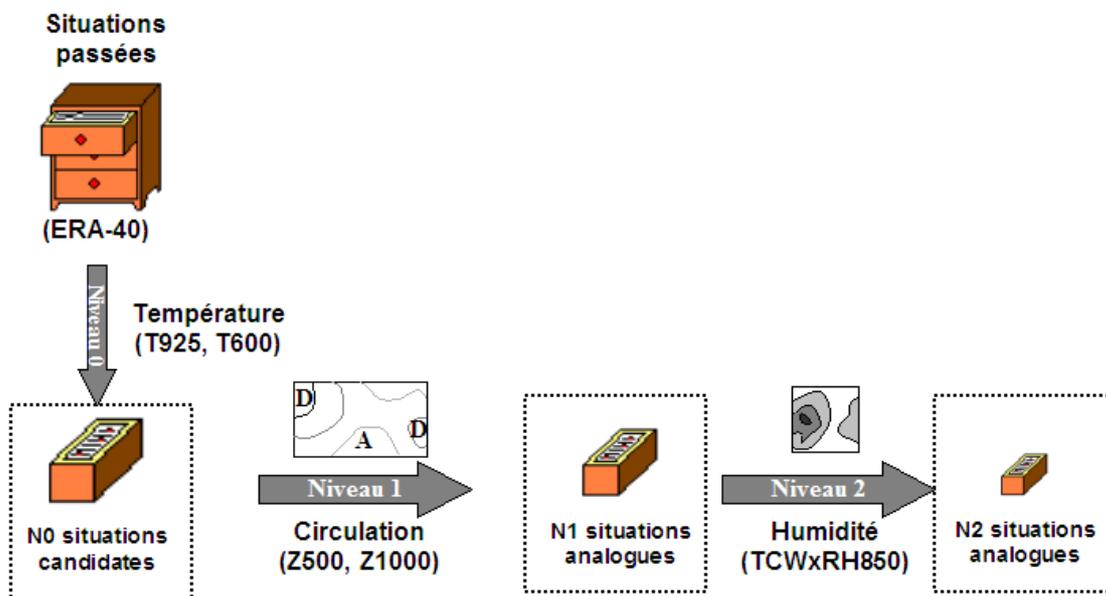


Figure 6.24 : Procédure schématisée de sélection d'analogues par la méthode A1.

7 Prise en compte des mouvements verticaux

L'algorithme de sélection d'analogues, baptisé A1, est à présent constitué de trois phases sélectives. La première (niveau 0) consiste à pré-sélectionner les situations candidates à l'analogie de circulation à échelle synoptique. Ce niveau préliminaire ne permet en aucun cas de discriminer les situations météorologiques génératrices de précipitations des situations calmes et stables. La seconde phase sélective (niveau 1) permet en revanche de ne retenir que les situations présentant les mêmes centres d'action dans des secteurs voisins. Enfin, la troisième phase (niveau 2) consiste à ne retenir que les situations dont l'état hygrométrique atmosphérique est proche de celui de la situation cible considérée.

Cependant, l'analogie sur les champs de géopotentiels à 1000 et à 500 hPa peut être confuse, si bien qu'en disposant de situations dont le champ en altitude est similaire, nous ne sommes pas certains que ces mêmes situations présentent une analogie sur les champs de surface, pourtant importants pour la localisation des perturbations frontales. Nous proposons d'introduire une nouvelle variable qui permet théoriquement de localiser les fronts : la vitesse verticale. L'étude des liaisons entre les variables, menée au chapitre 5, a montré que la vitesse verticale W est peu liée avec les variables déjà utilisées dans notre méthode A1. Celle-ci avait déjà été testée par Bontron (2004) dans le cadre de sa thèse, sans succès probant. Mais il ne disposait de la vitesse verticale qu'au niveau de pression 500 hPa. Dans l'archive de réanalyses ERA que nous nous sommes constituée, cette variable est à présent disponible à d'autres niveaux de pression.

Ce chapitre présente les résultats d'une série de tests visant à introduire la vitesse verticale dans l'algorithme de sélection A1. Le score de mesure des performances reste le gain G défini précédemment. A présent, il exprime le gain apporté par une méthode A par rapport à la méthode A1 ($G(A/A1)$).

Pour rappel, la sélection d'analogues maintenant appliquée se déroule en trois étapes, dont les paramètres, valables pour le bassin de la Saône, sont les suivants :

- niveau 0 : mesure de la température à 925 hPa à +36h et à 600 hPa à +12h, domaine d'analogie (5°E ; 47,5°N), critère d'analogie $RMSE$, 2900 situations retenues ;

- niveau 1 : mesure du géopotential à 1000 hPa à +12h et à 500 hPa à +24h, domaine d'analogie (-5°E/15°E ; 40°N/50°N), critère d'analogie *TWS*, 70 situations retenues ;
- niveau 2 : mesure du produit de l'eau précipitable par l'humidité relative à 850 hPa à +12h et +24h, domaine d'analogie constitué de quatre points de grille, critère d'analogie *RMSE*, 25 situations retenues.

7.1 Un niveau d'analogie supplémentaire

7.1.1 Objectif et choix de la variable

7.1.1.1 Comment détecter les fronts ?

Les fronts, générateurs de l'essentiel des précipitations sur nos secteurs, ne sont pas nécessairement détectés de façon objective dans l'algorithme A1 au travers des deux champs de géopotential. Par définition, un front représente la frontière entre deux masses d'air dont les états thermiques et hygrométriques sont différents. Ainsi, il existe des écarts parfois importants de température et d'humidité entre les zones situées à l'avant et à l'arrière de l'axe frontal. Ces deux variables permettent donc de caractériser un front. Une variable élaborée qui combine ces deux dernières est la température potentielle équivalente *TPE*.

Par ailleurs, l'air est un gaz dont la densité varie en fonction de la température et de la pression atmosphérique. La différence de densité de l'air entre deux zones de masse d'air différente provoque alors des mouvements ascendants. Il existe plusieurs variétés de front, mais les principaux sont le front froid et le front chaud. Dans le cas d'un front froid, pour lequel une masse d'air froid vient à l'encontre d'une masse d'air chaud, l'air froid, plus dense et donc plus lourd, soulève l'air chaud par l'avant. Dans le cas d'un front chaud, l'air chaud progresse au-dessus de l'air froid dans le sens du flux en le remplaçant progressivement par l'arrière. Dans les deux cas, des mouvements ascendants de l'air se produisent mécaniquement. La vitesse verticale de l'air *W* est par conséquent une variable également intéressante pour caractériser un front. Une variable jumelée à la vitesse verticale est la divergence *DIV* du vent, comme nous l'avons montré par l'étude de leur liaison dans le chapitre 5. En effet, le soulèvement vertical est provoqué par la convergence de l'air. Ainsi, un front peut également être caractérisé par une zone de divergence négative, notamment en surface et dans les basses couches de la troposphère.

Différentes variables météorologiques, en principe générées par les MNP et mises à disposition dans les réanalyses, permettent donc de caractériser un front :

- la température potentielle équivalente *TPE* ;
- la divergence horizontale *DIV* ;
- la vitesse verticale *W*.

Néanmoins, nous avons remarqué au chapitre 5 par l'analyse des liaisons entre les variables, que la température *TPE* mesurée dans les basses couches est assez fortement liée avec le géopotential *Z500*, mais aussi avec la température. Comme nous l'avons mentionné (cf. section 5.3), nous avons testé la température *TPE* avant mais aussi après le niveau 2 de sélection sur l'humidité, et ce avec divers critères d'analogie. Malheureusement, cette variable d'analogie n'a pas permis d'améliorer la prévision des précipitations. En revanche, nous avons montré que les variables *DIV* et *W* sont très peu liées avec les variables déjà utilisées dans l'algorithme A1. Mais la corrélation entre ces deux variables est très forte lorsqu'elles sont prises à des niveaux de pression assez proches. Notre objectif étant d'introduire l'une ou l'autre de ces variables dans notre algorithme de sélection, il est donc nécessaire de faire un choix.

7.1.1.2 Choix entre la divergence horizontale et la vitesse verticale

La première idée venant à l'esprit pour faire ce choix est d'effectuer un test simple et arbitraire : appliquer un niveau intermédiaire entre le niveau 1 et le niveau 2 avec chacune de ces variables. Par simplicité, certains paramètres ont été fixés *a priori* : le critère d'analogie, le nombre d'analogues à sélectionner et les heures de mesure de la variable d'analogie. Un critère lié à la co-fluctuation comme le critère *TWS* ou le coefficient de corrélation *R* ne nous paraît pas approprié, puisqu'il ne s'agit pas de faire une analogie sur la forme des champs ou sur leur variation spatio-temporelle, mais plutôt sur la valeur de la variable elle-même. En effet, si nous prenons une situation cible pour laquelle la valeur de la divergence ou de la vitesse verticale est négative à un instant donné en un point de notre zone d'intérêt, *i.e.* à l'aplomb du bassin, nous sélectionnerons les situations passées présentant une valeur négative du même ordre de grandeur à ce même instant et au même point. Par conséquent, le critère d'analogie choisi est le critère de distance *RMSE*. Les nombres d'analogues retenues aux niveaux 1 et 2 restent inchangés ($N1=70$, $N2=25$), et nous avons choisi de retenir 50 analogues pour ce niveau intermédiaire. Enfin, les heures de mesure ont été fixées à +12h et +24h, ce qui nous paraît être un bon compromis entre le temps de calcul et ce que l'on cherche. Ces instants sont d'une part inclus dans la période du cumul pluviométrique, et d'autre part suffisamment séparés pour tenir compte du mouvement vertical sur notre zone d'intérêt en début et en fin de période. La valeur finale du critère *RMSE* correspond à la somme des valeurs du critère *RMSE* obtenues lorsque la variable d'analogie est mesurée à +12h d'une part et à +24h d'autre part.

Nous avons ensuite choisi d'optimiser le niveau de pression auquel est mesurée la variable ainsi que le domaine d'analogie. Nous rappelons également que nous avons retenu quatre niveaux de pression situés dans les basses couches de la troposphère pour la variable *DIV* (1000, 925, 850 et 700 hPa) et six niveaux pour la variable *W* (925, 850, 700, 775, 600 et 500 hPa). Pour chacune des deux variables, la pertinence d'analogie a été évaluée en chaque point d'un domaine assez vaste en les mesurant aux différents niveaux de pression. Le point de grille le plus pertinent a été considéré comme point de départ pour l'optimisation du domaine.

Les gains G de performance, obtenus sur le bassin de la Saône par rapport à la méthode A1 lorsque le domaine optimal associé aux différents niveaux de pression est appliqué, sont reportés dans le Tableau 7.1 pour les deux variables DIV et W . Les performances de prévision sont globalement plus élevées si nous insérons un niveau intermédiaire de sélection entre les niveaux 1 et 2. Nous obtenons le gain le plus élevé lorsque la vitesse verticale est mesurée au niveau de pression 850 hPa. Nous retrouvons également le décalage en altitude entre le niveau optimal associé à la variable DIV , plus élevé, et le niveau optimal associé à la variable W . Par ailleurs, bien qu'ils ne soient pas indiqués ici puisque l'objectif n'est pas de déterminer et d'analyser le domaine optimal, les domaines d'analogie optimaux sont généralement de faible étendue, constitués de deux à six points de grille.

Ces résultats indiquent que la vitesse verticale semble à privilégier par rapport à la divergence horizontale. Notre choix est donc fait : nous conservons à présent la variable W pour la suite de nos travaux.

Tableau 7.1 : Gains G (%) de performance, obtenus sur le bassin de la Saône par rapport à la méthode A1, lorsque le domaine optimal associé aux différents niveaux de pression est appliqué, pour les deux variables DIV et W .

Variable	Niveau de pression (hPa)						
	1000	925	850	775	700	600	500
DIV	-0,01	1,75	1,24	-	-0,52	-	-
W	-	1,26	2,24	2,09	1,83	1,55	1,12

7.1.2 Un niveau d'analogie intermédiaire

La vitesse verticale est à présent considérée seule, l'objectif étant d'introduire cette variable dans un niveau supplémentaire de sélection. L'utilisation de cette variable avant le niveau 1 ne nous semble pas vraiment pertinente, puisque c'est la circulation synoptique qui agit d'abord sur les champs de vitesse verticale. Deux autres possibilités s'offrent à nous pour introduire un niveau supplémentaire avec cette variable : i) entre les niveaux 1 et 2, ii) après le niveau 2. Nous avons donc testé ces deux possibilités en procédant de la même manière que précédemment. Trois niveaux sont donc considérés par la suite : du niveau 1 au niveau 3.

Le critère d'analogie ($RMSE$) et les heures de mesure (+12h et +24h) restent les mêmes, et les nombres d'analogues à sélectionner sont identiques : 70 analogues sont sélectionnées au niveau 1, puis 50 analogues sont retenues à l'issue du niveau intermédiaire et enfin 25 analogues sont retenues au dernier niveau de sélection. Dans un cas, le niveau 2 permet de faire l'analogie sur la vitesse verticale et niveau 3 sur l'humidité ; dans l'autre cas le niveau 2 correspond à l'analogie sur l'humidité et le niveau 3 sur la vitesse verticale. Tous les niveaux de pression disponibles pour la variable W ont été appliqués, et le domaine d'analogie a été optimisé pour chacun d'entre eux (Tableau 7.2), sachant que le point de départ correspond toujours au point de grille le plus pertinent. C'est finalement entre les deux niveaux 1 et 2 initiaux que l'analogie sur la vitesse verticale s'avère la plus intéressante. Ainsi, la vitesse

verticale ne semble pas mieux discriminer les précipitations lorsqu'elle est utilisée après le niveau d'analogie sur l'humidité, sauf si elle est mesurée au niveau de pression 850 hPa, auquel cas le gain est légèrement positif. Néanmoins ce gain reste négligeable devant celui obtenu lorsque la variable W est utilisée dans un niveau de sélection intermédiaire.

Ce résultat montre finalement que les mouvements verticaux sont prépondérants devant l'hygrométrie dans la genèse des précipitations. Il semble donc préférable que l'analogie sur le caractère stable/instable de la basse troposphère soit recherchée juste après celle de l'analogie sur la circulation synoptique. C'est alors l'analogie sur l'humidité qui déterminera la quantité de précipitations susceptible d'être observée au cours de la situation cible, en fonction notamment de la quantité d'eau précipitable présente dans la colonne troposphérique.

Tableau 7.2 : Gains G (%) de performance, obtenus sur le bassin de la Saône par rapport à la méthode A1, lorsque l'analogie sur la variable W est effectuée au niveau 2 ou au niveau 3, en fonction du niveau de pression. Le domaine optimal associé à chaque niveau de pression a été appliqué.

Numéro du niveau d'analogie sur W	Niveau de pression (hPa)					
	925	850	775	700	600	500
2	1,26	2,24	2,09	1,83	1,55	1,12
3	-1,87	0,07	-0,27	-0,71	-1,5	-2,12

En conclusion, nous avons choisi une variable susceptible de mettre en évidence le passage d'un front : la vitesse verticale W . Nous allons à présent introduire cette variable dans un niveau intermédiaire entre les niveaux 1 et 2 initiaux en procédant à une optimisation plus complète des paramètres.

7.2 Optimisation des paramètres

Nous supposons que notre sélection d'analogues se déroule maintenant en quatre étapes. La numérotation des niveaux d'analogie existants diffère à présent de la numérotation initiale. Les niveaux 0 et 1 sont inchangés, tandis qu'à présent le niveau intermédiaire que nous introduisons ici correspond au niveau 2 ($N2$ situations analogues retenues), et le niveau de sélection par analogie sur l'humidité ($TCW\%$) correspond au niveau 3 ($N3$ situations retenues).

Lorsque nous désirons introduire un niveau d'analogie supplémentaire, plusieurs paramètres doivent être choisis ou optimisés :

- (a) critère d'analogie ;
- (b) niveau de pression auquel la variable doit être mesurée ;
- (c) heure(s) de mesure de la variable
- (d) domaine d'analogie ;
- (e) nombre d'analogues à retenir.

Il faut à nouveau faire des choix afin de déterminer ces paramètres étape par étape. Dans le cas de la vitesse verticale, nous avons vu que le choix du critère d'analogie (a) adapté était assez instinctif. Ainsi pour effectuer la précédente série de tests, nous avons choisi arbitrairement le critère de distance *RMSE*. Dans un premier temps, nous avons décidé de vérifier que celui-ci est bien adapté à la vitesse verticale en appliquant différents critères d'analogie. Nous en profiterons pour optimiser dans le même temps le niveau de pression auquel la variable *W* doit être mesurée et le domaine spatial sur lequel l'analogie doit être effectuée, et ce en considérant chacun des critères sélectionnés. Nous conserverons ici les heures de mesure (+12h et +24h) et le nombre *N2* (50) fixés précédemment.

Une fois le critère (a) et le niveau de pression (b) fixés, nous optimiserons les nombres d'analogues (e) puis les heures de mesure (c) tout en optimisant le domaine d'analogie (d), celui-ci dépendant théoriquement des instants auxquels la variable est mesurée.

7.2.1 Choix du critère d'analogie

La première phase de l'optimisation que nous avons adoptée consiste à appliquer les critères *RMSE*, *MAE*, *R* et *TWS*, et à évaluer la pertinence d'analogie des points de grille puis à optimiser le domaine d'analogie associé à chacun des critères, et ce en mesurant la vitesse verticale sur chacun des six niveaux de pression disponibles. Cela nous permettra de vérifier si un critère de distance tel que le critère *RMSE* est bien adapté à la vitesse verticale.

Pour évaluer la pertinence des points de grille, nous rappelons que le domaine élémentaire sur lequel est évalué le gain *G* de performance est constitué d'un point de grille pour les critères *RMSE* et *MAE*, et de quatre points de grille pour les critères *R* et *TWS*. La pertinence d'analogie des points de grille a été évaluée sur un vaste domaine (-15°E/20°E ; 35°N/60°N). Le Tableau 7.3 indique le gain *G* maximal obtenu sur le domaine, en appliquant chacun des critères et en mesurant la vitesse verticale à chacun des six niveaux de pression.

Tableau 7.3 : Gain *G* (%) de performance moyenne, obtenu sur le bassin de la Saône par rapport à la méthode A1, en mesurant *W* à chacun des niveaux de pression et en appliquant chacun des critères d'analogie, sur le domaine d'analogie élémentaire le plus pertinent (*G* maximal).

Critère	Niveau de pression (hPa)					
	925	850	775	700	600	500
<i>RMSE</i>	-0,05	1,24	1,37	1,24	1,12	0,81
<i>MAE</i>	0,32	1,13	1,24	1,12	0,94	0,94
<i>R</i>	-1,61	-0,88	-0,61	-0,64	-1,16	-1,37
<i>TWS</i>	-1,64	-0,83	-0,62	-0,98	-1,20	-1,22

Le meilleur gain de performance, qui s'élève à +1,37% par rapport à la méthode A1, est obtenu avec le critère *RMSE* en mesurant la vitesse verticale à 775 hPa. Quel que soit le critère appliqué, c'est d'ailleurs le niveau 775 hPa qui s'avère le plus pertinent. En revanche, les performances sont dégradées lorsque les critères *R* et *TWS* sont appliqués sur la vitesse

verticale, et ce quel que soit le niveau de pression. Mais ces derniers sont des critères de forme, pour lesquels un domaine plus étendu pourrait être plus adapté qu'un domaine élémentaire constitué seulement de quatre points de grille. Nous optimiserons le domaine un peu plus tard et vérifierons si le résultat est différent. En outre, la sélection d'analogues par le critère *RMSE*, lorsque la vitesse verticale est mesurée à 925 hPa, mène aussi à une légère perte de performance par rapport à A1.

La cartographie de la pertinence d'analogie, lorsque la vitesse verticale est mesurée à 775 hPa, est présentée sur la Figure 7.1 pour les quatre critères testés. Ces cartes nous permettent notamment d'estimer l'échelle d'espace à laquelle la vitesse verticale est pertinente en fonction du critère d'analogie utilisé.

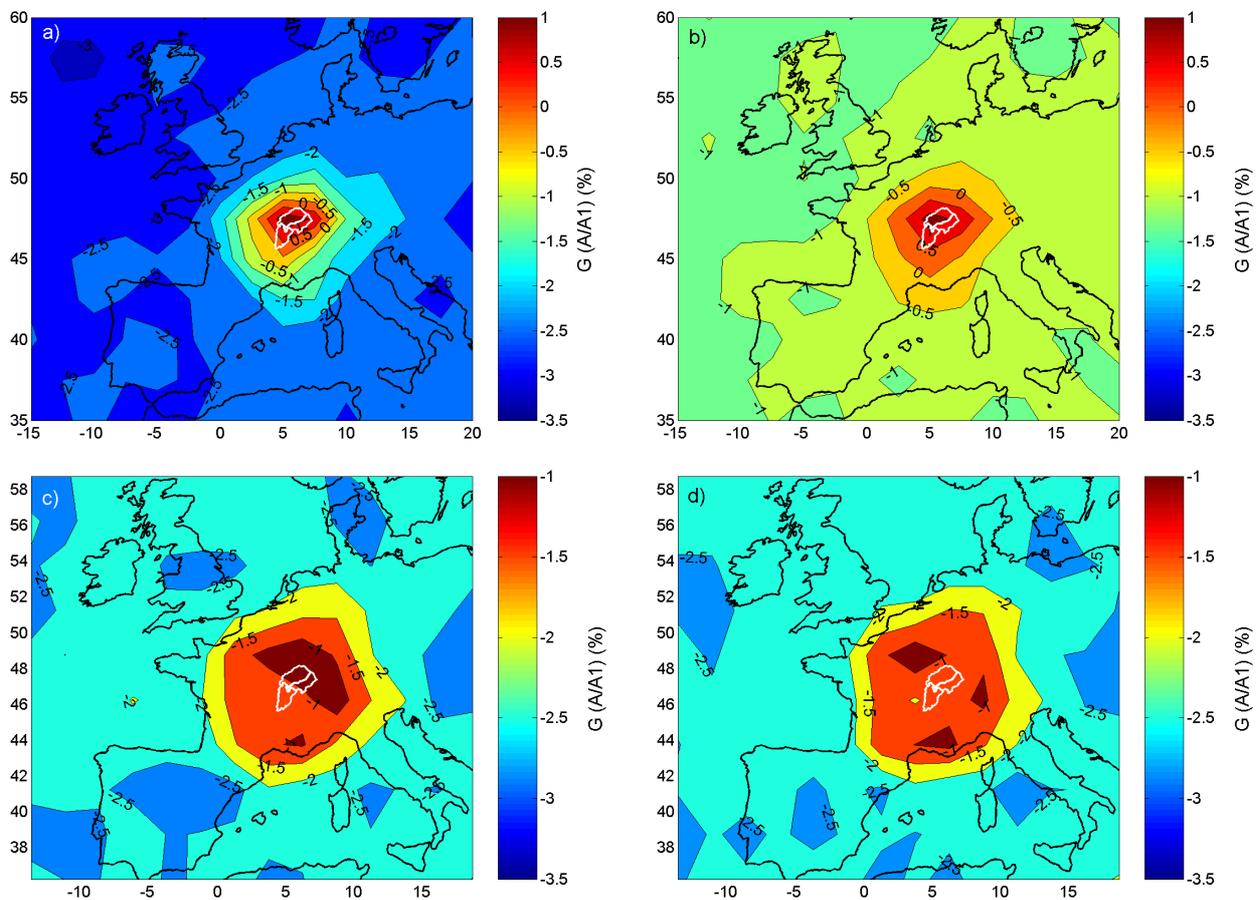


Figure 7.1 : Cartographie de la pertinence d'analogie exprimée en gain G (%) de performance, obtenu sur le bassin de la Saône par rapport à la méthode A1. Le critère d'analogie appliqué est : a) *RMSE*, b) *MAE*, c) *R* et d) *TWS*. La vitesse verticale est ici mesurée au niveau 775 hPa.

Nous remarquons que la zone la plus pertinente apparaît beaucoup plus locale lorsqu'un critère de distance est appliqué (*RMSE*, *MAE*). Aucune direction privilégiée n'apparaît, et ce quel que soit le critère appliqué. Si nous comparons la pertinence des points de grille obtenue par un critère de distance (Figure 7.1 a) et b)) à celle obtenue par un critère lié à la co-fluctuation (Figure 7.1 c) et d)), nous remarquons que les performances sont plus rapidement dégradées si on considère des points un peu plus éloignés du bassin. Notons toutefois que les

cartes c) et d) ne sont pas vraiment comparables avec les cartes a) et b) puisque les points de grille constituant chaque domaine élémentaire ne sont pas au même nombre. On ne peut donc pas tirer de conclusion hâtive à partir de ces éléments.

Par conséquent, le domaine d'analogie a été optimisé pour tous les cas considérés (critère d'analogie, niveau de pression) afin d'évaluer l'intérêt d'appliquer un critère lié à la co-fluctuation sur un plus grand domaine. Le domaine de départ choisi correspond au domaine élémentaire le plus pertinent, *i.e.* pour lequel le gain de performance G est le plus élevé. Les paramètres associés aux domaines optimaux sont reportés dans le Tableau 7.4.

Tableau 7.4 : Gain G (%) de performance, obtenu sur le bassin de la Saône par rapport à la méthode A1, avec chacun des critères lorsque le domaine d'analogie appliqué est optimal.

Critère d'analogie	Niveau optimal	Domaine optimal	$G(A/A1)$
<i>RMSE</i>	850	2,5°E/7,5°E ; 45°N/47,5°N	2,24
<i>MAE</i>	850	2,5°E/7,5°E ; 45°N/47,5°N	2,27
<i>R</i>	775	2,5°E/7,5°E ; 45°N/57,5°N	0,26
<i>TWS</i>	775	2,5°E/7,5°E ; 42,5°N/55°N	0,37

Après optimisation du domaine, le gain G se révèle positif lorsque les critères R et TWS sont utilisés. Néanmoins, les valeurs restent très faibles. En revanche le gain devient significatif lorsqu'un critère de distance est utilisé. Le niveau de pression optimal est finalement le niveau 850 hPa, et le domaine d'analogie optimal est de faible dimension, à peine plus grand que le domaine appliqué au niveau 3 (extension zonale d'un pas de 2,5°). Nous noterons aussi que les domaines optimaux obtenus avec les critères R et TWS sont nettement plus étendus dans la direction méridienne.

La différence des gains G obtenus par les deux critères de distance est très faible, ce qui ne nous permet pas, à ce stade, de privilégier l'un d'eux. Nous les conservons pour l'étape suivante de l'optimisation des paramètres. Nous allons à présent nous intéresser au nombre de situations analogues à retenir. Il n'est pas impossible que la combinaison optimale des nombres $N1-N2-N3$ soit bien différente de celle appliquée jusqu'à présent.

7.2.2 Nombres d'analogues à sélectionner

L'objectif de cette section est de ré-optimiser les nombres de situations analogues à retenir tout en appliquant chacun des deux critères de distance $RMSE$ et MAE . Le critère d'analogie associé au gain G maximal sera finalement retenu pour ce niveau intermédiaire.

Trois niveaux consécutifs de sélection d'analogues sont appliqués. Nous avons conservé les nombres $N1$ et $N3$ pour choisir le critère d'analogie approprié. Les trois nombres ($N1$, $N2$, $N3$) doivent donc être à présent optimisés. Afin d'alléger le temps de calcul, nous avons choisi d'optimiser simultanément $N1$ et $N2$ et de conserver un nombre $N3$ égal à 25, et ceci pour trois raisons :

- en théorie, le nombre $N3$ d'analogues sélectionnées au dernier niveau dépend

- essentiellement de la longueur d'archive de situations candidates (qui n'a pas changé) ;
- le nombre $N3$ est notamment dépend de la variable d'analogie utilisée au niveau 3, et il se trouve que dans le cas présent la variable d'analogie ($TCW\%$) est identique à celle utilisée au cours du dernier niveau de sélection de la méthode A0 ;
 - le nombre retenu au dernier niveau doit être identique à celui fixé pour le calcul des performances de notre méthode de référence A0 ($N2=25$) pour que la comparaison au travers du gain de performance ait un sens.

Seuls les nombres $N1$ et $N2$ ont donc été optimisés. Le domaine d'analogie appliqué ici est le domaine optimisé précédemment ($2,5^\circ E/7,5^\circ E$; $45^\circ N/47,5^\circ N$). Nous avons fait varier le nombre $N1$ entre 200 et 40 par pas de 10, et le nombre $N2$ entre $N1-10$ et 30, par pas de 10. Il s'avère que le gain G maximal, de notre algorithme A par rapport à l'algorithme A1, obtenu par le critère $RMSE$, est plus élevé que celui obtenu par le critère MAE . Par conséquent, nous retiendrons le critère d'analogie $RMSE$ pour ce niveau de sélection sur la vitesse verticale.

La Figure 7.2 présente l'évolution des gains de performance en fonction du nombre $N2$ pour les différentes valeurs de $N1$, lorsque le critère $RMSE$ est appliqué.

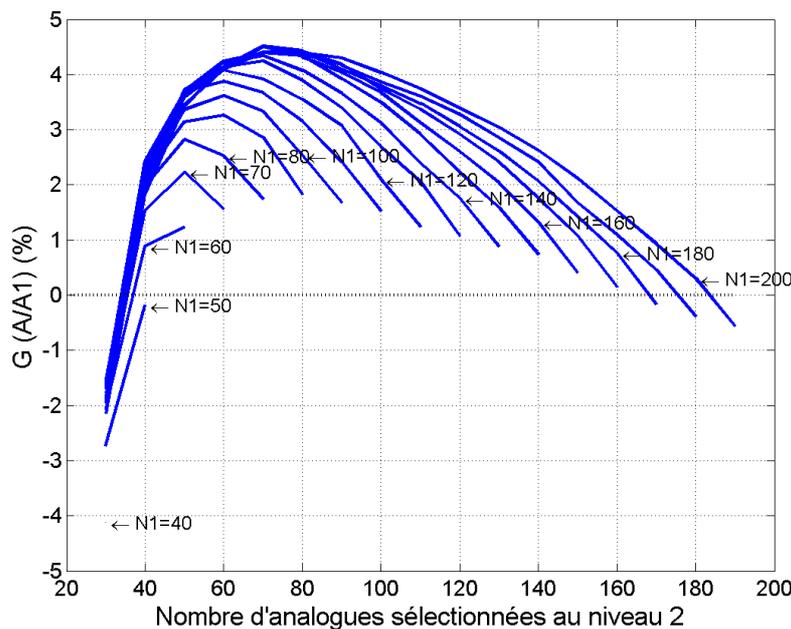


Figure 7.2 : Evolution du gain G (%) de performance, évalué sur le bassin de la Saône par rapport à la méthode A1, en fonction du nombre $N2$ pour différentes valeurs du nombre $N1$ ($N1=50$ à $N1=200$).

Le couple optimal ($N1=170$, $N2=70$) permet d'obtenir un gain G de 4,52 % par rapport à l'algorithme A1. Ce résultat démontre qu'il est important d'optimiser les nombres d'analogues à sélectionner : le gain n'était que de 2,24 % avant la ré-optimisation. En introduisant un niveau intermédiaire, nous devons donc retenir plus de situations analogues en termes de circulation atmosphérique, afin que des situations, peut-être moins analogues au niveau 1 que les 70 situations retenues initialement mais très voisines au niveau 2, soient finalement

retenues. Nous retrouvons d'ailleurs un nombre $N2$ identique au nombre $N1$ initialement appliqué dans la méthode A0 (70).

D'autre part, la Figure 7.2 révèle que si 30 analogues sont retenues à l'issue du niveau 2 (puis 25 au niveau 3), les performances sont dégradées quel que soit le nombre de situations retenues au niveau 1. Ce résultat montre que la vitesse verticale n'est pas suffisamment discriminante vis-à-vis des précipitations pour s'affranchir d'une sélection selon l'humidité. Si $N2$ est proche de $N3$, les situations analogues retenues au niveau 2 sont certes très ressemblantes au sens du mouvement vertical aux abords du bassin d'intérêt, mais ne présentent pas nécessairement de similitude en termes d'humidité. Il faut donc conserver un nombre suffisant de situations qui contiennent une information hygrométrique proche de celle de la situation cible pour laquelle la prévision est émise.

7.2.3 Heures de mesure et domaine d'analogie

7.2.3.1 Optimisation sur l'ensemble des situations cibles

Nous avons déterminé trois des cinq paramètres nécessaires pour l'introduction d'un nouveau niveau de sélection d'analogues (critère d'analogie, niveau de pression, nombre d'analogues à retenir). Seules les heures de mesure de la vitesse verticale, fixées jusqu'ici à +12h et +24h, n'ont pas encore été ré-optimisées, ainsi que le domaine d'analogie.

Afin d'avoir une idée de l'instant optimal auquel la vitesse verticale doit être mesurée par rapport à la période du cumul pluviométrique, nous avons évalué le gain G en mesurant cette variable à un seul instant, en faisant varier l'heure de mesure entre -12h et +42h par pas de 6h, ce qui représente au total dix heures de mesure. En sus, nous avons également considéré chacun des niveaux de pression disponibles (Figure 7.3). Or les résultats de l'optimisation des heures de mesure dépendent théoriquement du domaine d'analogie appliqué. En effet, il faut tenir compte du phénomène d'advection. Une zone de forte ascendance (ou de subsidence) se déplace dans le temps, si bien que la zone d'ascendance observée en un lieu de l'espace à l'instant t se retrouvera en un autre lieu à l'instant $t+1$. L'extension et la localisation du centre de gravité du domaine d'analogie dépendent ainsi des instants auxquels on mesure une variable d'analogie. Par conséquent, nous avons optimisé le domaine pour chaque heure de mesure et chaque niveau de pression.

A l'image de ce que Bontron (2004) avait obtenu avec la variable d'humidité $TCW\%$, la vitesse verticale doit être nécessairement mesurée à l'intérieur de la période du cumul pluviométrique. En effet, celle-ci traduit les mouvements verticaux responsables de la formation des précipitations. Il est donc logique de mesurer cette variable durant la période du cumul. L'heure optimale est +18h quel que soit le niveau de pression considéré. Par ailleurs, cette figure révèle l'intérêt de mesurer cette variable dans les couches basses et moyennes de la troposphère (850 à 700 hPa).

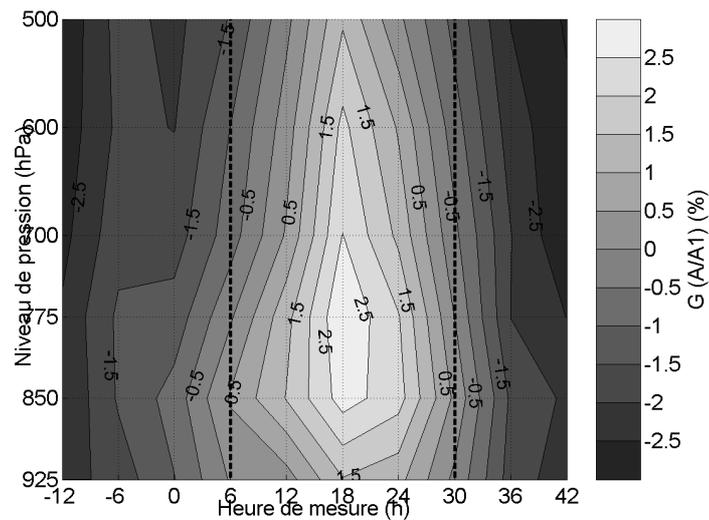


Figure 7.3 : Gain G (%) de performance, évalué sur le bassin de la Saône par rapport à la méthode A1, en fonction du niveau de pression (ordonnée) et de l'heure (abscisse) auxquels la vitesse verticale peut être mesurée.

Nous avons ensuite considéré successivement deux, trois, quatre puis cinq heures de mesure parmi les dix heures, en ayant fixé préalablement le niveau de pression à 850 hPa (niveau optimal avec deux heures de mesure, cf. section 7.1.2). De même que pour deux heures de mesure, la valeur finale du critère $RMSE$ est calculée en effectuant la somme des valeurs du critère $RMSE$ obtenues lorsque la variable W est mesurée à chaque instant de mesure. Les combinaisons et les gains associés sont indiqués sur la Figure 7.4.

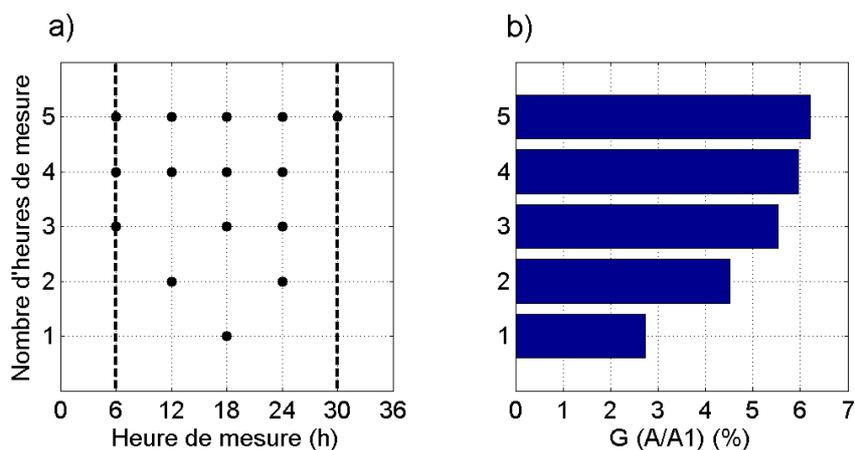


Figure 7.4 : a) Combinaison des instants de mesure optimaux de la vitesse verticale et b) gain de performance G (%) associé en fonction du nombre d'heures considéré, valables pour le bassin de la Saône. La vitesse verticale est ici mesurée au niveau de pression 850 hPa.

Il apparaît nécessaire de mesurer la vitesse verticale à des instants inclus dans la période du cumul journalier, ce qui est conforme à ce que nous avons obtenu avec une heure de mesure. Une symétrie apparaît dans la répartition des instants de mesure par rapport au milieu de la période de cumul. Ce résultat indique finalement que l'analogie sur cette variable ne doit pas nécessairement être faite à un moment privilégié de la période du cumul.

Si l'on mesure la vitesse verticale à trois ou quatre instants différents, l'heure de mesure +6h fait partie de la combinaison optimale, alors que l'heure +30h n'est pas privilégiée.

Bien que nous ne l'ayons pas indiqué précédemment, nous nous sommes aperçus que le domaine optimal était le même pour toutes les combinaisons optimales (2,5°E/7,5°E ; 45°N/47,5°N), à l'exception de la combinaison d'une heure de mesure (0°E/7,5°E ; 45°N/50°N).

Ce résultat traduit finalement l'importance de rechercher l'analogie sur un domaine restreint centré approximativement sur le bassin d'intérêt, indépendamment du nombre d'heures de mesure. Finalement, si nous comparons les domaines associés aux trois niveaux de sélection de l'algorithme (Figure 7.5), l'échelle à laquelle l'analogie sur la vitesse verticale doit être recherchée, représentée par la taille du domaine optimal, est comprise entre l'échelle considérée au niveau 1 et celle considérée au niveau 3. Le fait que les trois domaines s'emboîtent justifie finalement l'introduction de cette sélection intermédiaire.

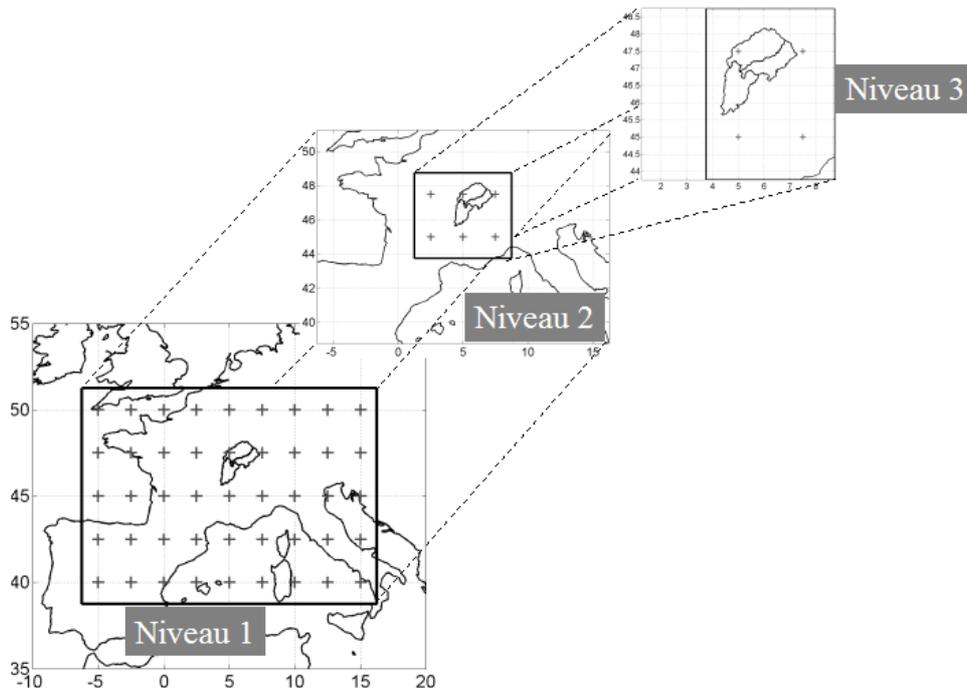


Figure 7.5 : Schéma montrant l'emboîtement des domaines d'analogie appliqués successivement aux niveaux 1, 2 et 3, optimisés pour le bassin de la Saône. Les croix représentent les centres des mailles carrées de 2,5°x2,5° qui constituent un domaine.

Si nous nous intéressons au gain de performance (Figure 7.4 b), nous remarquons qu'il augmente avec le nombre d'instantanés auxquels la vitesse verticale est mesurée, jusqu'à atteindre un gain maximal avec cinq heures de mesure. Cela démontre que la vitesse verticale est fortement variable dans le temps. Une zone d'ascendance ou de subsidence marquée peut traverser la zone d'étude à tout moment d'une journée, ce qui nécessite de rechercher l'analogie sur le champ de vitesse verticale à tous les pas de temps disponibles dans la période du cumul. Toutefois, l'augmentation du gain décroît progressivement avec celle du

nombre d'heures. Ainsi, le gain G augmente de 1,79 points si on mesure cette variable à deux instants plutôt qu'à un seul instant ; alors qu'il ne gagne que 1,01 points puis 0,42 points si on la mesure à trois instants au lieu de deux puis à quatre au lieu de trois. Néanmoins, le temps de calcul associé à cette sélection n'augmente pas sensiblement si la vitesse verticale est mesurée à plusieurs instants, puisque seules 170 situations retenues au premier niveau sont candidates à ce second niveau.

Au vu des résultats, il nous semble intéressant de considérer quatre ou cinq heures de mesure de la vitesse verticale. Avant de choisir définitivement le nombre d'instant de mesure, nous nous sommes tout de même interrogés sur l'impact du nombre d'heures de mesure sur les performances en fonction de la pluviosité des situations cibles.

7.2.3.2 Suivant la pluviosité des situations cibles

De la même manière que pour la température, nous avons considéré deux groupes distincts de situations : celui des journées non pluvieuses (groupe S) et celui des journées pluvieuses (groupe P). Pour rappel, ces groupes contiennent respectivement 26,9% et 56,5% des situations cibles de la période de calage pour le bassin de la Saône.

Pour chacun des deux groupes, nous avons appliqué les combinaisons d'heures de mesure optimales obtenues pour l'ensemble des situations, puis le gain G a été évalué (Figure 7.6). Les gains de performance obtenues pour les groupes S et P doivent être considérés séparément, puisque le score $CRPSS$ de référence (méthode A1) associé aux situations sèches (84,71%) est beaucoup plus élevé que celui associé aux situations pluvieuses (45,51%).

Nous nous intéressons ici à l'évolution du gain G en fonction du nombre d'heures auxquelles la vitesse verticale est mesurée. Le gain maximal obtenu pour le groupe P (resp. S) est obtenu pour quatre (resp. cinq) heures de mesure. Même si la différence de gain est négligeable entre quatre et cinq heures pour le groupe P, nous sommes tentés de ne considérer que quatre heures de mesure, afin de privilégier la prévision des situations pluvieuses ; la prévision des situations sèches étant déjà très satisfaisante.

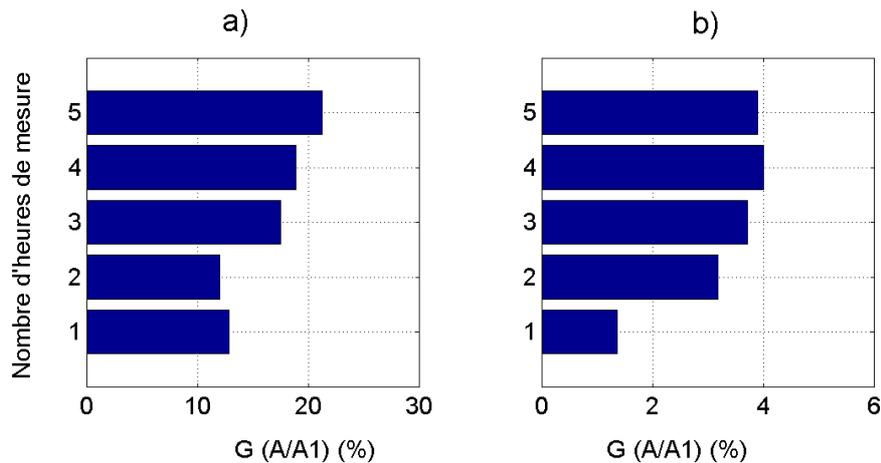


Figure 7.6 : Gains G (%) de performance moyenne, obtenus sur le bassin de la Saône par rapport à la méthode A1, en fonction du nombre d'heures de mesure : a) pour le groupe des situations sèches, b) pour celui des situations pluvieuses.

Nous avons à présent optimisé et choisi l'ensemble des paramètres associés à ce niveau d'analogie intermédiaire faisant intervenir la vitesse verticale. Les paramètres valables pour le bassin de la Saône sont les suivants :

- critère d'analogie : $RMSE$;
- niveau de pression auquel la variable W doit être mesurée : 850 hPa ;
- heure(s) de mesure : +6h, +12h, +18h et +24h ;
- domaine d'analogie : (2,5°E/7,5°E ; 45°N/47,5°N) ;
- nombre d'analogues à retenir : $N2=70$ parmi 170 analogues retenues au niveau 1.

Pour le bassin de la Saône, le gain de performance moyenne obtenu par l'introduction de cette nouvelle variable s'élève donc à 5,95 % par rapport à la méthode A1 sur la période de calibration. Ce gain est intéressant certes, mais doit à présent être validé sur le bassin de la Seine et évalué sur la période de validation.

7.3 Validation et diagnostic

Nous proposons tout d'abord de nommer ce nouvel algorithme A2, qui inclut quatre niveaux de sélection : un niveau préliminaire de sélection de situations candidates (niveau 0), puis un niveau de sélection d'analogues en terme de circulation générale (niveau 1), puis en terme de mouvement vertical (niveau 2) et enfin en terme d'humidité (niveau 3).

Nous avons adopté le même protocole que dans le chapitre précédent : nous validerons l'algorithme en effectuant les mêmes optimisations de paramètres sur le bassin de la Seine que sur le bassin de la Saône et en nous assurant que le gain de performance est également positif ; puis nous vérifierons que le gain positif est conservé sur la période de validation avant d'établir un diagnostic des résultats.

7.3.1 Adaptation au bassin de la Seine

Les résultats obtenus sur la Saône ont révélé la pertinence du critère *RMSE* comme critère d'analogie. Nous conservons ici ce critère pour l'adaptation au bassin de la Seine. En revanche, les autres paramètres (niveau de pression et heures de mesure, domaine d'analogie, nombres d'analogues à retenir) doivent être optimisés sur ce bassin.

De la même manière que sur le bassin de la Saône, nous avons d'abord fixé les heures de mesure de la vitesse verticale à +12h et +24h, et nous avons optimisé à la fois le niveau de pression, les nombres d'analogues retenues aux niveaux 1 et 2 ainsi que le domaine d'analogie du niveau 2.

La cartographie de la pertinence d'analogie, établie en fixant les nombres d'analogues à 70 et à 50 aux niveaux 1 et 2, nous a permis de sélectionner le point le plus pertinent associé à chacun des six niveaux de pression. Le Tableau 7.5 indique les coordonnées de ces points ainsi que le gain de performance associé, exprimé par rapport à la méthode A1. Ainsi en retenant 70 analogues au niveau 1 et 50 au niveau 2, le niveau de pression optimal est, comme pour le bassin de la Saône, le niveau 850 hPa.

Tableau 7.5 : Coordonnées du point de grille le plus pertinent pour chacun des niveaux de pression auxquels la vitesse verticale peut être mesurée, et gain G (%) de performance moyenne obtenu par rapport à la méthode A1 ; valables pour le bassin de la Seine.

Niveau de pression	Coordonnées du point	$G(A2/A1)$ (%)
925	0°E ; 47,5°N	0,31
850	2,5°E ; 47,5°N	0,85
775	2,5°E ; 47,5°N	0,52
700	2,5°E ; 47,5°N	0,36
600	2,5°E ; 47,5°N	0,06
500	2,5°E ; 47,5°N	-0,28

Les résultats de l'optimisation du niveau de pression, des nombres $N1$ et $N2$ ainsi que du domaine d'analogie montrent, comme pour la Saône, l'intérêt de mesurer la vitesse verticale à 850 hPa. La Figure 7.7 indique l'évolution du gain de performance par rapport à la méthode A1, en fonction des nombres $N1$ et $N2$. L'optimum est atteint pour un nombre $N1$ de 180 et un nombre $N2$ de 70, avec un gain G de 2,86%. Le domaine d'analogie optimal associé à ce couple est d'assez faible étendue (0°E/5°E ; 45°N/50°N). Comme pour la Saône, le domaine d'analogie optimal évolue relativement peu avec les nombres d'analogues sélectionnés.

Les nombres d'analogues retenues aux niveaux 1 et 2 ayant été fixés, nous avons optimisé les heures de mesure de la vitesse verticale pour le bassin de la Seine. Lorsque cette variable n'est mesurée qu'à un seul instant, nous retrouvons des résultats identiques à ceux obtenus sur la Saône, le niveau et l'heure de mesure optimaux étant 850 hPa et +18h (Figure 7.8).

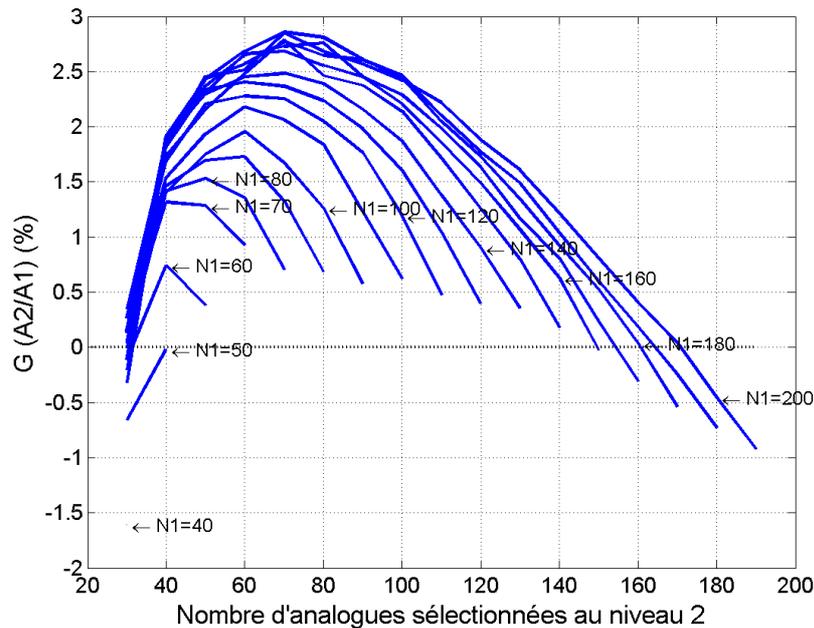


Figure 7.7 : Evolution du gain G (%) de performance, évalué par rapport à la méthode A1, en fonction du nombre N_2 pour différentes valeurs du nombre N_1 ($N_1=50$ à $N_1=200$), pour le bassin de la Seine.

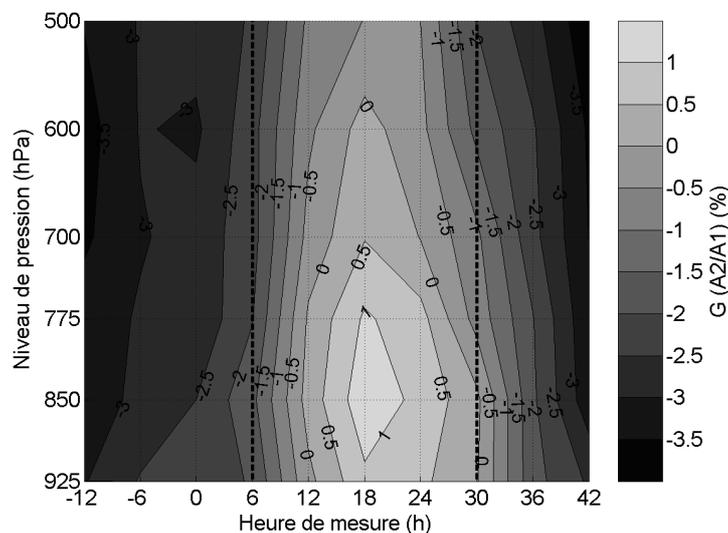


Figure 7.8 : Gain G (%) de performance, évalué par rapport à la méthode A1, en fonction du niveau de pression (ordonnée) et de l'heure (abscisse) auxquels la vitesse verticale peut être mesurée, pour le bassin de la Seine.

Toutes les combinaisons de deux, trois, quatre et cinq heures de mesure ont été testées, et pour chacune d'entre elles, le domaine d'analogie a été optimisé. La Figure 7.9 présente les combinaisons optimales et les gains associés pour chaque nombre d'heures de mesure considéré.

L'évolution du gain de performance en fonction du nombre d'instantanés auxquels la vitesse verticale est mesurée est identique à celle obtenue sur le bassin de la Saône. Le gain

augmente assez fortement entre un et deux instants, puis cette augmentation est plus faible avec celle du nombre d'heures de mesure. Le gain est d'ailleurs équivalent si l'on mesure cette variable à quatre ou cinq instants différents. Comme pour le bassin de la Saône, nous choisirons pour la suite la combinaison de quatre instants pour le bassin de la Seine. En revanche, la combinaison optimale diffère légèrement de celle obtenue sur la Saône : celle-ci inclut l'heure +30h au lieu de l'heure +6h.

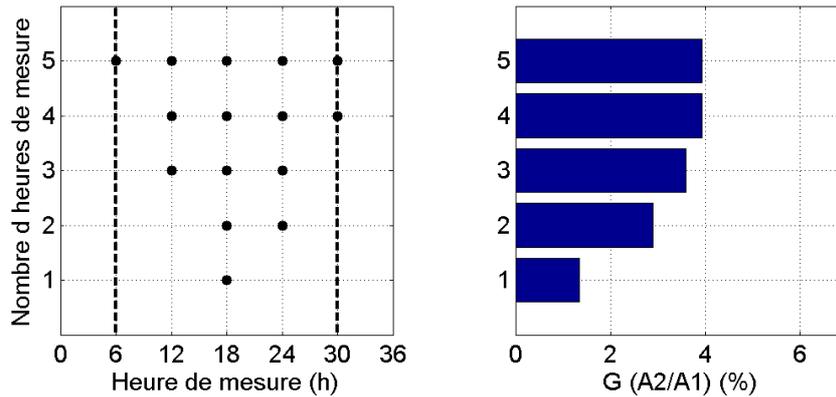


Figure 7.9 : Combinaison des instants de mesure optimaux de la vitesse verticale (gauche) et gain G (%) de performance associé (droite) en fonction du nombre d'heures considéré, valables pour le bassin de la Seine. La vitesse verticale est ici mesurée au niveau de pression 850 hPa.

Dans les deux cas, il est préférable de rechercher des situations dont le champ de vitesse verticale est similaire à plusieurs instants inclus dans la période du cumul. Le gain apporté est suffisamment intéressant pour conserver cette variable.

7.3.2 Validation sur les deux bassins tests

Nous proposons à présent de valider l'algorithme A2, pour les bassins de la Saône et de la Seine, par l'évaluation du gain G , obtenu par rapport à l'algorithme A1, sur la période de validation. Les gains obtenus sur chaque sous-bassin et sur l'ensemble du bassin sont reportés dans le Tableau 7.6 (Saône) et dans le Tableau 7.7 (Seine). A titre de comparaison, les gains G obtenus en calibration y figurent également.

Le gain de performance obtenu sur la période de calibration est bien conservé sur la période de validation, à l'exception des bassins Seine-Int et Loing-E pour lesquels le gain G est légèrement négatif. Ces valeurs entachent ainsi le gain de performance moyenne obtenu sur l'ensemble du bassin de la Seine, qui n'est plus que de 1,28 %. Cependant pour ce bassin, les valeurs de gain sont nettement plus faibles sur la période de validation que sur la période de calibration, contrairement au bassin de la Saône pour lequel les gains étaient du même ordre.

Tableau 7.6 : Gains G (%) de performance moyenne, obtenue par la méthode A2 par rapport à la méthode A1 sur le bassin de la Saône, évalués sur les deux périodes cibles indépendantes de validation et de calibration.

	Validation	Calibration
Doubs	7,07	6,29
Petite Saône	5,2	5,3
Saône Intermédiaire	5,45	6,22
Ensemble du bassin	5,91	5,94

Tableau 7.7 : Gains G (%) de performance moyenne, obtenue par la méthode A2 par rapport à la méthode A1 sur le bassin de la Seine, évalués sur les deux périodes cibles indépendantes de validation et de calibration.

	Validation	Calibration
Yonne-C	2,06	4,24
Seine-Int	-0,79	3,9
Seine-B	2,79	4,49
Marne-F	3,07	4,24
Loing-E	-0,43	2,78
Ensemble du bassin	1,28	3,91

Ce résultat provient d'une perte de performance dans la justesse des prévisions. En effet, sur la période de validation et pour l'ensemble du bassin de la Seine, le gain en justesse s'élève à -4,21 % par rapport à la méthode A1, alors que le gain en finesse atteint 7,81 %. Pour les deux bassins particuliers évoqués précédemment, les gains en justesse atteignent environ -6 % à -8 %. Une raison à ce résultat pourrait provenir de la climatologie du bassin de la Seine sur la période de validation. En effet, nous avons préalablement choisi nos sous-périodes de calibration/validation en vue du développement de la méthode d'analogie sur le bassin de la Saône. Nous nous étions assurés que les climatologies des deux périodes étaient similaires (cf. introduction à la partie II). En revanche, les climatologies des deux périodes sont différentes pour certains sous-bassins de la Seine. La Figure 7.10 montre les valeurs des quantiles de pluie non nulle obtenues sur les deux sous-périodes ainsi que sur la période totale, pour les bassins Marne-F et Loing-E.

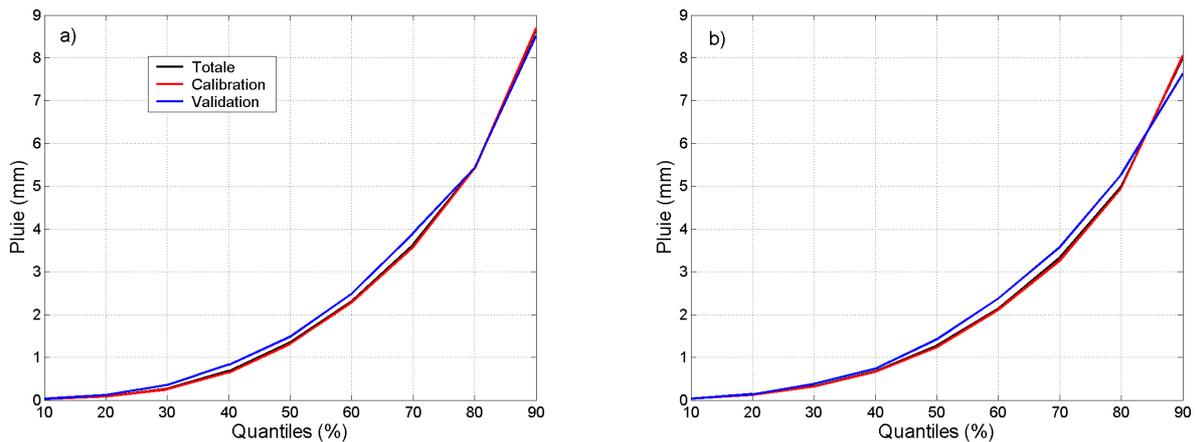


Figure 7.10 : Comparaison des quantiles de pluie non nulle d'ordre 10 à 90 % obtenus sur la période totale et sur les deux sous-périodes. a) pour le bassin Marne-F ; b) pour le bassin Loing-E.

Pour les cumuls de pluie inférieurs au quantile de pluie non nulle d'ordre 70 %, les différences entre les deux sous-périodes sont similaires pour les deux bassins. En revanche, des écarts non négligeables apparaissent pour les précipitations plus importantes. Pour le bassin Marne-F, les quantiles élevés sont du même ordre de grandeur sur les trois périodes, tandis que pour le bassin Loing-E, les quantiles d'ordre 70 % et 80 % sont plus élevés sur la période de validation et le quantile d'ordre 90 % est moins élevé. La différence de climatologie entre les deux sous-périodes pourrait donc être à l'origine du gain négatif obtenu pour le bassin Loing-E. Notons également que nous observons des climatologies légèrement différentes pour le bassin Seine-Int pour lequel le gain en validation est également négatif. Nous ne pouvons néanmoins affirmer que les résultats obtenus en validation dépendent uniquement de la climatologie.

En moyenne le gain de performance est conservé sur la période de validation. L'introduction de la vitesse verticale dans l'algorithme semble donc pertinente. Mais en quoi les distributions prévues sont-elles meilleures ? Le gain est-il systématique ? Pour répondre à ces questions, nous avons examiné les résultats de la même façon que dans le chapitre précédent.

7.3.3 Diagnostic approfondi des résultats

7.3.3.1 Gain en finesse et en justesse

Le score *CRPS* a été décomposé en finesse et justesse de prévision et le gain de performance respectif, par rapport à la méthode A1, a été évalué pour les deux bassins tests (Tableau 7.8 et Tableau 7.9). Les résultats révèlent que l'analogie sur les champs de vitesse verticale permet d'obtenir des distributions nettement plus fines, avec un gain en finesse par rapport à la méthode A1 de plus de 9% pour le bassin de la Saône et de près de 7% pour le bassin de la Seine. Les distributions obtenues par la méthode A2 sont aussi plus justes en moyenne que par la méthode A1, mais le gain en justesse est moins significatif. Si on

s'intéresse individuellement aux sous-bassins, on remarque que les gains sont très variables, dépendant notamment de la valeur du score de référence, *i.e.* de la valeur du score *CRPSS* obtenue par la méthode A1. Ainsi, le gain en justesse est positif pour chacun des sous-bassins à l'exception du bassin du Loing à Episy, pour lequel le gain est légèrement négatif.

Tableau 7.8 : Gains G (%) de performance, obtenus par la méthode A2 par rapport à la méthode A1, sur le bassin de la Saône, par décomposition du gain global en finesse et en justesse.

	Gain global (%)	Gain en finesse (%)	Gain en justesse (%)
Doubs	6,29	9,12	3,45
Petite Saône	5,30	8,99	1,76
Saône Intermédiaire	6,22	10,07	2,66
<i>Ensemble des bassins</i>	<i>5,94</i>	<i>9,40</i>	<i>2,61</i>

Tableau 7.9 : Gains G (%) de performance, obtenus par la méthode A2 par rapport à la méthode A1, sur le bassin de la Seine, par décomposition du gain global en finesse et en justesse.

	Gain global (%)	Gain en finesse (%)	Gain en justesse (%)
Yonne-C	4,24	7,08	1,12
Seine-Int	3,90	6,57	1,84
Seine-B	4,49	7,48	2,01
Marne-F	4,24	6,41	2,39
Loing-E	2,78	6,95	-0,45
<i>Ensemble des bassins</i>	<i>3,91</i>	<i>6,89</i>	<i>1,35</i>

La finesse moyenne des distributions, bien plus élevée que précédemment, est donc à l'origine de l'amélioration des performances lorsque la vitesse verticale est introduite. Il est vrai qu'il vaut mieux que les deux composantes de la qualité d'une prévision (finesse et justesse) soient améliorées. Dans notre cas, la finesse est assurément meilleure, et la justesse des prévisions obtenue pour l'ensemble de nos bassins n'est pas dégradée pour autant.

Nous retrouvons ici des résultats similaires à ceux obtenus lors de l'introduction de la température. En introduisant ainsi les deux niveaux de sélection supplémentaires (niveau 0 avec la température et niveau 2 avec la vitesse verticale), nous obtenons des prévisions beaucoup plus fines qu'avec la méthode de référence A0 (gains en finesse de 14,4 % et de 11,2 % respectivement pour l'ensemble du bassin de la Saône et de celui de la Seine), mais aussi un peu plus justes (gains en justesse de 5,6 % et de 4,6 %).

7.3.3.2 Robustesse du gain de performance

Comme pour la température, nous avons analysé la robustesse du gain de performance afin de vérifier si le gain est fréquemment positif. La distribution des écarts entre les valeurs de score *CRPS* calculées par les méthodes A1 et A2 est représentée sur la Figure 7.11 pour le bassin de la Marne à La Ferté. Le gain est à nouveau loin d'être systématique, avec une distribution centrée en une valeur proche de 0 mais positive. On devine que le gain est plus

fréquemment positif que négatif, contrairement au gain apporté par la méthode A1 par rapport à la méthode A0 où c'était peu évident. Ce résultat se retrouve pour tous les sous-bassins étudiés.

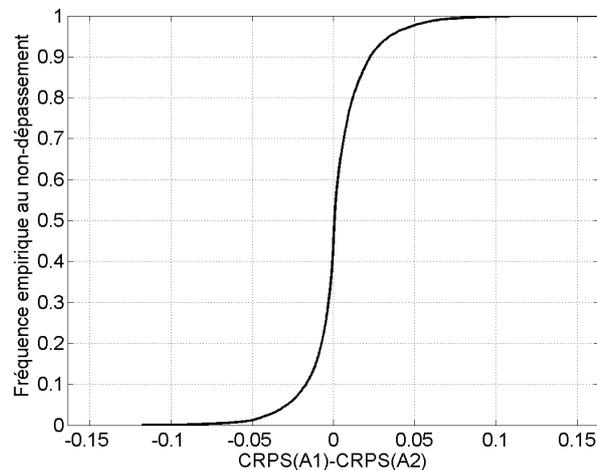


Figure 7.11 : Distribution empirique des écarts de score *CRPS* entre les méthodes A1 et A2 calculés sur la période de calibration, pour le bassin de la Marne à La Ferté.

Rappelons-nous que la vitesse verticale a été introduite pour caractériser la stabilité des masses d'air, et donc pour améliorer potentiellement la prévision des situations pluvieuses. Cela nous amène à nous poser la question suivante : gagne-t-on plus souvent à appliquer ce niveau intermédiaire pour les situations cibles pluvieuses ? Nous avons repris les groupes S et P des situations sèches et pluvieuses, séparément pour les bassins de la Saône et de la Seine. Pour chaque groupe, la distribution des écarts de score *CRPS* a été tracée. La Figure 7.12 présente, pour le bassin de la Marne à La Ferté, les distributions associées à chacun des deux groupes.

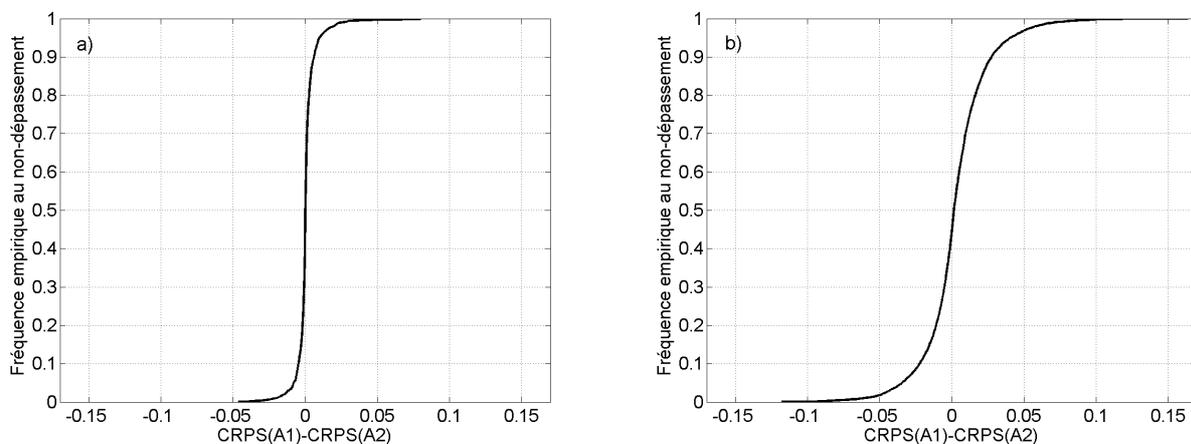


Figure 7.12 : Distribution empirique des écarts de *CRPS* entre les méthodes A1 et A2 calculés sur la période de calibration, pour le bassin de la Marne à La Ferté. a) situations du groupe S ; b) situations du groupe P.

Les deux distributions ne sont pas nécessairement comparables puisque, rappelons-le, la taille de chaque échantillon est différente (échantillon nettement plus petit pour le groupe S). Les résultats sont tout de même intéressants : la prévision des situations pluvieuses ne semble pas privilégiée par rapport à celle des situations sèches. On remarque néanmoins que la proportion de situations cibles, pour lesquelles les performances sont meilleures avec la méthode A2, est plus élevée pour le groupe P par rapport au groupe S. Par ailleurs le nombre d'écart très faibles est beaucoup plus élevé pour le groupe P, avec une courbe en forme de créneau plus prononcée. Cela s'explique par les très faibles valeurs de score *CRPS* pour les situations de ce type. En effet, comme nous l'avons déjà mentionné, les journées durant lesquelles il ne pleut pas sont mieux prévues que les autres. Enfin, nous constatons également que l'écart positif extrême est plus élevé que l'écart négatif extrême pris en valeur absolue.

Les résultats obtenus pour tous les sous-bassins ont été résumés dans le Tableau 7.10, dans lequel sont indiquées les proportions de situations cibles pour lesquelles le gain est positif ou négatif, et ce pour l'ensemble des situations mais aussi pour celles des groupes S et P. Ces résultats sont similaires pour tous les sous-bassins. On remarque toutefois que la perte de performance est plus fréquente pour les sous-bassins de la Seine, avec en moyenne 44 à 45 % des situations cibles de la période de calibration, contre 39 à 40 % pour les sous-bassins de la Saône. C'est en effet sur les jours sans pluie observée que les performances sont plus souvent moins dégradées pour les sous-bassins de la Saône (moins de 35 % des situations) que pour ceux de la Seine (42 à 47 % des situations). Ces résultats doivent néanmoins être nuancés en raison de l'échantillonnage, plus faible pour le bassin de la Seine (13,8 % de l'effectif) que pour celui de la Saône (26,9 % de l'effectif).

Tableau 7.10 : Proportions (%) de situations cibles pour lesquelles le gain de performance de la méthode A2 par rapport à la méthode A1 est positif (Gain +) et négatif (Gain -), pour toutes les situations confondues, pour les situations sèches (groupe S) et pour les situations pluvieuses (groupe P).

Bassin	Toutes situations confondues		Situations sèches (S)		Situations pluvieuses (P)	
	Gain +	Gain -	Gain +	Gain -	Gain +	Gain -
Doubs	57,3	39,0	55,8	31,0	56,8	43,2
Petite Saône	55,0	39,8	50,1	32,5	56,7	43,3
Saône Int.	56,9	39,9	54,7	34,2	57,0	42,9
Yonne-C	55,6	44,3	54,5	45,2	55,6	44,4
Seine-Int	54,9	45,0	52,6	47,1	54,8	45,2
Seine-B	55,7	44,0	53,2	45,5	56,1	43,9
Marne-F	55,7	44,2	55,2	44,3	55,7	44,3
Loing-E	54,0	44,9	52,3	42,2	54,4	45,5

7.4 Conclusion

Dans ce chapitre, nous avons cherché à introduire une variable d'analogie qui permette de

caractériser et de localiser les fronts, générateurs de l'essentiel des précipitations sur nos bassins tests. Nous avons suggéré plusieurs variables, mais c'est finalement la vitesse verticale qui s'est révélée la plus intéressante pour améliorer la sélection d'analogues.

Nous avons essayé d'introduire cette nouvelle variable de deux manières : soit entre les niveaux 1 et 2, soit après le niveau 2. Les résultats ont révélé l'intérêt de l'introduire dans un niveau de sélection intermédiaire. Par conséquent, dans le nouvel algorithme incluant la vitesse verticale, la sélection en termes d'humidité constitue la dernière étape (niveau 3), et la sélection en termes de mouvement vertical constitue le niveau 2.

Nous avons procédé ensuite à l'optimisation des paramètres de ce niveau de sélection supplémentaire. La première étape a consisté à fixer certains paramètres (nombres d'analogues, heures de mesure) et à appliquer plusieurs critères d'analogie en mesurant la variable W à différents niveaux de pression. Le critère qui est apparu le plus performant est le critère de distance *RMSE*.

Nous avons ensuite optimisé les nombres d'analogues à sélectionner aux niveaux 1 et 2, le nombre retenu au niveau 3 restant identique. Nous supposons de cette manière que le nombre $N3$ ne dépend pas de l'échantillon de situations retenues au cours des niveaux de sélection précédents. Le nombre $N2$ optimal est finalement identique au nombre $N1$ de la méthode A1 (70), en revanche le nombre $N1$ optimal s'avère bien plus élevé (170). Nous retenons ainsi beaucoup plus de situations analogues d'un point de vue de la circulation atmosphérique, le tri étant ensuite fait par analogie sur le mouvement vertical de l'air.

Les instants auxquels la variable W est mesurée ont été optimisés en faisant varier le nombre d'instants entre un et cinq. L'optimisation a montré que la vitesse verticale doit être mesurée à des heures incluses dans la période du cumul pluviométrique. Les performances sont nettement améliorées lorsque cette variable est mesurée à plusieurs instants au lieu d'un seul instant, mais elles semblent être équivalentes entre quatre et cinq heures de mesure. La distinction entre les situations pluvieuses et non pluvieuses a révélé que le gain de performance maximal n'est pas atteint pour la même combinaison. Ainsi pour les jours où il pleut, il semble préférable de mesurer la variable W à quatre instants. C'est ce choix qui a été finalement retenu.

L'échelle à laquelle l'analogie sur la vitesse verticale doit être recherchée, représentée par la taille du domaine optimal, est finalement comprise entre l'échelle considérée au niveau 1 et celle considérée au niveau 3. En effet, nous avons vu que les domaines successivement appliqués aux trois niveaux s'emboîtaient. La taille du domaine justifie ainsi l'introduction de cette sélection intermédiaire, ce qui constitue un résultat probant. Par ailleurs, nous avons remarqué que le domaine optimal ne dépend pas du choix des heures de mesure. Finalement, le gain de performance obtenu par rapport à la méthode A1 s'avère conséquent pour le bassin de la Saône.

Nous avons ensuite validé le nouvel algorithme, nommé A2, en effectuant les mêmes optimisations pour le bassin de la Seine. Les résultats sont similaires à ceux obtenus pour le bassin de la Saône. Seule une des quatre heures de mesure diffère par rapport à ce que nous

obtenons sur la Saône. Le gain de performance finalement obtenu sur cet autre bassin test n'est pas négligeable. Le gain de performance a ensuite été évalué sur la période de validation. En moyenne, il reste largement positif pour le bassin de la Saône, tandis qu'il est plus faible mais également positif pour celui de la Seine. En revanche, les gains évalués pour les sous-bassins de la Seine ne sont pas toujours positifs, ce qui pourrait être dû à la différence de climatologie des pluies observée entre les périodes de calibration et de validation. Enfin, au travers d'un diagnostic des résultats, nous nous sommes aperçus que les distributions des précipitations sont en moyenne beaucoup plus fines que celles obtenues par la méthode A1, et qu'elles sont également un peu plus justes, sauf pour un sous-bassin de la Seine, le Loing à Episy, pour lequel le gain en justesse est faiblement négatif. Néanmoins en moyenne, le gain en finesse et le gain en justesse sont tous deux positifs. En revanche, à l'image de ce que nous avons mis en évidence lors de l'introduction de la température, le gain de performance apporté par cette sélection supplémentaire n'est pas robuste. Ainsi, la prévision par la méthode A2 s'avère plus performante pour seulement 54 à 57 % des situations de la période de calibration.

Enfin, il faut garder à l'esprit que cette variable permet de caractériser les fronts à échelle synoptique (phénomène prépondérant dans la production de précipitations sur les secteurs d'étude), mais elle ne permet pas, à l'échelle considérée dans les réanalyses, de caractériser les phénomènes de convection estivale ou de convection d'air froid observés dans la traîne d'un front.

Finalement, la procédure de sélection par la méthode A2 se déroule maintenant en quatre étapes (Figure 7.13) : un niveau 0 de pré-sélection des situations candidates selon la température, un niveau 1 de sélection d'analogues en terme de circulation générale, un niveau 2 de sélection d'analogues en terme de mouvement vertical, et un niveau 3 de sélection d'analogues en terme d'humidité.

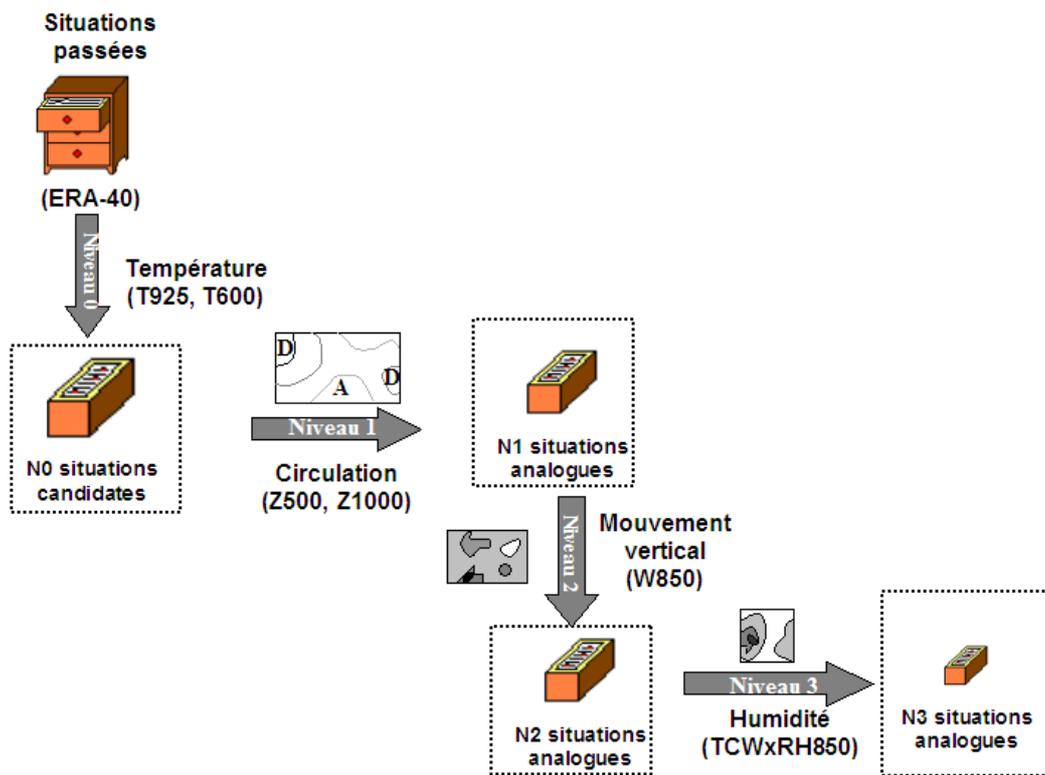


Figure 7.13 : Procédure schématisée de sélection d'analogues par la méthode A2.

8 Optimisation des paramètres de référence

Deux voies possibles d'amélioration d'une méthode de prévision par analogie ont été évoquées en introduction à la seconde partie : (1) le développement et (2) l'optimisation. Les chapitres précédents ont exploré la première voie, en introduisant deux nouvelles variables : la température et la vitesse verticale. Les variables d'analogie étant désormais définies, le présent chapitre emprunte la seconde voie en étudiant l'optimisation des procédures d'identification d'analogie selon les variables de pression et d'humidité (Z et $TCW\%$), communes à la méthode A0 et aux méthodes A1 et A2 développées respectivement dans les chapitres 6 et 7.

Deux pistes seront explorées :

- i) au premier niveau de sélection de la méthode A0, nous remettrons en question le domaine d'analogie fixe, actuellement identique quelle que soit la situation cible. Nous tenterons de le rendre plus flexible en le rendant dépendant à la situation cible.
- ii) aux deux niveaux de sélection de la méthode A0, les variables d'analogie sont mesurées à des instants et à des niveaux de pression définis par Bontron (2004) selon leur disponibilité dans l'archive LTHE. Or dans le cadre de ce travail, grâce à l'utilisation de l'archive de réanalyses ERA-40, des paramètres supplémentaires (heures de mesure et niveaux de pression) sont désormais accessibles. Nous étudierons donc les gains que peut éventuellement apporter l'utilisation de ces nouvelles données.

Les résultats de ces optimisations feront l'objet d'une synthèse à la fin de chapitre : ainsi, seuls seront retenus les éléments nouveaux les plus utiles en terme de performance de prévision.

Dans ce chapitre, la méthode d'analogie utilisée sera la méthode A2 développée précédemment. La sélection d'analogues se fait donc en quatre étapes, avec les paramètres suivants valables pour le bassin de la Saône :

- niveau 0 : mesure de la température à 925 hPa à +36h et à 600 hPa à +12h, domaine

- d'analogie constitué d'un seul point de grille (de coordonnées (5°E ; 47,5°N) pour le bassin de la Saône), critère d'analogie *RMSE*, 2900 situations retenues ;
- niveau 1 : mesure du géopotential à 1000 hPa à +12h et à 500 hPa à +24h, domaine d'analogie (-5°E/15°E ; 40°N/50°N), critère d'analogie *TWS*, 170 situations retenues ;
 - niveau 2 : mesure de la vitesse verticale à 850 hPa à +6h, +12h, +18h et +24h, domaine d'analogie (2,5°E/7,5°E ; 45°N/47,5°N), critère d'analogie *RMSE*, 70 situations retenues ;
 - niveau 3 : mesure du produit de l'eau précipitable par l'humidité relative à 850 hPa à +12h et +24h, domaine d'analogie constitué de quatre points de grille, critère d'analogie *RMSE*, 25 situations retenues.

Nous nous intéresserons ici aux paramètres de référence des niveaux 1 et 3 : ceux liés aux variables *Z* et *TCW%*.

8.1 Domaine d'analogie du niveau 1

Dans la procédure de sélection d'analogues du niveau 1, l'exploration est limitée à une fenêtre spatiale figée, représentative de l'ensemble des situations. Or il se trouve que la démarche du prévisionniste est différente. Considérons par exemple une situation synoptique de retour d'est. La dépression d'altitude est typiquement centrée plus au sud-est du bassin d'intérêt. C'est alors dans cette direction que le prévisionniste recherche l'information utile, *i.e.* les centres d'action auxquels est soumis le bassin. Bontron (2004) avait optimisé la variable et le domaine d'analogie associés au premier niveau de sélection. Par conséquent, il serait judicieux d'appliquer un domaine dynamique en fonction de chaque situation cible.

Or Bontron (2004) avait tenté d'adapter le domaine à la situation cible de deux manières. La première consistait à construire une fenêtre dynamique, dont les points de grille étaient dépendants du champ de géopotential. La seconde consistait à optimiser la fenêtre en fonction du type de temps des situations cibles. Mais les résultats ne s'étaient pas avérés concluants.

Nous proposons ici d'explorer une solution alternative. De manière générale, on constate que les précipitations se déplacent selon le sens du flux d'altitude. C'est donc dans la direction de la provenance du flux qu'il faudrait peut-être rechercher l'analogie sur les champs de géopotential. L'objectif est donc ici d'appliquer un domaine en fonction du flux d'altitude.

Nous avons décidé de diviser la provenance du flux d'altitude en huit classes (Figure 8.1) : Ouest/Sud-Ouest (OSO), Sud-Ouest/Sud (SOS), Sud/Sud-Est (SSE), Sud-Est/Est (SEE), Est/Nord-Est (ENE), Nord-Est/Nord (NEN), Nord/Nord-Ouest (NNO) et Nord-Ouest/Ouest (NOO). Le flux est ici défini par les composantes *U* (zonale) et *V* (méridienne) du vent horizontal mesuré à 500 hPa. Puisque l'analogie sur les champs de géopotential *Z1000* et *Z500* est recherchée respectivement à +12h et +24h, nous avons choisi de mesurer les composantes du vent à +18h. Enfin, le vent est mesuré au point de grille le plus proche du

bassin de la Saône (5°E ; 47,5°N).

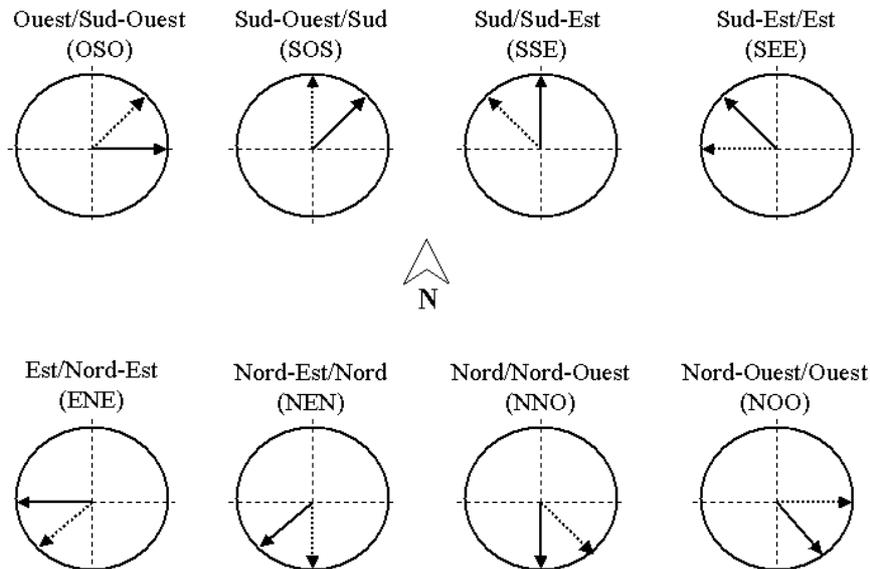


Figure 8.1 : Classification de la direction du flux d'altitude à 500 hPa. La direction du flux est comprise entre la flèche pleine et la flèche discontinue, à l'exclusion de la direction définie par cette dernière.

Les situations cibles ne sont pas réparties uniformément dans ces classes, comme l'illustre le Tableau 8.1, dans lequel est reportée la proportion de situations cibles de la période de calibration appartenant à chaque classe. Ainsi les régimes de composante Ouest sont prédominants, tandis que les classes de composante Est sont peu représentées. Afin de caractériser la climatologie, la Figure 8.2 représente les distributions empiriques des précipitations (transformées, cf. section 2.2.2.3) obtenues sur la période de calibration et pour les situations appartenant à chaque classe de flux, pour le bassin du Doubs. Les classes de composante Ouest, notamment OSO et SOS, sont les plus pluvieuses ; tandis que celles de composante Est sont plus sèches. Pour ces classes, certaines situations présentent toutefois des cumuls pluviométriques importants (lames d'eau transformées supérieures à 0,8) : c'est en particulier le cas lors d'une situation à retour d'est.

Tableau 8.1 : Proportion (%) de situations cibles appartenant à chacune des huit classes de provenance du vent à 500 hPa.

	Classe de provenance du vent à 500 hPa mesuré à +18h							
	OSO	SOS	SSE	SEE	ENE	NEN	NNO	NOO
Proportion (%)	25,0	13,6	4,1	3,3	5,1	9,1	16,7	23,3

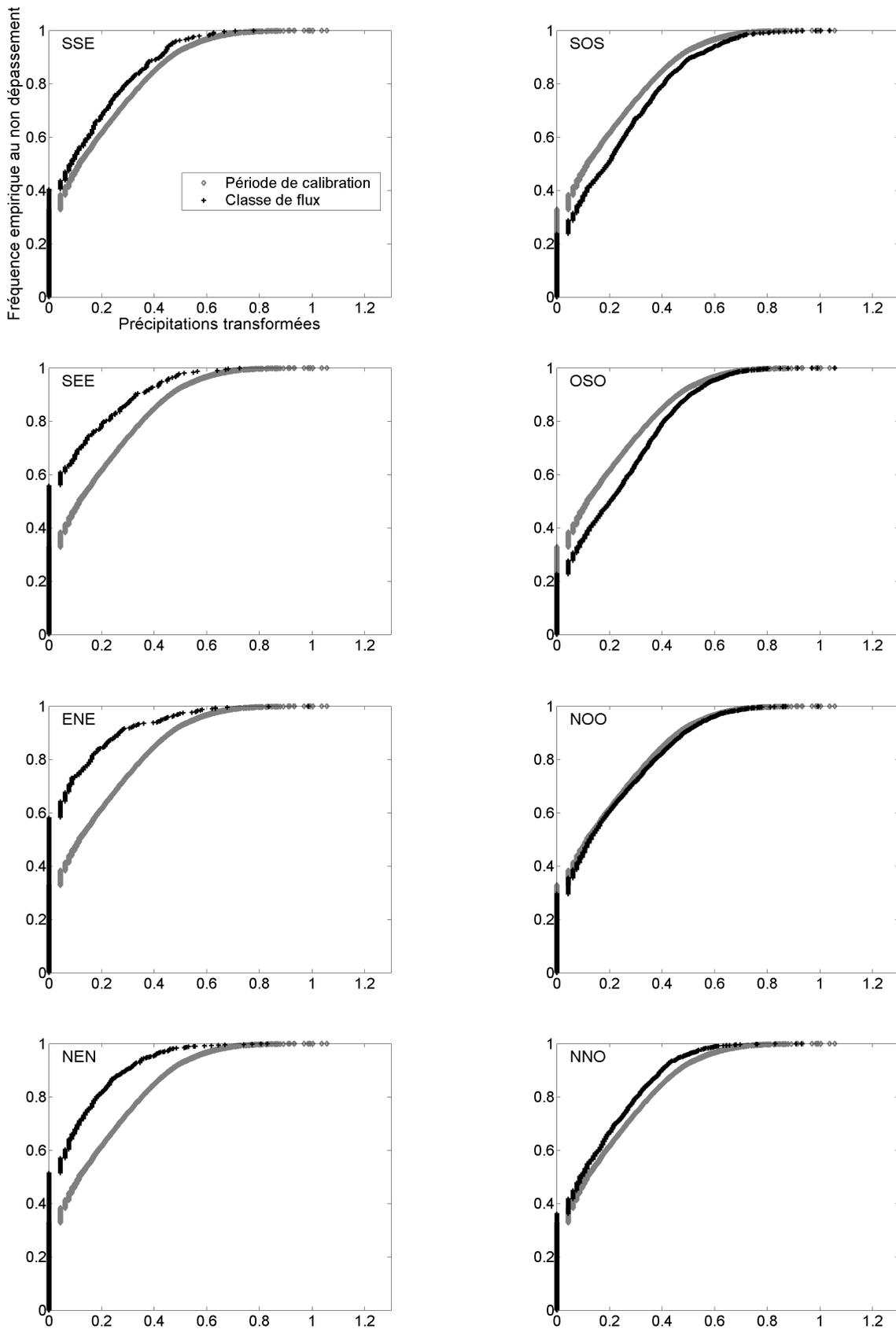


Figure 8.2 : Distributions empiriques des précipitations observées sur le bassin du Doubs, d'une part pour l'ensemble des situations la période de calibration, et d'autre part pour les situations appartenant à chaque classe de flux.

Pour chacune des classes, la pertinence d'analogie des points de grille a été cartographiée. La performance de prévision a été évaluée à l'issue du dernier niveau de sélection de la méthode A2, uniquement à l'aide du score *CRPS* afin d'analyser la répartition des zones pertinentes. Les cartes de pertinence sont représentées sur la Figure 8.3 pour chacune des huit classes prédéfinies.

En raison de l'échantillonnage (Tableau 8.1), il faut relativiser les résultats obtenus pour les classes les moins représentées (SSE, SEE et ENE). Cette cartographie de la pertinence d'analogie par classe révèle une forte liaison entre zone de pertinence et direction de provenance du vent d'altitude. En effet, pour les classes les plus représentées, la zone la plus pertinente se situe généralement en marge du bassin de la Saône et présente une forme étirée dans la direction de la provenance du vent. De plus, elle semble aussi être légèrement décalée dans l'espace par rapport au bassin. Ce décalage correspond à l'influence des systèmes dépressionnaires sur les précipitations à l'avant de ces derniers. Par exemple pour la classe OSO, dont la climatologie révélée sur la Figure 8.2 est très pluvieuse, le système dépressionnaire est généralement positionné au Nord-Ouest du bassin. La zone d'intérêt se situe alors en amont, au Sud-Est de ce système. En revanche, pour la classe ENE, plutôt représentative des situations anticycloniques (flux d'origine continentale), il n'y a pas de véritable cohérence. La répartition est plutôt homogène avec une zone de plus forte pertinence autour du bassin et sur l'Europe du nord. Cela traduit la présence récurrente, lorsque le vent a une composante principale Est, d'un anticyclone positionné à proximité du bassin ou au nord de celui-ci. Ces résultats témoignent de la sensibilité de la zone de plus forte pertinence au flux directeur.

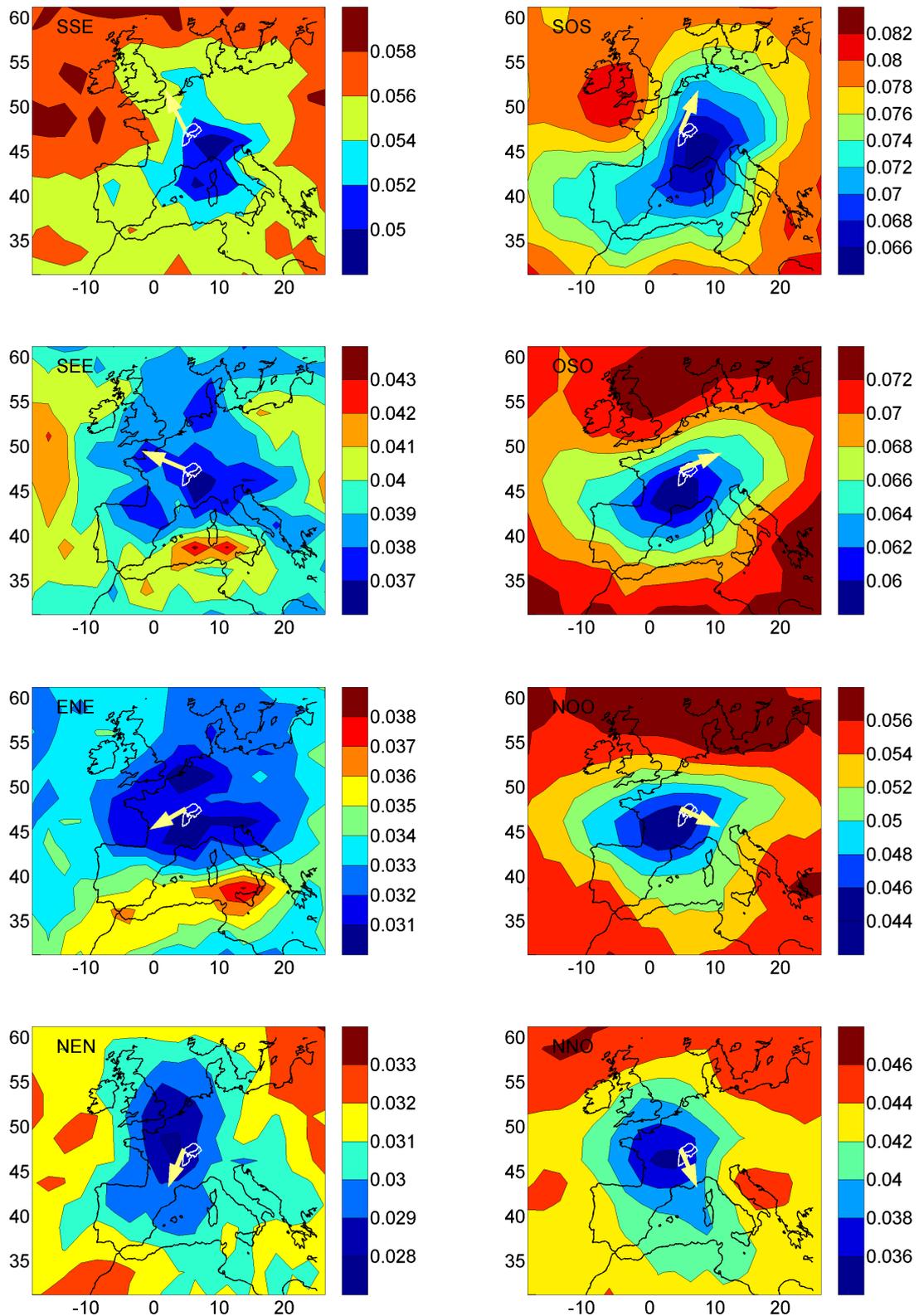


Figure 8.3 : Cartographie de la pertinence d'analogie (CRPS), obtenue sur le bassin de la Saône, pour chacune des huit classes de provenance du vent à 500 hPa. Les flèches jaunes indiquent la direction moyenne du vent de chaque classe.

Nous avons donc optimisé ensuite le domaine d'analogie associé à chaque classe, en partant du point de grille le plus pertinent. Les domaines optimaux obtenus sont très

variables en fonction des classes considérées, donc fortement dépendants des situations cibles. Les domaines obtenus pour les classes OSO et SEE sont représentés sur la Figure 8.4. Ils sont très différents, et celui de la classe SEE doit fortement dépendre de l'échantillon de situations cibles associées (seulement 3,3 % de l'effectif).

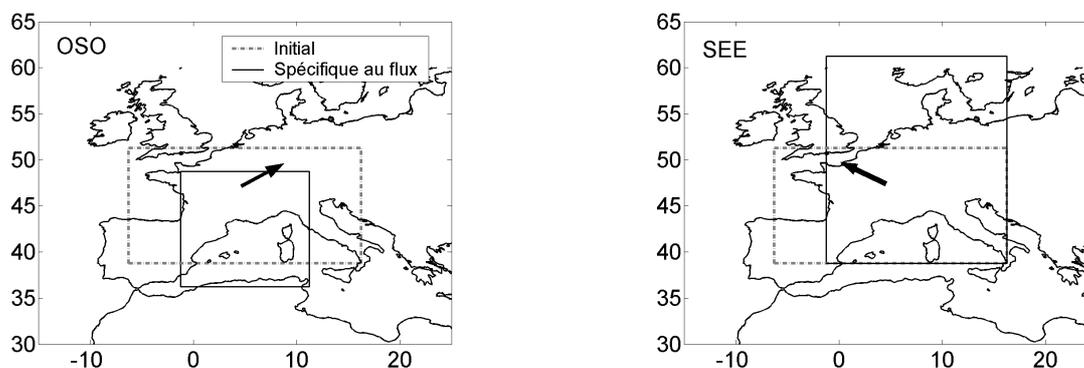


Figure 8.4 : Domaine optimal obtenu pour les classes OSO et SEE de provenance du vent à 500 hPa (trait plein), pour le bassin de la Saône. Le domaine fixe initial est également indiqué en trait discontinu.

Nous avons ensuite appliqué, pour chacune des situations cibles de la période de calibration, le domaine optimal associé à la classe à laquelle elle appartient. Le gain G de performance moyenne, obtenu par rapport à la méthode A2, s'élève à 0,75 % sur la période de calibration. Ce gain est donc faible par rapport à ceux obtenus jusqu'à présent.

Il nous semble toutefois intéressant de vérifier si le gain varie d'une classe à une autre. Les performances ont donc été évaluées pour chacune des classes et nous en avons déduit le gain G par rapport à la méthode A2 pour chaque classe de flux (Tableau 8.2).

Tableau 8.2 : Gain G (%) de performance, obtenu par rapport à la méthode A2 en fonction de la classe de flux, pour chaque sous-bassin et pour l'ensemble des bassins de la Saône. Les valeurs négatives sont indiquées en gras.

	Classe de flux							
	OSO	SOS	SSE	SEE	ENE	NEN	NNO	NOO
Doubs	0,97	1,77	0,63	0,87	1,43	2,31	-0,28	0,32
Petite S.	0,63	2,12	4,60	3,46	1,33	3,04	-0,82	0,43
Saône l.	0,45	1,38	-0,09	-1,12	1,92	1,65	0,42	1,12
Ensemble	0,68	1,75	1,71	1,07	1,56	2,33	-0,23	0,62

Pour l'ensemble des sous-bassins, les gains G s'avèrent positifs sauf pour la classe NNO. De manière générale, les gains diffèrent fortement d'un sous-bassin à un autre. Néanmoins, les scores associés aux sous-bassins sont souvent positifs, traduisant un certain intérêt d'appliquer un domaine spécifique à ces classes de flux.

Cependant, ce faible gain de performance n'a pas été conservé en validation, période sur laquelle on a mesuré un gain G négatif de -0,73 % pour l'ensemble du bassin de la Saône. Les gains G évalués par classe se révèlent très contrastés, dont certains sont largement négatifs,

à l'image de celui obtenu pour la classe SEE (-13,75 %).

8.2 Des heures de mesure supplémentaires

L'heure de mesure d'une variable d'analogie est un paramètre qu'il est possible d'optimiser dès lors que l'on dispose des variables d'analogie à différents instants. La méthode de référence A0 a été optimisée par Bontron (2004) avec une archive (l'archive LTHE) au pas de temps de 12 h. Disposant maintenant d'une archive au pas de temps de 6 h, nous avons donc décidé de mesurer l'apport des heures supplémentaires 6h TU et 18h TU en ce qui concerne les trois variables d'analogie *Z500*, *Z1000* et *TCW%*.

8.2.1 Variables du niveau 1

Dans la méthode A0, les variables *Z1000* et *Z500* sont mesurées respectivement aux instants +12h et +24h d'un jour cible *C*. Nous avons utilisé, en plus des heures 0 h TU et 12 h TU, les deux heures supplémentaires 6 h TU et 18 h TU en testant plusieurs combinaisons d'heures de mesure des géopotentiels *Z1000* et *Z500*. Nous espérons de cette manière caractériser la circulation générale avec plus de précision et ainsi prendre en compte des éléments synoptiques négligés au pas de temps de 12 h. Le nombre total d'instantes a été limité arbitrairement à quatre pour les deux variables, ce qui représente six cas différents :

- un instant de mesure pour chaque variable (1 cas) ;
- un instant pour l'une et deux pour l'autre (2 cas) ;
- deux instants pour chaque variable (1 cas) ;
- un instant pour l'une et trois instants pour l'autre (2 cas).

Pour une prévision de précipitations au pas de temps journalier, les heures de mesure sont généralement incluses dans la période du cumul pluviométrique, comme Bontron (2004) l'a démontré lors de ses développements et comme nous l'avons confirmé pour la vitesse verticale (cf. 7.2.3). Nous avons donc restreint l'optimisation aux instants de mesure +6h à +30h pour un jour cible *C*. Par ailleurs, précisons que la valeur du critère *TWS* finalement calculée correspond à la moyenne arithmétique des valeurs du critère *TWS* obtenues pour les différentes heures de mesure.

Toutes les combinaisons possibles ont été explorées pour chacun des six cas, et nous avons retenu la combinaison pour laquelle le meilleur gain de performance a été obtenu (Figure 8.5). Les gains de performance obtenus sont faiblement positifs voire négatifs dans certains cas. Nous nous apercevons que la combinaison optimale d'une heure de mesure pour chacune des variables (+6h et +24h) n'est pas la combinaison utilisée jusqu'à présent (+12h et +24h), mais le gain de performance est négligeable (+0,2 %). Le gain le plus élevé (+0,68 % par rapport à la méthode A2) est obtenu en mesurant chacune des variables *Z1000* et *Z500* à deux instants. Cependant, la valeur de ce gain est faible, non seulement par rapport à l'augmentation du temps de calcul induit, mais aussi par rapport à d'autres pistes explorées.

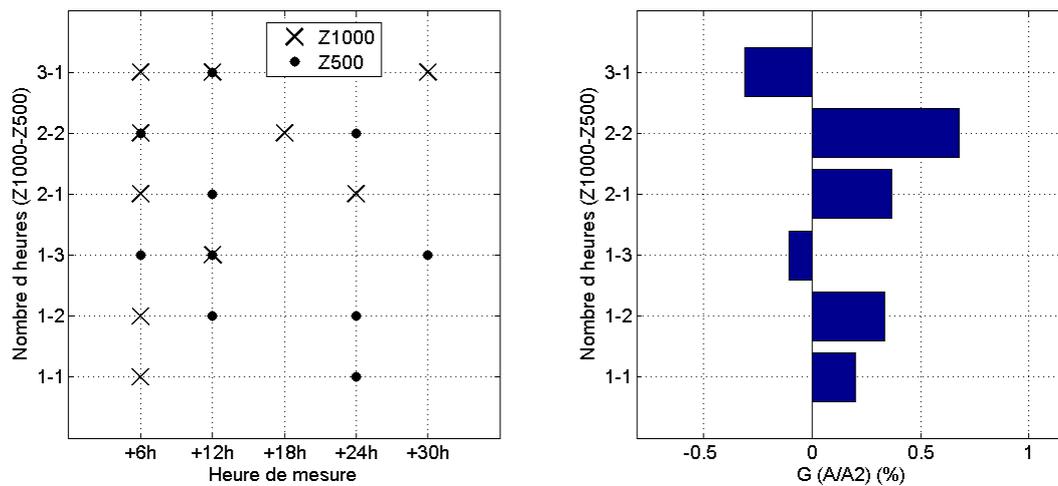


Figure 8.5 : Combinaisons optimales d'heures de mesure des deux variables d'analogie Z1000 et Z500 du niveau 1, lorsque chaque variable est mesurée à un, deux ou trois instants ; et gains G (%) de performance associés aux combinaisons, valables pour le bassin de la Saône.

8.2.2 Variable du niveau 3

La variable $TCW\%$ utilisée au cours du dernier niveau de sélection est mesurée à deux instants différents (+12h et +24h). De la même manière que pour les variables du niveau 1, nous avons envisagé toutes les combinaisons possibles d'une heure à cinq heures de mesure, ces instants étant toujours inclus dans la période du cumul pluviométrique. Précisons également que la valeur du critère $RMSE$ finalement calculée a été obtenue en sommant les valeurs du critère $RMSE$ associées aux différents instants de mesure.

Les combinaisons optimales et le gain G correspondant sont indiqués sur la Figure 8.6.

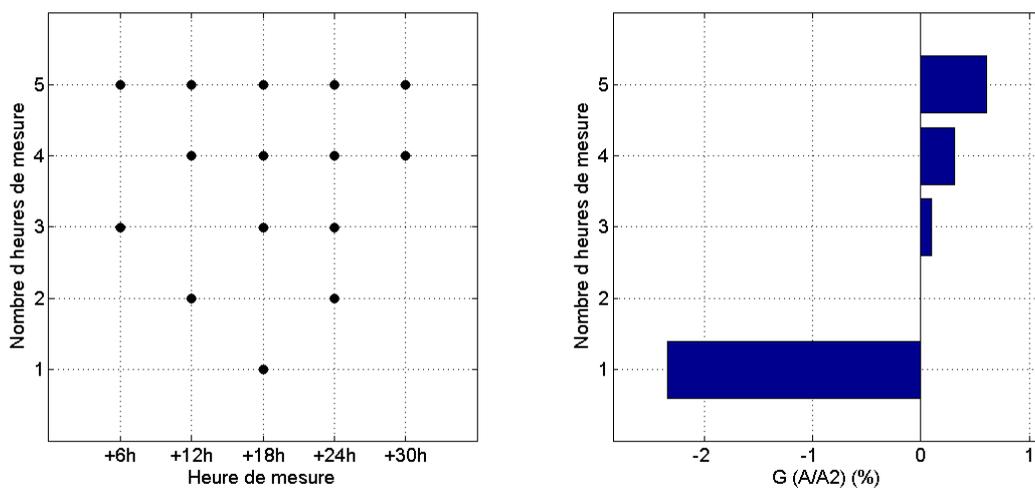


Figure 8.6 : Combinaisons optimales d'heures de mesure de la variable $TCW\%$ du niveau 3, pour un nombre d'instants compris entre un et cinq ; et gains G (%) de performance associés aux combinaisons, valables pour le bassin de la Saône.

C'est lorsque la variable $TCW\%$ est mesurée aux cinq heures de mesure que les

performances sont les plus élevées. Le gain apporté par la mesure à cinq instants au lieu de deux reste tout de même assez faible (+0,61 % par rapport à la méthode A2). En revanche, il apparaît clairement que la variable *TCW%* doit être mesurée à au moins deux instants. En effet en la mesurant à un seul instant (à +18h), la perte de performance est significative (-2,34 % par rapport à la méthode A2). Nous remarquons d'autre part que les heures supplémentaires 6 h TU et 18 h TU ne sont pas privilégiées lorsque *TCW%* est mesurée à deux instants : la combinaison optimale est identique à celle utilisée jusqu'à présent.

8.3 Des niveaux de pression supplémentaires

8.3.1 Constat préliminaire

Parmi les variables brutes utilisées dans la méthode de référence A0, deux sont mesurées à un niveau de pression : le géopotential *Z* (à 1000 et à 500 hPa) et l'humidité *RH* (à 850 hPa). Dans le cadre de nos travaux, nous avons eu accès à tous les niveaux de pression disponibles pour ces variables, mais seuls ont été retenus ceux qui nous semblent les plus pertinents vis-à-vis de l'information que peuvent fournir ces variables. La Figure 8.7 indique, pour les variables *Z* et *RH*, les niveaux de pression qui furent accessibles par Bontron (2004) dans l'archive LTHE et ceux que nous avons extraits des réanalyses ERA-40.

Les niveaux supplémentaires auxquels la variable *Z* pourrait être mesurée ne sont pas nombreux (400, 600 et 925 hPa), et ne devraient donc pas apporter d'information décisive par rapport à ceux de l'archive LTHE, qui représentent des niveaux de pression standard. En revanche, d'autres niveaux que ceux de l'archive LTHE sont désormais accessibles pour l'humidité et pourraient apporter une information complémentaire : à 925 hPa, la saturation de l'air en humidité est synonyme de nuages plutôt bas et potentiellement générateurs de pluie, en particulier en montagne¹ ; à 775 ou 700 hPa, l'humidité peut renseigner sur le potentiel d'orages organisés², mais aussi sur la présence de nuages moyens, également producteurs de précipitations.

1 Certains épisodes pluvieux peuvent être causés par la convergence d'un air humide sur le flanc exposé des montagnes. L'exemple le plus connu en France, mais aussi le plus intense, est celui qui s'abat sur le massif des Cévennes.

2 Les systèmes orageux puissants et organisés sont souvent favorisés par la présence d'air plus sec dans les couches moyennes de la troposphère (autour de 700 hPa).

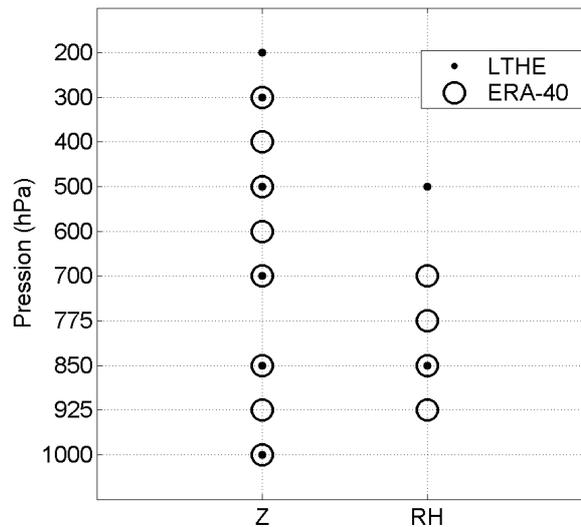


Figure 8.7 : Niveaux de pression aux quels les variables Z et RH sont disponibles : dans les archives LTHE et ERA-40, respectivement utilisées par Bontron (2004) et dans la présente thèse.

Nous allons à présent procéder à une optimisation afin d'évaluer l'intérêt d'utiliser d'autres niveaux de pression que le niveau 850 hPa pour l'humidité relative.

8.3.2 Application à l'humidité relative

Dans cette section, les niveaux 0, 1 et 2 de sélection de l'algorithme A2 sont appliqués avec les paramètres valables pour le bassin de la Saône. Les paramètres conservés pour le niveau 3 sont le critère d'analogie ($RMSE$) et les heures de mesure (+12h et +24h).

8.3.2.1 Un niveau de pression fixe

Nous considérons les quatre niveaux de pression aux quels la variable RH a été rapatriée. Le domaine d'analogie est peut-être dépendant du niveau de pression : nous avons donc choisi de l'optimiser pour chaque niveau auquel l'humidité relative peut être mesurée. La pertinence d'analogie des points de grille a également été cartographiée, et ce pour l'ensemble du bassin de la Saône (Figure 8.8).

Nous aurions pu nous attendre à voir une répartition de la pertinence différente d'un niveau de pression à un autre. Les résultats montrent, au contraire, que cette répartition est fortement similaire. La zone la plus pertinente est bien centrée sur le bassin, ce qu'avait également montré Bontron (2004) avec la variable $RH850$. Le domaine optimal associé à chacun des niveaux de pression devrait donc être identique. Cependant, le point de grille le plus pertinent, à partir duquel le domaine d'analogie sera optimisé, n'est pas tout à fait le même pour les quatre niveaux de pression : point de coordonnées (5°E ; 47,5°N) pour les niveaux 925 et 850 hPa ; point de coordonnées (7,5°E ; 47,5°N) pour les niveaux 775 et 700 hPa.

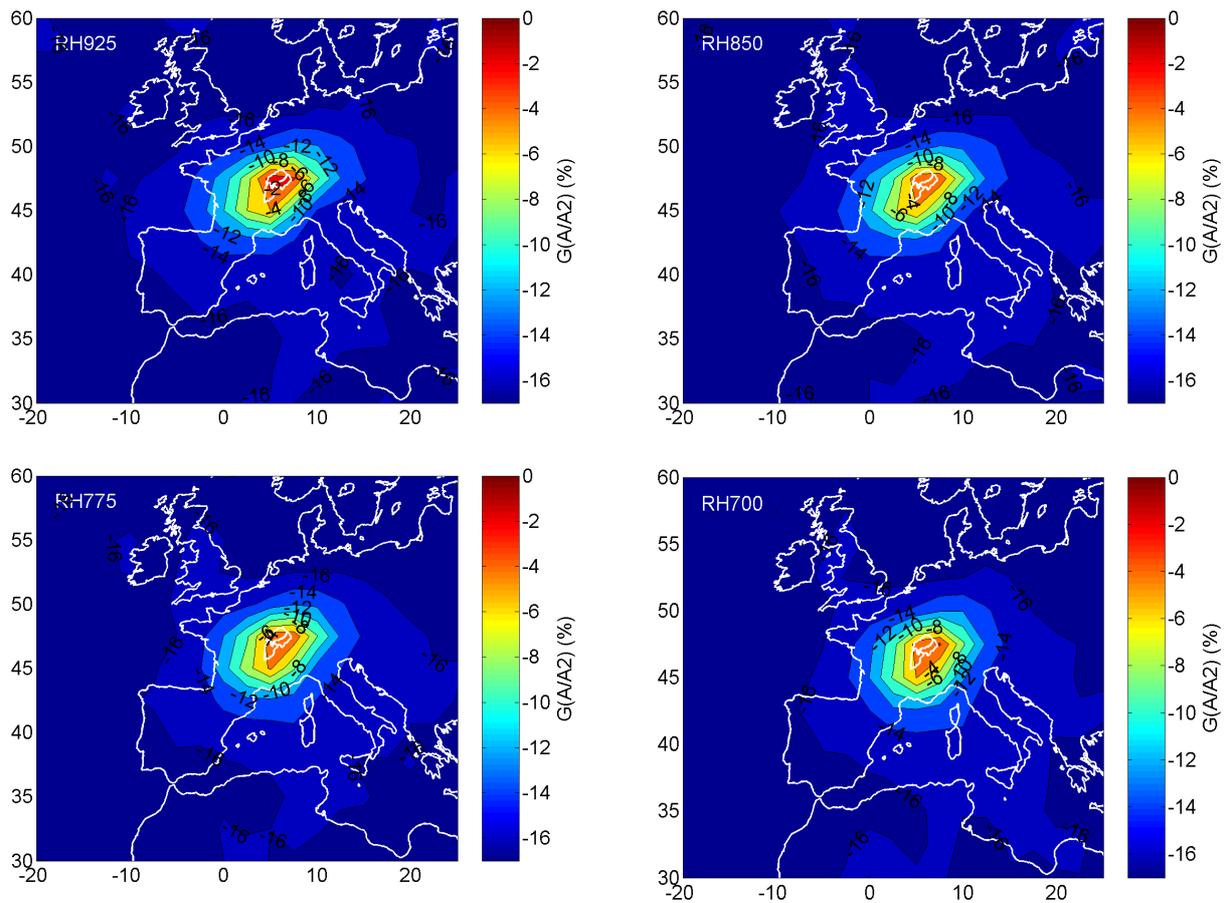


Figure 8.8 : Cartographie de la pertinence d'analogie à l'issue du niveau 3 pour le bassin de la Saône, exprimée en gain G (%) de performance par rapport à la méthode A2, lorsque l'humidité relative est mesurée aux niveaux de pression 925 hPa, 850 hPa, 775 hPa et 700 hPa.

Le domaine d'analogie a ensuite été optimisé pour l'ensemble du bassin de la Saône lorsque l'humidité relative est mesurée à chaque niveau de pression. Le domaine optimal est identique, quel que soit le niveau de pression considéré pour l'humidité relative. C'est donc le même que celui appliqué jusqu'à présent lorsque l'humidité était mesurée au niveau 850 hPa. Les gains de performance obtenus pour les différents niveaux de pression auxquels l'humidité relative est mesurée sont indiqués dans le Tableau 8.3. Le gain a également été évalué pour chaque sous-bassin, de manière à vérifier si le niveau optimal diffère d'un bassin à un autre.

Tableau 8.3 : Gain G (%) de performance moyenne, obtenu pour le bassin de la Saône par rapport à la méthode A2, pour les différents niveaux de pression auxquels l'humidité relative peut être mesurée. Les valeurs minimales et maximales figurent respectivement en italique et en gras.

	Niveaux de pression de l'humidité relative (hPa)			
	925	850	775	700
Doubs	0,30	<i>0,00</i>	0,80	0,30
Petite Saône	1,15	<i>0,00</i>	1,00	1,51
Saône Intermédiaire	0,66	<i>0,00</i>	0,47	1,12
Ensemble du bassin	0,71	<i>0,00</i>	0,76	0,98

Les résultats sont plutôt surprenants. En effet, le gain G est systématiquement positif si l'humidité relative est mesurée à un autre niveau de pression que le niveau de référence (850 hPa). Les meilleures performances sont obtenues lorsque l'humidité est mesurée à 700 hPa, sauf pour le bassin du Doubs pour lequel le niveau optimal se situe à 775 hPa. Ainsi pour l'ensemble des sous-bassins, le gain G par rapport à la méthode A2 s'élève à près de 1% lorsque la variable RH est mesurée à 700 hPa.

L'origine des précipitations étant diverse, le niveau de pression optimal auquel l'humidité doit être mesurée peut être théoriquement, comme nous l'avons déjà indiqué en début de section, dépendant du type de situation cible. Pour les précipitations convectives par exemple, l'humidité nécessaire au développement des nuages convectifs se situe dans les basses ou très basses couches de la troposphère, soit vers 950 à 850 hPa. Les précipitations stratiformes, d'origine frontale, sont généralement issues de nuages formés sur une grande partie de la couche troposphérique, et notamment dans les couches moyennes. Nous allons donc vérifier si les performances de prévision peuvent encore être améliorées si le niveau de pression auquel la variable RH est mesurée dépend du type de temps d'une situation cible.

8.3.2.2 Un niveau de pression par type de temps

Dans la littérature, il existe de nombreuses classifications par type de temps (cf. résultats du projet européen ACCORD¹). Nous avons opté pour la classification par types de temps construite par EDF (Paquet et al., 2006), qui a été spécifiquement adaptée au cadre français métropolitain pour l'application de la méthode SCHADEX. Huit classes ont été retenues à partir du champ de pression au niveau de la mer (Figure 8.9) : Onde Atlantique (OA, classe 1), Océanique Stationnaire (OS, classe 2), Sud-Ouest (SO, classe 3), Sud (SU, classe 4), Nord-Est (NE, classe 5), Retours d'Est (RE, classe 6), Dépression Centrale (DC, classe 7) et Anticyclonique (AN, classe 8).

La période sur laquelle la classification SCHADEX a été établie couvre 50 ans (1953-2002). Or notre période de calibration couvre seulement 25 années dans la période 1972-2002. Le Tableau 8.4 indique les proportions de situations cibles de notre période appartenant à chaque classe, ainsi que celles obtenues sur la période de 50 ans. On constate que la probabilité d'occurrence de chacun des types de temps est très similaire sur les deux périodes.

1 Site web du projet ACCORD (Atmospheric Circulation Classification and Regional Downscaling) : <http://www.cru.uea.ac.uk/cru/projects/accord>

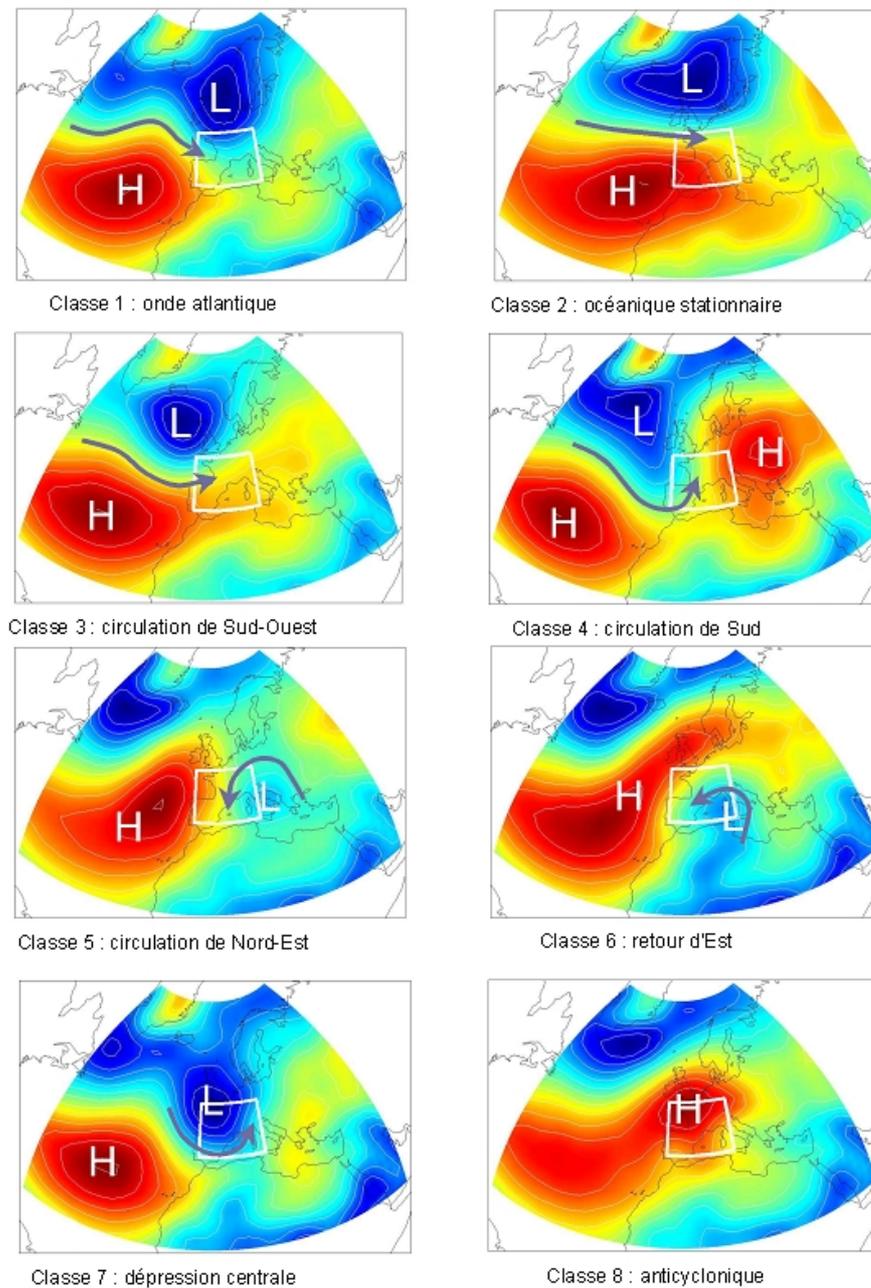


Figure 8.9 : Champs de pression moyens associés aux huit types de temps de la méthode SCHADEX (d'après Paquet et al., 2006).

Tableau 8.4 : Proportions de situations cibles appartenant à chacune des huit classes SCHADEX, obtenues sur la période de calibration et sur la période 1953-2002.

Classe	1 (OA)	2 (OS)	3 (SO)	4 (SU)	5 (NE)	6 (RE)	7 (DC)	8 (AN)
Période de calibration	7,6	23,8	7,8	17,9	6,2	6,6	3,4	26,7
Période 1953-2002	7,6	23,5	7,8	17,7	6,6	5,9	3,4	27,6

Dans un premier temps, notre objectif est de connaître le niveau optimal de pression de l'humidité relative associé à chaque type de temps et de vérifier si le niveau est dépendant de la classe SCHADEX. Pour chaque situation cible de notre période de calibration, la

performance moyenne (*CRPS*) sur le bassin de la Saône a été évaluée en considérant chacun des quatre niveaux de pression pour la mesure de l'humidité relative. La proportion de situations cibles pour lesquelles le score *CRPS* obtenu est le plus faible a ensuite été calculée pour un niveau de pression donné et pour chacune des huit classes (Figure 8.10).

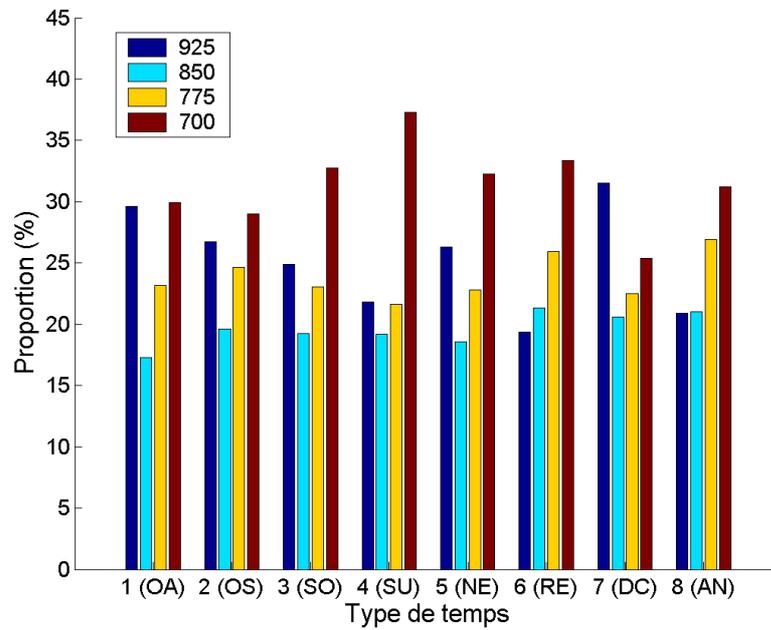


Figure 8.10 : Proportions de situations cibles pour lesquelles chaque niveau de pression est optimal, obtenues pour chacune des huit classes SCHADEX.

Il est préférable de relativiser les résultats pour les classes 1, 3, 5, 6 et 7, pour lesquels les proportions de situations cibles sont particulièrement faibles (moins de 732 journées parmi les 9161 que couvre notre période). La Figure 8.10 montre d'une part qu'il n'y a pas de réelle discrimination du niveau de pression suivant le type de temps d'une situation cible. En effet, la plus faible proportion obtenue est tout de même de 17 % pour la classe 2 (*RH* mesurée à 850 hPa). D'autre part, les niveaux de pression 775 hPa et 850 hPa sont généralement moins privilégiés que les deux autres, sauf pour les types 6 (retour d'est) et 8 (anticyclonique). Compte tenu de ce résultat et du précédent (Tableau 8.3), nous pouvons d'ores et déjà exclure l'utilisation de l'humidité relative au niveau de pression 850 hPa, mais aussi au niveau 775 hPa, qui n'est par ailleurs pas un niveau standard en météorologie. Le niveau 700 hPa est visiblement préféré par la majorité des situations cibles de chaque classe, à l'exception de la classe 7 (dépression centrale), pour laquelle l'humidité est préférentiellement mesurée à 925 hPa. L'inconvénient pour cette classe est le faible échantillonnage (3,4 % des situations cibles). Par ailleurs pour les types des classes 1 (onde atlantique) et 2 (océanique stationnaire), il n'y a pas vraiment de préférence pour l'un des niveaux 925 et 700 hPa.

En conclusion, au vu de ces résultats valables pour le bassin de la Saône, aucun niveau de pression n'est privilégié pour un type de temps donné. Bien que nous puissions mesurer l'humidité relative à l'unique niveau de pression 700 hPa pour l'ensemble des situations (gain

de 0,98 % par rapport à A2), nous avons tout de même appliqué la sélection d'analogues en mesurant l'humidité relative au niveau de pression optimal associé au type de chaque situation cible, à savoir le niveau 925 hPa pour les journées du type 7 (dépression centrale), et le niveau 700 hPa pour les journées restantes. Le gain G obtenu par rapport à A2 (1,12 %) est à peine supérieur au gain évalué lorsque l'humidité est mesurée à 700 hPa pour l'ensemble des situations.

8.3.2.3 Deux niveaux de pression fixes

La mesure de l'humidité à un niveau de pression dépendant du type de temps semble légèrement améliorer les performances de prévision. Cependant, nous avons remarqué que les deux niveaux de pression 925 et 700 hPa sont souvent privilégiés au détriment des deux autres niveaux 850 et 775 hPa. Nous avons ainsi envisagé de mesurer l'humidité relative aux deux niveaux de pression 925 et 700 hPa. Pour cela, nous avons simplement calculé les critères d'analogie de ce troisième niveau de sélection ($RMSE$) associés aux variables $TCW700$ et $TCW925$ définies par :

$$\begin{aligned} TCW925 &= TCW \times RH925 \\ TCW700 &= TCW \times RH700 \end{aligned} \tag{8.1}$$

De cette manière, nous renommons implicitement la variable d'analogie $TCW\%$, utilisée au dernier niveau de sélection, par $TCW850$, afin de définir deux nouvelles variables élaborées : $TCW925$ et $TCW700$.

Pour chacune de ces variables, la valeur du critère d'analogie $RMSE$ est calculée et les deux valeurs correspondantes sont simplement sommées pour obtenir une valeur totale du critère $RMSE$. Le domaine d'analogie a alors été optimisé : il reste finalement identique à celui appliqué jusqu'à présent. Le gain G de performance moyenne, obtenu par rapport à la méthode A2, a ensuite été évalué. Le Tableau 8.5 indique les résultats obtenus dans cette section lorsque l'humidité est mesurée : au niveau de pression 925 hPa seul, au niveau de pression 700 hPa seul, au niveau de pression 925 hPa ou au niveau 700 hPa suivant le type de temps, et aux deux niveaux de pression 925 hPa et 700 hPa.

La mesure de l'humidité relative aux niveaux 925 et 700 hPa conduit au gain de performance le plus élevé pour chaque sous-bassin (2,19 % par rapport à la méthode A2 pour l'ensemble du bassin). Ce résultat nous a incités à tester la mesure de l'humidité relative à trois niveaux différents, en fixant deux des trois niveaux à 925 hPa et 700 hPa. Les deux combinaisons possibles (850 hPa ou 775 hPa comme troisième niveau) n'ont cependant pas mené à un gain de performance plus élevé que celui obtenu avec les deux niveaux.

Tableau 8.5 : Gain G (%) de performance moyenne, obtenu sur le bassin de la Saône par rapport à la méthode A2, lorsque l'humidité relative est mesurée à un seul niveau (925 ou 700 hPa), au niveau 925 ou 700 hPa en fonction du type de temps, et aux deux niveaux 925 hPa et 700 hPa. Les valeurs minimales et maximales figurent respectivement en italique et en gras.

	Niveaux de pression de l'humidité relative			
	925	700	Selon le type de temps	925 et 700
Doubs	<i>0,30</i>	<i>0,30</i>	0,42	1,66
Petite Saône	<i>1,15</i>	1,51	1,68	2,62
Saône Intermédiaire	<i>0,66</i>	1,12	1,23	2,27
Ensemble du bassin	<i>0,71</i>	0,98	1,12	2,19

La mesure de la variable RH à 925 hPa et à 700 hPa conduit finalement à des gains significatifs, sans toutefois surpasser ceux obtenus grâce aux développements précédents. Nous allons à présent faire une synthèse des optimisations effectuées dans ce chapitre, retenir celles qui mènent au meilleur gain de performance de prévision, et enfin valider les résultats.

8.4 Synthèse, validation et diagnostic

L'objectif de ce chapitre fut d'optimiser certains paramètres des niveaux d'analogie de la méthode de référence A0 en appliquant la méthode développée A2. Aucun niveau d'analogie supplémentaire ni aucune nouvelle variable n'a donc été introduit. L'ensemble des résultats d'optimisation est synthétisé dans le Tableau 8.6. Les gains les plus élevés sont obtenus par la mesure de la variable RH aux niveaux de pression 925 et 700 hPa. Nous allons donc conserver ces nouveaux paramètres et ne tiendrons finalement pas compte des autres résultats d'optimisation.

Tableau 8.6 : Gains G de performance obtenus sur le bassin de la Saône par rapport à la méthode A2, par les différentes optimisations effectuées.

	Optimisation effectuée			
	Domaine du niveau 1 flexible	Heures de mesure au niveau 1	Heures de mesure au niveau 3	RH aux niveaux 925 et 700 hPa
Doubs	0,85	0,53	0,5	1,66
Petite Saône	0,94	0,6	0,65	2,62
Saône Intermédiaire	0,48	0,93	0,74	2,27
Ensemble du bassin	0,75	0,68	0,61	2,19

Nous avons volontairement choisi de ne pas nommer A3 cette nouvelle méthode. Nous la nommerons A2bis par la suite. Ceci permet de nous rappeler que seul le niveau de pression de mesure de la variable d'humidité élaborée a été modifié par rapport à la méthode A2.

De la même manière que pour les méthodes déjà développées, l'algorithme A2bis va être validé, d'une part en effectuant les mêmes optimisations sur le bassin de la Seine et en vérifiant que le gain de performance obtenu est également positif, et d'autre part en

s'assurant que le gain positif est conservé sur la période de validation. Nous établirons ensuite un diagnostic des résultats.

8.4.1 Application au bassin de la Seine

Les principaux résultats obtenus pour le bassin de la Seine vont être présentés ici, lorsque l'humidité est mesurée à un seul niveau de pression puis à deux niveaux de pression, en reprenant également l'approche par type de temps.

Un seul niveau de pression

Les quatre niveaux de pression auxquels peut être mesurée l'humidité relative ont chacun été considérés. Le domaine d'analogie appliqué reste le même que celui appliqué jusqu'à présent au dernier niveau de sélection (2,5°E/5°E ; 47,5°N/50°N). Le gain G de performance moyenne, obtenue par rapport à la méthode A2 en considérant chaque niveau de pression, est reporté dans le Tableau 8.7. Le gain a été évalué pour chaque sous-bassin et également pour l'ensemble du bassin de la Seine.

Tableau 8.7 : Gain G de performance moyenne, obtenu sur le bassin de la Seine par rapport à la méthode A2, pour les différents niveaux de pression auxquels l'humidité relative peut être mesurée. Les valeurs minimales et maximales figurent respectivement en italique et en gras.

	Niveaux de pression de l'humidité relative			
	925	850	775	700
Yonne-C	1,01	<i>0</i>	0,56	0,79
Seine-Int	<i>-0,62</i>	0	1,36	1,73
Seine-B	0,73	<i>0</i>	0,80	1,06
Marne-F	0,46	<i>0</i>	0,72	0,82
Loing-E	0,28	<i>0</i>	0,64	1,20
Ensemble des bassins	0,36	<i>0</i>	0,82	1,13

Nous obtenons les mêmes résultats que sur le bassin de la Saône. Le niveau privilégié est le niveau 700 hPa, sauf pour le bassin de l'Yonne-C pour lequel la mesure de RH au niveau 925 hPa conduit au gain G maximal. Les performances évaluées lorsque l'humidité est mesurée au niveau 850 hPa sont les plus faibles, à l'exception du bassin de la Seine-Int (gain G négatif obtenu pour le niveau 925 hPa).

Deux niveaux de pression

Nous avons également appliqué le niveau de pression adapté à chacune des classes SCHADEX. A l'image des résultats obtenus pour le bassin de la Saône, le niveau de pression privilégié est le niveau 700 hPa pour tous les types de temps à l'exception du type de la classe 4 (Sud). Les performances ont été évaluées en appliquant le niveau de pression optimal associé à chaque classe SCHADEX. Enfin, l'humidité a également été mesurée à deux niveaux de pression distincts. Lorsqu'elle est mesurée à un seul niveau de pression, nous avons vu que

les deux niveaux les plus informatifs sont 700 hPa et 775 hPa. Il ne serait cependant pas intéressant de mesurer l'humidité à ces deux niveaux puisqu'ils sont très proches. L'information que chacun apporterait serait alors très similaire. Nous avons donc décidé de mesurer la variable *RH* aux mêmes niveaux que pour le bassin de la Saône (925 hPa et 700 hPa). Les gains obtenus par rapport à la méthode A2 dans chacun des cas sont reportés dans le Tableau 8.8.

Comme pour le bassin de la Saône, la mesure de l'humidité relative aux niveaux 925 hPa et 700 hPa conduit au gain de performance le plus élevé pour chaque sous-bassin (2,19% par rapport à la méthode A2 pour l'ensemble du bassin de la Seine).

Tableau 8.8 : Gain *G* de performance moyenne, obtenu sur le bassin de la Seine par rapport à la méthode A2, lorsque l'humidité relative est mesurée au niveau 925 hPa seul, au niveau 700 hPa seul, au niveau 925 hPa ou 700 hPa en fonction du TDT, aux deux niveaux 925 hPa et 700 hPa. Les valeurs minimales et maximales figurent respectivement en italique et en gras.

	Niveaux de pression de l'humidité relative			
	925	700	Selon TDT	925 et 700
Yonne-C	1,01	<i>0,79</i>	1,27	2,67
Seine-Int	<i>-0,62</i>	1,73	1,14	2,51
Seine-B	<i>0,73</i>	1,06	1,25	2,96
Marne-F	<i>0,46</i>	0,82	1,09	2,65
Loing-E	<i>0,28</i>	1,20	1,06	2,44
Ensemble des bassins	<i>0,36</i>	1,13	1,16	2,64

En conclusion, les performances obtenues par la méthode A2 sont améliorées lorsque l'humidité est mesurée à deux niveaux de pression.

8.4.2 Validation sur les deux bassins tests

Nous proposons à présent de valider la méthode A2bis sur les bassins de la Saône et de la Seine. Le gain *G*, obtenu par rapport à l'algorithme A2, a été évalué sur la période de validation. Les gains obtenus sur chaque sous-bassin sont reportés dans le Tableau 8.9. A titre de comparaison, les gains *G* obtenus en calibration y figurent également.

Tableau 8.9 : Gains *G* de performance, obtenus par la méthode A2bis par rapport à la méthode A2 sur les sous-bassins de la Saône et de la Seine, évalués sur les deux périodes cibles indépendantes de validation et de calibration (en italique).

Période	Bassin versant et gain <i>G</i> (A2bis/A2) (%)							
	Doubs	Petite S.	Saône I.	Yonne-C	Seine-Int	Seine-B	Marne-F	Loing-E
<i>Calibration</i>	<i>1,7</i>	<i>2,62</i>	<i>2,19</i>	<i>2,72</i>	<i>2,46</i>	<i>3,02</i>	<i>2,75</i>	<i>2,42</i>
Validation	1,64	3,11	2,21	2,22	2,91	1,81	2,2	2,49

Les résultats obtenus sur la période de validation sont rassurants. Le gain est bien conservé sur tous les sous-bassins étudiés. On remarque à nouveau que pour le bassin de la Seine, les gains obtenus en validation sont plus faibles que ceux obtenus en calibration.

8.4.3 Diagnostic des résultats

De la même manière que lors des développements précédents, les résultats ont été étudiés plus en détail, en s'intéressant au gain en finesse et justesse, puis à la robustesse du gain de performance apporté par la méthode A2bis par rapport à la méthode A2.

8.4.3.1 Gain en finesse et en justesse

Le score *CRPS* a été décomposé en finesse et justesse de prévision et le gain de performance respectif, par rapport à la méthode A2, a été évalué pour les deux bassins tests (Tableau 8.10 et Tableau 8.11).

Tableau 8.10 : Gains de performance, obtenus par la méthode A2bis par rapport à la méthode A2, sur le bassin de la Saône, par décomposition du gain global en finesse et en justesse.

	Gain global	Gain en finesse	Gain en justesse
Doubs	1,7	2,05	1,36
Petite Saône	2,62	3,36	1,97
Saône Intermédiaire	2,19	3,38	1,17
<i>Ensemble des bassins</i>	<i>2,17</i>	<i>2,93</i>	<i>1,5</i>

Les gains obtenus sont positifs sur l'ensemble des sous-bassins, ce qui traduit la pertinence de la mesure de l'humidité à deux niveaux de pression. La finesse et la justesse des distributions sont toutes deux améliorées, avec quelques nuances suivant les bassins étudiés. Pour le bassin de la Saône, le gain en finesse est plus élevé que le gain en justesse, tandis que c'est l'inverse qui est observé pour le bassin de la Seine. Néanmoins, l'essentiel est que la mesure de l'humidité aux deux niveaux ne dégrade ni la finesse ni la justesse des prévisions.

Tableau 8.11 : Gains de performance, obtenus par la méthode A2bis par rapport à la méthode A2, sur le bassin de la Seine, par décomposition du gain global en finesse et en justesse.

	Gain global	Gain en finesse	Gain en justesse
Yonne-C	2,72	2,25	3,09
Seine-Int	2,46	2,84	2,18
Seine-B	3,02	2,16	3,7
Marne-F	2,75	2,15	3,23
Loing-E	2,42	1,68	2,95
<i>Ensemble des bassins</i>	<i>2,66</i>	<i>2,22</i>	<i>3,02</i>

8.4.3.2 Robustesse du gain de performance

Nous procédons toujours de la même manière pour évaluer la robustesse du gain de performance. Le score *CRPS* a été calculé pour chaque situation cible de la période de calibration en appliquant les algorithmes A2 et A2bis. La distribution empirique des écarts de score *CRPS* entre les deux algorithmes a ensuite été construite. Nous rappelons que si l'écart

est positif, alors la performance de prévision est améliorée par la méthode A2bis.

Les résultats sont identiques à ceux obtenus pour les méthodes A1 et A2. La positivité du gain de performance n'est pas robuste, à l'image de la distribution des écarts de score *CRPS* représentée sur la Figure 8.11 pour le bassin du Doubs. Il y a presque autant de situations pour lesquelles la prévision est dégradée que de situations pour lesquelles elle est améliorée.

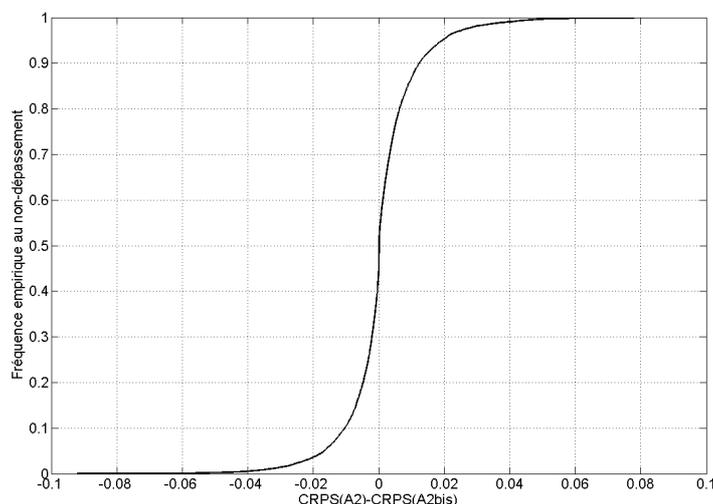


Figure 8.11 : Distribution empirique des écarts de *CRPS* entre les méthodes A2 et A2bis calculés sur la période de calibration, pour le bassin du Doubs.

8.4.4 Application à la méthode A1

Dans la méthode A1 que nous avons développée dans le chapitre 5, l'humidité est mesurée uniquement au niveau 850 hPa. Il nous paraît légitime de vérifier si la mesure de la variable *RH* à deux niveaux différents peut aussi améliorer les performances de prévision obtenues par la méthode A1. Dans le cadre de cette vérification, aucune optimisation n'a été effectuée. Les variables *TCW925* et *TCW700* ont simplement été utilisées à la place de la variable *TCW850*, puis les performances ont été comparées. Le Tableau 8.12 indique les gains de performance obtenus par la méthode A1bis par rapport à la méthode A1 sur les périodes de calibration et de validation, pour chacun des sous-bassins de la Saône et de la Seine.

Tableau 8.12 : Gain *G* de performance obtenu par la méthode A1bis par rapport à la méthode A1, évalué sur chaque sous-bassin de la Saône et de la Seine, sur les périodes de calibration et de validation.

Période	Bassin versant et gain <i>G</i> (A1bis/A1) (%)							
	Doubs	Petite S.	Saône I.	Yonne-C	Seine-Int	Seine-B	Marne-F	Loing-E
<i>Calibration</i>	2,13	3	2,37	2,68	3,11	2,81	2,64	2,45
<i>Validation</i>	2,53	3,42	3,2	1,81	2	1,89	2,16	1,57

La mesure de l'humidité à deux niveaux de pression permet également d'augmenter les performances de la méthode A1. Les gains obtenus, que ce soit sur la période de calibration ou sur la période de validation, sont du même ordre de grandeur que ceux obtenus

précédemment entre les méthodes A2bis et A2. Nous nommerons A1bis la méthode A1 ainsi modifiée.

8.5 Conclusion

Une approche d'optimisation a été empruntée dans ce chapitre, appliquée aux procédures d'identification des analogues selon les variables de pression et d'humidité de la méthode de référence A0. Pour ce faire, la méthode A2, constituée d'un niveau pré-sélectif suivi de trois niveaux de sélection (niveaux 1, 2 et 3), a été appliquée.

Nous avons essayé d'appliquer, au niveau 1, un domaine d'analogie dépendant du flux d'altitude. La direction du vent à 500 hPa a été divisée en huit classes ; et pour chacune des classes, un domaine a été optimisé. En dépit des résultats encourageants obtenus par cartographie de la pertinence d'analogie par classe, l'intérêt d'appliquer un domaine spécifique à une classe de flux, dont la taille et l'étendue dépendent de la direction du flux d'altitude, apparaît bien faible. Son application peut même s'avérer dangereuse. En effet, l'utilisation d'un domaine spécifique relève de la sur-optimisation de l'algorithme, si bien que les performances risquent d'être fortement dégradées pour certaines situations.

D'autre part dans le cadre de cette thèse, nous disposons, pour les variables *Z500*, *Z1000* et *TCW%* exploitées dans la méthode A0, d'heures de mesure supplémentaires (6 h TU et 18 h TU) par rapport à celles dont Bontron disposait en 2004 (00 h TU et 12 h TU). Cette opportunité a été saisie en identifiant les instants de mesure optimaux pour ces variables. La recherche des analogies aurait pu être améliorée, notamment au niveau 1, mais les gains de performance apportés ne se sont pas révélés suffisants. Les analogies sur les trois variables *Z500*, *Z1000* et *TCW%* seront donc recherchées aux mêmes instants que précédemment.

Enfin, nous avons exploité d'autres niveaux de pression auxquels peuvent être mesurées les variables brutes *Z* et *RH* utilisées dans la méthode A0, d'heures de mesure et de niveaux de pression. Or les niveaux supplémentaires par rapport à ceux de l'archive LTHE utilisée par Bontron (2004) sont peu nombreux pour le géopotential *Z*, tandis que dans les basses couches de la troposphère, l'archive ERA contient, en plus du niveau 850 hPa initialement disponible dans l'archive LTHE, les niveaux 925, 775 et 700 hPa. Nous avons donc envisagé la mesure de la variable *RH* à un niveau puis à deux niveaux de pression, et une approche par type de temps a également été suggérée. Finalement, la mesure simultanée de l'humidité relative aux deux niveaux de pression 925 hPa et 700 hPa permet d'obtenir un gain substantiel de performance. Nous avons nommé *TCW925* et *TCW700* les variables d'analogie correspondantes. Bien que l'approche par type de temps conduise à un gain légèrement inférieur, celle-ci n'est pas à écarter pour autant si on désire adapter la méthode à un autre bassin.

Les mêmes essais ont ensuite été effectués sur le bassin de la Seine, et les résultats obtenus ont permis de valider cette variante de la méthode A2, nommée A2bis pour la suite de nos travaux. Le gain de performance a également été évalué sur la période de validation pour

tous les sous-bassins étudiés : il est finalement conservé, ce qui justifie l'optimisation entreprise. Le diagnostic des performances établi ensuite a montré qu'en moyenne, finesse et justesse de la prévision sont améliorées par rapport à la méthode A2. En revanche, le gain de performance n'est pas systématique. Cette conclusion est d'ailleurs la même que celles tirées lors des précédents développements. Ainsi en introduisant de nouvelles variables dans un algorithme de recherche d'analogues, il semble que même si en moyenne le gain de performance est conséquent, sa robustesse n'en est pas moins assurée.

Enfin, puisque cette amélioration est *a priori* indépendante de l'algorithme A2¹, nous avons également mesuré la variable d'humidité aux deux niveaux de pression dans la méthode A1. Les gains obtenus s'avèrent du même ordre de grandeur que ceux obtenus avec la méthode A2. Ainsi pour la suite, en plus des méthodes A1, A2 et A2bis, nous considérerons aussi la méthode A1bis, qui correspond à la méthode A1 dans laquelle la variable *RH* est mesurée aux niveaux 925 hPa et 700 hPa.

¹ La variable d'humidité n' a effectivement pas été introduite dans la méthode A2, il s'agit d'une variable déjà utilisée dans la méthode de référence A0.

Troisième partie

De la prévision des pluies à la prévision des débits

Introduction à la partie III

Les quatre nouveaux algorithmes de sélection d'analogues (A1, A1bis, A2 et A2bis), ont leurs caractéristiques désormais bien établies (tableau de synthèse en page suivante, dont les valeurs des paramètres sont valables pour le bassin de la Saône). Ils ont fait l'objet d'une évaluation préliminaire des performances dans la partie précédente. Nous proposons d'approfondir l'évaluation par le biais d'autres scores, aux vertus « plus opérationnelles », présentés dans le chapitre 3.

Cette partie vise également à utiliser les méthodes en entrée d'un modèle de transformation pluie-débit et à analyser le comportement de la chaîne de prévision ainsi constituée :

- lorsque les précipitations sont prévues dans un contexte de prévision parfaite ;
- lorsque cette chaîne est alimentée en scénarios de prévision opérationnelle des précipitations, *i.e.* lorsque l'analogie est recherchée sur des champs météorologiques non plus analysés ou observés mais prévus. Ceci permettra de mettre en évidence la dégradation de la prévision des variables d'analogie et d'évaluer son impact d'une part sur les performances de la prévision des précipitations adaptée par analogie et d'autre part sur celles de la prévision des débits ;

Le premier chapitre présentera les outils nécessaires à la mise en place de la chaîne de prévision hydrométéorologique, et la méthodologie employée pour la génération de prévisions probabilistes de débit. Le chapitre suivant s'intéressera à l'évaluation des performances de prévisions de pluie, et de prévision de débit lorsque la chaîne est alimentée en scénarios de précipitations obtenus dans un contexte de prévision parfaite. Enfin le dernier chapitre permettra de nous placer en conditions réelles, les champs des variables d'analogie étant alors prévus par un modèle numérique. Les prévisions de pluie et de débit seront analysées et évaluées, et nous tenterons d'identifier la part des différentes sources d'erreur dans l'erreur finalement obtenue sur la prévision de débit en bout de chaîne. Des cas de crue seront également analysés en fin de chapitre, afin d'étudier notamment la possibilité, en conditions réelles, d'anticiper une crue sur les bassins tests de la Saône et de la Seine jusqu'à une semaine à l'avance.

		Algorithmes de sélection				
		A0	A1	A1bis	A2	A2bis
Pré-sélection des situations candidates (Niveau 0)	Variable		Température	Température	Température	Température
	Niveaux et heures de mesure	Critère calendaire figé (fenêtre temporelle glissante de quatre mois centrée sur le jour cible)	925 hPa à +36h 600 hPa à +12h			
	Domaine spatial		Un seul point de grille (5°E ; 47,5°N)			
	Critère		<i>RMSE</i>	<i>RMSE</i>	<i>RMSE</i>	<i>RMSE</i>
	Nombre d'analogues		<i>N0=2900</i>	<i>N0=2900</i>	<i>N0=2900</i>	<i>N0=2900</i>
Variable			Géopotential			
Niveau 1	Niveaux et heures de mesure		1000 hPa à +12h 500 hPa à +24h			
	Domaine spatial		Echelle synoptique (-5°E/15°E ; 40°N/50°N)			
	Critère		<i>TWS</i>			
	Nombre d'analogues	<i>N1=70</i>	<i>N1=70</i>	<i>N1=70</i>	<i>N1=170</i>	<i>N1=170</i>
Niveau 2	Variable	Eau précipitable et humidité relative	Eau précipitable et humidité relative	Eau précipitable et humidité relative	Vitesse verticale	Vitesse verticale
	Niveaux et heures de mesure	850 hPa à +12h et +24h	850 hPa à +12h et +24h	925 hPa et 700 hPa à +12h et +24h	850 hPa à +6h, +12h, +18h et +24h	850 hPa à +6h, +12h, +18h et +24h
	Domaine spatial	Echelle très locale (5°E/7,5°E ; 45°N/47,5°N)	Echelle très locale (5°E/7,5°E ; 45°N/47,5°N)	Echelle très locale (5°E/7,5°E ; 45°N/47,5°N)	Echelle locale (2,5°E/7,5°E ; 45°N/47,5°N)	Echelle locale (2,5°E/7,5°E ; 45°N/47,5°N)
	Critère	<i>RMSE</i>	<i>RMSE</i>	<i>RMSE</i>	<i>RMSE</i>	<i>RMSE</i>
	Nombre d'analogues	<i>N2=25</i>	<i>N2=25</i>	<i>N2=25</i>	<i>N2=70</i>	<i>N2=70</i>
Niveau 3	Variable	-	-	-	Eau précipitable et humidité relative	Eau précipitable et humidité relative
	Niveaux et heures de mesure	-	-	-	850 hPa à +12h et +24h	925 hPa et 700 hPa à +12h et +24h
	Domaine spatial	-	-	-	Echelle très locale (5°E/7,5°E ; 45°N/47,5°N)	Echelle très locale (5°E/7,5°E ; 45°N/47,5°N)
	Critère	-	-	-	<i>RMSE</i>	<i>RMSE</i>
	Nombre d'analogues	-	-	-	<i>N3=25</i>	<i>N3=25</i>

9 Outils et méthodologie

Ce chapitre se propose de décrire les outils nécessaires pour émettre une prévision hydrologique probabiliste à partir d'une prévision de pluie adaptée par analogie.

Dans la première section, nous présenterons le modèle hydrologique mis en œuvre pour la transformation de la pluie en débit et décrirons la procédure de calage du modèle. Afin d'évaluer la capacité du modèle à bien simuler le débit, nous étudierons ses performances en conditions parfaitement connues.

La méthodologie choisie pour la génération des prévisions probabilistes de débit sera ensuite présentée dans la seconde section. Nous définirons de cette manière la chaîne de prévision, constituée d'une méthode de prévision de pluie par analogie couplée au modèle hydrologique.

Enfin, nous décrirons dans la troisième section deux méthodes de prévision naïves, qui serviront de référence pour juger la plus value d'une prévision issue de la chaîne.

9.1 Modèle hydrologique

9.1.1 Description

Pour générer une prévision de débit, il existe plusieurs types de modèles hydrologiques, dont l'utilisation dépend des attentes du modélisateur, des données disponibles et du bassin versant. Nous avons retenu le modèle GR3P, qui a été appliqué avec succès à la Seine à Paris (Perrin et al., 2005). Il s'agit d'un modèle global à réservoirs de transformation pluie-débit de la famille des GR développés par le Cemagref (Perrin et al., 2001, 2003 ; Mathevet, 2005). GR3P (Génie Rural à 3 paramètres adapté à la Prévision) proposé par Tangara (2005) a été spécifiquement conçu pour la prévision des crues.

Ce modèle admet deux composantes, que l'on retrouve classiquement en modélisation hydrologique (Figure 9.1) : un module de production et un module de routage. Le module de production détermine un bilan quantitatif en eau, c'est-à-dire qu'il transforme la pluie brute P en pluie nette P_n en tenant compte de l'évapotranspiration E . Une partie de celle-ci, P_s ,

s'infiltré dans le réservoir de production, puis en ressort ultérieurement par percolation. L'eau percolée durant le pas de temps s'ajoute à la quantité $P_n - P_s$. La résultante gagne le réseau hydrographique après transformation via un hydrogramme unitaire HU injecté dans le réservoir de routage. Le module de production assure le suivi de l'humidité du sol et la mémoire des conditions antérieures. Le module de routage assure quant à lui la répartition temporelle et la quantité d'eau transitant dans le cours d'eau.

GR3P est adapté à la prévision, c'est-à-dire qu'il introduit une correction adaptative du niveau du dernier réservoir pour mieux représenter les conditions de saturation antérieures. Le modèle utilisé dans cette thèse fonctionne au pas de temps journalier. Dans la version horaire de GR3P (thèse de L. Berthet, soutenance prévue en février 2010), il existe une seconde étape de la mise à jour (report d'erreur) (Berthet et al., 2009).

Ainsi dans notre cas, avant émission des prévisions, la procédure d'assimilation modifie le niveau du réservoir R en fonction du dernier débit observé avant la prévision, au travers d'une relation analytique. Elle impose strictement l'égalité entre le débit observé et le débit simulé en sortie de réservoir, mais n'est active qu'au moment de la date d'émission de la prévision.

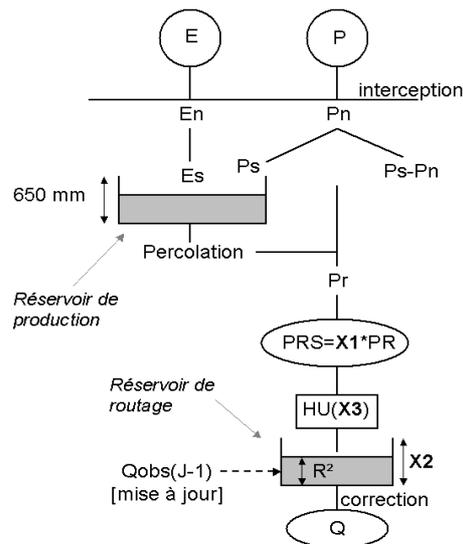


Figure 9.1 : Schéma structurel du modèle hydrologique de prévision GR3P (d'après Tangara, 2005).

Pour émettre une prévision, le modèle dans sa version actuelle a besoin :

- des chroniques des pluies de bassin, des débits, des évapotranspirations potentielles (ETP), avec *a minima* une année d'observations avant la date de prévision ;
- des scénarios de pluies prévues sur l'ensemble des échéances de prévision ;
- des données de débits des jours précédant la prévision pour l'assimilation ;
- des paramètres $X1$, $X2$, $X3$ du modèle, préalablement fixés par la procédure de calage du modèle. Ces derniers figurent en gras sur la Figure 9.1.

Le modèle GR3P peut également fonctionner en mode simulation, sans mise à jour des niveaux des réservoirs.

Il est important de noter que GR3P est un modèle hydrologique global. Il ne gère pas les disparités de fonctionnement. En particulier, il peut être en défaut lorsque un événement pluvieux est fortement localisé. Pour remédier à ce problème, il est possible de combiner des modèles globaux, calés sur des bassins de tête, à des outils de propagation hydraulique, mais cette approche n'a pas été exploitée dans cette thèse.

9.1.2 Procédure de calage du modèle

Pour fonctionner, le modèle doit être préalablement calé sur chacun des bassins d'intérêt. Les paramètres sont optimisés de telle sorte que sur la période de calage, les débits issus de GR3P en simulation se rapprochent le plus de débits observés. La procédure de calage est automatique. Le jeu des paramètres ($X1, X2, X3$) est identifié en parcourant l'espace R^3 et en cherchant à minimiser la valeur du score $RMSE$ obtenue pour différentes échéances E :

$$RMSE(E) = \sqrt{\sum_{J=1}^{nT} (Q_{prévu}(J+E) - Q_{observé}(J+E))^2} \quad (9.1)$$

où J désigne l'instant du début de la J ème prévision et nT le nombre de jours de prévision. Ce score est calculé sur l'ensemble des débits de la chronique disponible. Il faut noter que le module de mise à jour est inactif au moment de la mise en route du modèle, dans la première année de calage. Finalement, l'utilisateur choisira le jeu de paramètres optimal pour l'échéance souhaitée. Ainsi, un jeu de trois paramètres est optimisé numériquement pour chaque échéance de prévision, de J à $J+13$. Par convention, nous noterons J le premier jour de prévision, la dernière observation datant de $J-1$.

Les trois paramètres du modèle ont été calés, sur chacun des sous-bassins de la Saône et de la Seine, sur la plus longue période possible. Celle-ci correspond à la période commune sur laquelle les données de pluie (SAFRAN) et de débit (HYDRO) sont disponibles, c'est-à-dire du 01/08/1970 au 30/07/2008.

Le calage se fait en deux temps et reprend partiellement le principe de calage-contrôle de Klemes (1986). La première étape consiste à séparer la période totale en deux sous-périodes de même longueur : une période de calage et une période de contrôle. Dans notre cas, les paramètres du modèle sont alors optimisés sur la période de calage 01/08/1970-31/07/1990, puis le modèle est évalué sur la période de contrôle 01/08/1990-30/07/2008. Nous vérifions alors, par exemple à l'aide des valeurs de $RMSE$, que les erreurs de prévision du modèle obtenues pour chaque échéance sont du même ordre de grandeur sur les deux périodes. A titre d'illustration, l'évolution du $RMSE$ (en m^3/s) obtenu sur les deux sous-périodes est tracée sur la Figure 9.2 pour la Saône à Le Chatelet, en fonction de l'échéance (de J à $J+6$).

Les valeurs de $RMSE$ sont assez proches, ce qui traduit une certaine robustesse des paramètres optimisés. Elles sont néanmoins plus faibles sur la période de contrôle, ce qui pourrait provenir de la différence de climatologie des débits. Cette phase d'optimisation sert à vérifier si les paramètres obtenus sont robustes.

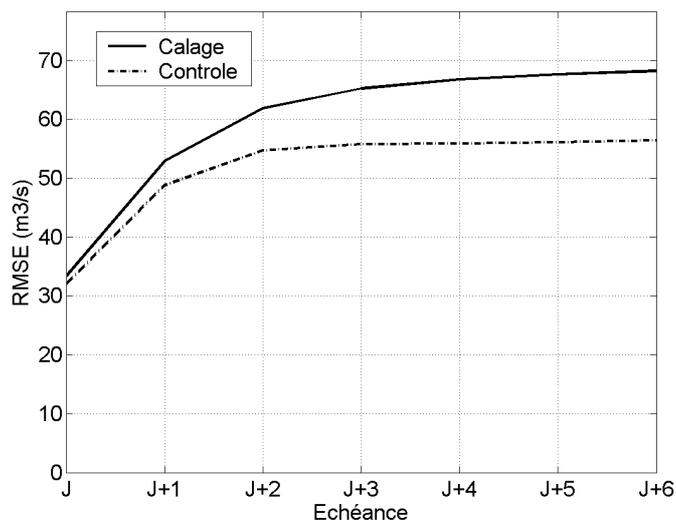


Figure 9.2 : Evolution du score $RMSE$, obtenu entre débit observé et débit simulé par le modèle pour la Saône à Le Chatelet, en fonction de l'échéance, sur les périodes de calage et de contrôle.

La seconde étape de calage consiste à optimiser, pour chaque échéance, les paramètres du modèle sur toute la période. Les tableaux 9.1 et 9.2 indiquent les valeurs minimales de $RMSE$ obtenues pour chacune des échéances, pour chacun des sous-bassins de la Saône et de la Seine.

Tableau 9.1 : Valeurs de $RMSE$ entre débit observé et débit simulé obtenues pour chacune des sept échéances de prévision, aux exutoires des bassins versants de la Saône.

Echéance	Le Châtelet (Saône)	Neublans (Doubs)	Couzon (Saône)
J	32,4	33,4	62,1
$J+1$	50,4	50,0	105,5
$J+2$	58,0	57,3	133,9
$J+3$	60,5	61,2	150,6
$J+4$	61,6	63,7	159,5
$J+5$	62,3	65,4	164,2
$J+6$	62,8	66,6	166,7

Les valeurs de $RMSE$ ne sont pas comparables entre les bassins en raison des différents régimes hydrologiques et surfaces de bassin. Néanmoins ces valeurs nous assurent bien que les erreurs de prévision croissent avec l'échéance. Elles tendent à se stabiliser pour les dernières échéances.

Il est important de noter que, pour une raison de simplicité, les paramètres qui seront utilisés pour la suite de nos travaux sont ceux optimisés pour la première échéance J . Maintenant que le calage est effectué, nous pouvons compléter l'évaluation de la capacité prédictive du modèle dans des conditions parfaitement connues, *i.e.* lorsque les pluies prévues sont observées.

Tableau 9.2: Valeurs de *RMSE* entre débit observé et débit simulé obtenues pour chacune des sept échéances de prévision, aux exutoires des bassins versants de la Seine.

Echéance	Bazoches (Seine)	Courlon (Yonne)	Episy (Loing)	La Ferté (Marne)	Paris (Seine)
<i>J</i>	4,5	13,0	3,4	10,0	27,4
<i>J</i> +1	7,8	19,0	4,9	16,5	38,6
<i>J</i> +2	10,5	22,1	5,5	18,3	46,3
<i>J</i> +3	12,7	23,9	6,0	20,7	52,0
<i>J</i> +4	14,6	25,2	6,2	22,2	56,4
<i>J</i> +5	16,0	26,2	6,5	23,2	60,0
<i>J</i> +6	17,1	27,0	6,7	24,2	63,2

9.1.3 Evaluation en conditions parfaitement connues

L'analyse des débits calculés par le modèle GR3P alimenté par les pluies observées permet d'examiner la qualité du modèle en conditions parfaitement connues. Elle est utile pour identifier les sources d'erreur dans la chaîne de prévision des débits.

Comme nous le verrons dans le chapitre 10, les prévisions hydrologiques seront évaluées lorsque les pluies introduites sont générées par les méthodes d'analogie dans un contexte de prévision parfaite (champs météorologiques réanalysés), soit sur la période allant du 01/08/1970 au 30/08/2002. Or pour évaluer le modèle, il est nécessaire de supprimer la période de chauffe, c'est-à-dire la période nécessaire pour que les débits simulés ne soient plus influencés par les conditions initiales imposées aux niveaux des réservoirs. Nous avons retenu une période de deux ans, la première prévision considérée dans l'évaluation datera donc du 02/08/1972. Pour que la taille de l'échantillon soit identique pour toutes les échéances, le dernier jour de prévision considéré sera le 17/08/2002, la dernière échéance correspondant ainsi au 30/08/2002. Par conséquent, la période retenue pour l'évaluation s'étend du 02/08/1972 au 17/08/2002.

9.1.3.1 Analyse mensuelle à l'aide du critère de Nash

Cette section vise à évaluer le comportement et les performances du modèle GR3P suivant le mois de l'année lorsque les pluies futures sont connues (observées).

Pour chaque mois de l'année et pour chacune des échéances, nous avons calculé le critère de Nash (Nash and Sutcliffe, 1970), défini par l'équation 9.2 :

$$Nash = 1 - \frac{\sum_{J=1}^{nT} (Q_{obs}(J) - Q_{sim}(J))^2}{\sum_{J=1}^{nT} (Q_{obs}(J) - \overline{Q_{obs}})^2} \quad (9.2)$$

où $Q_{obs}(J)$ est le débit observé le jour J , $Q_{sim}(J)$ est le débit simulé à partir des pluies

observées le jour J , $\overline{Q_{obs}}$ est la valeur moyenne des débits observés sur la période et nT le nombre de jours de la période.

La Figure 9.3 présente, pour chacun des huit bassins de la Saône et de la Seine, l'évolution du critère de Nash en fonction du mois de l'année, pour les sept premières échéances de prévision.

Puisque les paramètres du modèle retenus sont optimaux pour la première échéance, les valeurs du critère de Nash sont très proches de 1 pour le jour J . Ces valeurs diminuent logiquement avec l'échéance, les erreurs étant combinées au fil des jours malgré la connaissance des pluies. Les valeurs diminuent plus fortement avec l'échéance en hiver (durant la saison pluvieuse), mais surtout en été. En effet, le modèle a des difficultés à faire mieux qu'une estimation naïve donnée par la moyenne mensuelle interannuelle. La dynamique des débits, lorsque le bassin est en phase de soutien d'étiage, est différente de la dynamique du bassin en crue. Or cette différentiation n'est pas possible dans le modèle hydrologique dans sa version actuelle. En outre, le modèle ne prend pas en compte les influences humaines qui se manifestent fortement en étiages. Enfin, le modèle a été optimisé sur les débits bruts. La reconstitution des forts débits a été implicitement privilégié au détriment des valeurs faibles, ce qui induit des performances plus faibles durant la saison des étiages. Néanmoins en dehors de la saison d'été, les performances du modèle sont satisfaisantes, avec des valeurs de Nash généralement supérieures à 0,8. Pour les bassins de la Seine à Bazoches et de la Marne à La Ferté, on remarque un comportement différent. Pour la Marne à la Ferté, les valeurs sont très irrégulières ; tandis que pour la Seine à Bazoches, les valeurs de Nash sont presque équivalentes tout au long de l'année. Pour ces bassins, cela peut être dû à un problème d'optimisation puisque les séries débitométriques comportent beaucoup de lacunes (cf. 2.2.3).

En raison de la présence de nombreuses lacunes dans les séries, nous écarterons désormais ces deux bassins lorsque nous nous intéresserons aux prévisions de débit.

9.1.3.2 Comparaison aux prévisions de débit supposant les pluies futures nulles

Nous avons comparé les débits simulés par le modèle aux débits prévus lorsque les pluies futures sont nulles. Pour chaque bassin, le critère de persistance $C4$ (Wallis and Todini, 1975) a été calculé pour chaque échéance E de prévision. Il est défini par :

$$C4 = 100. \left(1 - \frac{\sum [Q_{obs}(t+E) - Q_{prev}(t+E)]^2}{\sum [Q_{obs}(t+E) - Q_{obs}(t)]^2} \right) \quad (9.3)$$

où Q_{obs} est le débit observé, Q_{prev} est le débit prévu, J est l'instant de prévision et E l'échéance.

Ce critère permet d'évaluer la qualité d'une prévision par rapport à une prévision naïve qui consisterait à prendre comme débit prévu le débit observé à l'instant de la prévision.

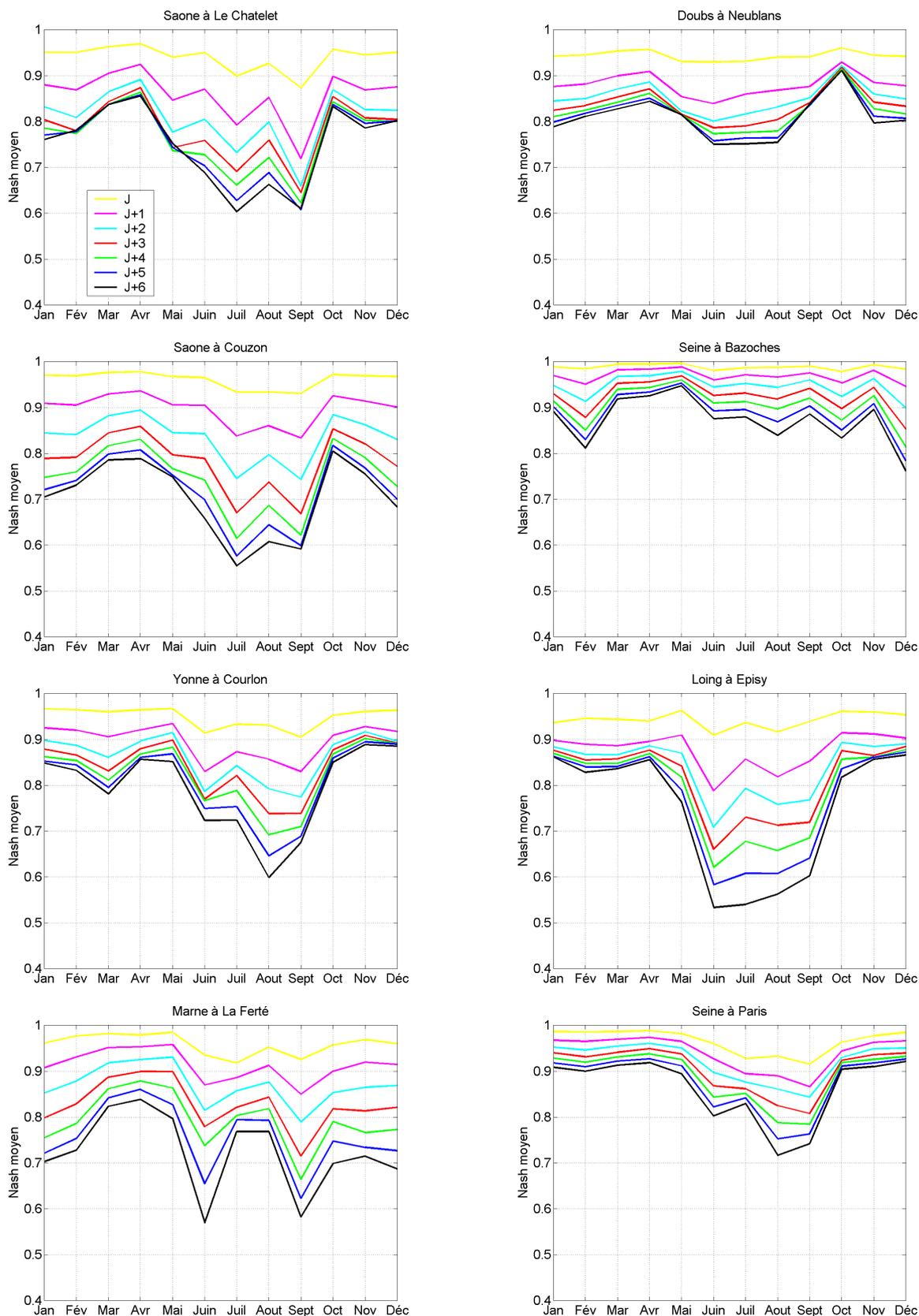


Figure 9.3 : Evolution de la valeur moyenne mensuelle du critère de Nash sur les débits simulés, pour les huit bassins de la Saône et de la Seine.

La Figure 9.4 présente l'évolution du critère $C4$ en fonction de l'échéance de prévision,

lorsque les pluies injectées dans le modèle hydrologique sont connues ou nulles.

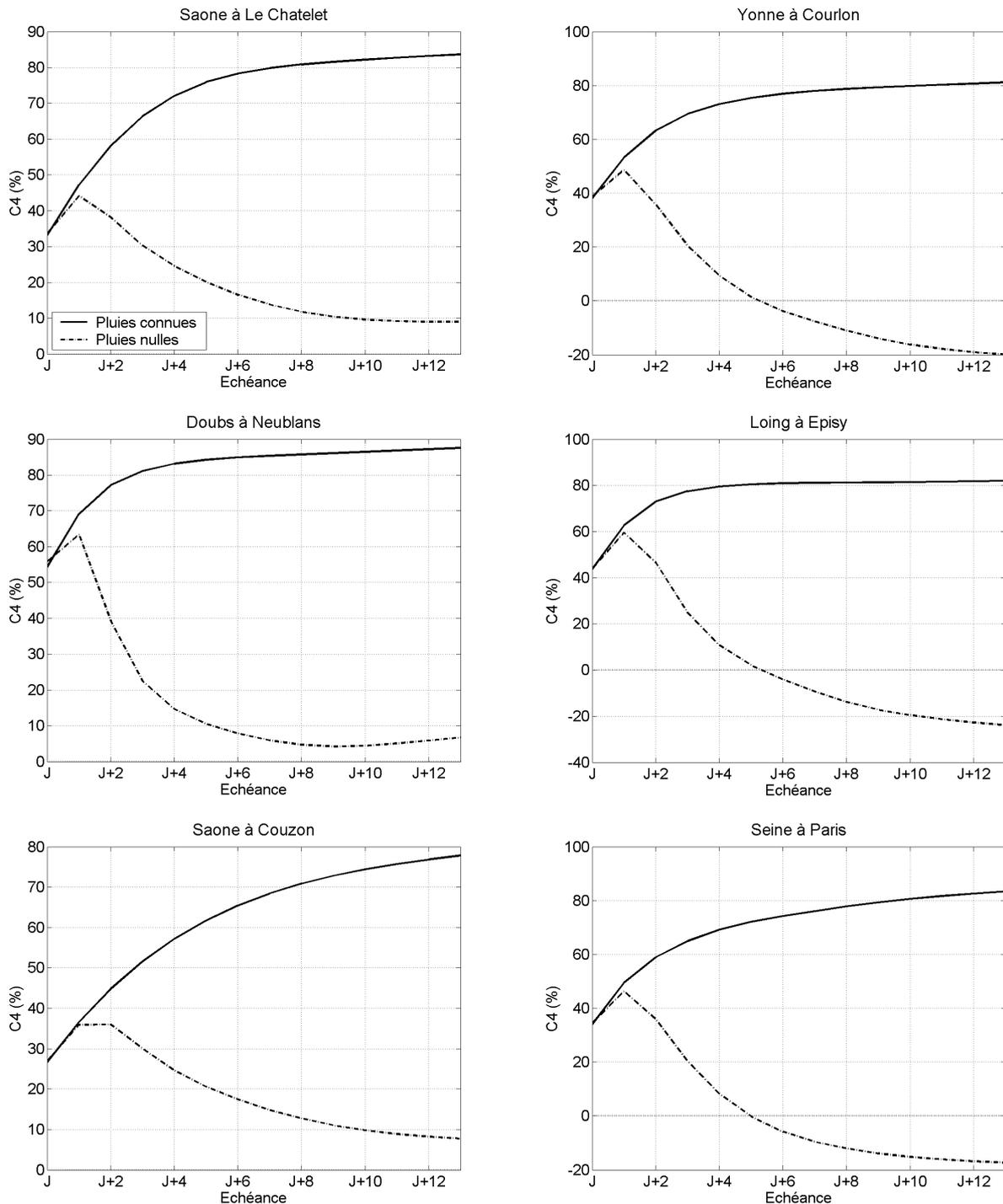


Figure 9.4 : Evolution du critère $C4$ en fonction de l'échéance, lorsque les pluies futures sont connues (trait plein) et lorsqu'elles sont nulles (trait discontinu), sur chacun des six sous-bassins étudiés.

Ces courbes permettent de connaître l'échéance de prévision à partir de laquelle il est nécessaire d'injecter dans le modèle une information sur les pluies afin d'obtenir une prévision de débit correcte. En d'autres termes, ils nous indiquent l'effet de mémoire des bassins. On remarque ici que cet effet mémoire n'excède pas deux jours (échéance $J+1$) pour

tous les bassins à l'exception de la Saône à Couzon (trois jours). D'autre part, nous retrouvons ici la réactivité aux pluies de certains bassins, à l'image de celui du Doubs à Neublans. La valeur du critère $C4$ décroît plus rapidement que pour la Saône à Le Chatelet et la Saône à Couzon.

Ces graphiques permettent également d'identifier l'échéance à partir de laquelle la prévision du débit obtenu par injection d'une pluie nulle devient moins performante que la prévision naïve. Le modèle semble mieux adapté aux bassins de la Saône puisque la valeur du critère $C4$ reste positive jusqu'à la dernière échéance. En revanche pour les bassins de la Seine, la valeur devient négative dès l'échéance $J+3$, ce qui signifie qu'au-delà de cette échéance, la prévision naïve basée sur la persistance apporte plus d'information que la prévision obtenue lorsque les pluies prévues sont nulles.

9.2 Génération des prévisions de débit

Dans le cadre de cette application hydrologique, nous avons choisi de conserver l'aspect probabiliste des prévisions jusqu'au débit. L'objectif final est d'établir, comme pour les précipitations, une distribution statistique empirique des débits prévus. La principale difficulté à générer des scénarios probabilistes de débits, à partir des prévisions issues d'une méthode d'analogie par adaptation statistique, réside dans la construction de scénarios cohérents dans le temps pour les échéances désirées. En effet, en adaptation statistique, les situations analogues retenues chaque jour sont indépendantes dans le temps, puisque les sorties d'un MNP sont adaptées indépendamment pour chaque échéance. En revanche, en prévision statistique simple autrefois utilisée (Lorenz, 1969) (cf. 1.2.1), le problème ne se pose pas : l'analogie est uniquement recherchée sur la situation observée le jour courant ; les situations analogues retenues permettant ainsi d'obtenir des séquences pluviométriques temporelles pour les jours suivants. C'est par exemple le cas de l'étude de Diomede et al. (2008). Cette technique est limitée et ne nous semble pas adaptée à nos objectifs, puisqu'il s'agit d'une méthode de prévision statistique simple.

Par conséquent, il n'est pas évident de construire directement des scénarios à partir des prévisions par analogie adaptées statistiquement des sorties d'un MNP. Afin de remédier à ce problème, nous avons mis en place une méthode alternative (Figure 9.5).

Sachant que nous disposons de N pluies observées lors des situations analogues pour chaque échéance de prévision, nous avons construit N scénarios en tirant aléatoirement, à chaque échéance, la pluie associée à une des N situations analogues retenues. Il est important de noter que le tirage aléatoire ne tient pas compte de la corrélation entre plusieurs journées successives : les journées analogues tirées successivement pour un scénario sont donc supposées indépendantes¹. Nous disposons ainsi, pour chaque jour de prévision, de N scénarios d'évolution temporelle des précipitations *a priori* représentatifs des distributions

1 Il s'avère qu'en moyenne, la corrélation entre les pluies journalières observées sur deux journées successives est de l'ordre de 0,1.

issues des analogues. Ils ont ensuite été injectés dans le modèle GR3P et des distributions empiriques de débits prévus ont été établies par examen statistique des débits générés pour chaque échéance.

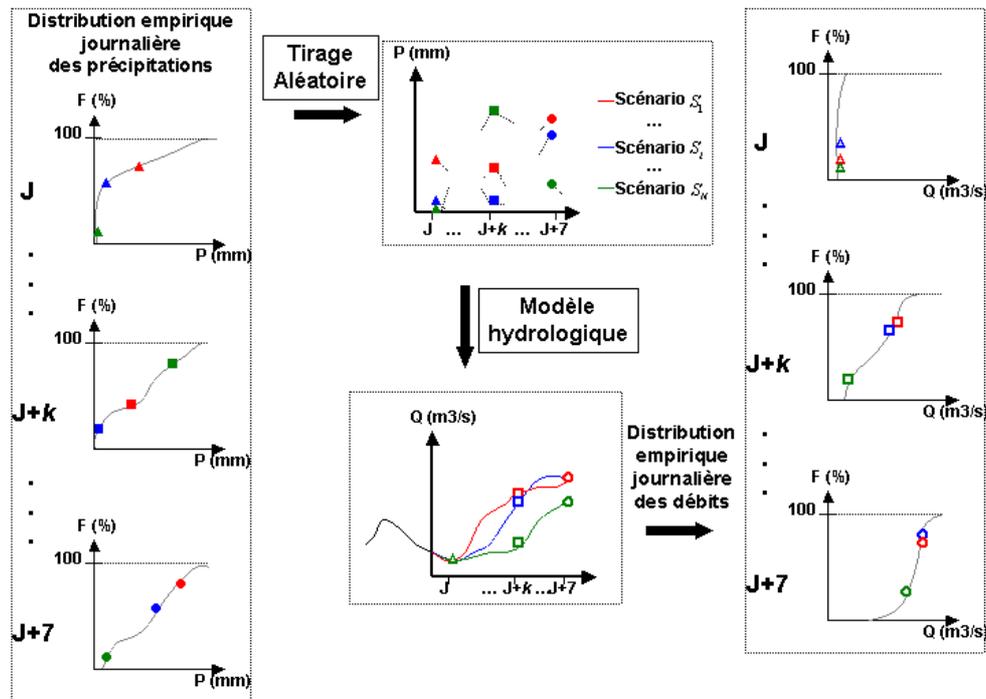


Figure 9.5 : Procédure schématisée de la génération de N scénarios de débits pour les échéances J à $J+7$, à partir des distributions empiriques journalières des précipitations issues d'une méthode d'analogie.

Outre ces N scénarios, nous avons généré trois scénarios supplémentaires représentatifs des distributions des pluies : pour chaque échéance, les scénarios sont constitués en prenant respectivement les quantiles de prévision d'ordre 20, 60 et 90 % de chaque distribution ($Pdis20$, $Pdis60$ et $Pdis90$). Cela ne permet pas de construire des scénarios réalistes, puisqu'il s'agit finalement de scénarios moyens très différents : le scénario associé au quantile $Pdis20$ est alors un scénario optimiste (faibles débits prévus), celui associé au quantile $Pdis90$ est alors un scénario pessimiste (forts débits prévus), et celui associé au quantile $Pdis60$ constitue un scénario médian. Ces scénarios ont été utilisés par Plasse (2008) dans le cadre d'une application hydrologique préliminaire sur le bassin de la Seine, pour construire des courbes enveloppes et un scénario moyen.

9.3 Prévisions naïves

L'objectif principal des deux chapitres suivants est d'évaluer les prévisions probabilistes de débit dans les contextes de prévision parfaite et de prévision opérationnelle des précipitations. Dans les deux cas, plusieurs critères d'évaluation de la prévision de débit ont été exploités. Etant donnée la forme probabiliste de la prévision hydrologique, nous avons

décidé d'utiliser notamment le même score d'évaluation que pour les pluies : le *CRPS*. Nous avons également défini deux prévisions naïves, que nous utiliserons comme prévision de référence dans le calcul de la compétence (*CRPSS*) du *CRPS* : la première est la persistance et la seconde est donnée par le régime hydrologique, référence que nous appellerons plus simplement « climatologie » par la suite.

9.3.1 Persistance

Cette prévision naïve consiste, pour chaque jour J de prévision, à prévoir pour chaque échéance un débit égal à débit $Q(J-1)$ observé la veille du jour d'émission de la prévision. Le score *CRPS* est alors calculé sur les N valeurs du même débit. Ainsi la distribution prévue pour chaque échéance est parfaitement fine, mais la justesse de la distribution est évidemment dégradée, et ce d'autant plus que l'échéance augmente. Le schéma d'élaboration de cette prévision naïve est indiqué par la Figure 9.6.

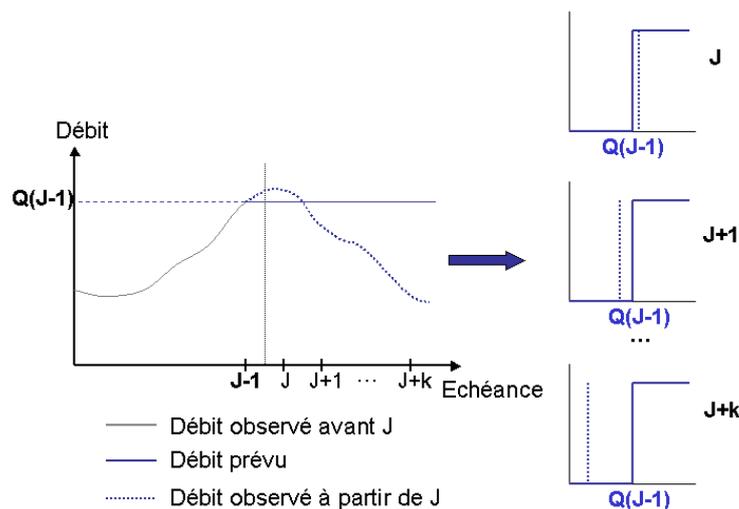


Figure 9.6 : Schéma d'élaboration de la prévision naïve basée sur la persistance.

9.3.2 Climatologie

Nous nous sommes inspirés de la référence choisie pour les pluies. Pour rappel, on extrait aléatoirement autant de valeurs que d'analogues dans les pluies observées autour d'une fenêtre temporelle de largeur quatre mois centrée sur le jour cible.

Pour les débits, il faut tenir compte de l'auto-corrélation entre deux jours successifs : on ne peut faire le tirage de manière totalement aléatoire. Le schéma d'élaboration de cette prévision naïve est indiqué par la Figure 9.7.

Pour chaque jour J de prévision, on recherche ainsi dans l'archive historique au maximum $2N$ jours pour lesquels le débit est le plus proche du débit $Q(J-1)$ observé la veille, N étant égal au nombre d'analogues. Cette recherche est saisonnalisée : on conserve, pour le mois M

du jour J de la prévision, les jours inclus dans la fenêtre temporelle fixe $[M-1 ; M+1]$, avec au maximum N jours pour lesquels le débit observé est inférieur à $Q(J-1)$, et au maximum N jours pour lesquels le débit observé est supérieur à $Q(J-1)$. Si le débit $Q(J-1)$ est extrême (par exemple, le plus faible des débits observés sur la chronique), alors on ne pourra que retenir N jours pour lesquels le débit observé est supérieur à $Q(J-1)$. De plus, la recherche de jours dans la même année est interdite dans une fenêtre de largeur seize jours centrée sur le jour J . Il est par conséquent possible, dans le pire des cas, de ne retenir que $N-16$ jours.

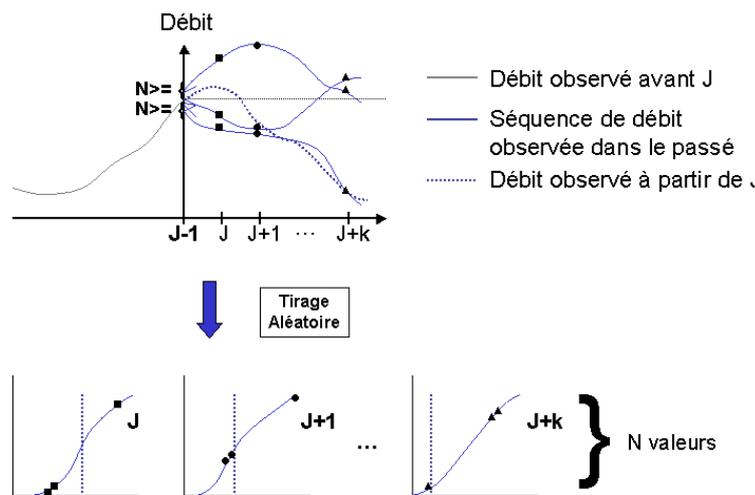


Figure 9.7 : Schéma d'élaboration de la prévision naïve basée sur la climatologie.

Une fois les jours sélectionnés, on extrait les séquences débimétriques des E jours suivants (E étant l'échéance maximale). Pour chaque échéance de prévision, on dispose alors d'une distribution empirique de débits. On tire ensuite aléatoirement N valeurs de débit pour chaque échéance dans ces séquences. On dispose ainsi de N scénarios de débit, à partir desquels une distribution empirique des valeurs probables est extraite pour chaque échéance, et ceci compte tenu du débit de la veille. Une valeur du score $CRPS$ est alors calculée pour chaque échéance.

10 Dans un contexte de prévision parfaite

La construction des algorithmes retenus à l'issue de la seconde partie a été basée sur le score global *CRPS* et sa compétence *CRPSS* par rapport à la prévision climatologique. Nous proposons dans ce chapitre d'évaluer de manière approfondie ces algorithmes et leurs performances à l'aide d'autres scores définis dans le chapitre 3. Ces scores nous permettront de connaître des qualités et des défaillances qui n'ont pas pu être mises en évidence jusqu'à présent. Ceci fera l'objet de la première section de ce chapitre. Dans la seconde section, le comportement de la chaîne de prévision hydrologique alimentée par les scénarios de précipitations sera analysé. L'analyse s'effectue ici dans un contexte de prévision parfaite, *i.e.* les champs météorologiques d'une situation cible sont extraits des réanalyses ERA-40.

10.1 Prévision des précipitations

L'objectif de cette section est d'évaluer et de comparer les méthodes de prévision des précipitations par analogie dans un contexte de prévision parfaite. Outre le score global *CRPSS*, nous avons notamment utilisé les scores issus des tableaux de contingence. Ces derniers permettent d'évaluer par exemple la capacité des méthodes à ne pas générer de fausse alerte ou à détecter des événements supérieurs à un seuil. Cette mesure de performance intéresse particulièrement l'utilisateur potentiel de telles méthodes.

Dans cette partie, nous avons généré, sur les bassins de la Saône et de la Seine, les scénarios probabilistes de prévision de pluie dans un contexte de prévision parfaite, sur la période la plus longue possible. Cette période s'étire à présent du 01/08/1970 (borne inférieure de la période SAFRAN en fin de thèse) au 30/8/2002 (borne supérieure de la période couverte par l'archive ERA-40).

10.1.1 Ré-optimisation des nombres d'analogues

Durant la phase de développement, nous avons construit les algorithmes A0, A1, A1bis, A2 et A2bis en retenant à chaque niveau d'analogie un nombre de situations que nous avons

préalablement optimisé sur la période de calibration, longue de 25 années. Or à présent, notre période d'étude couvre 32 années. Il est donc préférable de ré-optimiser ces nombres sur cette période afin d'obtenir les *meilleures* prévisions possibles. Cependant, la combinatoire devient très conséquente pour les algorithmes A1/A1bis (espace à trois dimensions) et A2/A2bis (espace à quatre dimensions). Nous avons donc adopté une stratégie d'optimisation des nombres d'analogues sélectionnées à chaque niveau d'analogie, qui pourra par ailleurs être reprise par un nouvel utilisateur lorsqu'il désirera adapter une méthode à un nouveau bassin versant.

10.1.1.1 Stratégie d'optimisation

Les nombres à optimiser sont $(N1, N2)$ pour la méthode A0, $(N0, N1, N2)$ pour les méthodes A1 et A1bis, et $(N0, N1, N2, N3)$ pour les méthodes A2 et A2bis (cf. tableau de synthèse en introduction à la partie III).

Pour la méthode A0, les nombres $N1$ et $N2$ peuvent être aisément optimisés de façon simultanée. En émettant l'hypothèse que le nombre d'analogues retenues à l'issue du dernier niveau de sélection dépend essentiellement de la longueur de la période d'archive, les nombres $N2$ et $N3$ du dernier niveau de sélection des méthodes respectives A1/A1bis et A2/A2bis peuvent être fixés, étant égaux au nombre $N2$ de la méthode A0.

Pour les méthodes A1 et A1bis, restent les nombres $N0$ et $N1$ à optimiser. Nous avons cherché à optimiser le nombre $N0$, mais il s'avère qu'en plus du lourd temps de calcul, la sensibilité du score *CRPSS* à ce nombre est faible, d'où le risque de se situer vers un optimum local. En revanche, lors de l'optimisation de ce nombre sur la période de calibration (section 5.2.3), nous avons remarqué que la valeur optimale était à peu près équivalente au nombre de situations candidates retenues par sélection selon la saison dans l'algorithme de référence. De cette manière, le nombre $N0$ optimal, sur une période donnée et pour un jour cible, est supposé correspondre au nombre de situations contenues dans une fenêtre temporelle de plus ou moins deux mois centrée sur le jour cible de chaque année candidate. Dans notre cas, avec une période longue de 32 années, le nombre $N0$ choisi s'élève ainsi à 3900. Pour les méthodes A1/A1bis, les nombres $N0$ et $N2$ étant alors fixés, le nombre $N1$ peut être facilement ré-optimisé.

Pour les méthodes A2 et A2bis, nous avons procédé de la même manière pour les nombres $N0$ et $N3$. Par conséquent, seuls les nombres $N1$ et $N2$ doivent être ré-optimisés, ce qui réduit considérablement le temps de calcul.

10.1.1.2 Méthode A0

Les nombres $N1$ et $N2$ ont donc été optimisés sur chaque bassin test sur la période longue de 32 années dans un contexte de prévision parfaite. Pour chacune des combinaisons $(N1, N2)$, les performances de prévision ont été exprimées en termes de *CRPSS*. La combinaison optimale et le score *CRPSS* associé sont indiqués dans le Tableau 10.1 pour les bassins de la

Saône et de la Seine.

Tableau 10.1 : Combinaisons ($N1, N2$) optimales et scores $CRPSS$ associés, obtenus pour les deux bassins tests.

Bassin	Nombre $N1$	Nombre $N2$	$CRPSS$ (%)
Saône	80	25	49,78
Seine	80	30	47,59

Pour chacun des quatre algorithmes, le nombre d'analogues retenues à l'issue du dernier niveau de sélection s'élève donc à 25 pour le bassin de la Saône et à 30 pour celui de la Seine. Seul le nombre $N2=25$ pour le bassin de la Saône est identique à celui obtenu initialement sur la période de 30 ans (1972-2002).

10.1.1.3 Méthodes A1 à A2bis

Le nombre $N0$ de situations candidates est préalablement fixé à 3900, et le nombre d'analogues retenues à l'issue du dernier niveau de sélection ($N2$ ou $N3$ selon la méthode) s'élève à 25 pour la Saône et à 30 pour la Seine. La dernière étape consiste donc à ré-optimiser :

- le nombre d'analogues retenues au niveau 1 pour les algorithmes A1/A1bis à $N0$ et $N2$ fixés ;
- les nombres d'analogues retenues aux niveaux 1 et 2 pour les algorithmes A2/A2bis à $N0$ et $N3$ fixés.

Pour les méthodes A1/A1bis, nous avons fait varier $N1$ entre 30 et 300 par pas de 10. La Figure 10.1 présente l'évolution de la performance moyenne obtenue avec l'algorithme A1 sur le bassin de la Saône, exprimée en termes de $CRPSS$, en fonction du nombre $N1$.

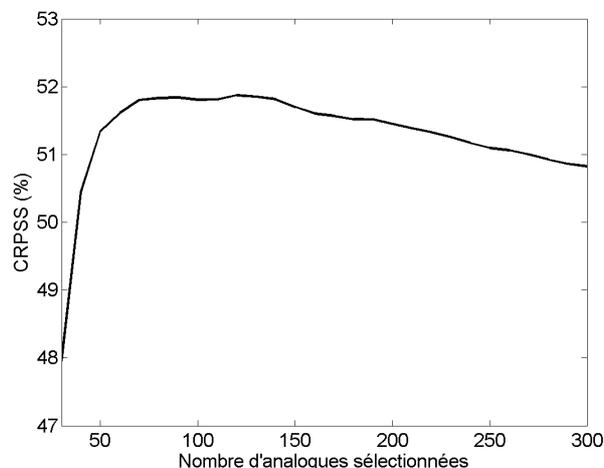


Figure 10.1 : Evolution des performances ($CRPSS$ (%)) obtenues sur l'ensemble bassin de la Saône par l'algorithme A1, en fonction du nombre $N1$ d'analogues retenues au premier niveau de sélection.

Bien qu'il existe un palier pour lequel la performance est relativement stable (entre 70 et 140), le nombre $N1$ optimal s'élève à 120 ($CRPSS$ de 51,87 %). Pour les algorithmes A2 et A2bis, nous avons fait varier $N1$ entre 150 et 500, et $N2$ entre 50 et $N1-10$ pour chaque valeur de $N1$. Les combinaisons optimales obtenues pour les algorithmes A1, A1bis, A2 et A2bis, sur chacun des deux bassins tests, sont indiquées dans le Tableau 10.2. La hiérarchie obtenue entre les algorithmes sur les deux sous-périodes retenues pour le développement de la méthode (calibration et validation) est respectée sur cette période plus longue. Par ailleurs, pour les méthodes A2/A2bis, le nombre d'analogues retenues au niveau 1 devient très élevé, valant plus double du nombre $N1$ optimisé pour les méthodes A0 à A1bis.

Tableau 10.2 : Nombres d'analogues optimaux et performances associées ($CRPSS$ (%)), obtenus pour les différents algorithmes retenus sur l'ensemble des bassins de la Saône et de la Seine.

Bassin	Algorithme	Nombre de situations retenues				$CRPSS$ (%)
		$N0$	$N1$	$N2$	$N3$	
Saône	A1	3900	120	25	-	51,87
	A1bis	3900	100	25	-	53,07
	A2	3900	320	70	25	54,84
	A2bis	3900	330	90	25	55,97
Seine	A1	3900	130	30	-	49,41
	A1bis	3900	130	30	-	50,76
	A2	3900	280	80	30	51,20
	A2bis	3900	280	100	30	52,4

Les nombres d'analogues à sélectionner à chaque niveau ayant été ré-optimisés, nous avons pu générer, pour chaque jour de la période allant du 01/08/1970 au 30/08/2002, des prévisions probabilistes des précipitations journalières sur chacun des deux bassins tests, dans un contexte de prévision parfaite. Ces données ont ensuite servi à alimenter le modèle hydrologique GR3P dans le but d'évaluer le modèle. Avant d'analyser les prévisions de débit, nous proposons d'évaluer les différents algorithmes de sélection d'analogues à l'aide d'autres scores de vérification, notamment en fonction de seuils de précipitations.

10.1.2 Etude statistique et critique des prévisions de précipitations

10.1.2.1 Finesse et justesse

Les scores $CRPSS$ indiqués précédemment donnent déjà une idée du gain de performance que peut apporter l'un ou l'autre algorithme par rapport à l'algorithme de référence A0. Nous avons également séparé finesse et justesse de la prévision et évalué le gain G apporté par chacun des trois algorithmes par rapport à l'algorithme A0 (Figure 10.2), en termes de finesse et de justesse des distributions prévues. Cela permet d'apprécier les différences entre les méthodes et ainsi de connaître quels sont les impacts de l'introduction de nouvelles variables sur la finesse et la justesse des prévision.

Nous retrouvons ici un gain de performance élevé en finesse que nous avons déjà obtenu sur la période de calibration. Les gains en finesse sont comparables entre les méthodes, puisque le nombre de situations analogues finalement retenues est identique pour chacune des méthodes. En effet, si pour l'une des méthodes, nous n'avions retenu qu'une analogue, alors la finesse de la distribution serait parfaite, ce qui donnerait un gain de 100 %. Ainsi, ces résultats montrent que le gain apporté par la méthode A2 est marqué, avec une valeur de plus de 15 % pour la Saône et de 10 % pour la Seine par rapport à la méthode A0.

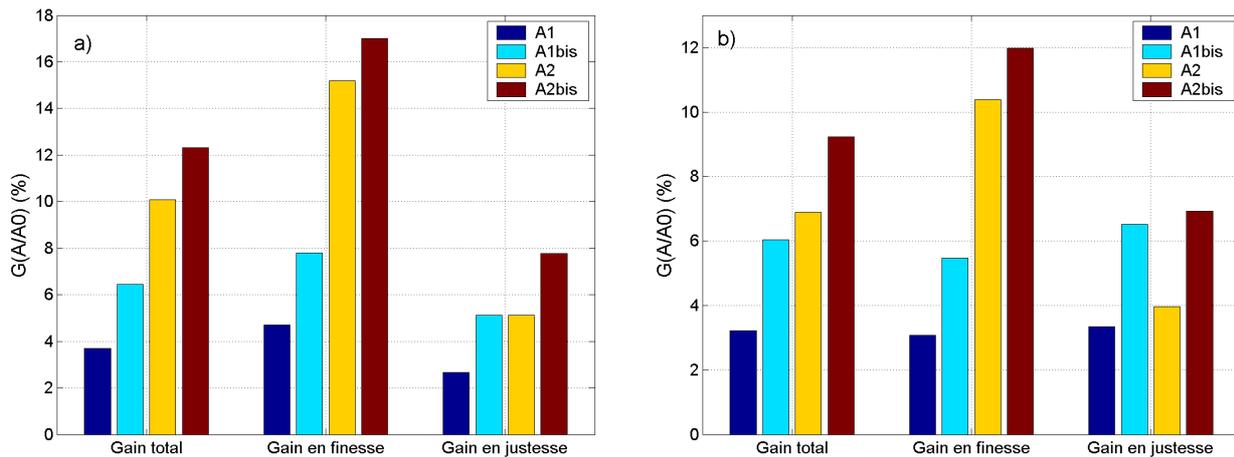


Figure 10.2 : Gains G (%) de performance moyenne, obtenus par les méthodes A1 à A2bis par rapport à la méthode A0 (gain total et gain décomposé en finesse et en justesse). a) bassin de la Saône ; b) bassin de la Seine

Le gain en justesse s'avère plus faible, même s'il atteint près de 7 à 8 % lorsque la méthode A2bis est appliquée. Par ailleurs, on remarque que pour le bassin de la Seine, le gain en justesse est très faible entre les méthodes A1 et A2 (moins de 1 %), tandis qu'il est plus marqué pour le bassin de la Saône (plus de 2 %). Pour le bassin de la Seine, c'est surtout la mesure de l'humidité relative aux nouveaux niveaux de pression qui contribue à un gain en justesse. On remarque par ailleurs que pour les deux bassins, les gains en justesse apportés d'une part par la vitesse verticale et d'autre part par les nouveaux niveaux de pression, par rapport à la méthode A1, s'ajoutent parfaitement. Cela indique que les gains en justesse apportés par ces deux nouveaux éléments sont indépendants.

10.1.2.2 Scores issus des tableaux de contingence

Diagramme ROC

Dans cette section, les scores POD et $POFD$ (cf. sections 3.2.1 et 3.2.2) ont été évalués pour chacune des N valeurs de précipitations prévues formant une distribution empirique (soit 25 valeurs pour le bassin de la Saône et 30 pour le bassin de la Seine), et ce pour différents seuils de pluie journalière. Les seuils de précipitations choisis sont le seuil 0 mm (pluie/non pluie), et les quantiles de pluie non nulle d'ordre 10 à 90 % par pas de 10 % ($PNN10$ à

PNN90), ainsi que les quantiles d'ordre 95 et 99 % (*PNN95* et *PNN99*). Les valeurs de ces quantiles sont reportées dans le Tableau 10.3, pour les sous-bassins de la Saône et de la Seine.

Tableau 10.3 : Valeurs de cumul pluviométrique associés aux différents quantiles de pluie non nulle, pour chaque sous-bassin de la Saône et de la Seine.

Sous-bassin	Ordre (%) du quantile de pluie non nulle										
	10	20	30	40	50	60	70	80	90	95	99
Doubs	0,2	0,4	0,9	1,7	2,9	4,4	6,5	9,1	13,6	18,4	28,7
P. Saône	0,2	0,4	0,8	1,5	2,4	3,6	5,1	7,1	10,7	14,3	24,4
Saône I.	0,1	0,3	0,7	1,2	2,1	3,2	4,6	6,7	10,7	14,5	25,7
Yonne-C	0,1	0,4	0,7	1,3	2,1	3,1	4,5	6,3	9,4	12,7	19,8
Seine-Int	0,1	0,3	0,5	1	1,6	2,4	3,6	5,2	7,9	11	18,3
Seine-B	0,1	0,3	0,7	1,2	1,9	2,9	4,2	5,8	8,6	11,8	19,3
Marne-F	0,1	0,3	0,7	1,3	2	3	4,3	6,2	9,3	12,4	21,4
Loing-E	0,1	0,3	0,6	1,1	1,8	2,6	3,8	5,5	8,5	11,6	19,3

Le diagramme ROC associé à chacun de ces seuils a ensuite été construit, en appliquant chacune des méthodes d'analogie. Le nombre d'analogues retenues à l'issue du dernier niveau étant différent entre le bassin de la Saône et celui de la Seine, les diagrammes ont été construits séparément pour chaque bassin. Par souci de simplicité, nous avons évalué les scores pour l'ensemble des sous-bassins. Les diagrammes obtenus sur chacun des bassins sont similaires. Pour cette raison, seuls les diagrammes associés au bassin de la Saône sont indiqués sur la Figure 10.3, pour les seuils 0 mm, *PNN20*, *PNN60* et *PNN90*.

Les courbes s'écartent fortement de la première bissectrice, ce qui traduit un pouvoir discriminant des prévisions important. Considérons le seuil 0 mm. Pour les deux bassins, les quatre courbes sont relativement bien symétriques, ce qui indique que les taux de détection et de fausse détection sont équivalents. On note un léger avantage aux méthodes développées dans la thèse (A1 à A2bis) par rapport à la méthode de référence A0, avec un score *POD* plus élevé à score *POFD* fixé, et un score *POFD* plus faible à score *POD* fixé. Cet avantage est matérialisé par une aire sous la courbe ROC (score *AROC*) plus élevée.

Si l'on s'intéresse à un seuil un peu plus élevé tel que le seuil *PNN20*, les courbes présentent une symétrie parfaite. Les taux de détection et de fausse détection sont ainsi équilibrés. Par ailleurs, bien que ce ne soit pas illustré ici, l'avantage aux méthodes développées dans la thèse est plus prononcé pour le bassin de la Saône que pour le bassin de la Seine.

Enfin, si on considère des seuils plus élevés tels que les seuils *PNN60* et *PNN90*, le pouvoir discriminant des prévisions reste très satisfaisant. En revanche, toutes les courbes ROC sont dissymétriques : les méthodes de prévision peuvent effectivement nous avertir de l'occurrence d'un événement, cependant une fausse détection n'est pas exclue.

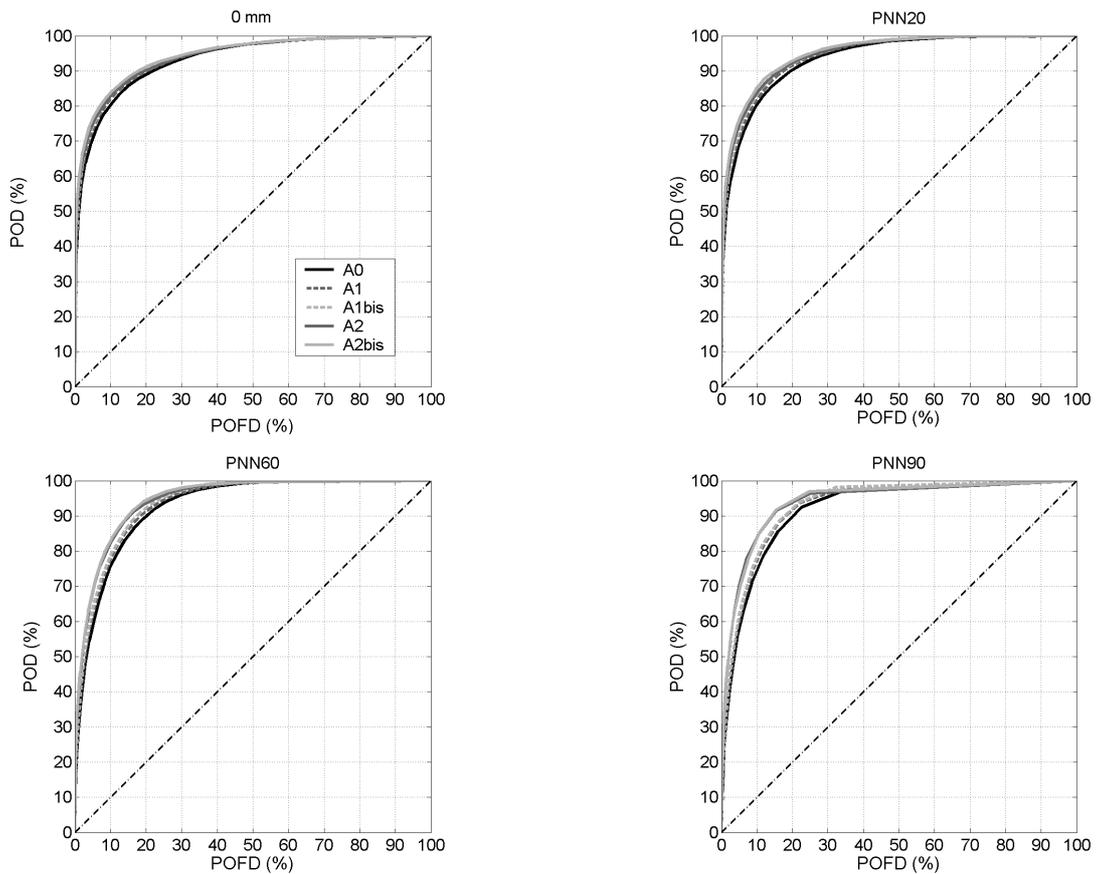


Figure 10.3 : Diagrammes ROC construits pour quatre seuils de lame d'eau (0 mm (pluie/non pluie), PNN20, PNN60 et PNN90), obtenus par application des cinq algorithmes (A0, A1, A1bis, A2, A2bis) sur le bassin de la Saône.

Ces résultats peuvent être résumés par les valeurs de score *AROC* (aire sous la courbe ROC), présenté dans le chapitre 3 (section 3.2.5). La Figure 10.4 présente, a) pour le bassin de la Saône et b) pour le bassin de la Seine, l'évolution du score *AROC* en fonction du seuil de pluie.

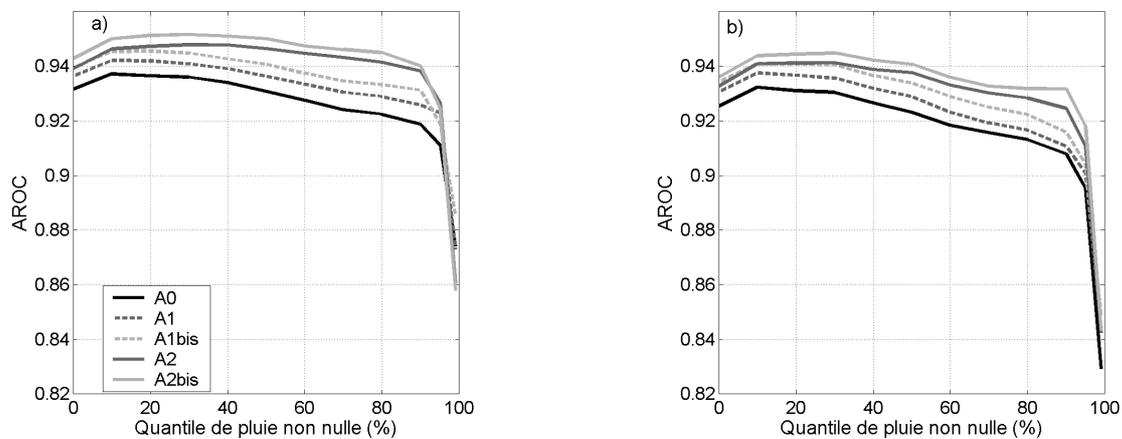


Figure 10.4 : Evolution du score *AROC* en fonction du seuil de pluie (quantile de pluie non nulle) lorsque chacune des méthodes (A0, A1, A1bis, A2, A2bis) est appliquée. a) Bassin de la Saône ; b) Bassin de la Seine.

La hiérarchie entre les méthodes est bien conservée, avec une valeur du score *AROC* croissante de la méthode A0 à la méthode A2bis jusqu'au seuil *PNN95*. Au-delà, les résultats sont très relatifs du fait d'un échantillon réduit. Les algorithmes A1/A1bis se démarquent quelque peu de l'algorithme A0 avec une aire sous la courbe plus grande, mais ce sont surtout les algorithmes A2/A2bis qui s'avèrent plus performants, avec un net avantage à mieux détecter les événements mais aussi à émettre moins de fausses alertes. Cet avantage est moins prononcé pour le bassin de la Seine que pour celui de la Saône. Par ailleurs, le seuil de prévision optimal diffère selon la méthode appliquée. Ainsi la valeur maximale du score *AROC* est obtenue pour le seuil *PNN10* avec les méthodes A0, A1 et A1bis, et pour le seuil *PNN20* avec les méthodes A2 et A2bis. L'analogie supplémentaire sur la vitesse verticale permet donc une prévision optimale d'événements un peu plus importants.

Score de Hanssen-Kuipers (*TSS*)

Le score de Hanssen and Kuipers (*TSS*) a également été défini au chapitre 3 (section 3.2.4). Il a été évalué avec chaque méthode d'analogie lorsque la prévision est définie par l'un des trois quantiles *Pdis20*, *Pdis60* et *Pdis90* d'ordre 20, 60 et 90 % habituellement extraits d'une distribution d'analogues (Figure 10.5). Par simplicité, le score a été évalué pour l'ensemble des sous-bassins de la Saône et de la Seine.

Cette figure montre des résultats plus nuancés que ceux obtenus précédemment avec les gains et le score *AROC*. Ainsi pour le quantile de prévision *Pdis20*, qui représente la borne inférieure de l'intervalle de confiance, c'est la méthode A1bis qui présente la valeur du score *TSS* la plus élevée pour les précipitations inférieures au seuil *PNN70*. L'introduction de la température (A1) puis de la vitesse verticale (A2) permettent d'obtenir un score *TSS* plus élevé. En revanche, c'est la mesure de l'humidité aux niveaux 925 hPa et 700 hPa qui permet d'améliorer le plus sensiblement le score *TSS*. Pour les précipitations supérieures au seuil *PNN70*, les écarts entre les courbes se réduisent logiquement dans la mesure où le quantile *Pdis20* ne permet de détecter ce type d'événement qu'assez rarement. On note simplement que les scores *TSS* obtenus par les méthodes A1bis et A2bis restent légèrement supérieurs aux autres pour ces événements un peu plus rares.

Pour les quantiles de prévision *Pdis60* et *Pdis90*, c'est généralement avec la méthode A2bis que le score *TSS* est le plus élevé, à l'exception des précipitations supérieures au seuil *PNN60* pour le quantile *Pdis60* où la méthode A1bis se révèle plus performante en termes de *TSS*. Ainsi lorsque le quantile de prévision est *Pdis90*, l'écart de valeurs de *TSS* atteint près de 9 % entre les méthodes A0 et A2bis pour les précipitations faibles et moyennes (*PNN20* à *PNN50*).

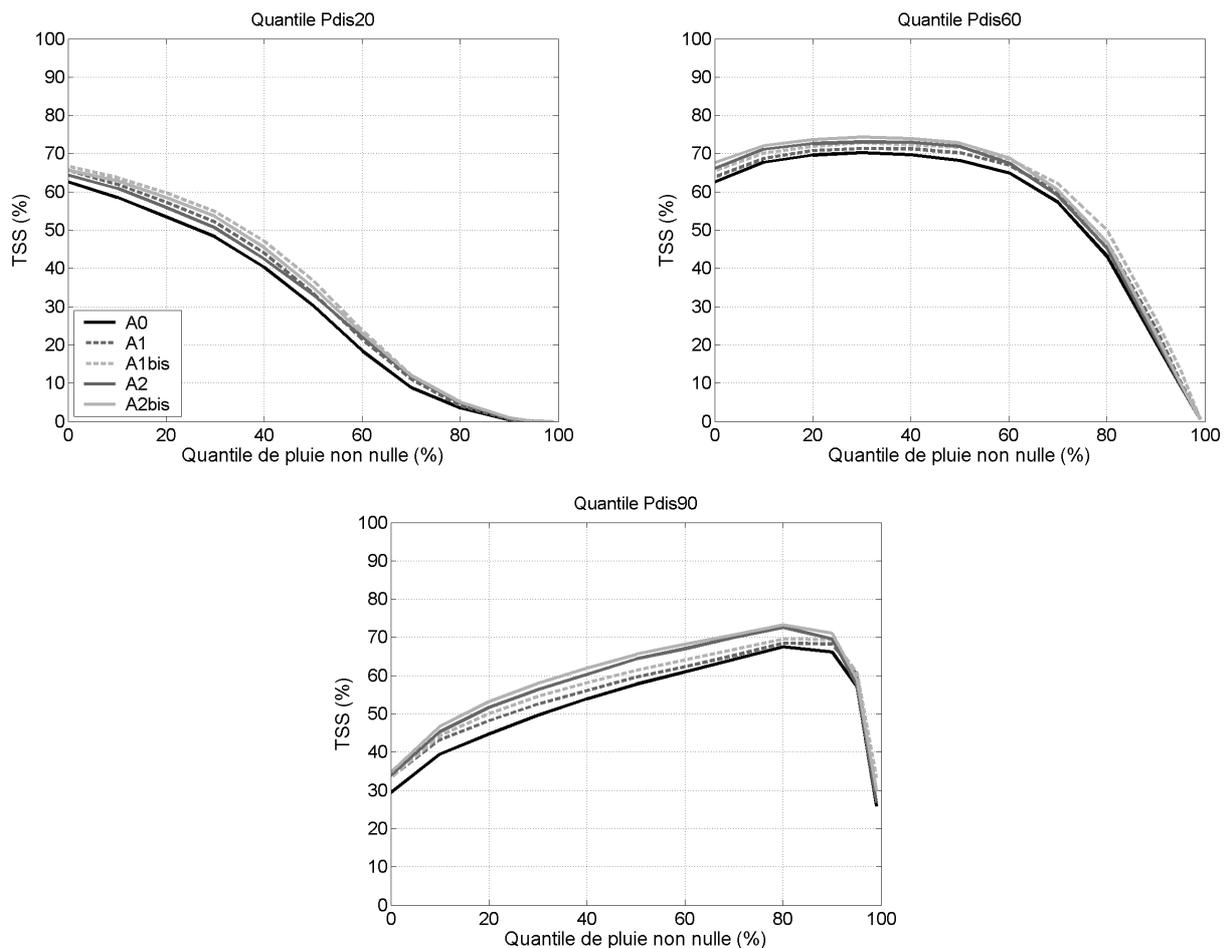


Figure 10.5 : Evolution du score TSS (%) obtenu avec chaque méthode d'analogie (A0, A1, A1bis, A2, A2bis) en fonction du seuil de précipitations, lorsque la prévision est définie par les quantiles au non-dépassement $Pdis20$, $Pdis60$ et $Pdis90$ d'une distribution prévue.

Taux de détection et fausses alertes

Jusqu'à présent, les outils d'évaluation mis en œuvre ont donné l'avantage à la dernière méthode développée A2bis. Pour autant, cette méthode permet-elle d'anticiper fréquemment des événements et génère-elle moins de fausses alertes que les autres méthodes ?

Nous nous intéressons aux scores POD et FAR définis dans le chapitre 3 (sections 3.2.1 et 3.2.3). Un système de prévision parfait n'émettrait jamais de fausse alerte et ne manquerait aucun événement. Mais en réalité, il est nécessaire de faire un compromis. Il faut évidemment que le taux de détection soit élevé pour les événements pluvieux, mais aussi que le taux de fausses alertes soit assez faible. Nous avons évalué les scores POD et FAR pour les différents seuils de précipitations, et en considérant comme prévision un des trois quantiles $Pdis20$, $Pdis60$ et $Pdis90$ d'une distribution empirique des pluies. La Figure 10.6 présente l'évolution des scores POD (gauche) et FAR (droite), évalués pour l'ensemble des sous-bassins de la Saône et de la Seine, en fonction du seuil de pluie.

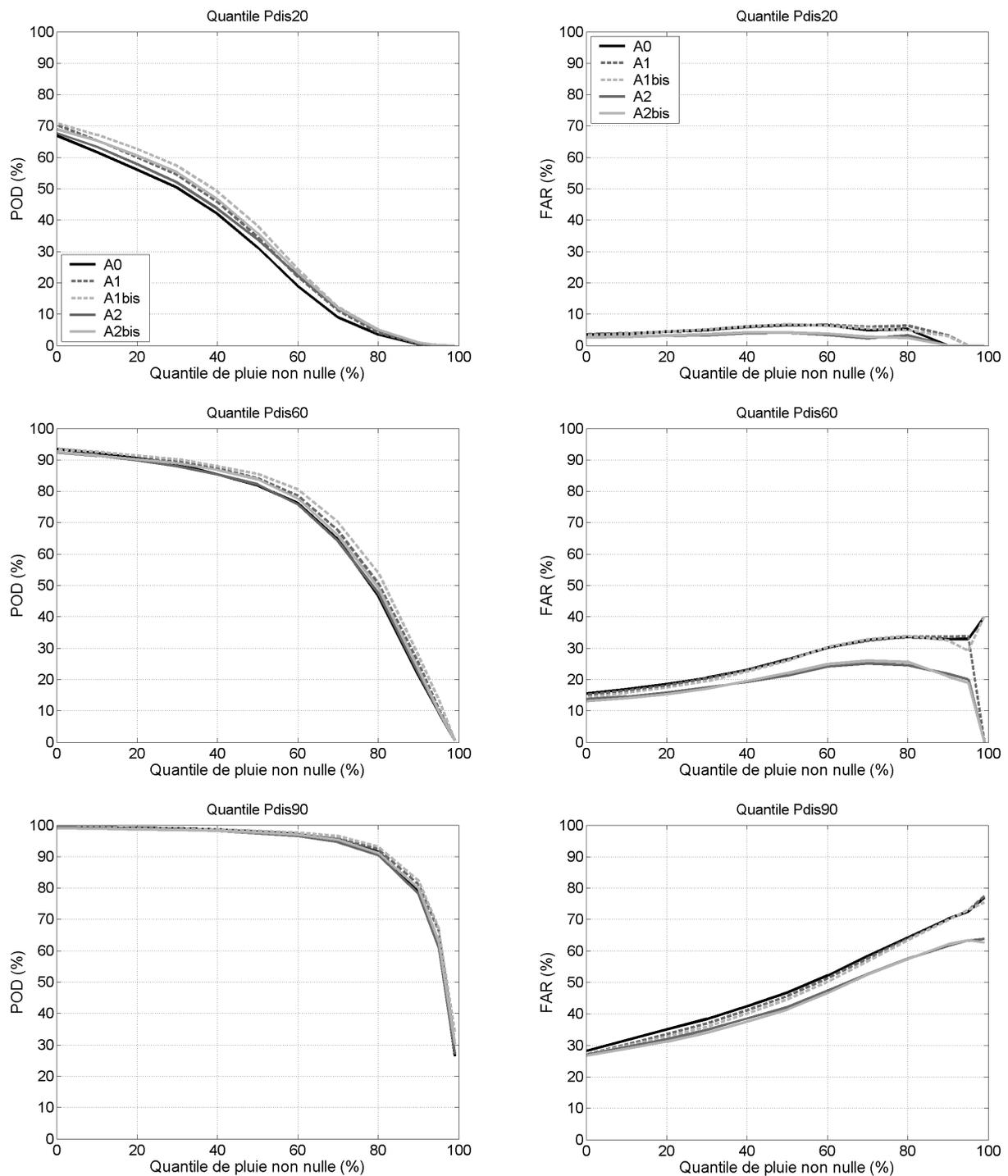


Figure 10.6 : Evolution des scores *POD* (gauche) et *FAR* (droite) évalués avec chacune des méthodes d'analogie (A0, A1, A1bis, A2, A2bis) en fonction du seuil de précipitations (quantile de pluie non nulle), lorsque la prévision est définie par les quantiles au non-dépassement *Pdis20*, *Pdis60* et *Pdis90* d'une distribution prévue.

Ces résultats révèlent des différences surprenantes entre les méthodes. En effet, en considérant comme prévision le quantile *Pdis20*, les graphiques de l'évolution du *POD* montrent que :

- la méthode à privilégier pour des précipitations inférieures au *PNN60* est A1bis ;

- pour les faibles précipitations (inférieures au seuil $PNN40$), la méthode A2 est équivalente à la méthode de référence A0.

En utilisant le quantile $Pdis60$ comme prévision, l'équivalence entre les méthodes A0 et A2 en termes de POD se retrouve pour tous les seuils de précipitation. La méthode A1bis apparaît plus adaptée à détecter les événements. Ce résultat indique que l'introduction de la vitesse verticale diminue la capacité de détection. En revanche, la sélection des situations candidates par la température semble l'améliorer sensiblement.

Les résultats pour le quantile de prévision $Pdis90$ sont similaires à ceux obtenus pour le quantile $Pdis60$.

Si on s'intéresse au score FAR , les résultats sont en revanche bien différents. Quel que soit l'événement pluvieux prévu et quel que soit le quantile utilisé pour la prévision, l'analogie sur la vitesse verticale permet de diminuer sensiblement le taux de fausses alertes, et ce d'autant plus que l'événement pluvieux est marqué. Les méthodes A2 et A2bis présentent des valeurs de score FAR équivalentes. Ce résultat est, de manière générale, aussi valable pour les méthodes A0, A1 et A1bis. D'autre part, on remarque que pour le score FAR , les écarts entre les courbes sont beaucoup plus importants que ceux obtenus pour le score POD . Les méthodes A2 et A2bis privilégient donc un taux de fausse alerte bien plus faible au détriment d'une probabilité de détection un peu moins élevée.

10.1.2.3 Des prévisions biaisées ?

La limitation de l'archive de situations candidates notamment induit un biais dans la prévision des événements importants, comme nous l'avons indiqué dans le chapitre 1 (section 1.4.1). Il serait donc utile de quantifier les biais induits par les méthodes développées. Pour ce faire, nous avons étudié le diagramme de Talagrand et le diagramme de fiabilité, introduits au chapitre 3 (section 3.3).

Diagramme de Talagrand

Nous avons construit un diagramme pour chaque algorithme et pour chaque sous-bassin. Leurs formes diffèrent quelque peu d'un bassin à un autre. Nous proposons d'analyser les diagrammes obtenus pour les bassins du Doubs et du Loing-E (Figure 10.7).

Les histogrammes obtenus sur les deux sous-bassin du Doubs et du Loing-E sont différents. Ceux associés au bassin du Doubs ont plutôt une forme en dôme (ou « n »), et présentent une légère dissymétrie, notamment pour les méthodes A2 et A2bis. Pour le bassin du Loing-E, tous les histogrammes sont fortement dissymétriques, ce qui traduit un biais dans la prévision. La fréquence est bien plus élevée pour les derniers rangs : la valeur observée se situe donc bien souvent dans la gamme supérieure des distributions, voire au-delà de la valeur la plus forte. Par conséquent, la prévision est en moyenne sous-estimée et n'est pas centrée sur l'observation.

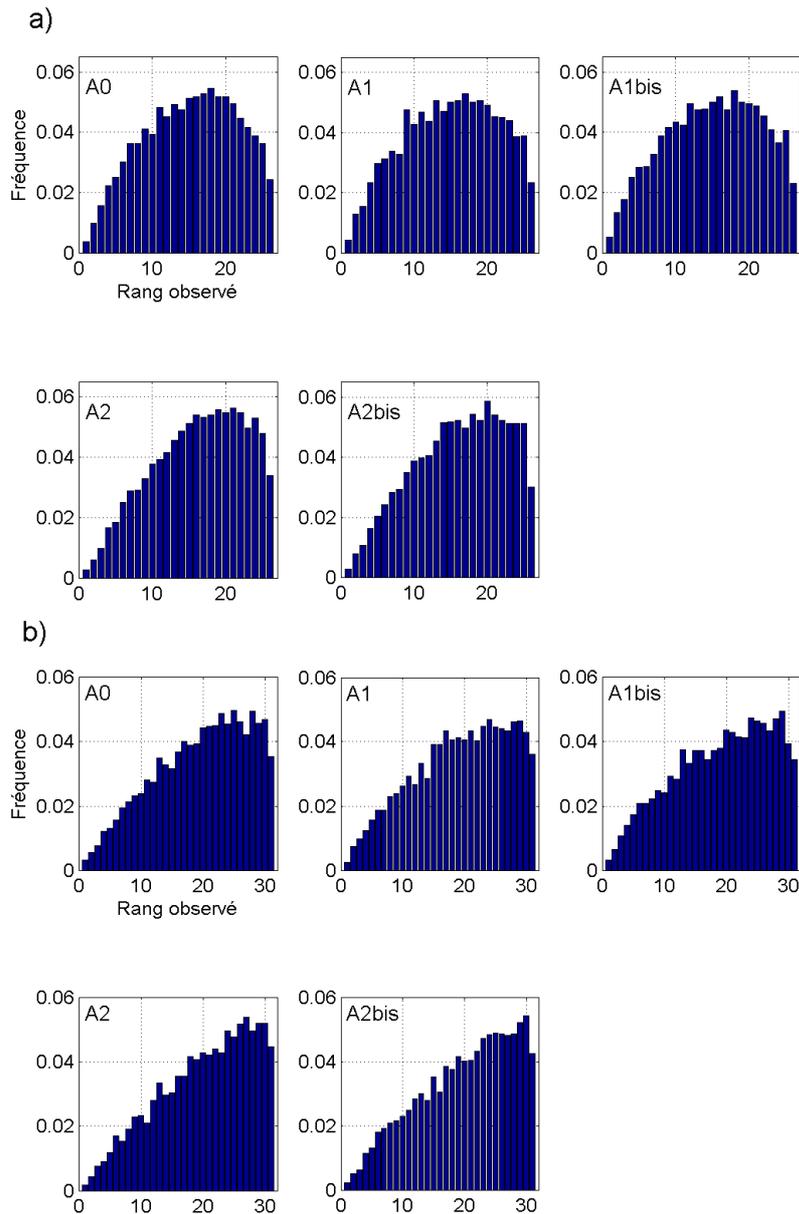


Figure 10.7 : Diagramme de Talagrand construit pour chaque méthode d'analogie (A0, A1, A1bis, A2, A2bis). a) bassin du Doubs ; b) bassin du Loing-E.

Pour les deux bassins, ce biais apparaît plus fort lorsque les méthodes A2 et A2bis sont appliquées. L'introduction de la vitesse verticale entraîne ainsi une sous-estimation plus fréquente des précipitations.

Diagramme de fiabilité

Les diagrammes de fiabilité, définis au chapitre 3 (section 3.3.2), ont été tracés pour chaque sous-bassin en considérant les N fréquences empiriques associées aux N valeurs formant la distribution des pluies. La Figure 10.8 présente, pour le bassin du Doubs, les diagrammes de fiabilité obtenus en appliquant les différentes méthodes de prévision, pour les événements supérieurs aux seuils 0 mm (pluie/non pluie), $PNN50$ et $PNN90$; assortis de leurs histogrammes des fréquences. Les diagrammes obtenus pour les autres sous-bassins sont

similaires.

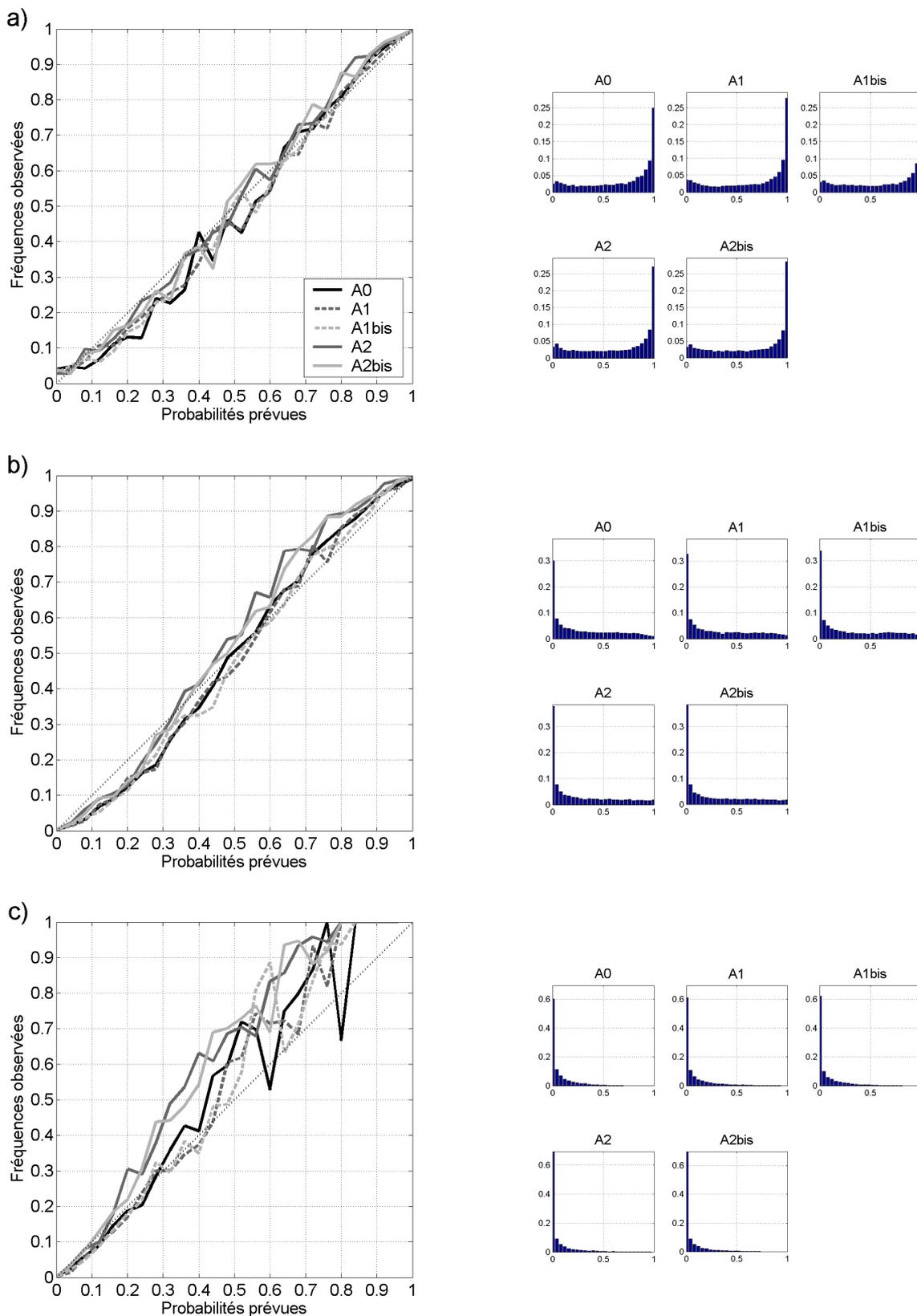


Figure 10.8 : Diagrammes de fiabilité assortis de leurs histogrammes des fréquences, construits pour les différentes méthodes d'analogie (A0, A1, A1bis, A2, A2bis). a) seuil 0 mm ; b) seuil *PNN50* ; c) seuil *PNN90*.

Si l'on considère des événements d'intensité faible ou modérée (seuils 0 mm et *PNN50*), les courbes sont très proches de la première bissectrice, ce qui traduit une fiabilité très satisfaisante, et ce quelle que soit la méthode d'analogie appliquée. De manière générale, les méthodes sont équivalentes pour ce type d'événement. Pour les faibles probabilités (*i.e.* pour les fortes fréquences au non-dépassement), les courbes se situent en-dessous de la première bissectrice, ce qui induit une sur-estimation de ces événements par ces valeurs de probabilité. Ce résultat est en particulier valable pour les méthodes A0 à A1bis. Au contraire, pour les fortes probabilités (*i.e.* pour les faibles fréquences au non-dépassement), les courbes se situent au-dessus pour les événements supérieurs au seuil *PNN50*, ce qui traduit une sous-estimation de ces événements par ces valeurs de probabilité. Par ailleurs, les courbes associées aux méthodes A2 et A2bis s'écartent plus sensiblement de la bissectrice : nous retrouvons ici la sous-estimation de ce type d'événement par ces méthodes développées.

Si nous nous intéressons aux événements plus forts, tels que ceux supérieurs au seuil *PNN90*, nous nous heurtons à un problème d'échantillonnage (Rousset Regimbeau, 2007). En effet, les fortes probabilités ne sont pas beaucoup utilisées pour un seuil élevé de précipitations tel que le seuil *PNN90*. Les diagrammes indiquent cependant un biais quasi-systématique pour ce type d'événement : les prévisions sont très souvent sous-estimées puisque les courbes se situent généralement au-dessus de la première bissectrice, en particulier pour les probabilités prévues supérieures à 0,4. La sous-estimation des faibles probabilités prévues (inférieures à 0,4) est à nouveau mise en évidence lorsque les méthodes A2 et A2bis sont appliquées ; tandis que la fiabilité est nettement meilleure pour ces probabilités lorsque les prévisions sont issues des méthodes A0 à A1bis (notamment A1 et A1bis).

Nous allons à présent discuter de l'ensemble des résultats présentés précédemment.

10.1.3 Discussion et conclusion

Dans cette section, les méthodes de sélection d'analogues développées dans la seconde partie ont été mises en œuvre dans un contexte de prévision parfaite sur la période la plus longue possible (01/08/1970-30/08/2002). Plusieurs scores ont été exploités afin d'évaluer de manière approfondie les prévisions de précipitations.

La décomposition du score global *CRPS* en finesse et en justesse a révélé un gain important en finesse des prévisions issues des méthodes développées par rapport à la méthode de référence A0, notamment lorsque la vitesse verticale est introduite (algorithme A2). En revanche, les gains en justesse sont moindres. Ceux apportés d'une part par la vitesse verticale et d'autre part par les nouveaux niveaux de pression s'avèrent indépendants. Ce résultat montre ainsi la pertinence de l'introduction de ces variables.

L'évaluation des performances de prévision en fonction du seuil de précipitations a permis de mettre en évidence plusieurs éléments. Selon les diagrammes ROC et le score *TSS*, qui tiennent compte à la fois du taux de détection et du taux de fausse détection, les méthodes A2

et A2bis sont les plus performantes si l'on considère comme prévision des quantiles de la distribution moyens ou hauts. Pour des événements pluvieux de faible ou de moyenne ampleur, la méthode A1bis s'avère la plus performante en termes de *TSS* si la prévision est définie par un faible quantile de la distribution (à l'image du quantile P_{dis20}). Si on s'intéresse séparément au taux de détection (*POD*) et au taux de fausses alertes (*FAR*), on comprend mieux les résultats obtenus par le score *TSS*. Contrairement à ce que l'on présageait initialement, l'analogie sur la vitesse verticale semble pénaliser la probabilité de détection d'événements pluvieux ; tandis que la méthode A1bis semble plus adaptée à la détection d'événements, notamment d'intensité faible à modérée. En revanche, l'introduction de la vitesse verticale dans l'algorithme permet de diminuer assez nettement le taux de fausses alertes. Par conséquent, il est normal que l'on obtienne de tels résultats sur les scores globaux (*G*, *AROC*), puisque ces derniers tiennent compte, finalement, de ces deux indicateurs. Le gain de performance apporté par les méthodes A2 et A2bis serait donc principalement dû à la diminution du taux de fausses alertes. L'utilisateur d'une méthode d'analogie comme celles évaluées ici se reportera donc sur l'une ou l'autre méthode en fonction de ses attentes.

Enfin, le biais des prévisions a pu être identifié à l'aide du diagramme de Talagrand et du diagramme de fiabilité. Les résultats diffèrent d'un bassin à un autre, mais de manière générale, deux éléments principaux sont à retenir :

- la prévision des événements forts est très souvent sous-estimée par les méthodes d'analogie. Cette sous-estimation est notamment importante lorsque la vitesse verticale est introduite, *i.e.* lorsque les méthodes A2/A2bis sont appliquées ;
- pour les faibles événements, les faibles quantiles d'une distribution des pluies peuvent également conduire à une sur-estimation, en particulier lorsque les méthodes A0 à A1bis sont appliquées.

D'autre part, la fiabilité, qui représente une des qualités d'une méthode prévision, s'avère généralement bonne pour des événements de faible ou de moyenne ampleur. Ces résultats rejoignent ceux obtenus lors de l'étude des scores *POD* et *FAR* : une sous-estimation se traduit bien par un taux de détection plus faible mais aussi un taux de fausses alertes plus faible. Bien que pour les forts événements, les pluies soient plus souvent sous-estimées lorsque l'analogie sur la vitesse verticale est ajoutée, il faut également garder à l'esprit que le taux de fausses alertes est plus faible.

10.2 Prévision des débits

Dans cette section, l'objectif est d'évaluer statistiquement les performances des systèmes formés par le couplage méthode d'analogie-modèle hydrologique. Pour générer les prévisions de débit, nous avons procédé de la façon décrite dans le chapitre précédent (section 9.2). Nous avons donc utilisé en entrée du modèle hydrologique GR3P les prévisions de pluie

obtenues par les différentes méthodes d'analogie dans un contexte de prévision parfaite.

10.2.1 Pertinence des scénarios obtenus par tirage aléatoire

Avant de diagnostiquer les résultats, nous nous sommes assurés que le tirage aléatoire est bien uniforme, c'est-à-dire qu'il respecte bien les distributions de pluie initiales des analogues. En théorie, la représentativité de la dynamique obtenue par une méthode d'analogie est d'autant meilleure que la distribution des pluies prévue est fine. Il est clair que si nos prévisions ne sont pas suffisamment fines, le risque de construire des scénarios incohérents avec la dynamique est plus élevé.

En appliquant chaque méthode sur chaque sous-bassin, le score *CRPS* a été évalué sur les distributions obtenues par tirage aléatoire pour la première échéance J et moyenné sur la période de prévision (02/08/1972-17/08/2002). Cette valeur a été comparée à la valeur moyenne du score *CRPS* évalué sur les distributions initiales générées sur la même période. Les valeurs obtenues sur le bassin de la Petite Saône sont indiquées dans le Tableau 10.4.

La qualité des distributions prévues construites de manière aléatoire est détériorée par rapport à celle des distributions initiales, avec des valeurs de *CRPS* plus élevées. Ce résultat est plutôt logique, puisqu'il s'agit d'un tirage avec remise. Néanmoins, la hiérarchie entre les algorithmes est toujours respectée, ce résultat étant valable pour l'ensemble des sous-bassins étudiés.

Tableau 10.4 : Valeurs moyennes du score *CRPS* évalué sur la période 02/08/1972-17/08/2002 pour le bassin de la Petite Saône, à partir des distributions initiales et de celles obtenues par tirage aléatoire.

	A0	A1	A1bis	A2	A2bis
Distributions initiales	1,16	1,13	1,09	1,06	1,03
Distributions aléatoires	1,21	1,17	1,14	1,09	1,07

Nous avons ensuite calculé la valeur moyenne sur les sous-bassins des gains G de performance obtenus par les méthodes A1 à A2bis par rapport à la méthode A0. Ces valeurs moyennes, calculées sur les distributions initiales et sur les distributions obtenues par le tirage aléatoire, sont reportées dans le Tableau 10.5. Les gains sont équivalents : par conséquent, le tirage aléatoire respecte globalement bien les distributions initiales des analogues.

Tableau 10.5 : Valeurs du gain G (%) de performance moyenne, apporté par les méthodes A1 à A2bis par rapport à la méthode A0, obtenues sur les distributions initiales et sur les distributions construites par tirage aléatoire.

	$G(A1/A0)$	$G(A1bis/A0)$	$G(A2/A0)$	$G(A2bis/A0)$
Distributions initiales	3,16	5,91	8,36	10,44
Distributions aléatoires	3,24	5,82	8,73	10,74

Néanmoins ce résultat n'assure pas que la hiérarchie entre les méthodes est identique pour des cumuls sur plusieurs jours. Ainsi nous avons calculé, en appliquant chaque méthode, les

cumuls pluviométriques obtenus sur sept jours pour chaque scénario d'un jour de prévision. Nous en avons déduit la valeur médiane et l'avons comparée à la valeur observée.

La Figure 10.9 illustre, pour le bassin de la Petite Saône (Saône à Le Chatelet), les nuages de points assortis de la droite de régression lorsque chaque méthode appliquée. La corrélation entre la médiane des valeurs de pluie prévue cumulée sur sept jours et la valeur observée est de plus en plus forte avec l'amélioration de la méthode de sélection. La hiérarchie est, à ce niveau, également respectée. On remarque cependant que le coefficient directeur de la droite est plus faible avec les méthodes A2 et A2bis qu'avec les méthodes A0 à A1bis, avec une valeur maximale obtenue par la méthode A1bis. Ceci traduit une sous-estimation moyenne de la pluie cumulée sur sept jours plus prononcée avec les deux dernières méthodes, cohérente avec les analyses menées sur la pluie journalière. On retrouve par ailleurs l'ensemble de ces résultats sur les autres sous-bassins.

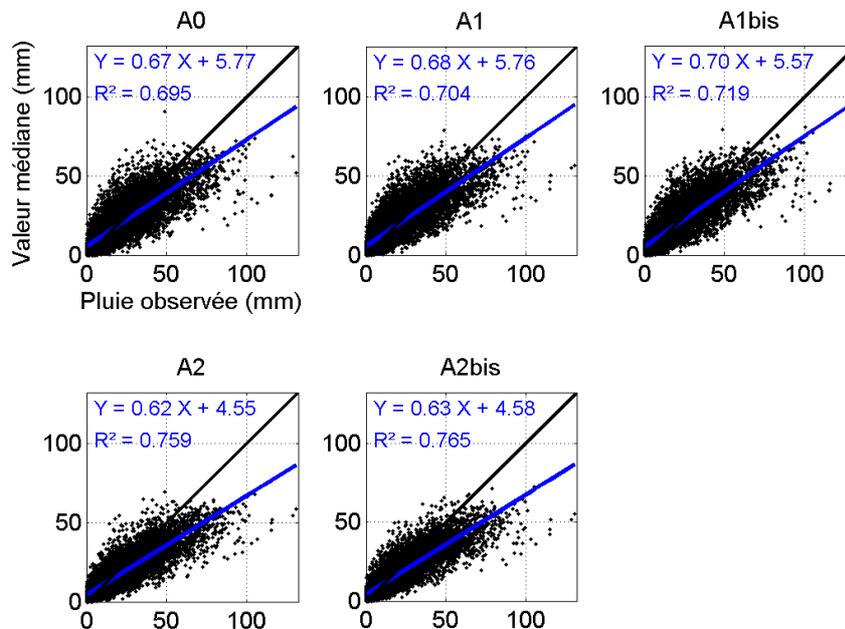


Figure 10.9 : Corrélation entre valeurs observées et valeurs médianes de la distribution des précipitations cumulées sur sept jours, lorsque chaque méthode d'analogie (A0, A1, A1bis, A2, A2bis) est appliquée au bassin de la Saône à Le Chatelet (droite tracée en bleu).

10.2.2 Evaluation des prévisions de débit

La procédure d'évaluation des prévisions hydrologiques est identique à celles des prévisions de précipitations. Dans un premier temps, nous allons évaluer globalement les prévisions à l'aide de la compétence *CRPSS* du *CRPS* par rapport à une prévision de référence. Nous avons défini dans le chapitre précédent deux prévisions de référence (section 9.3) avec lesquelles le score *CRPSS* peut être évalué : la persistance et la climatologie. Dans un second temps, nous reprendrons certains scores issus des tableaux de contingence, afin d'évaluer la capacité du modèle, lorsqu'il est alimenté en scénarios générés par les différentes

méthodes d'analogie, à anticiper des événements.

10.2.2.1 Comparaison à la persistance

Nous avons évalué, pour chaque sous-bassin, la valeur moyenne sur la période du score *CRPS* obtenue par les différentes méthodes, ainsi que celle obtenue par la persistance, et ce pour chacune des 14 échéances de prévision. La compétence *CRPSS* en a été déduite. La Figure 10.10 présente, pour chaque station de la Saône et de la Seine, l'évolution du *CRPSS* en fonction de l'échéance de prévision.

Nous constatons que, de manière générale, les scores *CRPSS* obtenus par les différentes méthodes sont assez proches. Les écarts sont logiquement de plus en plus importants avec l'augmentation de l'échéance de prévision. Les compétences sont largement positives pour chacune des échéances, atteignant des valeurs de l'ordre de 40 à 60 %, sauf pour le bassin du Loing à Episy à partir de $J+12$ (échéance de 13 jours) pour lequel les valeurs deviennent légèrement négatives. Nous pouvons donc conclure que la prévision de débits à partir des prévisions de pluie par analogie apporte bien plus d'information qu'une prévision naïve. Logiquement, nous retrouvons des valeurs très proches entre les méthodes pour les échéances de 1 à 3 ou 4 jours, même si les *CRPSS* obtenus par les méthodes A0 à A1bis sont très légèrement supérieurs à ceux évalués lorsque le modèle est alimenté par les méthodes A2 ou A2bis. Pourtant pour ces échéances, l'effet des pluies sur le débit calculé par le modèle est très faible.

Mais dès l'échéance de 4 jours, on observe des différences entre bassins. Pour le bassin du Doubs à Neublans, la prévision hydrologique obtenue par les méthodes A2 et A2bis est meilleure que celle obtenue par les méthodes A0 à A1bis pour les échéances allant de 3 à 8 jours. En revanche, pour d'autres bassins, la hiérarchie est différente. Bien souvent, la méthode A1bis fournit les meilleures prévisions hydrologiques. Ce résultat se retrouve à partir de 9 jours d'échéance pour l'ensemble des bassins. Il s'explique par la raison suivante. Les cumuls de pluie injectés dans le modèle à partir de l'échéance de 8 jours ($J+7$) sont nuls. L'effet de ces pluies nulles sur les débits prévus est en moyenne détecté à partir du lendemain ou du surlendemain, comme nous l'avons vu précédemment avec le critère de persistance *C4* (section 9.1.3.2). Or nous avons remarqué que les pluies sont plus souvent surestimées par les méthodes A0 à A1bis, avec un taux de fausses alertes plus élevé (section 10.1.2.2). Il est donc fort probable que leur impact sur les débits prévus soit détecté sur les échéances suivantes. Le niveau du réservoir du modèle hydrologique est alors, en moyenne, pour les dernières échéances, plus élevé lorsque les scénarios de pluie sont générés par les méthodes A0 à A1bis que par les méthodes A2 et A2bis.

Par ailleurs, l'échéance de prévision pour laquelle la compétence est maximale est de 9 ou 10 jours pour chacune des méthodes et pour l'ensemble des bassins, à l'exception du Loing à Episy. Au-delà de 9 jours d'échéance, la combinaison des erreurs de prévision des méthodes d'analogie, des pluies nulles injectées et des erreurs du modèle hydrologique induit une baisse progressive de la compétence. Ceci signifie que les pluies prévues apportent de

l'information au modèle durant les deux jours qui suivent la dernière échéance de prévision de pluie (7 jours), conformément à l'effet mémoire du bassin. La rupture un peu plus tardive pour le bassin de la Saône à Couzon est réaliste, indiquant un temps de réaction plus long.

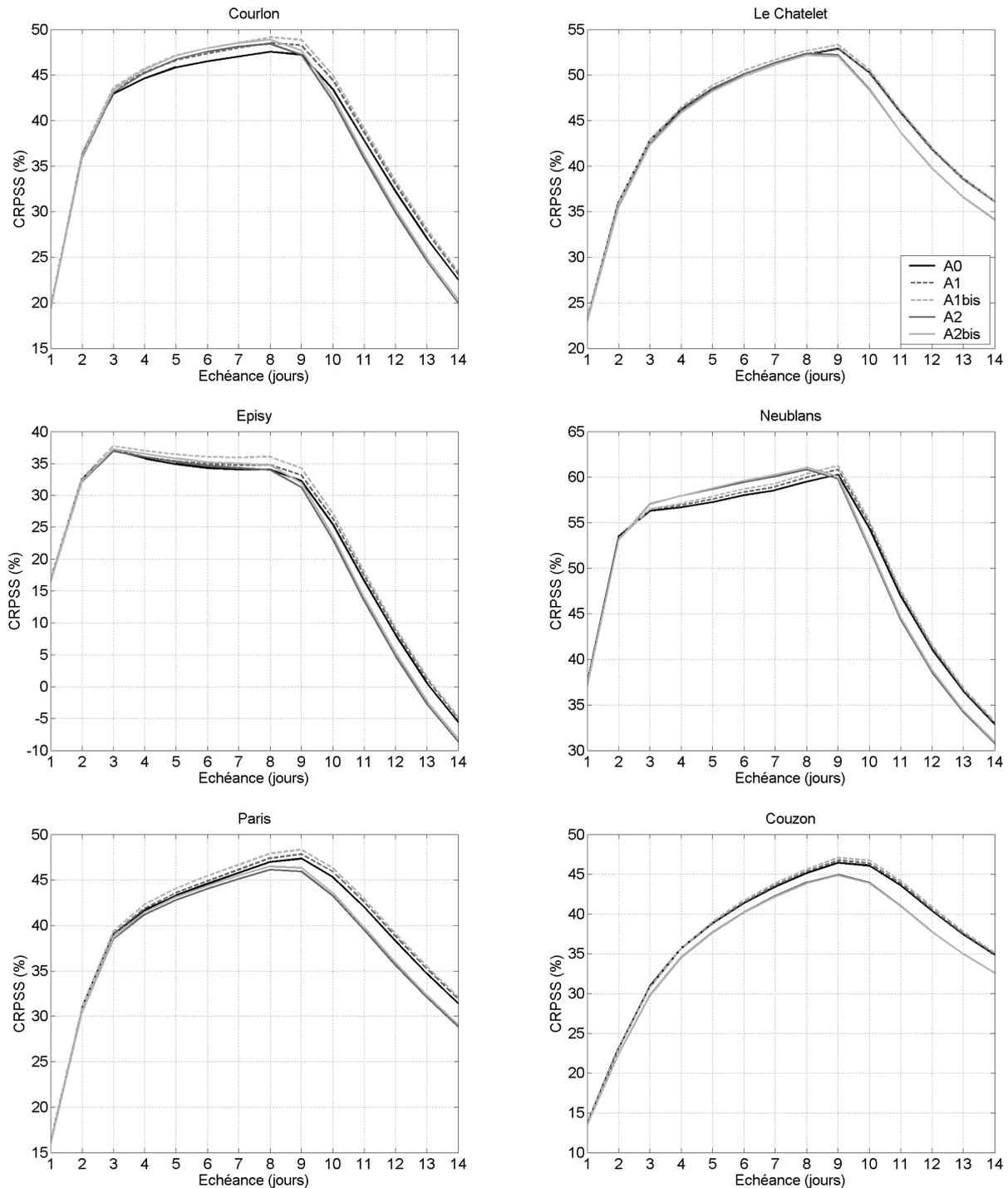


Figure 10.10 : Evolution du score CRPSS (%) en fonction de l'échéance de prévision, pour chaque station retenue, lorsque le modèle GR3P est alimenté en scénarios issus de chacune des méthodes d'analogie (A0, A1, A1bis, A2, A2bis).

10.2.2.2 Comparaison à la climatologie

Pour chaque bassin, nous avons évalué, pour les sept premières échéances (J à $J+6$), la valeur mensuelle moyenne du score $CRPS$ des prévisions de débit générées par chacune des méthodes, ainsi que la valeur mensuelle moyenne du $CRPS$ de la prévision naïve correspondant à la référence climatologique. Nous en avons déduit la compétence $CRPSS$ pour chaque mois de l'année. Les résultats obtenus avec les différentes méthodes sont très similaires. La Figure 10.11 a) présente, pour le bassin de l'Yonne à Courlon, l'évolution du $CRPSS$ en fonction du mois de l'année, pour chacune des échéances, lorsque l'algorithme A0 de sélection d'analogues est appliqué.

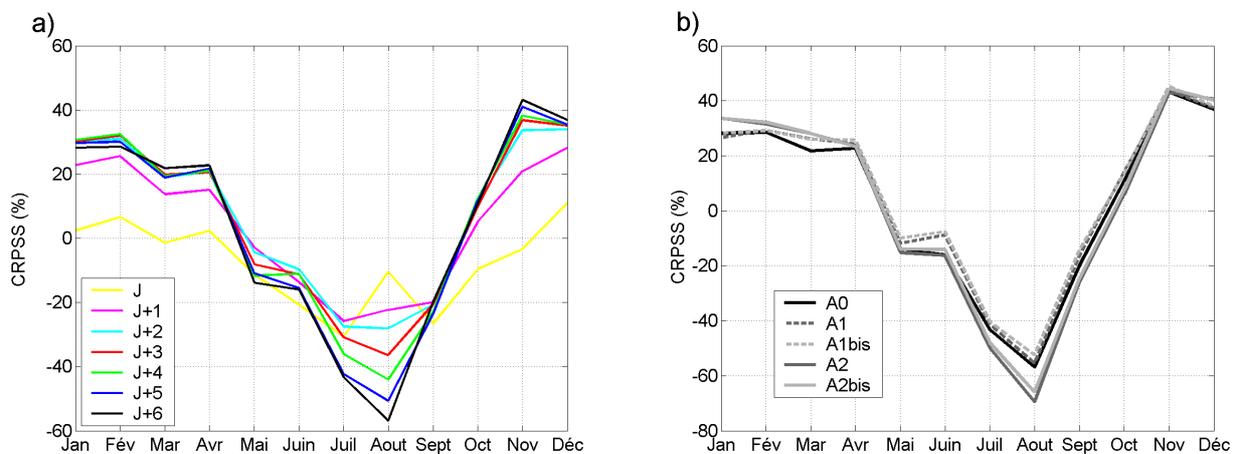


Figure 10.11 : Evolutions du score $CRPSS$ (%) en fonction du mois de l'année, pour le bassin de l'Yonne à Courlon. a) pour chacune des sept premières échéances de prévision (J à $J+6$) lorsque le modèle GR3P est alimenté par la méthode A0 ; b) pour l'échéance $J+6$ lorsque le modèle est alimenté par les différentes méthodes d'analogie (A0, A1, A1bis, A2, A2bis).

L'échéance à partir de laquelle les pluies ont un effet sur le débit est $J+2$ pour le bassin de l'Yonne à Courlon (cf. critère de persistance). Le comportement du score $CRPSS$ suivant l'échéance provient essentiellement de la prévision naïve. En hiver et au printemps, à partir de l'échéance $J+2$, la prévision issue du modèle s'avère plus utile que la prévision naïve puisque l'évolution du débit dépend des pluies prévues, bien qu'elles soient entachées d'erreurs. Le score $CRPSS$ est alors maximal à mi-échéance (vers $J+3$ ou $J+4$), tandis que le score associé aux premières échéances est plus faible, avec des valeurs proches de 0. L'auto-corrélation est en effet forte sur ces bassins ; la modélisation pluie-débit a des difficultés à faire mieux qu'une prévision naïve (débit proche du débit observé la veille du jour de prévision) pour des bassins au temps de réaction de l'ordre de la journée ou de quelques jours.

La compétence est également négative durant la fin du printemps et l'été (de mai à septembre), quelle que soit l'échéance. En effet, durant la période d'étiage, la variabilité du débit est faible. Par conséquent, la prévision climatologique est plus fiable qu'une prévision basée sur des scénarios de précipitations. En outre, les prévisions obtenues par analogie sont généralement moins performantes en été, saison durant laquelle les précipitations sont

fréquemment d'origine convective.

A courte échéance ainsi que durant l'été (mai à septembre), la prévision naïve apporte donc plus d'information que la prévision émise par le modèle hydrologique alimenté par les scénarios issus de nos méthodes de prévision.

Pour comparer les résultats obtenus avec les différentes méthodes, nous avons représenté sur la Figure 10.11 b) l'évolution des scores *CRPSS* en fonction du mois de l'année, pour l'échéance $J+6$. Les différences entre les méthodes sont finalement assez faibles, même si nous remarquons que les performances de prévision sont plus élevées avec les méthodes A2 et A2bis en hiver, tandis qu'elles sont plus faibles à la fin du printemps et en été.

10.2.2.3 Scores issus des tableaux de contingence

De la même manière que pour les pluies, les prévisions probabilistes de débit ont été évaluées à l'aide des scores issus des tableaux de contingence. Nous avons défini onze seuils de débit correspondant aux quantiles $Q10$ à $Q90$ (d'ordre 10 % à 90 % par pas de 10 %) de la distribution des débits observés sur la période couverte par l'archive hydrométrique, et aux quantiles $Q95$ et $Q99$ (d'ordre 95 % et 99 %). Pour chacun de ces seuils, les neuf quantiles $Qdis10$ à $Qdis90$ d'une distribution prévue ont été considérés comme prévision pour calculer les scores *POD*, *POFD*, *TSS* et *FAR*. Nous avons étudié ces scores pour les dix premières échéances de prévision (J à $J+9$).

Par souci de simplicité et pour éviter d'analyser une quantité trop importante de graphiques, nous avons décidé d'évaluer les scores pour l'ensemble des bassins. Puisque l'impact sur les débits prévus des pluies générées par les différentes méthodes n'est perceptible qu'à partir du troisième jour, nous n'étudierons pas les deux premières échéances de prévision (J et $J+1$).

Score de Hanssen-Kuipers (*TSS*)

Dans un premier temps, nous nous intéressons à la dernière échéance ($J+9$) pour laquelle l'intérêt d'utiliser des prévisions de pluie est maximale. La Figure 10.12 présente l'évolution du score *TSS* évalué pour l'ensemble des bassins lorsque les quantiles $Qdis10$, $Qdis50$ et $Qdis90$ sont utilisés comme prévision.

Les différences de score obtenues pour la prévision de débit sont moins perceptibles entre les méthodes, ce qui rend difficile la lecture de ces graphiques. Lorsque la prévision est définie par le quantile $Qdis10$, les valeurs de *TSS* sont équivalentes entre les méthodes. En revanche pour le quantile médian, nous commençons à percevoir des écarts. Pour les bas débits, les prévisions issues des méthode A1bis et A2bis présentent des valeurs de *TSS* très légèrement supérieures aux autres. Pour des débits supérieurs au débit seuil $Q30$, les valeurs de *TSS* obtenues par la méthode A1bis sont plus élevées. A partir du débit seuil $Q50$, nous retrouvons des résultats similaires à ceux obtenus sur les pluies. Les prévisions de débit obtenues par les méthodes A0 à A1bis se révèlent plus performantes en termes de *TSS*. Pour

le quantile de prévision Q_{dis90} , la supériorité de la prévision obtenue par A2bis est plus prononcée pour les bas débits, tandis que la tendance ne s'inverse qu'au-delà du débit seuil Q_{70} .

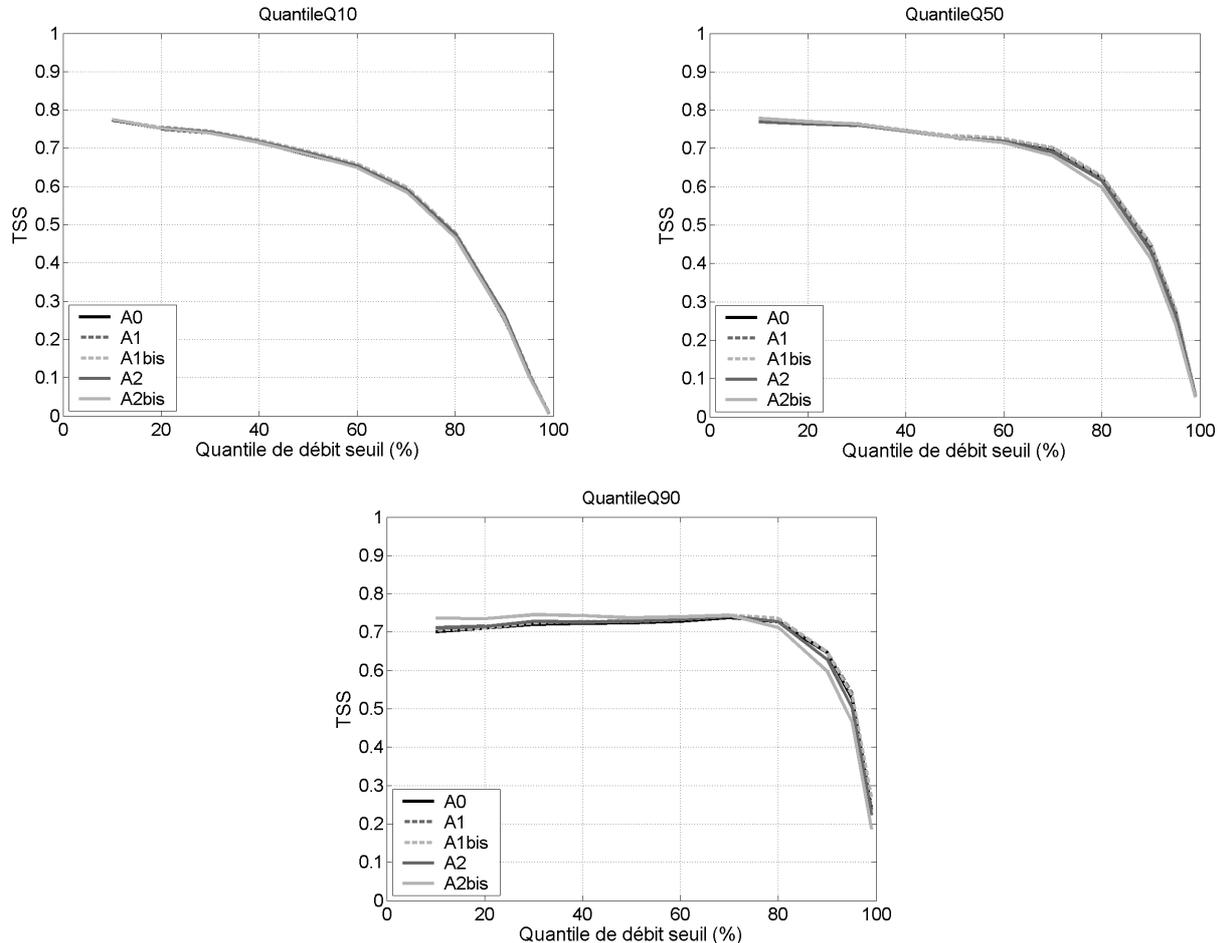


Figure 10.12 : Evolution du score TSS en fonction du quantile de débit seuil pour l'échéance $J+9$, lorsque la prévision est définie par les quantiles Q_{dis10} , Q_{dis50} et Q_{dis90} d'une distribution empirique.

Mais finalement, le prévisionniste qui utilise un tel système de prévision est en droit de se demander quel quantile de débit est le mieux adapté à la prévision d'un événement donné. Nous avons retenu la méthode A1bis et nous nous intéressons toujours à l'échéance $J+9$. L'évolution du score TSS a été tracée pour chaque quantile de prévision en fonction du seuil de débit (Figure 10.13).

Ainsi le quantile le plus adapté n'est pas le même selon l'événement considéré. Pour les très bas débits, le quantile Q_{dis40} présente la valeur de TSS la plus élevée. Le quantile Q_{dis60} se révèle plus intéressant pour les bas débits proches du quantile Q_{30} . Pour les débits moyens et hauts, les quantiles les plus élevés d'une distribution sont les plus adaptés. Pour des débits supérieurs au débit seuil Q_{70} , le quantile Q_{dis90} présente le meilleur score TSS . Ce résultat traduit directement la sous-estimation moyenne des précipitations importantes par une technique de prévision par analogie.

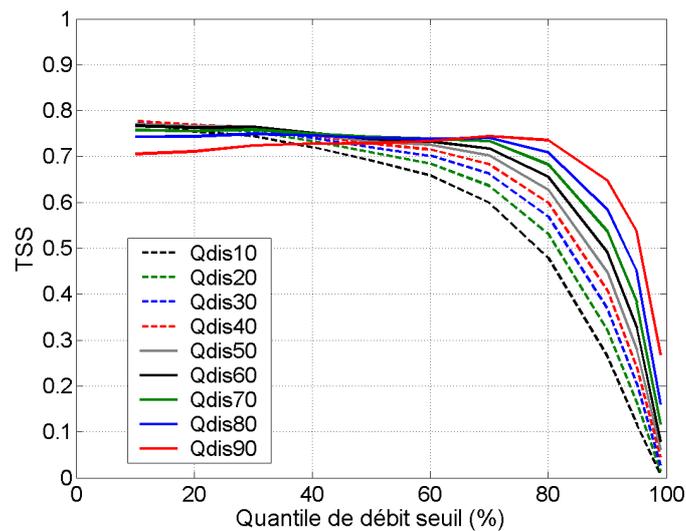


Figure 10.13 : Evolution en fonction du quantile de débit seuil du score *TSS* évalué lorsque la prévision, obtenue par la méthode *A1bis*, est définie par chacun des neuf quantiles de débit d'une distribution empirique.

Taux de détection et fausses alertes

Si nous nous intéressons aux scores *POD* et *FAR* associés au quantile de prévision *Qdis90* (Figure 10.14), nous retrouvons exactement les mêmes résultats que sur les pluies. Les valeurs du score *POD* obtenues par la méthode *A2bis* sont systématiquement plus faibles que celles obtenues par les autres méthodes, et c'est avec la méthode *A1bis* que les événements sont les mieux détectés. De même, les taux de fausses alertes sont plus élevés avec *A0* à *A1bis* qu'avec les méthodes *A2* et *A2bis*, en particulier pour les forts débits.

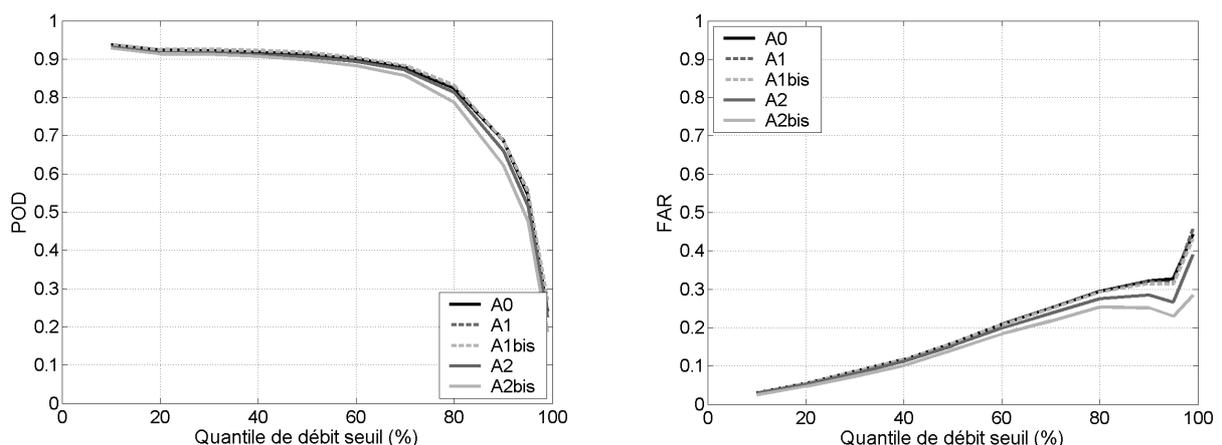


Figure 10.14 : Evolution des scores *POD* et *FAR* en fonction du quantile de débit seuil, pour l'échéance de prévision *J+9*, lorsque les scénarios de précipitations sont générés avec chacune des méthodes d'analogie (*A0*, *A1*, *A1bis*, *A2*, *A2bis*).

10.2.2.4 Evaluation des scénarios associés aux quantiles des distributions des analogues

Dans le chapitre précédent, nous avons évoqué la construction de scénarios de débit associés aux quantiles $Pdis20$, $Pdis60$ et $Pdis90$ des distributions empiriques des pluies (section 9.2). Bien que ces scénarios ne rendent pas nécessairement compte des incertitudes sur les pluies puisqu'ils sont relativement lissés, nous avons tout de même calculé, pour chaque méthode et chaque bassin, les écarts entre les débits observés et les débits prévus à partir des scénarios $Pdis20$, $Pdis60$ et $Pdis90$ des distributions des analogues. La Figure 10.15 présente, a) pour le bassin du Doubs à Neublans et b) pour celui de la Saône à Le Chatelet, l'évolution des valeurs de $RMSE$ (mm) en fonction de l'échéance lorsque chacune des méthodes est appliquée. Sur ces bassins, nous avons remarqué que la hiérarchie en termes de débit entre les différentes méthodes n'était pas identique (Figure 10.10). Nous retrouvons ici cette même hiérarchie pour les scénarios de débit obtenus par les quantiles $Pdis20$ et $Pdis60$. Ainsi pour le Doubs à Neublans, ce sont avec les méthodes A2 et A2bis que les valeurs de $RMSE$ sont les plus faibles ; alors que pour la Saône à Le Chatelet, la méthode la plus performante en termes de $RMSE$ est A1bis. Néanmoins pour ce bassin, les prévisions les plus dégradées sont obtenues par la méthode A0, ce qui n'était pas le cas pour les prévisions de débit générées par les scénarios de pluie construits par tirage aléatoire.

D'autre part, pour le scénario de débit obtenu avec le quantile $Pdis90$, les écarts entre les méthodes A0 à A1bis et les méthodes A2/A2bis se creusent très nettement avec l'augmentation de l'échéance. Or nous avons vu, par l'analyse des précipitations prévues, que le taux de fausses alertes est plus élevé avec les méthodes sans analogie sur la vitesse verticale. Certains débits étant largement surestimés avec le quantile $Pdis90$, il n'est donc pas étonnant d'obtenir de tels écarts de $RMSE$ puisque ce score donne plus de poids aux écarts importants.

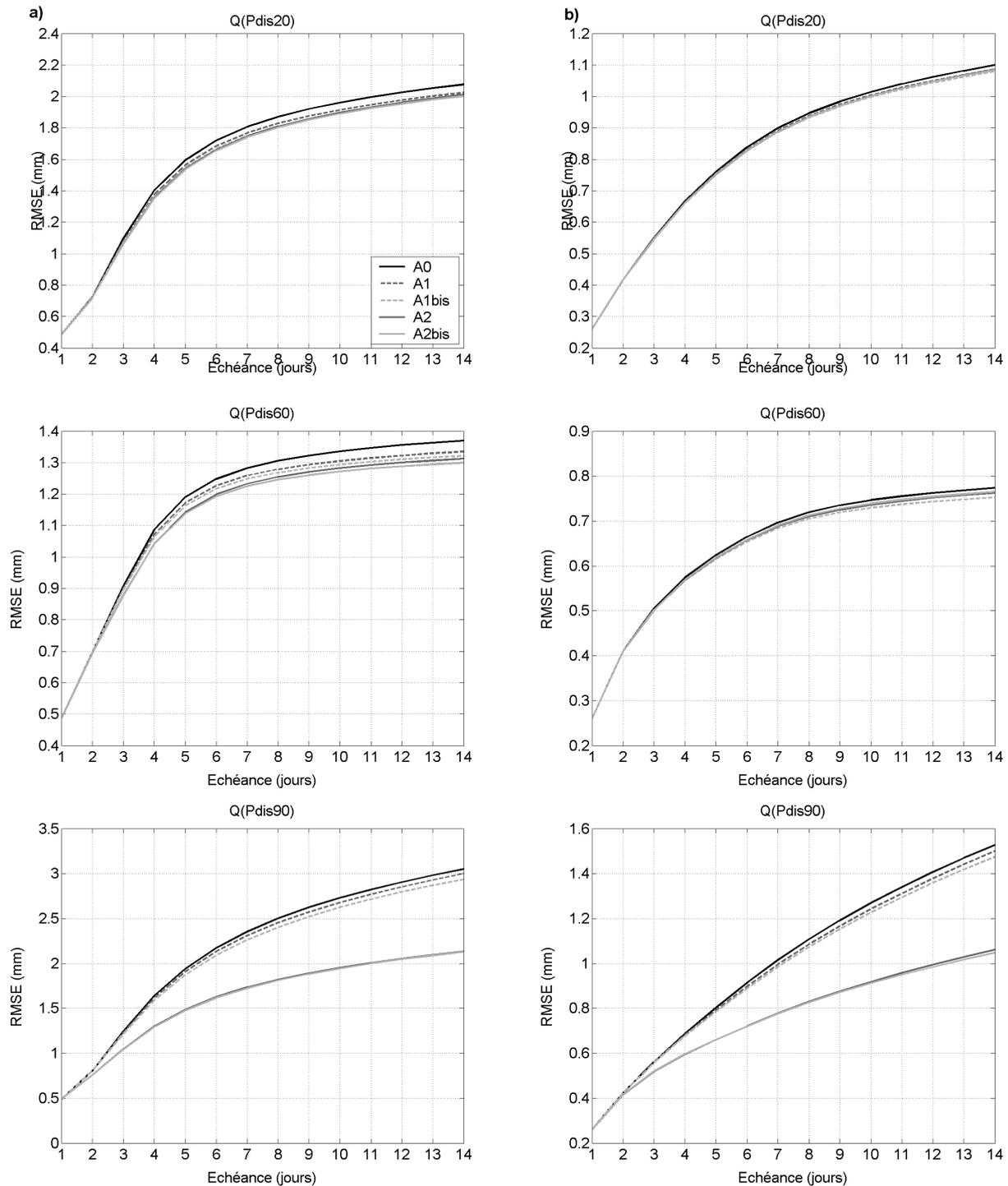


Figure 10.15 : Evolution des erreurs $RMSE$ entre le débit observé et les débits $Q(Pdis20)$, $Q(Pdis60)$ et $Q(Pdis90)$ prévus à partir des scénarios $Pdis20$, $Pdis60$ et $Pdis90$ des distributions des analogues, générées par chacune des méthodes d'analogie (A0, A1, A1bis, A2, A2bis). a) bassin du Doubs à Neublans ; b) bassin de la Saône à Le Chatelet.

10.2.3 Conclusion

Dans la seconde section de ce chapitre, nous avons généré les prévisions probabilistes de débit au droit de six stations des bassins de la Saône et de la Seine. Nous nous sommes d'abord assurés que le tirage aléatoire respecte bien les distributions initiales fournies par les

méthodes d'analogie. Ensuite les prévisions hydrologiques ont été évaluées à l'aide de scores identiques à ceux utilisés pour les prévisions de précipitations.

A l'aide du score *CRPSS*, nous avons comparé les prévisions obtenues à des prévisions de référence, définies dans le chapitre 9 (section 9.3) : la persistance et la climatologie. Ainsi, la prévision de débits à partir d'une méthode d'analogie s'avère bien plus utile qu'une prévision naïve au-delà de quelques jours, principalement pendant la saison des hautes eaux. A partir d'une échéance de quatre jours, des différences assez marquées ont été observées sur certains bassins, avec une hiérarchie assez aléatoire entre les méthodes. Cependant, les différences de performance observées entre les méthodes sont faibles : le modèle hydrologique semble ainsi amortir l'impact des différences de prévision de pluie sur les prévisions de débit, même pour une échéance de l'ordre de la semaine.

L'analyse des scores issus des tableaux de contingence n'a pas permis de révéler des différences marquées pour la même raison que précédemment. Seuls les quantiles élevés d'une distribution probabiliste de débit permettent de percevoir des nuances. Pour les débits bas et moyens, les prévisions hydrologiques à partir des scénarios de pluie générés par la méthode A2bis sont les plus performantes en termes de *TSS* ; tandis que cette tendance s'inverse pour des débits plus élevés. L'examen des scores *POD* et *FAR* a mené aux mêmes conclusions que pour les pluies : le taux de détection d'un événement est systématiquement plus faible lorsque les scénarios de pluie injectés dans le modèle sont générés par la méthode A2bis ; mais les fausses alertes sont plus rares lorsque les scénarios sont fournis par une méthode avec analogie sur la vitesse verticale, et ce surtout pour les forts débits.

Par ailleurs, nous avons étudié la sensibilité des performances mesurées par le score *TSS* au choix d'un quantile de débit particulier. Nous avons vu que plus l'événement est faible, plus le quantile au non-dépassement à retenir est faible. Ainsi pour des événements importants comme les crues, le quantile de prévision *Qdis90* s'avère le plus adapté.

Enfin, nous avons également évalué, à l'aide du score *RMSE*, les scénarios $Q(Pdis20)$, $Q(Pdis60)$ et $Q(Pdis90)$ définis par les quantiles *Pdis20*, *Pdis60* et *Pdis90* d'une distribution des analogues. Les résultats sont globalement similaires à ceux obtenus lorsque les prévisions de débit sont générées à partir de scénarios construits par tirage aléatoire, à quelques nuances près.

11 Dans un contexte de prévision opérationnelle

La chaîne de prévision, présentée au cours des deux précédents chapitres, a pour but de fonctionner en conditions réelles, et c'est bien ce qui intéresse particulièrement l'utilisateur quotidien potentiel. Jusqu'à présent, des diagnostics de performance ont été établis dans un contexte de prévision parfaite : les situations cibles étaient décrites par des champs météorologiques réanalysés, c'est-à-dire que les variables d'analogie n'étaient pas prévues par un modèle météorologique, comme c'est le cas en contexte opérationnel. Or dans ce dernier cas, les champs prévus sont entachés d'incertitudes : plus l'échéance de prévision augmente, plus les erreurs de prévision augmentent également.

L'objectif principal de ce chapitre est d'analyser le comportement de la chaîne de prévision lorsqu'elle est alimentée en scénarios de prévision opérationnelle des précipitations. Ceci nous permettra de mettre en évidence la dégradation de la prévision des variables d'analogie et d'évaluer son impact d'une part sur les performances de la prévision des précipitations adaptée par analogie et d'autre part sur celles de la prévision des débits. Nous chercherons notamment à évaluer la part des différentes sources d'incertitude de prévision (MNP, méthode d'analogie, modèle hydrologique) dans l'erreur finalement observée dans la prévision probabiliste de débit.

Pour répondre à ces objectifs, nous nous placerons dans un contexte de prévision opérationnelle : il est donc nécessaire de disposer des champs des variables d'analogie produits par un MNP opérationnel. Dans un premier temps, nous évoquerons les choix préliminaires et les contraintes sur le modèle météorologique et la période d'étude. Nous évaluerons ensuite, à l'aide des scores présentés dans le chapitre 4, la qualité des prévisions. Enfin, nous terminerons ce chapitre par une étude de cas afin d'illustrer, en conditions opérationnelles, la capacité de la chaîne à prévoir des dépassements de seuils.

11.1 Choix préliminaires

11.1.1 Période d'étude et modèle de prévision

Dans le cadre de la thèse, le CEPMMT nous a permis d'accéder aux prévisions archivées du MNP opérationnel, modèle nommé CEP par la suite. Ces archives sont limitées en durée pour certaines variables, ce qui limite leur usage dans nos travaux. Par exemple, la variable *TCW* n'est produite que depuis la fin des années 1990. De plus, le modèle est initialisé deux fois par jour (00 h TU et 12 h TU) depuis le début des années 2000, alors qu'avant le modèle n'était initialisé qu'à 12h TU. Si nous désirons émettre une prévision datée de 1999, la prévision de la variable *W850* à +6h le jour *J* n'est donc pas disponible. Par ailleurs, la structure du MNP n'a cessé d'évoluer au cours du temps. La résolution du système de prévision du CEP a ainsi été améliorée à plusieurs reprises depuis la création du système opérationnel¹ (Tableau 11.1).

Tableau 11.1 : Principaux changements dans la résolution du système opérationnel de prévision du centre européen CEPMMT, ayant eu lieu depuis la création du premier système en 1985.

Année du changement	Troncature	Nombre de niveaux verticaux
1987	106	19
1993	213	31
2000	319	60
2001	511	60
2006	799	91

Pour les raisons évoquées précédemment, nous avons dû restreindre notre étude à une période plus récente et suffisamment longue, du 01/10/2001 au 01/10/2004 (trois ans), pour permettre un traitement statistique pertinent mais aussi un retour d'expérience sur quelques événements importants.

Dans le cadre de cette étude, le modèle CEP est initialisé à 00h TU. Pour chaque jour de la période, l'échéance maximale de prévision des variables d'analogie a été fixée à +204h, ce qui permet d'émettre une prévision probabiliste des précipitations sur les bassins d'intérêt, des échéances *J* à *J*+7.

11.1.2 Disponibilité des variables d'analogie

Toutes les variables utilisées dans les méthodes d'analogie retenues sont disponibles sur cette période, à l'exception de la température au niveau de pression 600 hPa, utilisée au cours du niveau pré-sélectif des situations candidates. En revanche, la prévision de température est disponible aux niveaux 700 hPa et 500 hPa.

Afin d'appliquer les méthodes A1 à A2bis, il existe deux solutions alternatives :

- recalculer la température au niveau 600 hPa à l'aide d'autres variables disponibles à ce

¹ La liste des changements du système opérationnel de prévision du CEPMMT est consultable à l'adresse suivante : http://www.ecmwf.int/products/data/operational_system/evolution/index.html.

niveau et à l'aide des relations thermodynamiques qui les lient ;

- la seconde est de mesurer la température au niveau inférieur ou au niveau supérieur.

Pour une raison de simplicité de mise en œuvre, nous avons opté pour la seconde solution. Afin de faire le meilleur choix, les performances obtenues par les algorithmes A1 à A2bis ont été évaluées sur la période 1970-2002 lorsque la température est mesurée à 700 hPa à +12h d'une part, et à 500 hPa à +12h d'autre part (Tableau 11.2), au lieu de 600 hPa à +12h. La performance est ici exprimée par la valeur moyenne des scores *CRPSS* évalués pour chacun des huit sous-bassins de la Saône et de la Seine. A titre de comparaison, nous avons également indiqué les *CRPSS* obtenus initialement avec la température mesurée à 600 hPa.

Tableau 11.2 : Performances moyennes (*CRPSS* (%)) obtenues par les méthodes A1, A1bis, A2 et A2bis ; lorsque la température est mesurée à 700 hPa ou à 500 hPa à +12h au lieu de 600 hPa à +12h. Le score *CRPSS* est évalué pour l'ensemble des huit sous-bassins tests.

Méthode	Niveau de pression de la température à +12h		
	700 hPa	500 hPa	600 hPa
A1	50,10	50,06	50,16
A1bis	51,57	51,56	51,61
A2	52,51	52,45	52,57
A2bis	53,73	53,66	53,75

Quelle que soit la méthode appliquée, les meilleures performances sont obtenues en mesurant la température au niveau de pression 700 hPa. Dans ce chapitre, la température sera donc mesurée à 700 hPa à +12h au niveau de pré-sélection des situations candidates, en complément de la température à 925 hPa à +36h.

11.2 Prévision des précipitations

Pour chaque jour de la période 01/10/2001-01/10/2004, les prévisions de précipitations adaptées par analogie ont été générées jusqu'à une échéance de huit jours ($J+7$). Dans cette section, l'objectif est d'évaluer les prévisions sur toute la période à partir des mêmes scores que ceux utilisés en prévision parfaite.

Afin de comparer les scores obtenus pour chaque échéance de prévision, le même échantillon de situations cibles doit être traité pour chaque échéance. Etant donné que l'échéance maximale de prévision est huit jours, la période cible sur laquelle les prévisions ont été évaluées s'étire donc du 08/10/2001 (journée cible pour laquelle nous disposons d'une prévision émise à J et à $J+7$, soit respectivement le 08/10/2001 et le 01/10/2001) au 01/10/2004 (journée cible pour laquelle nous disposons d'une prévision émise à J et à $J+7$, soit respectivement le 01/10/2004 et le 24/09/2004), soit 1090 jours.

11.2.1 Evaluation globale

11.2.1.1 A l'aide du score *CRPSS*

Nous proposons dans un premier temps d'évaluer globalement les prévisions de précipitations générées par les méthodes d'analogie retenues. Le score *CRPS* a été retenu non seulement pour tenir compte de l'ensemble des types d'événements (pluvieux et non pluvieux), mais aussi pour déterminer les gains en finesse et justesse des prévisions issues des méthodes développées dans cette thèse par rapport à la méthode de référence A0. La seule différence avec le calcul du *CRPS* réalisé dans le contexte de prévision parfaite réside dans l'utilisation des pluies brutes au lieu des pluies transformées (section 2.2.2.3), du fait que le modèle GR3P doit être alimenté en lames d'eau brutes. Les lames d'eau importantes ont ainsi plus de poids dans la valeur du score *CRPS*. Il est donc préférable de ne pas comparer les valeurs entre les bassins, notamment en raison des différences de climatologie.

Pour chaque sous-bassin, le score *CRPSS* a été évalué pour chaque échéance de prévision, la référence étant toujours la climatologie. La Figure 11.1 présente l'évolution des scores *CRPSS* en fonction de l'échéance pour chaque sous-bassin. On constate que les méthodes A2/A2bis se révèlent généralement plus performantes que les autres méthodes pour les premières échéances, généralement jusqu'à l'échéance $J+2$ voire $J+3$ pour les sous-bassins de la Saône. Les écarts de scores diffèrent fortement d'un bassin à un autre. Pour les bassins de l'Yonne-C et de la Seine-Int, les performances obtenues par l'application des méthodes A1bis et A2 sont souvent très proches, avec même un léger avantage pour la méthode A1bis. Par ailleurs, la dispersion des courbes diminue nettement à partir de l'échéance $J+3$, traduisant la baisse de fiabilité des prévisions des variables d'analogie avec l'augmentation de l'échéance. Toutes les méthodes sont alors équivalentes pour les dernières échéances. Ainsi, bien que la prévision des champs de vitesse verticale soit entachée d'erreurs, la fiabilité de la prévision des champs de géopotentiels semble telle que l'échantillon d'analogues retenues au niveau 1 n'est plus suffisamment pertinent. En effet, si toutes les situations similaires à la situation prévue en termes de circulation générale ne sont finalement pas analogues à la situation qui sera réellement observée, il est impossible de retenir de « bonnes » analogues à l'issue des niveaux suivants.

Par ailleurs, on remarque sur la Figure 11.1 que la prévision simplement basée sur la climatologie devient meilleure que les prévisions par analogie à partir de l'échéance $J+6$ ($CRPSS < 0$), et ce quel que soit le bassin. Il n'y a donc pas nécessairement d'intérêt à appliquer une telle méthode pour une échéance supérieure à six ou sept jours avec les prévisions du modèle CEP.

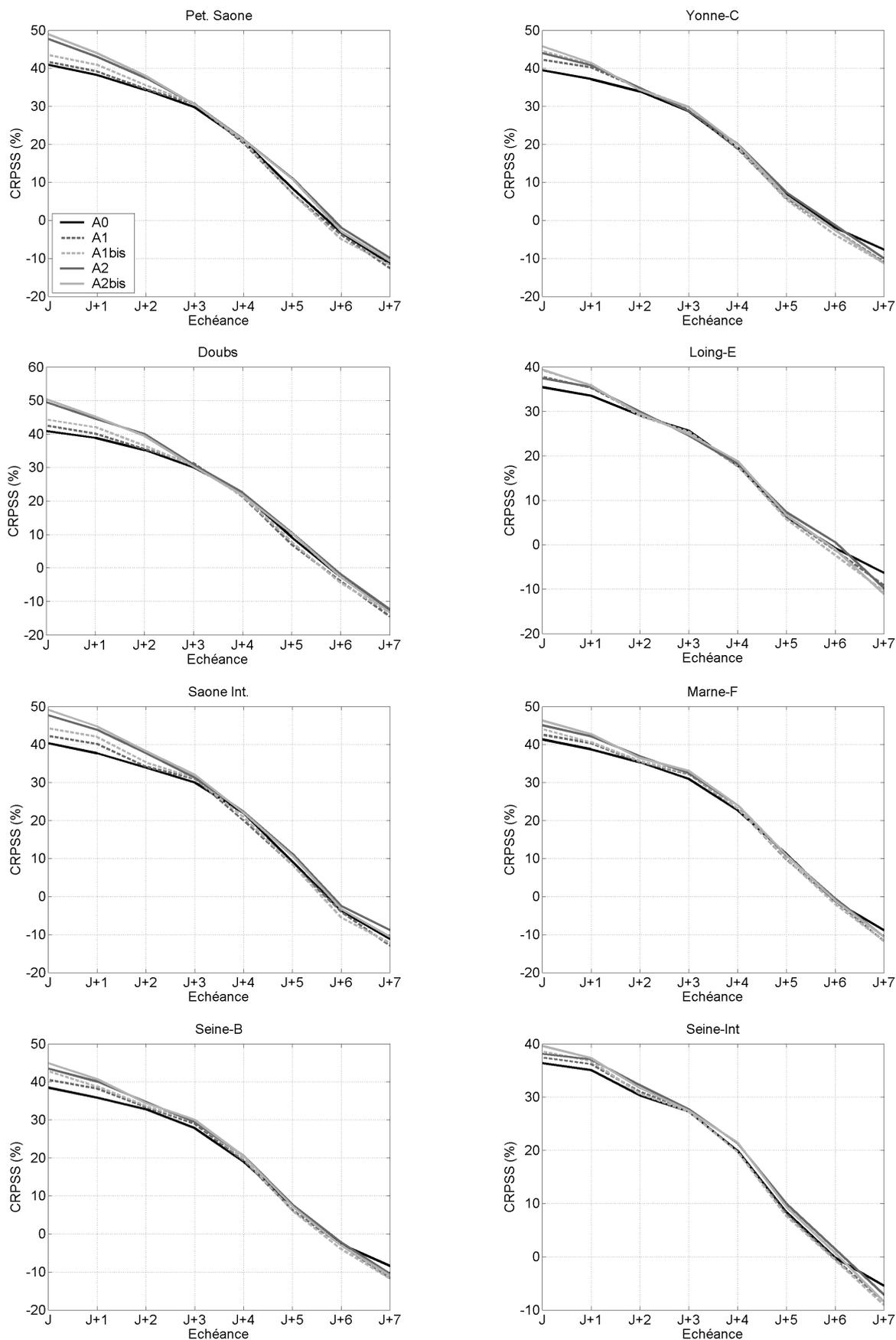


Figure 11.1 : Evolution des scores $CRPSS$ (%) en fonction de l'échéance de prévision, obtenus par les méthodes d'analogie appliquées à chacun des huit sous-bassins de la Saône et de la Seine.

Comparaison à la prévision parfaite

La baisse de la fiabilité avec l'échéance de la prévision adaptée par analogie de la prévision provient inéluctablement de celle des champs météorologiques fournis par le MNP. Afin de le vérifier, nous avons considéré, comme champs prévus à chaque échéance, les champs analysés par le même modèle. Nous nous replaçons donc ici dans un contexte de prévision parfaite. Cela permet de connaître l'échéance à partir de laquelle la dégradation de la prévision des champs des variables a un impact sur les performances des méthodes d'analogie. D'autre part, cela permet aussi de voir quelles sont les méthodes les plus sensibles à la qualité de la prévision du modèle CEP, et ainsi quelles sont les variables d'analogie dont la prévision se dégrade le plus sensiblement avec l'échéance.

Nous avons évalué, de la même façon que précédemment, le score *CRPSS* basé sur les prévisions adaptées par analogie, qui ont été produites à partir des champs analysés. En outre, le score *CRPS* a été décomposé en finesse et en justesse. La Figure 11.2 permet de comparer, pour le bassin du Doubs, les scores (*CRPSS*, compétence de la finesse et compétence de la justesse) obtenus par les méthodes A0 et A2bis dans les contextes de prévision parfaite et opérationnelle.

La fiabilité des prévisions générées pour la première échéance (jour J) est très bonne avec de faibles écarts de scores entre les prévisions obtenues dans les deux contextes. Dès l'échéance suivante ($J+1$), la fiabilité de la prévision baisse, que ce soit en considérant le score global d'utilité (*CRPSS*) ou en considérant la compétence de la justesse de la prévision. En revanche, la compétence de la finesse n'évolue pas au fil des échéances. La finesse des prévisions n'est visiblement pas influencée par la capacité du modèle CEP à bien prévoir les champs météorologiques. Les prévisions s'avèrent même généralement plus fines dans le contexte opérationnel, que ce soit avec la méthode A0 ou avec la méthode A2bis. Nous pouvons conclure de ces résultats que cette composante de la qualité d'une prévision ne dépend pas des sorties du MNP : elle n'est influencée que par l'algorithme de sélection d'analogues. Par ailleurs, nous retrouvons ici les résultats obtenus dans un contexte de prévision parfaite (section 10.1.2.1) : les prévisions obtenues par la méthode A2bis sont bien plus fines que celles obtenues par A0.

C'est donc sur la justesse des prévisions que la qualité des prévisions du MNP a un impact. Dans le cas présent, la justesse de prévision obtenue avec A2bis, bien que supérieure à celle obtenue par A0 pour la première échéance, devient équivalente l'échéance suivante puis s'avère moins bonne dès l'échéance $J+2$. Cela prouve que la vitesse verticale est une variable moins robuste que les géopotentiels ou les variables d'humidité, dont la prévision est plus rapidement entachée d'erreurs. Cet inconvénient avait déjà été souligné dans l'introduction à la seconde partie : nous en avons à présent la preuve. En revanche, l'écart en justesse entre les deux méthodes se stabilise à partir de l'échéance $J+4$, traduisant une baisse également importante de la fiabilité de la prévision des variables utilisées dans la méthode de référence A0. Nous retrouvons ici un résultat déjà obtenu par Plasse (2008) dans le cadre de

l'application de la méthode A0 au bassin de la Seine. Sur les quelques cas étudiés, il avait en effet remarqué que le biais dû au MNP dans la prévision s'observait surtout à partir de $J+4$.

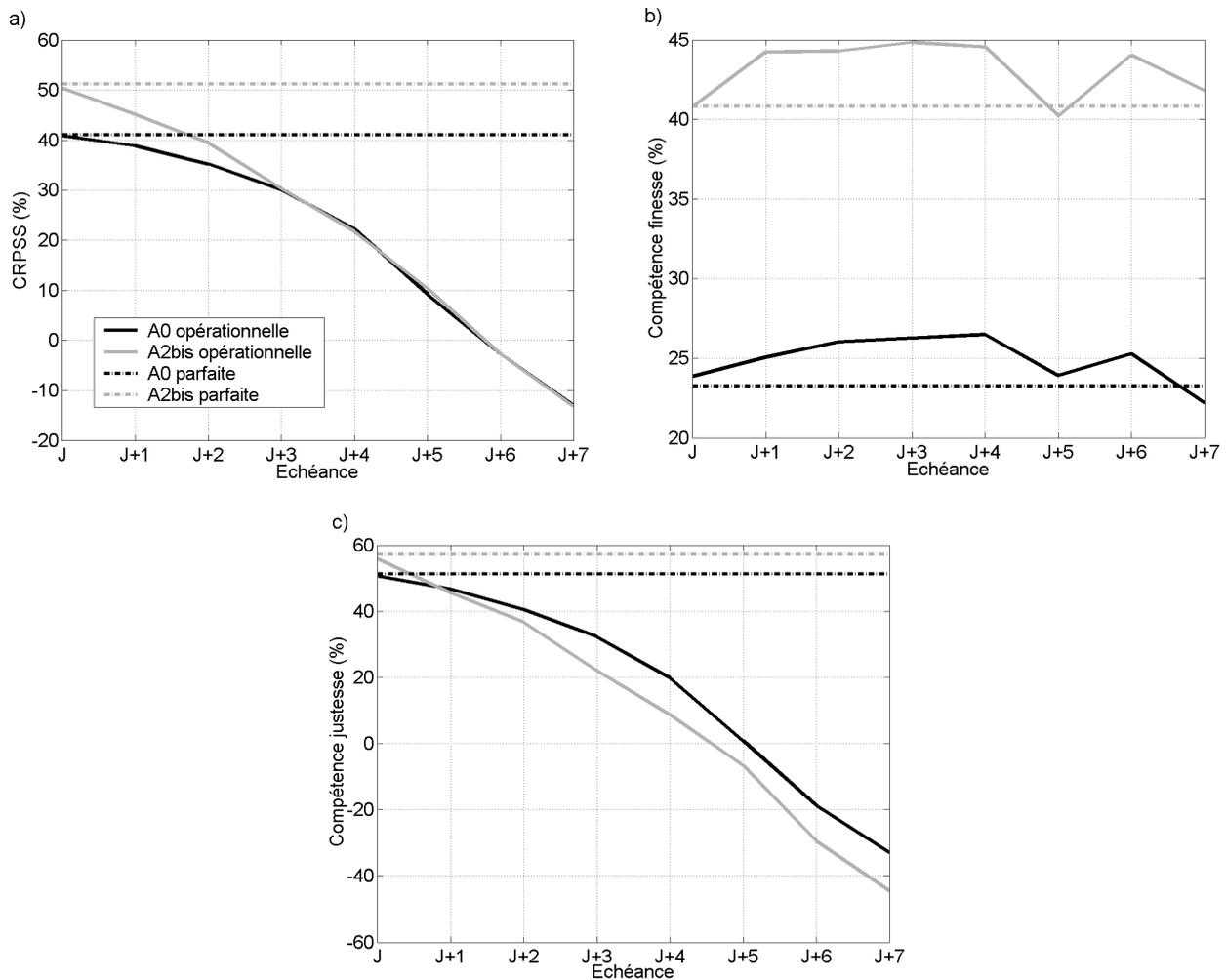


Figure 11.2 : Evolution en fonction de l'échéance : a) du score *CRPSS* (%), b) de la compétence de la finesse du *CRPS* (%) et c) de la compétence de la justesse du *CRPS* (%) ; évalués lorsque les méthodes A0 et A2bis sont appliquées au bassin du Doubs dans les contextes de prévision opérationnelle et parfaite.

11.2.1.2 Corrélation entre prévision et observation

Les différences de performance des méthodes en fonction de l'échéance ont également été étudiées par la corrélation entre la pluie observée et la pluie prévue, définie par le quantile de pluie P_{dis60} de la distribution des analogues. Les graphiques obtenus pour le bassin du Loing-E, pour la première et la dernière échéance (J et $J+7$), sont indiqués sur la Figure 11.3.

Pour l'échéance J , les valeurs du coefficient de détermination R^2 sont bien croissantes de la méthode A0 à la méthode A2bis, avec une hausse de la valeur du R^2 de près de 0,09 entre les méthodes A0 et A2bis. En revanche, pour l'échéance $J+7$, les valeurs de R^2 sont décroissantes, la méthode A0 permettant d'obtenir une corrélation plus forte entre la pluie observée et le quantile de prévision P_{dis60} qu'avec les autres méthodes. La valeur reste néanmoins très

faible, traduisant la déficience d'une méthode d'analogie pour cette échéance de prévision. Nous retrouvons également ici, quelle que soit la méthode appliquée, une sous-estimation systématique du quantile $Pdis60$ pour les événements pluvieux conséquents, dont la lame d'eau journalière excède 10 mm pour l'exemple du bassin du Loing-E. En revanche en-dessous de ce seuil, il n'y a visiblement pas de biais systématique.

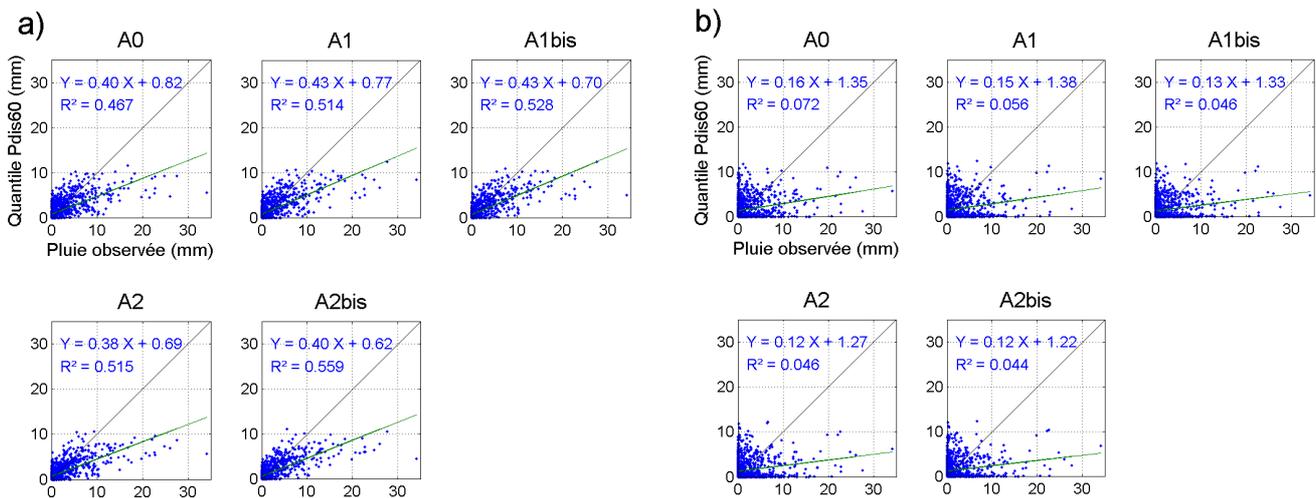


Figure 11.3 : Corrélation entre les lames d'eau observées et les valeurs du quantile $Pdis60$ prévues par chacune des méthodes d'analogie sur le bassin du Loing-E. a) échéance J ; b) échéance $J+7$.

Pour analyser plus globalement l'évolution en fonction de l'échéance de la corrélation entre la pluie observée et la valeur du quantile $Pdis60$, nous avons évalué la valeur moyenne du coefficient R^2 sur les sous-bassins pour chacune des méthodes d'analogie. La Figure 11.4 présente l'évolution des coefficients R^2 en fonction de l'échéance.

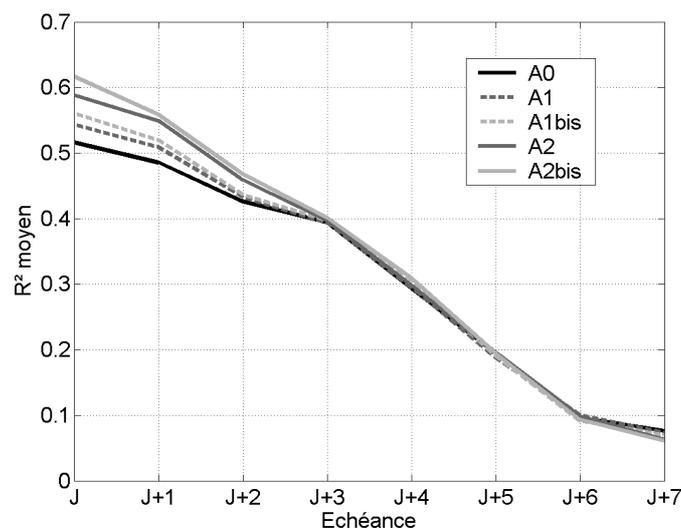


Figure 11.4 : Evolution de la valeur moyenne sur les sous-bassins du coefficient R^2 en fonction de l'échéance, évaluée lorsque chacune des méthodes d'analogie est appliquée.

Nous constatons que la valeur moyenne du coefficient R^2 décroît presque linéairement avec

l'échéance. L'écart entre les valeurs obtenues par les méthodes A0 et A2bis est assez élevé pour l'échéance J (près de 0,1), puis cet écart diminue jusqu'à l'échéance $J+3$. Conformément aux résultats obtenus en termes de compétence (cf. Figure 11.1) pour les échéances supérieures à $J+3$, les méthodes d'analogie semblent équivalentes : les pluies prévues par le quantile $Pdis60$ ne semblent donc pas être mieux corrélées aux observations avec l'une ou l'autre des méthodes, même si la méthode A2bis reste légèrement privilégiée pour les échéances $J+3$ et $J+4$.

11.2.2 Evaluation en fonction d'un seuil

L'utilisateur d'un système de prévision probabiliste peut notamment se demander si un événement est finalement bien anticipé dans un contexte de prévision opérationnelle, mais aussi s'il n'est pas fréquemment alerté sans que l'événement n'ait finalement lieu. Pour répondre à ce type d'interrogation, nous avons évalué pour chacune des méthodes les scores POD et FAR sur la période d'étude (situations cibles du 08/10/2001 au 24/09/2004), et ce pour l'ensemble des sous-bassins de la Saône et de la Seine. Cinq seuils ont été retenus parmi les douze étudiés dans un contexte de prévision parfaite : le seuil 0 mm (pluie/non pluie), et les seuils définis par les quantiles de pluie non nulle $PNN40$, $PNN60$, $PNN80$ et $PNN95$. Rappelons que les cumuls de pluie associés à ces seuils ont été indiqués dans le tableau 10.3 (section 10.1.2.2).

11.2.2.1 Evolution des scores POD avec l'échéance de prévision

La Figure 11.5 présente, pour trois seuils de précipitations (0 mm, $PNN60$ et $PNN95$), l'évolution en fonction de l'échéance des scores POD obtenus lorsque la prévision est définie par le quantile $Pdis90$ d'une distribution des analogues.

Le quantile $Pdis90$ représente la borne supérieure de l'intervalle de confiance de la prévision des précipitations. Par conséquent, le taux de détection des événements pluvieux (*i.e.* dont la lame d'eau est supérieure au seuil 0 mm) est bien évidemment très élevé, proche de 1 pour la première échéance, et ce quelle que soit la méthode appliquée. En revanche, le taux de détection ne diminue que très peu avec l'augmentation de l'échéance de prévision. Les méthodes ne se différencient que pour les dernières échéances. Ainsi à $J+7$, le taux de détection obtenu par les méthodes développées (A1 à A2bis) est légèrement inférieur au taux obtenu par la méthode A0.

Si on s'intéresse aux autres événements, on s'aperçoit que plus le seuil de précipitations est élevé, plus le taux de détection baisse rapidement avec l'échéance de prévision. De même, entre J et $J+4$, plus l'échéance augmente, plus l'écart entre les scores POD évalués avec les différentes méthodes d'analogie augmente. Ce résultat est notamment perceptible pour les cumuls pluviométriques supérieurs au seuil $PNN60$. Ainsi, les taux de détection obtenus par les méthodes A0 et A2bis sont proches de 1 pour la première échéance, puis l'écart se creuse pour atteindre près de 0,1 à $J+7$. Nous retrouvons ici le résultat déjà mis en évidence dans un

contexte de prévision parfaite : l'introduction de la vitesse verticale dans l'algorithme induit une baisse de la capacité de détection de ce type d'événement.

Pour les situations dont la lame d'eau observée est supérieure au seuil *PNN95*, dès la première échéance, le taux de détection est plus faible avec les méthodes A2/A2bis qu'avec les autres méthodes. L'écart de score *POD* n'augmente pas pour autant avec l'échéance, il tend plutôt à se stabiliser. Cependant, quelle que soit la méthode, le taux de détection de ces événements chute rapidement pour n'atteindre que 0,1 à 0,3 à J+7.

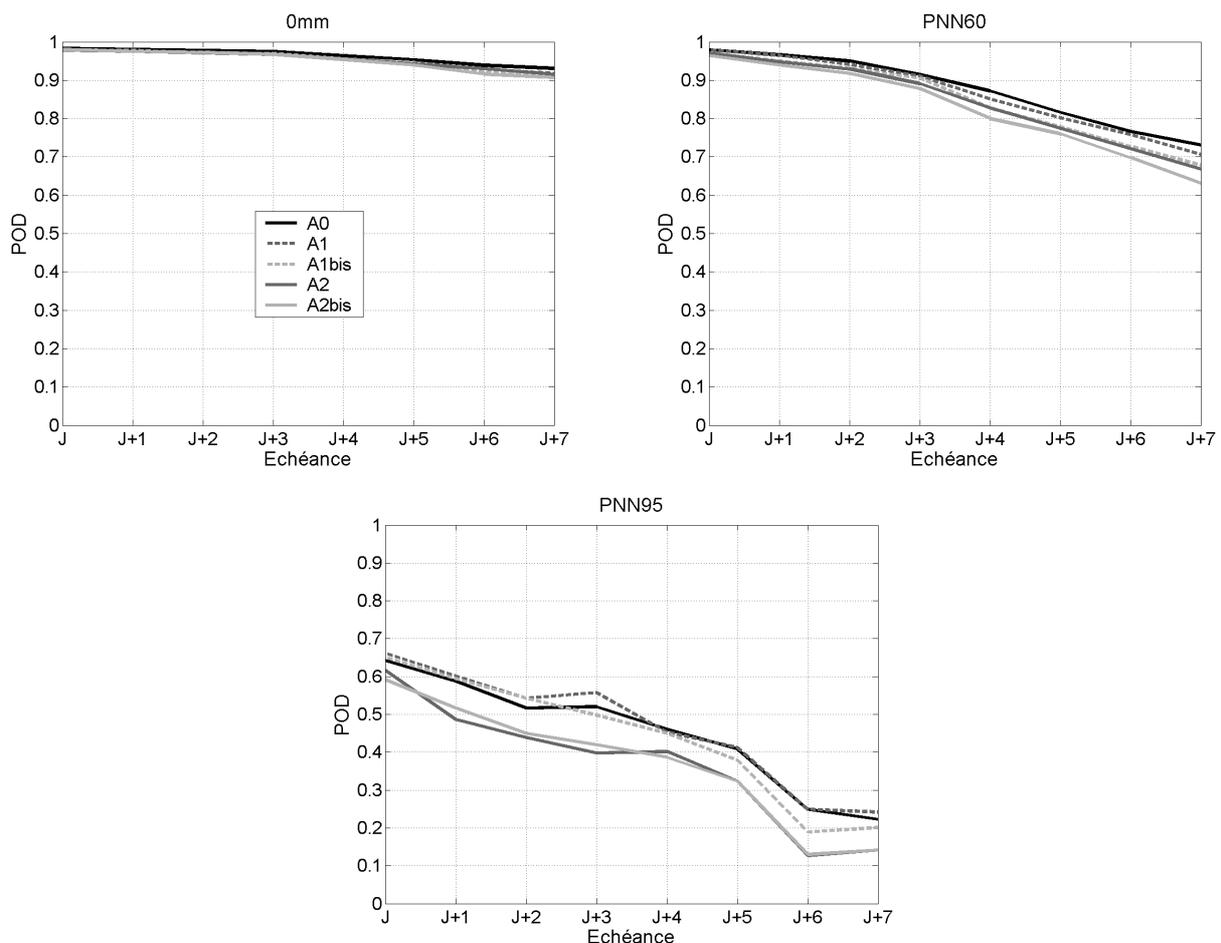


Figure 11.5 : Evolution du score *POD* en fonction de l'échéance, évalué lorsque chacune des méthodes d'analogie est appliquée, pour trois seuils de précipitations (0 mm, *PNN40* et *PNN95*).

Enfin pour un événement donné, la capacité de détection évolue-t-elle différemment avec l'échéance selon le quantile de prévision retenu ? Pour y répondre, nous avons étudié l'évolution du score *POD* en fonction de l'échéance lorsque la prévision est définie par chacun des neuf quantiles de la distribution (*Pdis10* à *Pdis90*). La Figure 11.6 indique l'évolution avec l'échéance des scores obtenus avec ces quantiles pour le seuil *PNN80*, lorsque la méthode de référence A0 est appliquée.

Logiquement, pour chaque échéance de prévision, le taux de détection augmente avec l'ordre du quantile au non-dépassement. L'écart est très élevé pour la première échéance (près de 0,9 entre les quantiles *Pdis10* et *Pdis90*), puis il faiblit au fil des échéances jusqu'à

atteindre seulement 0,6 à $J+7$. Le score *POD* diminue effectivement avec l'échéance pour les quantiles moyens et hauts d'une distribution (*Pdis30* à *Pdis90*), tandis que le score est stable pour les faibles quantiles (*Pdis10* et *Pdis20*). Il est également intéressant de voir que le taux de détection est à peu près stable jusqu'à $J+3$ pour les quantiles *Pdis30* à *Pdis50*.

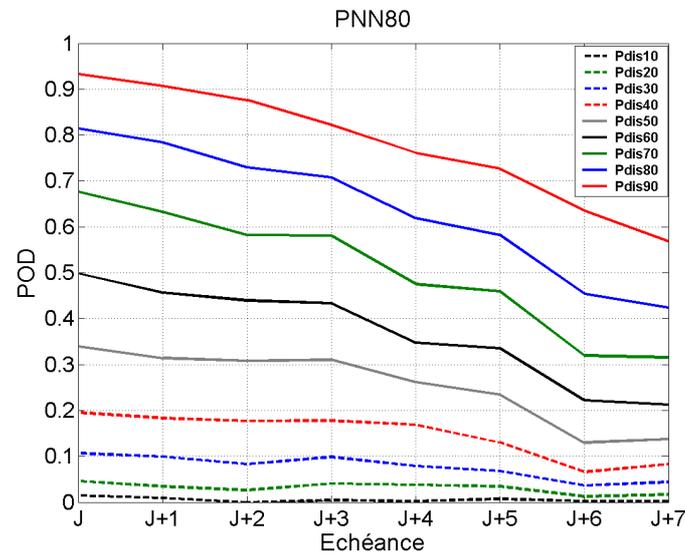


Figure 11.6 : Evolution du score *POD* en fonction de l'échéance, évalué pour chaque quantile d'une distribution prévue lorsque la méthode A0 est appliquée.

11.2.2.2 Evolution des scores FAR avec l'échéance

La Figure 11.7 indique, pour le seuil *PNN60*, l'évolution des scores *FAR* en fonction de l'échéance lorsque les quantiles *Pdis20*, *Pdis60* et *Pdis90* sont utilisés comme prévision.

De manière générale, le score *FAR* augmente logiquement avec l'échéance de prévision, et ceci est d'autant plus important que le quantile de prévision considéré est faible. Ainsi, le taux de fausses alertes augmente de près de 0,4 à 0,5 points entre la première et la dernière échéance pour le quantile *Pdis20*. La hausse du score *FAR* augmente moins sensiblement si l'on utilise des quantiles de prévision plus élevés, à l'image du quantile *Pdis60* (0,3 points en plus entre J et $J+7$) et du quantile *Pdis90* (0,1 à 0,2 points). Par ailleurs, les scores *FAR* sont en moyenne plus faibles avec les méthodes A2/A2bis, en particulier pour les premières échéances. Ce résultat rejoint celui obtenu dans un contexte de prévision parfaite. En revanche à partir de $J+5$, les méthodes émettent toutes presque autant de fausses alertes, ce qui implique là encore que les prévisions des champs de vitesse verticale se dégradent probablement plus que les autres avec l'échéance.

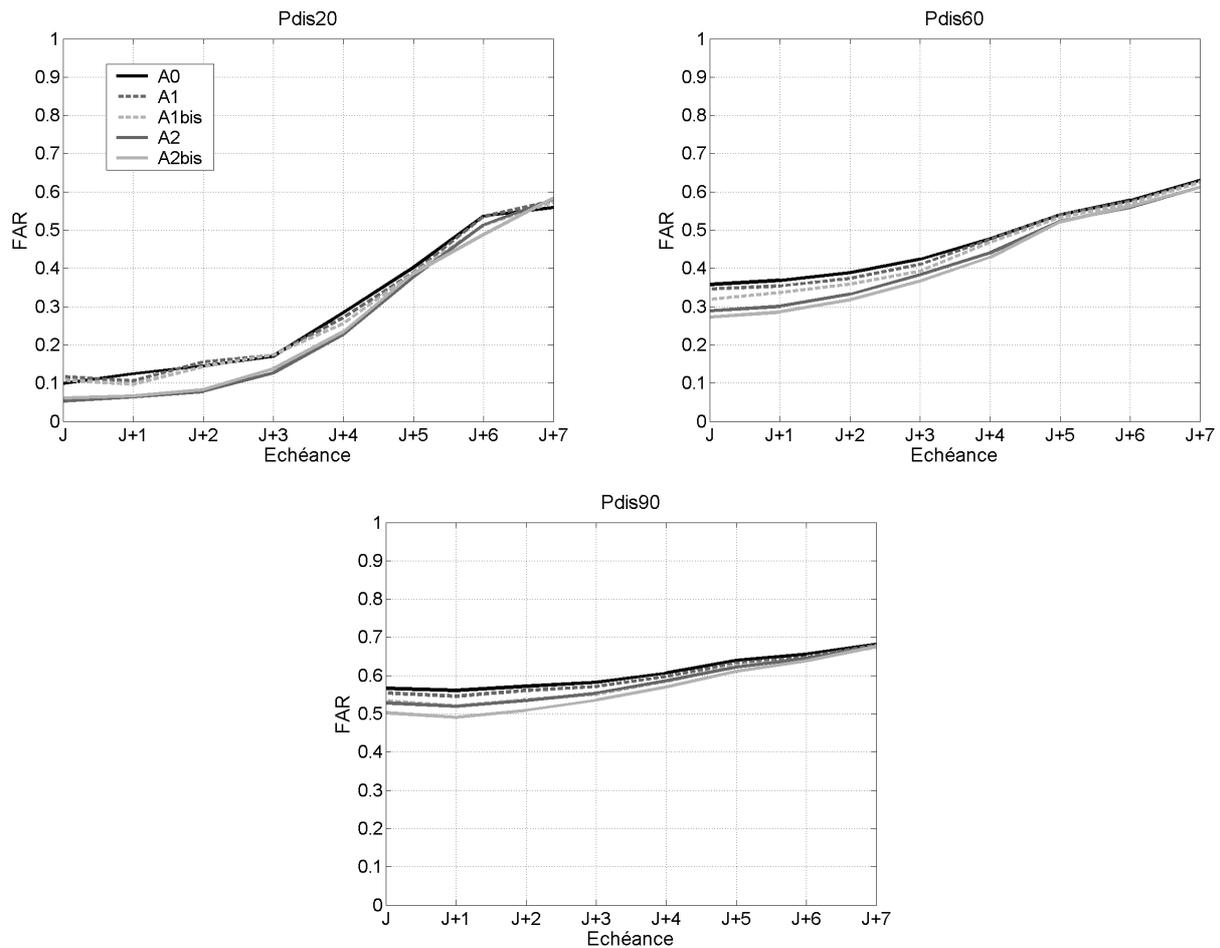


Figure 11.7 : Evolution du score *FAR* en fonction de l'échéance, évalué pour les quantiles de prévision *Pdis20*, *Pdis60* et *Pdis90* ; lorsque chacune des méthodes d'analogie est appliquée.

11.3 Prévision des débits

Nous rappelons ici que les scénarios de débit ont été générés par le modèle GR3P alimenté par les scénarios de pluie produits par nos cinq méthodes d'analogie après tirage aléatoire.

11.3.1 Evaluation globale

Comme pour les précipitations, les prévisions de débit ont d'abord été évaluées par le score *CRPSS*, dans lequel la prévision de référence considérée est la persistance.

11.3.1.1 A l'aide du score *CRPSS*

Pour chaque échéance de prévision de J à $J+7$, les scores *CRPSS* ont été évalués pour chaque méthode d'analogie. Nous nous contentons d'analyser ceux obtenus sur les bassins de la Saône à Le Chatelet et du Doubs à Neublans (Figure 11.8).

Les performances ne présentent pas le même comportement entre ces deux bassins. Pour le bassin de la Saône à Le Chatelet, elles augmentent jusqu'à l'horizon de prévision $J+6$, puis

elles diminuent légèrement entre les échéances $J+6$ et $J+7$; tandis que pour le Doubs à Neublans, le score $CRPSS$ augmente jusqu'à $J+2$ puis il diminue. Pour ce bassin, le gain apporté par la prévision à partir de pluies prévues par rapport à la persistance diminue donc dès le quatrième jour. Par ailleurs les différences de performance entre les prévisions émises à partir des méthodes d'analogie sont faibles pour la Saône à Le Chatelet, un léger avantage étant donné aux méthodes A0 à A1bis entre les échéances $J+4$ et $J+6$. Pour le Doubs à Neublans, les valeurs de $CRPSS$ sont plus élevées pour les prévisions émises à partir des méthodes A2/A2bis, tandis que les prévisions obtenues par la méthode A1 se révèlent généralement moins bonnes que les autres.

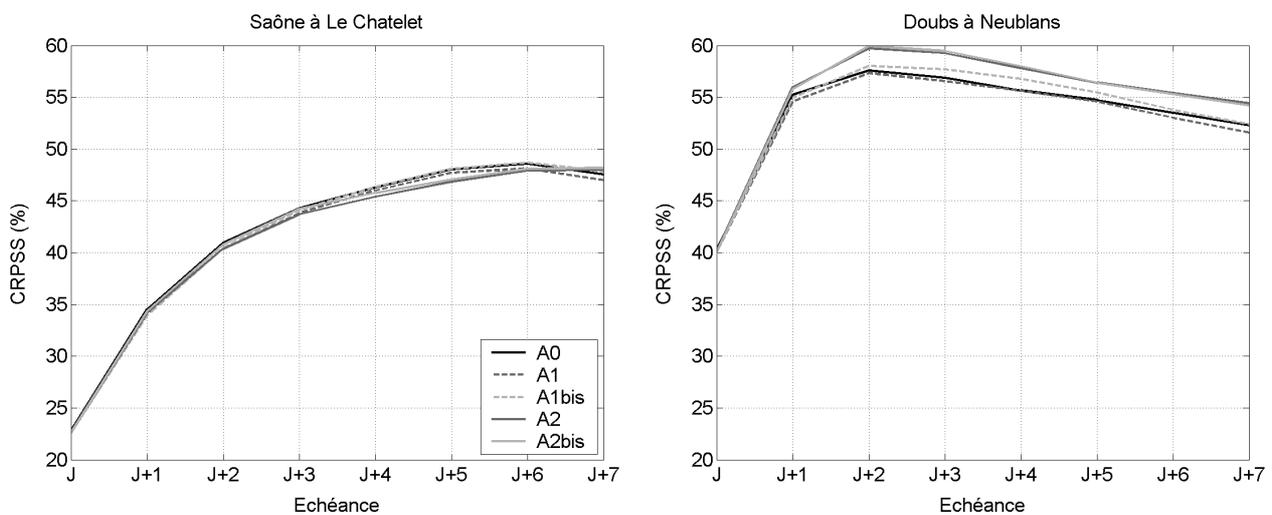


Figure 11.8 : Evolution du score $CRPSS$ (%) en fonction de l'échéance, évalué lorsque le modèle GR3P est alimenté en scénarios issus de chacune des méthodes d'analogie (A0, A1, A1bis, A2, A2bis) ; pour les bassins de la Saône à Le Chatelet et du Doubs à Neublans.

Comparaison à la prévision parfaite

Les erreurs de prévision des variables d'analogie ont très probablement un impact direct sur la qualité des prévisions hydrologiques. Afin de quantifier cet impact, nous avons procédé de la même manière que pour les pluies. Les prévisions de débit ont été obtenues à partir des prévisions de pluie adaptées des champs analysés par le modèle CEP (contexte de prévision parfaite). De cette manière, nous nous affranchissons de la part des erreurs de prévision dues au MNP. L'évolution des scores $CRPSS$ en fonction de l'échéance est indiquée sur la Figure 11.9, lorsque les méthodes A0 et A2bis sont appliquées en contexte de prévision parfaite et opérationnelle, pour les bassins de la Saône à Le Chatelet et du Doubs à Neublans.

Avec chacune des deux méthodes, le biais dû au MNP est détecté dans les prévisions de débit à partir de l'échéance $J+3$. Les performances obtenues dans un contexte de prévision parfaite augmentent jusqu'à la dernière échéance pour le bassin de la Saône à Le Chatelet, mais aussi pour celui du Doubs à Neublans, même si pour ce bassin une légère baisse du score $CRPSS$ s'observe entre $J+2$ et $J+5$ lorsque la méthode A0 est appliquée. Pour les dernières échéances, la faiblesse de la méthode A2bis est à nouveau mise en évidence, pour

laquelle l'écart de score entre les prévisions émises dans les deux contextes devient de plus en plus important. En effet, puisque la prévision des précipitations par la méthode A2bis se dégrade plus rapidement, l'erreur se propage fortement dans la chaîne de prévision. Les erreurs de la prévision des variables d'analogie fournies par le MNP se répercutent plus sensiblement et plus rapidement sur les prévisions de débits obtenues à partir de la méthode A2bis. Pour le bassin de la Saône à Le Chatelet, l'écart des scores *CRPSS* obtenus dans les deux contextes est perceptible dès $J+4$ pour la méthode A2bis, tandis qu'il ne l'est qu'à partir de $J+6$ pour la méthode A0. L'erreur est alors maximale pour la dernière échéance, avec un écart bien plus important lorsque la méthode A2bis est appliquée.

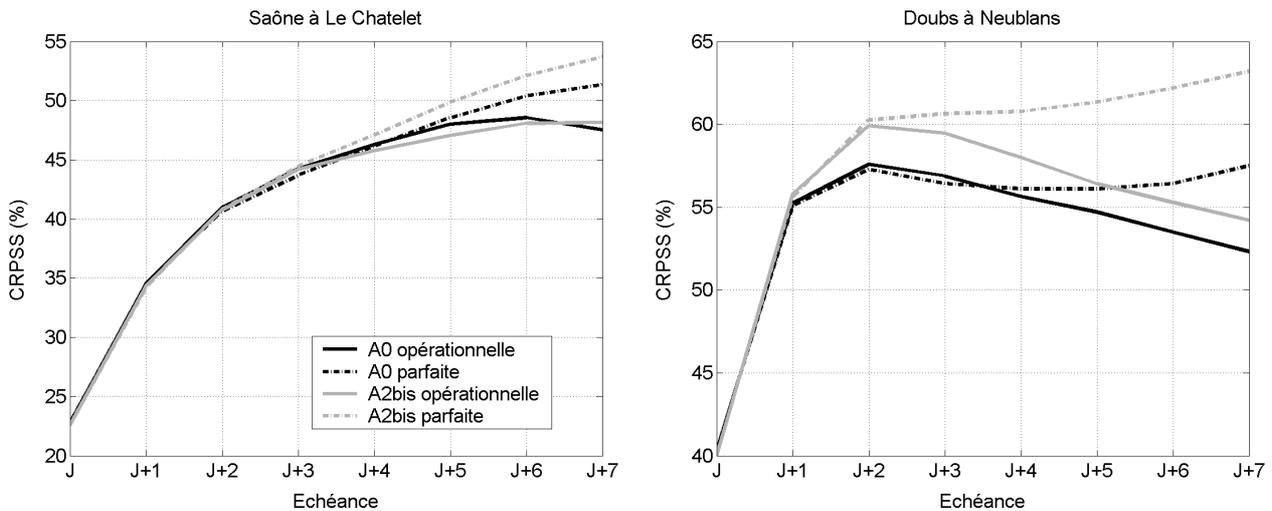


Figure 11.9 : Evolution en fonction de l'échéance du score *CRPSS* (%) évalué lorsque le modèle GR3P est alimenté en scénarios issus des méthodes A0 et A2bis dans les contextes de prévision opérationnelle et parfaite ; pour les bassins de la Saône à Le Chatelet et du Doubs à Neublans.

11.3.1.2 A l'aide du score *RMSE*

Un autre moyen d'établir une évaluation globale est de calculer le score *RMSE* entre débits observés et débits prévus. Pour un bassin donné, ce score a été évalué sur l'ensemble des débits prévus pour chaque échéance E selon l'équation 11.1 :

$$RMSE(E) = \sqrt{\frac{1}{nT \times N} \cdot \sum_{iT=1}^{nT} \sum_{iS=1}^N (Q_{obs}(iT, E) - Q_{prev}(iT, iS, E))^2} \quad (11.1)$$

où $Q_{obs}(iT, E)$ représente le débit observé à la iT ème date cible de la période pour l'échéance E , $Q_{prev}(iT, iS, E)$ est le débit prévu du iS ème scénario à la iT ème date cible pour l'échéance E , N est le nombre de scénarios (identique au nombre d'analogues finalement retenues) et nT le nombre de dates cibles considérées.

Pour illustrer les résultats obtenus, la Figure 11.10 présente, pour le bassin de la Saône à Le Chatelet, l'évolution en fonction de l'échéance des scores *RMSE* calculés à partir des

différentes méthodes d'analogie.

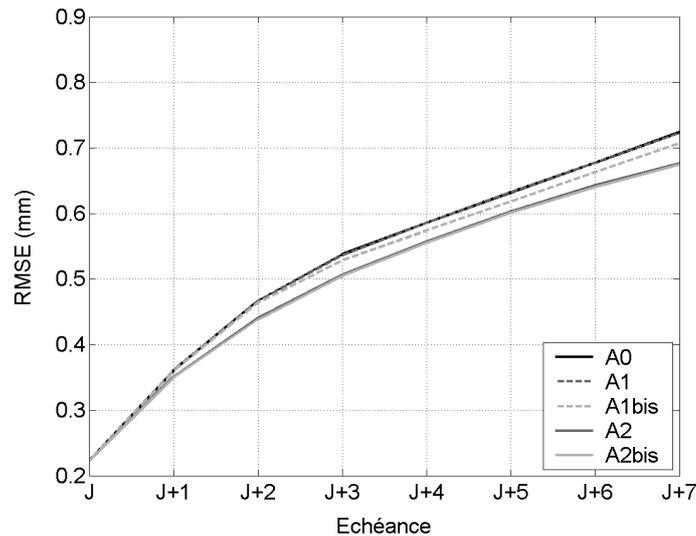


Figure 11.10 : Evolution du score $RMSE$ évalué lorsque le modèle GR3P est alimenté en scénarios issus de chacune des méthodes d'analogie (A0, A1, A1bis, A2, A2bis), pour le bassin de la Saône à Le Chatelet.

Si nous comparons ce graphique avec celui de la Figure 11.8, nous remarquons que la hiérarchie n'est plus la même : en terme de $RMSE$, les prévisions de débit générées à partir des méthodes A2 et A2bis sont meilleures que celles obtenue par les autres méthodes ; alors que si l'on exprime les performances par la compétence $CRPSS$, la tendance est inversée. Les résultats de l'évaluation des prévisions sont donc dépendants du score utilisé. Par ailleurs, l'analyse du score $RMSE$ confirme la surestimation des débits par les méthodes A0 et A1 (et A1bis dans une moindre mesure), puisque ce score est sensible aux écarts importants.

11.3.1.3 Sources d'erreurs dans la chaîne de prévision

Nous venons de mettre en évidence les erreurs de la prévision de débit émise dans un contexte opérationnel. Il est cependant impossible, au vu des résultats précédents, de connaître la part des erreurs dues au modèle numérique qui génère les prévisions des champs des variables d'analogie ($E1$), aux méthodes d'analogie ($E2$) et au modèle hydrologique ($E3$).

Afin d'identifier la part des erreurs de ces différentes sources, seul l'algorithme A2bis a été appliqué, et l'erreur quadratique moyenne MSE , définie par l'équation 11.2, a été évaluée dans différents cas.

$$MSE(E) = \frac{1}{nT \times N} \cdot \sum_{iT=1}^{nT} \sum_{iS=1}^N (Q_{obs}(iT, E) - Q_{prev}(iT, iS, E))^2 \quad (11.2)$$

Outre l'erreur totale entre i) débit observé et débits prévus à partir des prévisions de pluie obtenues dans un contexte de prévision opérationnelle (*Obs-A2bis opér.*), nous avons également évalué les erreurs :

- ii) $E3$ entre débit simulé avec pluies futures connues et débit observé (*Sim-Obs*) ;
- iii) $E2$ entre débit simulé avec pluies futures connues et débits prévus à partir des prévisions de pluie obtenues dans un contexte de prévision parfaite (*Sim-A2bis parf.*) ;
- iv) $E1 * E2$ entre débit simulé avec pluies futures connues et débits prévus à partir des prévisions de pluie obtenues dans un contexte de prévision opérationnelle (*Sim-A2bis opér.*).

La Figure 11.11 présente l'évolution en fonction de l'échéance des valeurs du score MSE (mm^2) obtenues dans chacun des cas pour le bassin versant de la Saône à Le Chatelet (Petite Saône).

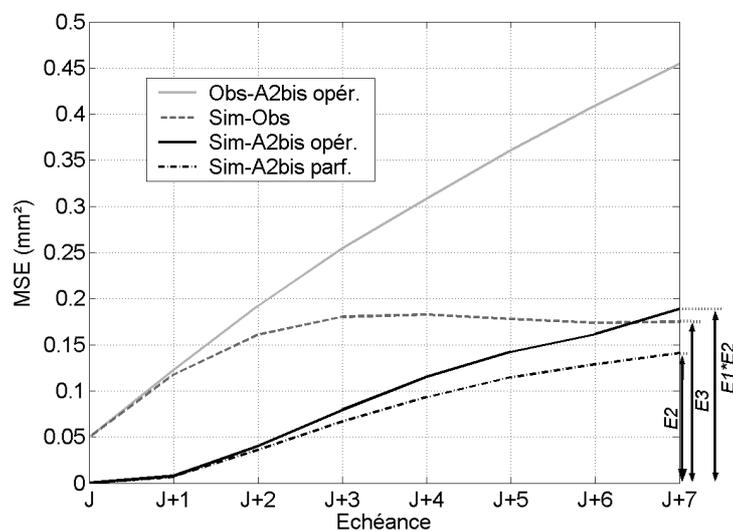


Figure 11.11 : Evolution du score MSE (mm^2) en fonction de l'échéance, entre les valeurs de débit obtenues dans différents cas.

L'interprétation de l'évolution des erreurs permet de montrer que celle due au modèle GR3P (*Sim-Obs*) est importante dès la première échéance, puis elle augmente jusqu'à $J+3$ avant de se stabiliser jusqu'à $J+7$. Ainsi, l'augmentation continue de l'erreur totale (*Obs-A2bis opér.*) entre les échéances $J+3$ et $J+7$ serait principalement due aux erreurs de la méthode d'analogie et du MNP. On constate également que l'erreur totale est supérieure à la somme des erreurs $E1 * E2 + E3$: les erreurs se combinent et ne sont donc pas indépendantes. Par ailleurs, aux échéances J et $J+1$, aucune différence notable n'est observée entre les valeurs de MSE *Obs-A2bis opér.* et *Sim-Obs*, et entre les valeurs de MSE *Sim-A2bis opér.* et *Sim-A2bis parf.*, du fait de l'effet mémoire du bassin.

L'identification de la part des erreurs de chacune des sources n'est pas aisée. Il est nécessaire de comparer certaines courbes entre elles afin d'isoler parfaitement la part d'une source d'erreur. Nous allons nous concentrer sur les valeurs de l'erreur MSE obtenue à l'échéance $J+7$. Celles-ci tiennent évidemment compte des erreurs des échéances précédentes (combinaison des erreurs de prévision produites à chaque échéance).

L'erreur due au modèle GR3P, $E3$, est identifiée directement sur la courbe *Sim-Obs*. Ainsi à $J+7$, la valeur du score MSE vaut environ $0,175 \text{ mm}^2$.

L'erreur $E2$ due à la méthode d'analogie peut être identifiée par la courbe *Sim-A2bis parf.* : elle est manifestement non nulle, l'erreur MSE augmentant avec l'échéance pour atteindre $0,141 \text{ mm}^2$ à $J+7$.

Enfin, l'erreur $E1$ due au MNP ne peut être mise en évidence directement sur la Figure 11.11. En effet, bien qu'elle soit prise en compte dans l'erreur MSE du cas *Sim-A2bis opér.*, elle est forcément combinée à l'erreur $E2$ due à la méthode A2bis. Néanmoins, l'erreur combinée $E1 * E2$ n'atteint que $0,189 \text{ mm}^2$ à $J+7$. Si l'on compare cette valeur à celle de l'erreur $E2$ ($0,141 \text{ mm}^2$), on peut en déduire que l'erreur due au MNP est négligeable devant l'erreur due à la méthode d'analogie.

En conclusion, au vu des éléments mis en évidence ici, l'erreur de la chaîne de prévision est en grande partie due au modèle GR3P et à la méthode d'analogie. L'erreur due à cette dernière est cependant inférieure à celle du modèle hydrologique, tandis que celle due à la prévision des variables d'analogie par le MNP est négligeable devant les deux autres.

11.3.2 Etude des performances en fonction du seuil de débit

Les prévisions hydrologiques ont été évaluées en fonction de seuils de débit afin d'analyser la capacité de la chaîne à détecter des événements de crue mais aussi à ne pas émettre de fausses alertes. Dans cette section, pour être au plus près des attentes des services de prévision des crues, nous avons choisi des seuils de débit proches de ceux définissant les niveaux de vigilance.

Les services de prévision des crues Rhône amont-Saône, Seine moyenne-Yonne-Loing et Seine amont-Marne amont nous ont transmis, pour la station située à l'exutoire de chacun des sous-bassins étudiés, les seuils approximatifs associés aux niveaux de vigilance crue de couleur jaune (niveau 1) et orange (niveau 2). Le débit associé au niveau de vigilance de couleur rouge (niveau 3) nous a été transmis pour certaines stations. Les différents seuils de débit, ici exprimés en m^3/s et notés $S1$, $S2$ et $S3$, sont reportés dans le Tableau 11.3.

Tableau 11.3 : Débits seuils $S1$, $S2$, $S3$ associés aux niveaux d'alerte 1, 2 et 3 ; et fréquences empiriques associées.

Exutoire	Débit seuil (m^3/s) [période de retour (ans)]		
	$S1$	$S2$	$S3$
Neublans (Doubs)	600 [$T < 1$]	1250 [$2 < T < 5$]	-
Le Chatelet (Saône)	600 [$T < 1$]	1200 [$5 < T < 10$]	-
Couzon (Saône)	1000 [$T < 1$]	2350 [$10 < T < 20$]	-
Courlon (Yonne)	550 [$2 < T < 5$]	700 [$10 < T < 20$]	900 [$T > 20$]
Paris (Seine)	974 [$1 < T < 2$]	1600 [$10 < T < 20$]	2050 [$T > 20$]
Bazoches (Seine)	300 [$5 < T < 10$]	350 [$T \sim 20$]	-
La Ferté (Marne)	350 [$2 < T < 5$]	500 [$T > 20$]	-
Episy (Loing)	90 [$1 < T < 2$]	200 [$10 < T < 20$]	260 [$T > 20$]

Pour information, à chaque valeur de débit est associée entre crochets la période de retour (en années), estimée lors de l'étude du régime hydrologique (chapitre 4, section 4.2.2). Ceci permet de situer les seuils de débits par rapport à la fréquence d'observation de dépassement de ces seuils.

Le Tableau 11.4 indique le nombre de jours où chaque débit seuil a été dépassé sur la période d'étude (08/10/2004-24/09/2004), sachant que la période totale compte 1090 jours, soit le nombre de jours communs aux huit échéances de prévision. Seule la station de Bazoches ne compte que 1079 jours sur cette période en raison de la présence de lacunes. Sur la période d'étude, le nombre de jours où le débit *S1* a été dépassé varie fortement d'un bassin à un autre.

Tableau 11.4 : Nombre de jours où les seuils de débit *S1*, *S2*, *S3* ont été dépassés sur la période 01/10/2001-08/10/2004.

Exutoire	Nombre de jours de dépassement de seuil		
	<i>S1</i>	<i>S2</i>	<i>S3</i>
Neublans (Doubs)	40	3	0
Le Chatelet (Saône)	37	0	0
Couzon (Saône)	116	0	0
Courlon (Yonne)	0	0	0
Paris (Seine)	39	0	0
Bazoches (Seine)	15	0	0
La Ferté (Marne)	9	0	0
Episy (Loing)	19	0	0

Les événements de crue dont le débit dépasse le seuil *S1* sont nettement plus nombreux sur le bassin de la Saône que sur le bassin de la Seine. Le débit seuil *S1* a ainsi été dépassé plus d'un jour sur dix sur la période à la station de Couzon (Saône), contre un jour par mois en moyenne à la station de Paris (Seine), ce qui n'est pas étonnant vu les fréquences associées à ce seuil. Le débit seuil *S2* n'a été dépassé qu'à la station de Neublans (Doubs) durant trois jours consécutifs, tandis que le seuil *S3* n'a été dépassé à aucune station. Il s'agit là d'un débit très rarement observé (période de retour de plusieurs dizaines d'années, cf. section 4.2.2, tableau 4.6 et figure 4.6), il n'est donc pas surprenant que celui-ci n'ait pas été franchi au cours de notre période.

A partir de ces seuils, nous avons évalué la capacité du modèle hydrologique, lorsqu'il est alimenté en pluies prévues par les méthodes d'analogie, à bien anticiper des événements de crue mais aussi à ne pas générer de fausses alertes. Nous nous sommes concentrés par la suite sur les bassins pour lesquels le nombre d'événements est significatif.

Dans un premier temps, nous nous intéressons au seuil *S1* pour lequel les événements sont les plus nombreux. Les scores *POD* et *FAR* ont été calculés lorsque la prévision est définie par un des quantiles au non-dépassement *Qdis10*, *Qdis50* et *Qdis90* d'ordre 10 %, 50 % et 90 %, extraits de la distribution. La Figure 11.12 présente l'évolution de ces scores (*POD* à gauche

et *FAR* à droite) pour le bassin de la Saône à Couzon (116 événements).

Les valeurs de score *POD* baissent logiquement avec l'augmentation de l'échéance de prévision. La baisse est cependant plus significative pour le quantile de prévision *Qdis10*, alors que pour les quantiles *Qdis50* et surtout *Qdis90*, la baisse est certes marquée entre *J* et *J+3* puis le taux de détection tend à se stabiliser entre *J+4* et *J+7*. Cela signifie que sur la période, la capacité à détecter un événement avec le quantile *Qdis90* est du même ordre, que l'échéance de prévision soit de quatre ou de huit jours. Il est intéressant de voir que l'on a pu anticiper près de la moitié des événements de crue avec le quantile *Qdis10*, pourtant représentatif de la borne inférieure de l'intervalle de confiance. Par ailleurs, l'impact des performances de prévision obtenues par les différentes méthodes d'analogie sur le taux de détection semble atténué puisque les scores sont similaires, avec un écart de score *POD* n'excédant pas 5 % entre les méthodes. Le modèle hydrologique semble donc amortir l'impact des erreurs de prévision des précipitations sur la prévision de débit. C'est tout de même lorsque les pluies injectées dans le modèle sont générées par les méthodes A0, A1 ou A1bis que le modèle permet de détecter le mieux les événements.

L'évolution des scores *FAR* évoque une certaine stabilité des fausses alertes en fonction de l'échéance pour les quantiles de prévision *Qdis10* et *Qdis50*. Les valeurs du score *FAR* sont étonnamment identiques jusqu'à l'échéance *J+5* pour ces deux quantiles. C'est seulement à *J+6* et *J+7* que le taux de fausses alertes devient plus élevé avec le quantile *Qdis50*. L'élément marquant que nous pouvons retenir est lié aux valeurs du score *FAR* lorsque le quantile *Qdis90* est utilisé comme prévision. Nous remarquons que le taux de fausses alertes augmente sensiblement avec l'échéance lorsque les méthodes A0 à A1bis sont appliquées (passant de 20 à 30 %), tandis qu'il se stabilise à 20 % avec les méthodes A2/A2bis. Nous avons bien remarqué, dans un contexte de prévision parfaite des précipitations, que l'usage de la vitesse verticale permettait de diminuer le taux de fausses alertes. Nous en avons également la preuve en prévision opérationnelle.

Les résultats obtenus sur les autres bassins ont moins d'intérêt étant donné le peu d'événements observés. Les scores *POD* et *FAR* évalués sur le bassin du Doubs à Neublans pour les événements supérieurs au seuil *S2* (trois événements observés) ne sont pas exploitables étant donné le faible nombre d'événements. Nous retiendrons cependant que, pour les événements rares, les prévisions peuvent fortement varier d'un jour à l'autre. Il est donc bien important de comparer les prévisions émises chaque jour avec celles émises les jours précédents pour se faire une idée du potentiel d'occurrence d'un événement, mais aussi de la pertinence de l'alerte lorsque celle-ci est émise. Afin d'illustrer ces éléments, nous proposons maintenant d'étudier quelques cas de crue rencontrés sur la période de prévision.

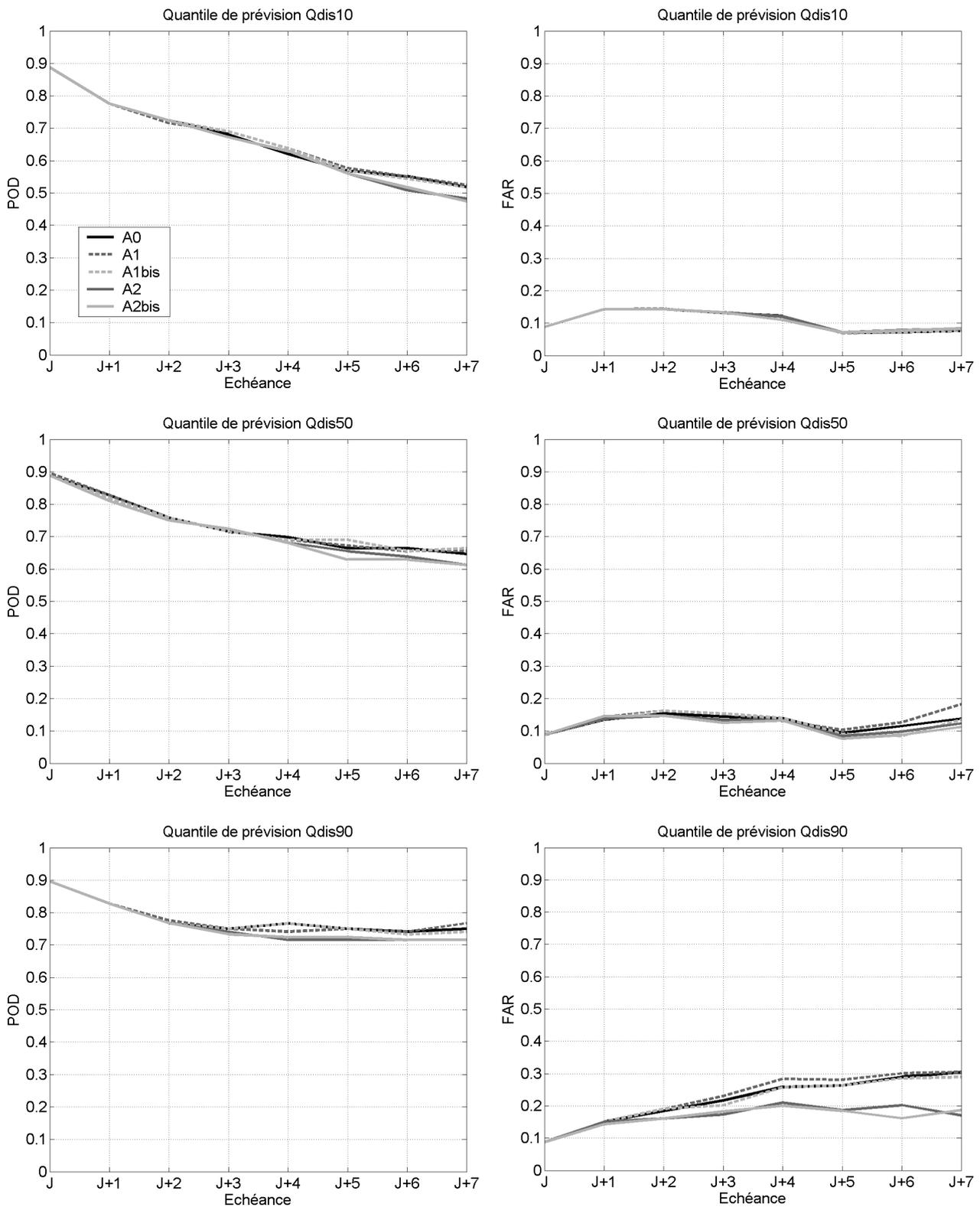


Figure 11.12 : Evolution en fonction de l'échéance des scores *POD* et *FAR* évalués pour la station de Couzon (Saône) lorsque la prévision est définie par un des quantiles *Qdis10*, *Qdis50* et *Qdis90* d'une distribution des débits ; lorsque le modèle GR3P est alimenté en scénarios issus de chacune des méthodes d'analogie (A0, A1, A1bis, A2, A2bis). Le seuil de débit considéré ici est le seuil *S1* défini précédemment.

11.4 Etude de cas

En complément de l'analyse statistique précédente, nous avons souhaité illustrer le comportement de la chaîne de prévision selon la méthode d'analogie sur des crues identifiées dans la période étudiée sur différents bassins. Afin d'envisager tous les cas de figure, ces études recouvrent des vraies alertes, des fausses alertes et des défauts d'alerte. Pour ces deux dernières configurations, nous rechercherons les origines probables des défaillances du système.

En outre, nous proposerons une comparaison entre les prévisions d'une méthode d'analogie et celles fournies par Météo-France aux services de prévision des crues, afin de montrer l'intérêt d'une technique d'analogie pour la prévision opérationnelle des précipitations.

11.4.1 Remarques préliminaires

11.4.1.1 Représentation graphique des prévisions

Les prévisions de précipitations seront représentées, pour chacune des huit premières échéances (J à $J+7$), par les valeurs des quantiles habituellement extraits d'une distribution : $Pdis20$, $Pdis60$ et $Pdis90$.

Les prévisions de débit seront représentées par les valeurs des quantiles $Qdis10$, $Qdis50$ et $Qdis90$ extraits de la distribution empirique des débits prévue pour chaque échéance (Figure 11.13), que nous qualifierons de « scénarios » par simplicité.

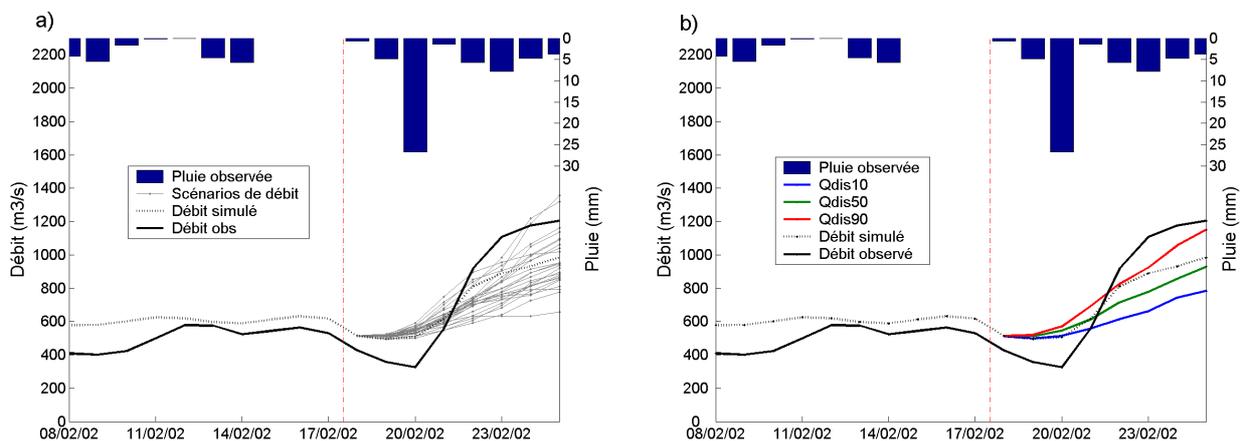


Figure 11.13 : Exemple de prévision probabiliste de débit. a) Ensemble des scénarios de débit obtenus à partir des scénarios de précipitations adaptés par analogie et construits par tirage aléatoire. b) Représentation de la prévision à l'aide des quantiles au non-dépassement $Qdis10$, $Qdis50$ et $Qdis90$ d'une distribution empirique des débits.

L'enveloppe constituée par les valeurs des quantiles $Qdis10$ et $Qdis90$ forme ainsi un intervalle de confiance, tandis que l'hydrogramme constitué des valeurs du quantile $Qdis50$ peut être assimilé à un scénario médian. En outre, sur chaque graphique seront tracées les évolutions du débit observé (trait plein noir) et du débit prévu lorsque le modèle est alimenté

par les pluies observées (trait pointillé noir). Par simplicité, nous parlerons de débit simulé au lieu du débit prévu avec les pluies observées. Notons également qu'une mise à jour du débit est systématiquement effectuée le jour où la prévision est émise, lorsque le débit du jour précédent est connu. Par ailleurs, le trait vertical en tiretés rouges indiqué sur chaque graphique représentera l'instant de prévision, ici le 18/02/02 à 00 h TU.

11.4.1.2 Méthodes d'analogie appliquées

Nous disposons de cinq algorithmes de sélection d'analogues. Nous avons choisi de retenir les méthodes A0, A1 et A2bis pour les cas de crue que nous proposons d'étudier. Ceci permettra de mesurer l'apport dans la prévision adaptée par analogie :

- de la pré-sélection dynamique des situations candidates à l'aide de la température (algorithme A1) ;
- de l'analogie sur les champs de vitesse verticale et de la mesure de l'humidité relative à deux niveaux de pression distincts.

11.4.2 Exemple de défaut d'alerte

Nous proposons d'analyser un exemple de défaut d'alerte ayant été détecté lorsque chacune des méthodes A0, A1 et A2bis est appliquée sur le Doubs à Neublans. L'épisode de crue étudié est celui du début du mois de novembre 2002, dont le débit de pointe a été observé le 5 novembre. Le dépassement du débit seuil $S1$ n'a pu être anticipé que la veille de l'événement. L'objectif est de déterminer quels sont les éléments à l'origine de ce défaut d'alerte.

Nous nous plaçons donc dans un contexte opérationnel : nous sommes le 28 octobre 2002 au matin. Les scénarios de débit prévus à partir des méthodes A0, A1 et A2bis sont indiqués sur la Figure 11.14. Le trait en tiretés gris correspond au seuil $S1$ de débit associé au premier niveau de vigilance crues du Doubs à Neublans.

Le dépassement du seuil $S1$ a eu lieu le 4 novembre 2002, soit à l'échéance $J+7$. Les scénarios obtenus sont fortement sous-estimés et n'indiquent aucune hausse significative du niveau du Doubs. L'épisode pluvieux responsable de la montée observée, qui débute le 1er novembre, ne semble pas être prévu par les méthodes d'analogie.

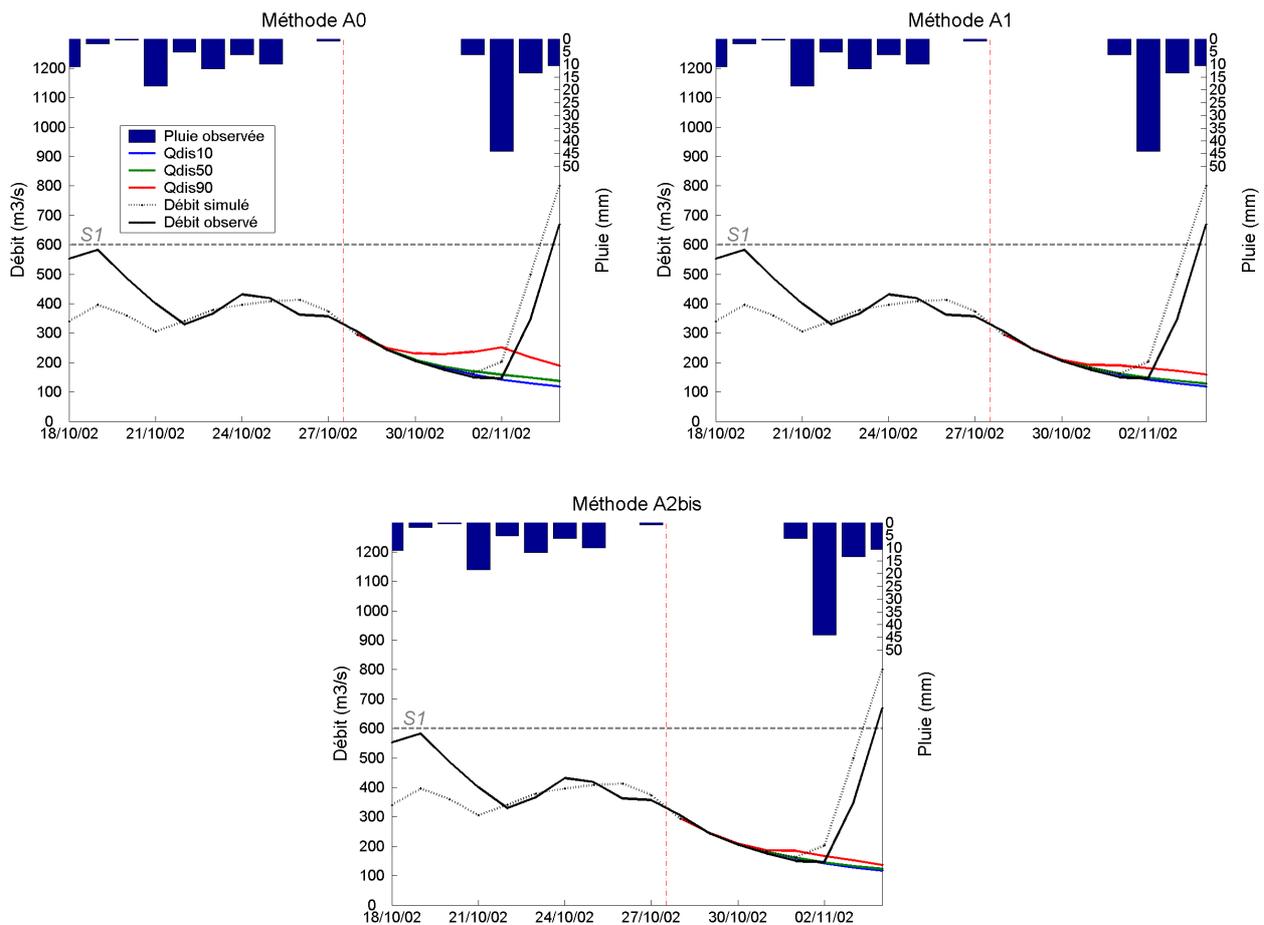


Figure 11.14 : Prévisions probabilistes de débit du Doubs à Neublans émises le 28 octobre 2002 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus des méthodes A0, A1 et A2bis.

Les jours se succèdent. Cinq jours plus tard, nous obtenons les prévisions indiquées sur la Figure 11.15 lorsque les méthodes A0, A1 et A2bis sont appliquées.

Le dépassement du seuil $S1$ par le débit observé a lieu à l'échéance $J+2$. Les prévisions hydrologiques issues des différentes méthodes d'analogie sont similaires : les courbes enveloppes formant l'intervalle de confiance se situent bien en-deçà du débit simulé. On notera simplement que les débits associés aux quantiles $Qdis50$ et $Qdis90$ sont plus élevés par les méthodes A1 et A2bis que ceux obtenus par la méthode de référence A0, et que le quantile $Qdis90$ obtenu par la méthode A2bis atteint presque le débit seuil $S1$. La hausse du niveau du Doubs est tout de même anticipée, ce qui traduit une cohérence temporelle des prévisions de pluie.

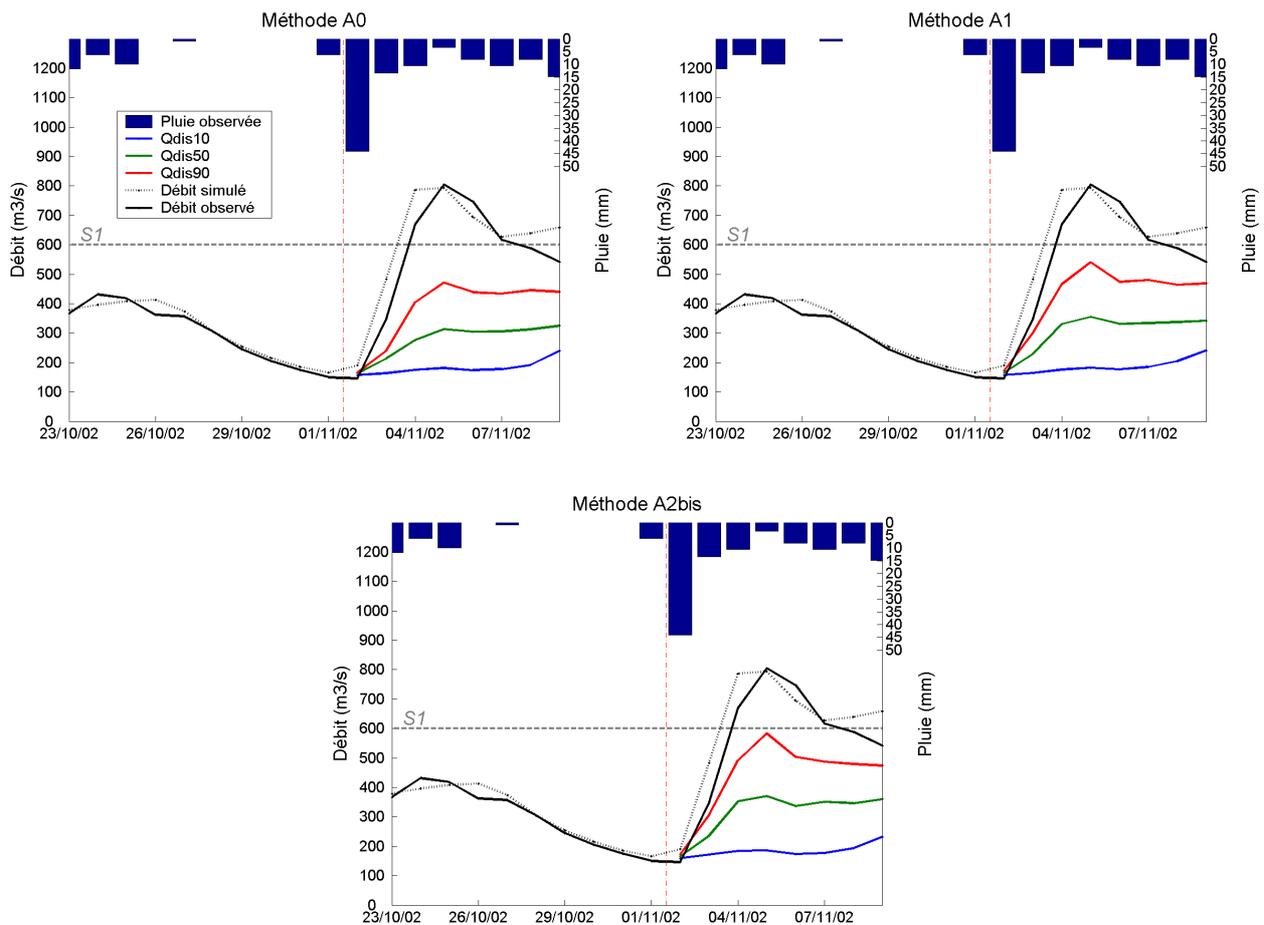


Figure 11.15 : Prévisions probabilistes de débit du Doubs à Neublans, émises le 2 novembre 2002 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus des méthodes A0, A1 et A2bis.

Le débit simulé par le modèle, lorsqu'il est alimenté par les pluies réellement observées, est très proche du débit observé. Les erreurs de prévision ne proviennent donc pas du modèle GR3P. Pour identifier la source, nous proposons d'analyser, par la Figure 11.16, les prévisions de débit obtenues lorsque les prévisions de précipitations sont générées dans un contexte de prévision parfaite, *i.e.* lorsque les champs des variables d'analogie adaptés sont analysés et non prévus par le modèle CEP. Les scénarios *Qdis10 PARF*, *Qdis50 PARF* et *Qdis90 PARF* figurant en pointillés sont ceux obtenus dans un contexte de prévision parfaite des précipitations.

Les scénarios obtenus dans un contexte de prévision parfaite des pluies diffèrent finalement peu des scénarios obtenus en contexte opérationnel. Les erreurs de prévision du modèle CEP sont donc très faibles. L'événement pluvieux principalement responsable de cette crue (plus de 44 mm en 24 h, soit un cumul proche du quantile quinquennal), se produit le jour de la prévision, soit le 2 novembre. Ce défaut d'alerte est donc dû principalement aux méthodes d'analogie, dont la performance est particulièrement limitée par la longueur de l'archive des situations passées.

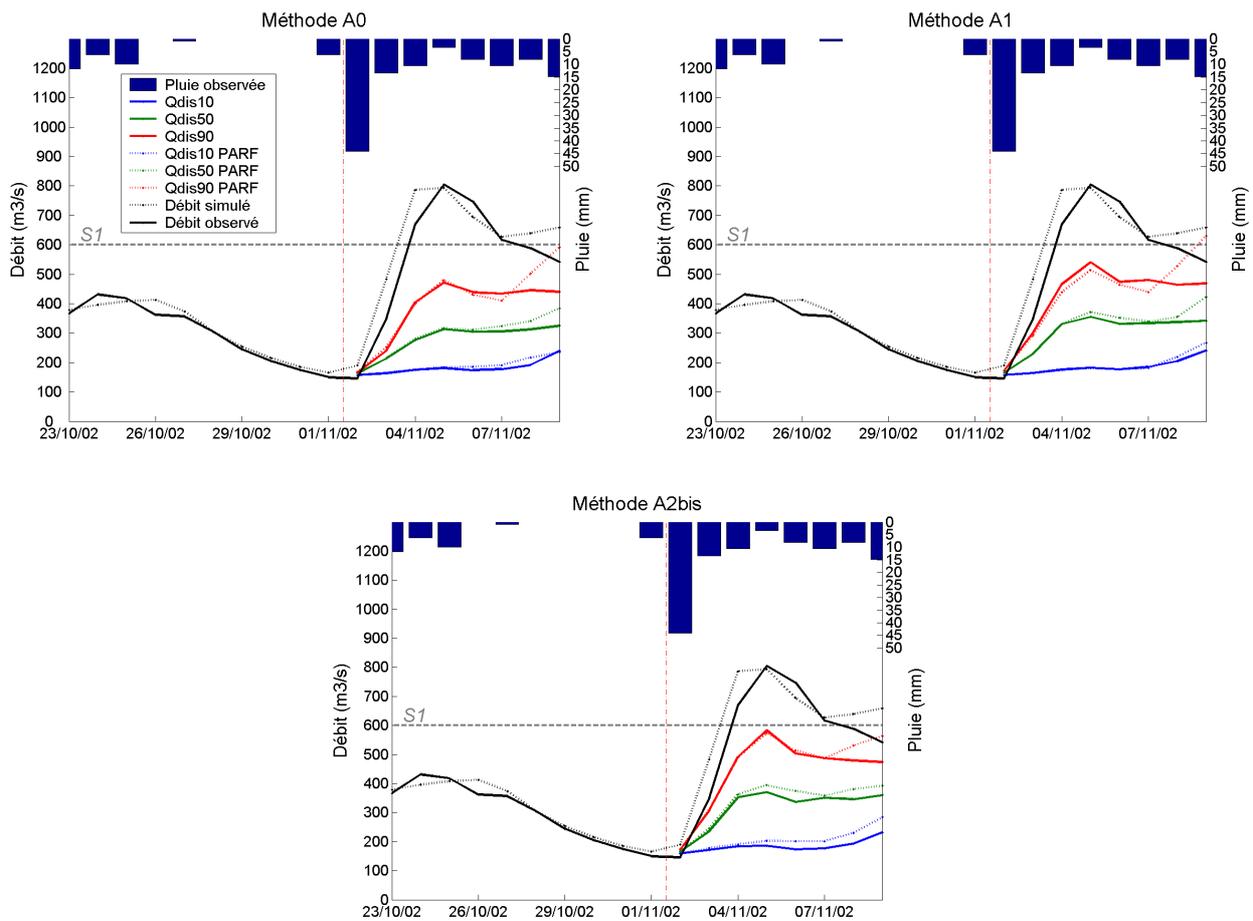


Figure 11.16 : Prévisions probabilistes de débit du Doubs à Neublans, émises le 2 novembre 2002 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus des méthodes A0, A1 et A2bis en contexte de prévision opérationnelle (trait plein) et parfaite (trait discontinu).

Intéressons-nous alors aux prévisions de pluie issues des méthodes d'analogie. Les distributions prévues pour le jour même (2 novembre 2002) obtenues par chacune des méthodes sont indiquées sur la Figure 11.17. D'une part, toutes les méthodes ont sélectionné des situations analogues dont la lame d'eau observée est faible. Les distributions sont assez proches pour les faibles fréquences au non-dépassement. Elles s'écartent ensuite pour des fréquences plus élevées : la distribution obtenue par la méthode A0 sous-estime le plus les précipitations : ainsi près de 80 % des analogues retenues par cette méthode présentent des cumuls inférieurs à 12 mm. En revanche, les distributions obtenues par A1 et A2bis sont moins biaisées et demeurent assez proches. Seule la lame d'eau maximale diffère fortement entre les deux méthodes. La méthode A2bis a ainsi permis de sélectionner une situation analogue dont la lame d'eau observée atteint 37,3 mm. Ce résultat montre que la valeur maximale peut également être considéré comme un indicateur pour les événements pluvieux les plus intenses. Ainsi, si une analogue présente une lame d'eau importante, le potentiel d'un épisode marqué est accentué.

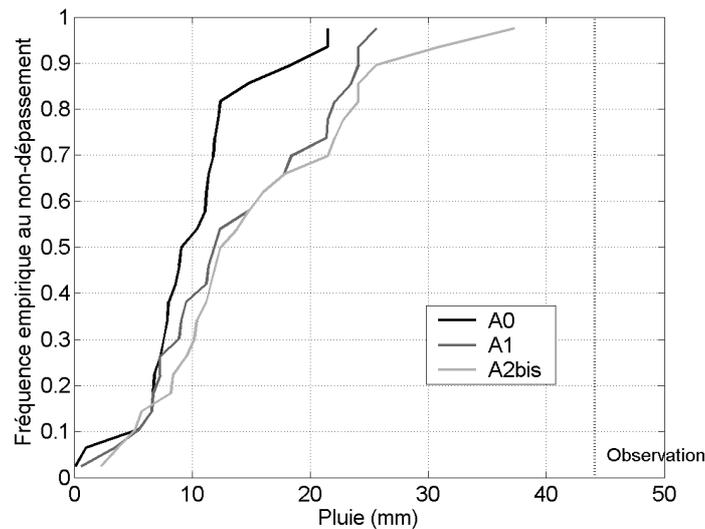


Figure 11.17 : Distributions empiriques des précipitations obtenues par les méthodes A0, A1 et A2bis, prévues le 2 novembre 2002 pour le même jour (échéance J), pour le bassin du Doubs à Neublans.

Cependant, une semaine avant cet événement pluvieux, nous avons sous-entendu lors de l'examen de la Figure 11.14 que les méthodes n'avaient pas pu permettre de l'anticiper. Nous pouvons nous en assurer en analysant l'évolution de la prévision de pluie émise pour ce 2 novembre lors des huit jours de prévision précédents. La Figure 11.18 présente l'évolution des quantiles $Pdis20$, $Pdis60$ et $Pdis90$ extraits des distributions des analogues sélectionnées par les méthodes A0 (a) et A2bis (b) en fonction du jour de l'émission de la prévision.

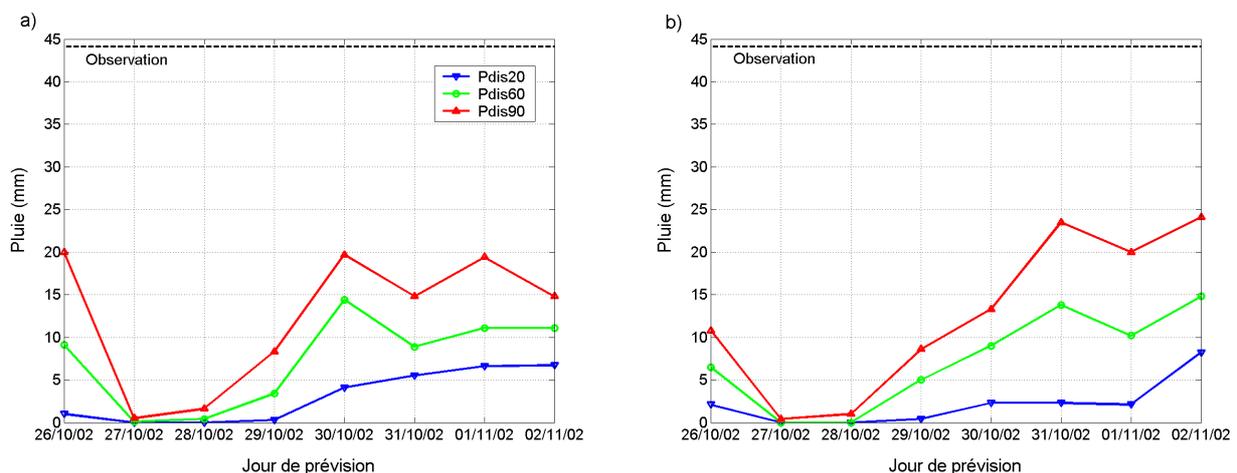


Figure 11.18 : Evolution de la prévision probabiliste des précipitations adaptée par analogie émise pour le 2 novembre 2002 sur le bassin du Doubs à Neublans, en fonction du jour de l'émission de la prévision ($J-7$ à J). a) Par la méthode A0 ; b) par la méthode A2bis.

Huit jours avant l'événement, les distributions prévues ne sont pas si mauvaises puisque les valeurs des trois quantiles sont supérieures à 0 mm, indiquant une forte probabilité de pluie pour le 2 novembre. Six à sept jours avant l'événement, le diagnostic change : les valeurs proposées par les méthodes A0 et A2bis sont très basses, traduisant sans aucun doute

une défaillance dans la prévision des variables d'analogie. Ces résultats traduisent une instabilité de la prévision du MNP. En revanche, cinq jours avant l'événement, les quantiles amorcent en hausse, traduisant de nouveau un risque fort d'événement pluvieux. Les jours suivants, l'épisode se confirme puisque les trois quantiles issus de chacune des méthodes poursuivent leur hausse. Les distributions issues de la méthode A0 sont plus fines que celles issues de la méthode A2bis, ce qui réduit les incertitudes de prévision par la méthode A0 ; mais le fait que le quantile Q_{dis90} obtenu par la méthode A2bis se stabilise autour de 25 mm les trois jours de prévision précédant l'événement indique également un potentiel non négligeable.

11.4.3 Exemple de fausse alerte

Nous proposons d'analyser un exemple de fausse d'alerte ayant été émise lorsque l'une des méthodes A0, A1 et A2bis est appliquée. La période étudiée se situe à la fin du mois de décembre 2002. Durant cette période, des fausses alertes ont été générées par les méthodes d'analogie à plusieurs reprises à l'échéance $J+7$. L'exemple le plus singulier est celui du 25 décembre 2002 sur la Saône à Le Chatelet. La Figure 11.19 présente l'évolution des scénarios de débit prévus (Q_{dis10} , Q_{dis50} et Q_{dis90}) à partir des méthodes A0, A1 et A2bis.

Peu de différences sont observées entre les prévisions émises à partir des méthodes A0 et A1. Les débits prévus sont fortement surestimés : ainsi la valeur du quantile Q_{dis90} atteint près de $950 \text{ m}^3/\text{s}$ à $J+7$ (le 01/01/03), alors qu'en réalité le débit simulé par le modèle n'atteint que $500 \text{ m}^3/\text{s}$. Une petite partie de la surestimation provient sans doute du modèle hydrologique, puisque le débit simulé à $J+7$ est légèrement supérieur au débit observé, d'environ $50 \text{ m}^3/\text{s}$. En revanche, les prévisions obtenues à partir de la méthode A2bis sont moins surestimées. Nous obtenons ici un exemple qui va dans le sens de la diminution du taux de fausses alertes mis en évidence précédemment. Ainsi, le quantile Q_{dis90} prévu en appliquant cette méthode d'analogie ne dépasse pas $800 \text{ m}^3/\text{s}$. Par ailleurs, l'enveloppe représentée par les courbes Q_{dis10} et Q_{dis90} obtenues par A2bis est moins large que celles obtenues par les deux autres méthodes. Ceci conforte également le résultat précédent lors de l'étude de la finesse des prévisions.

Vérifions à présent si les prévisions obtenues par A2bis sont plus fines que celles obtenues par A0. La Figure 11.20 présente, pour la prévision émise le 25 décembre 2002, l'évolution des trois quantiles de prévision en fonction de l'échéance, obtenus a) par la méthode A0 et b) par la méthode A2bis. La prévision émise dans un contexte de prévision parfaite est également indiquée (en pointillés) afin d'estimer la part de l'erreur du modèle CEP dans la surestimation de la prévision des précipitations.

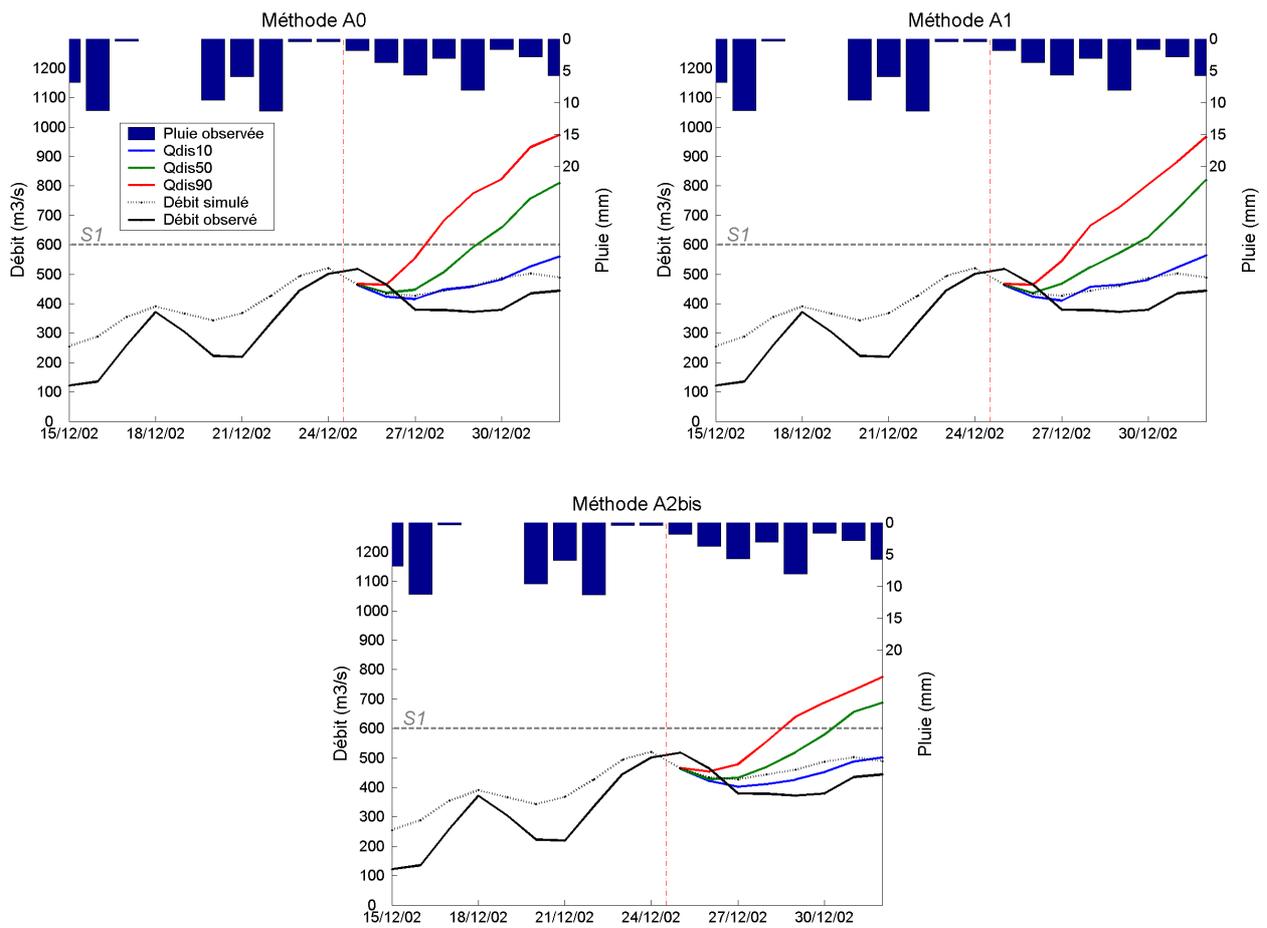


Figure 11.19 : Prévisions probabilistes de débit de la Saône à Le Chatelet, émises le 25 décembre 2002 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus des méthodes A0, A1 et A2bis.

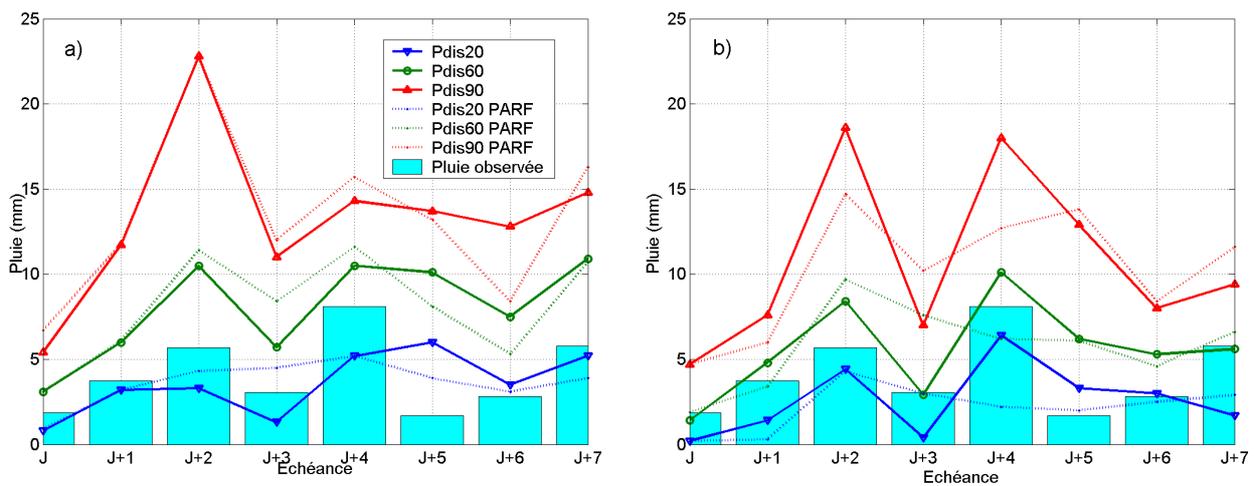


Figure 11.20 : Evolution de la prévision probabiliste des précipitations adaptée par analogie émise le 25 décembre 2002 sur le bassin de la Saône à Le Chatelet, en fonction de l'échéance de la prévision. a) Par la méthode A0 ; b) par la méthode A2bis.

L'enveloppe formée par les courbes *Pdis20* et *Pdis90* est effectivement plus étroite en

moyenne lorsque la méthode A2bis est appliquée, ce qui explique la plus faible largeur de l'enveloppe des scénarios de débit. La prévision des précipitations est bien surestimée par les deux méthodes, mais le biais est bien plus prononcé avec la méthode A0. La succession d'épisodes de faible ampleur systématiquement surestimés est sans doute à l'origine des scénarios de fort débit indiqués sur la Figure 11.19. Par ailleurs, en comparant les quantiles obtenus dans les contextes de prévision opérationnelle et parfaite, il s'avère que l'erreur de prévision des variables d'analogie est faible pour la méthode A0. Les valeurs des quantiles sont très proches pour les premières échéances. En revanche, les erreurs sont plus importantes dès $J+1$ lorsque la méthode A2bis est appliquée. Ce résultat confirme bien la difficulté à bien prévoir les champs de vitesse verticale. On remarque cependant que, contrairement à ce que l'on observe généralement, la prévision des pluies en contexte opérationnel peut s'avérer plus juste que celle obtenue en contexte de prévision parfaite. C'est notamment le cas pour les échéances $J+2$ et $J+3$ avec la méthode A2bis. En outre, les quantiles obtenus dans un contexte de prévision parfaite avec les deux méthodes sont également surestimés. Finalement, au vu de ces résultats, les prévisions probabilistes de débit auraient également été surestimées si le modèle avait été alimenté par les scénarios de précipitations générés dans un contexte de prévision parfaite. Par conséquent, cet exemple de fausse alerte serait principalement dû aux méthodes d'analogie.

11.4.4 Exemples de bonne alerte

Deux exemples de bonne alerte sont proposés. Le premier concerne le Loing à Episy au mois de janvier 2004. Le seuil $S1$ a été dépassé le 14 janvier et le débit de pointe a été observé le 15 janvier. L'événement a pu être anticipé huit jours à l'avance grâce aux méthodes d'analogie.

Nous sommes le 7 janvier 2004 au matin. La chaîne de prévision permet d'obtenir les prévisions représentées sur la Figure 11.21, lorsque chacune des trois méthodes d'analogie est appliquée.

Le seuil a été dépassé à l'échéance $J+7$ (huit jours). Le débit simulé par le modèle lorsqu'il est alimenté par les pluies réellement observées est très proche du débit mesuré à la station d'Episy : l'erreur du modèle GR3P est donc négligeable. Les scénarios obtenus par les différentes méthodes, bien que sous-estimés, notamment avec la méthode A2bis, montrent une hausse fort probable du niveau du Loing pour les dernières échéances. Ainsi la dynamique restituée peut déjà avertir le prévisionniste utilisateur de la chaîne.

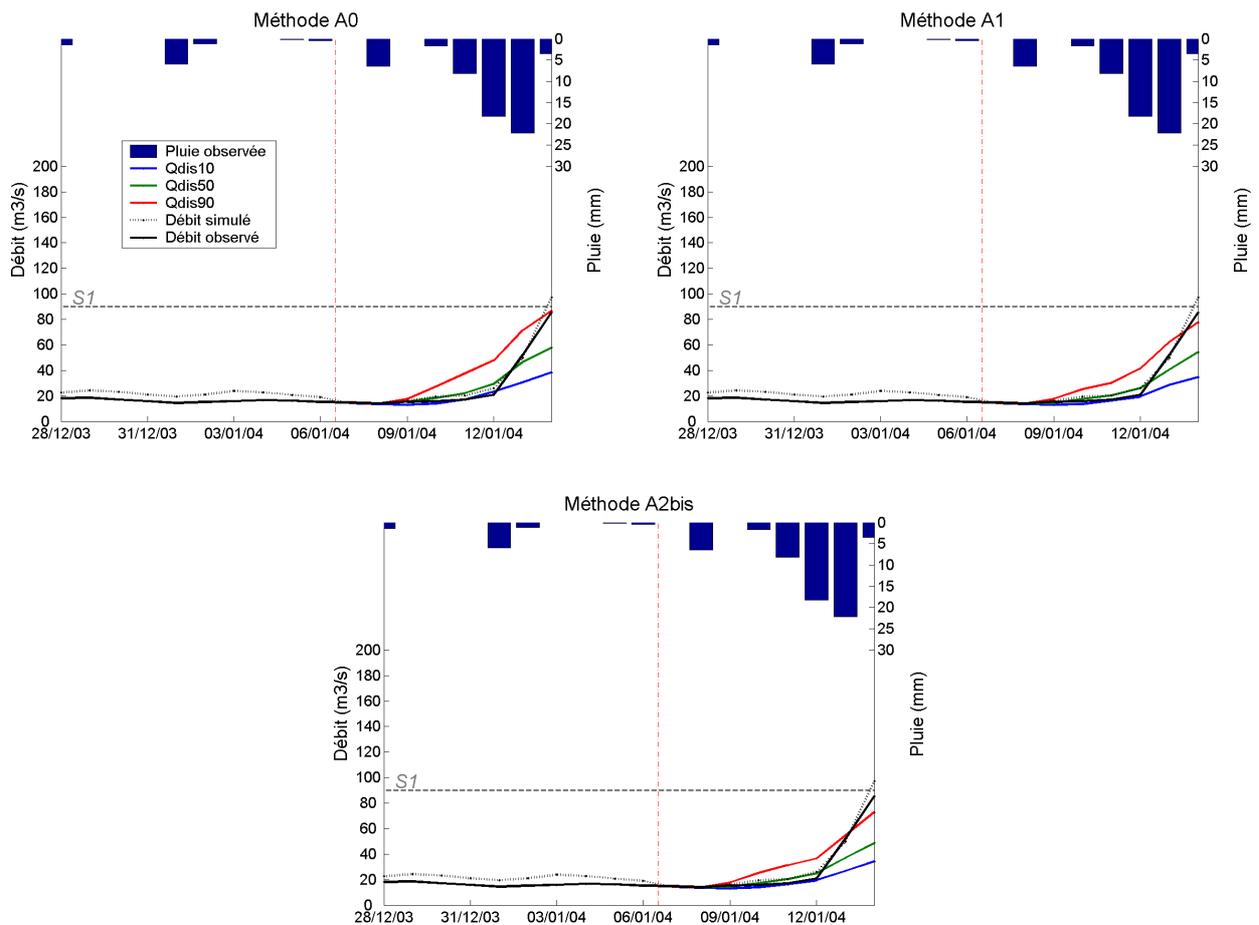


Figure 11.21 : Prévisions probabilistes de débit du Loing à Episy, émises le 7 janvier 2004 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus des méthodes A0, A1 et A2bis.

Le lendemain, les prévisions obtenues confirment un potentiel de crue du Loing (Figure 11.22). Certes une grande partie des scénarios de débit ne dépassent pas le seuil S1, mais la borne supérieure représentée par le quantile Qdis90 dépasse le seuil quelle que soit la méthode appliquée. L'élément à retenir est bien la concordance des scénarios, qui dans l'ensemble indiquent une hausse assez rapide du débit entre le 12 et le 15 janvier. Ceci devrait avertir l'utilisateur de la chaîne d'une hausse fort probable du niveau du Loing. Par ailleurs, la méthode A0 semble d'une part sous-estimer les précipitations : ainsi à J+7 le quantile de débit Qdis90 reste inférieur au débit simulé, ce qui traduit une sous-estimation, même avec les analogues les plus pluvieuses. D'autre part, cette méthode ne semble pas capable de reproduire correctement la répartition temporelle des précipitations en termes quantitatifs. La courbe associée au quantile Qdis90 n'a en effet pas la même pente que la courbe du débit simulé, contrairement aux deux autres méthodes.

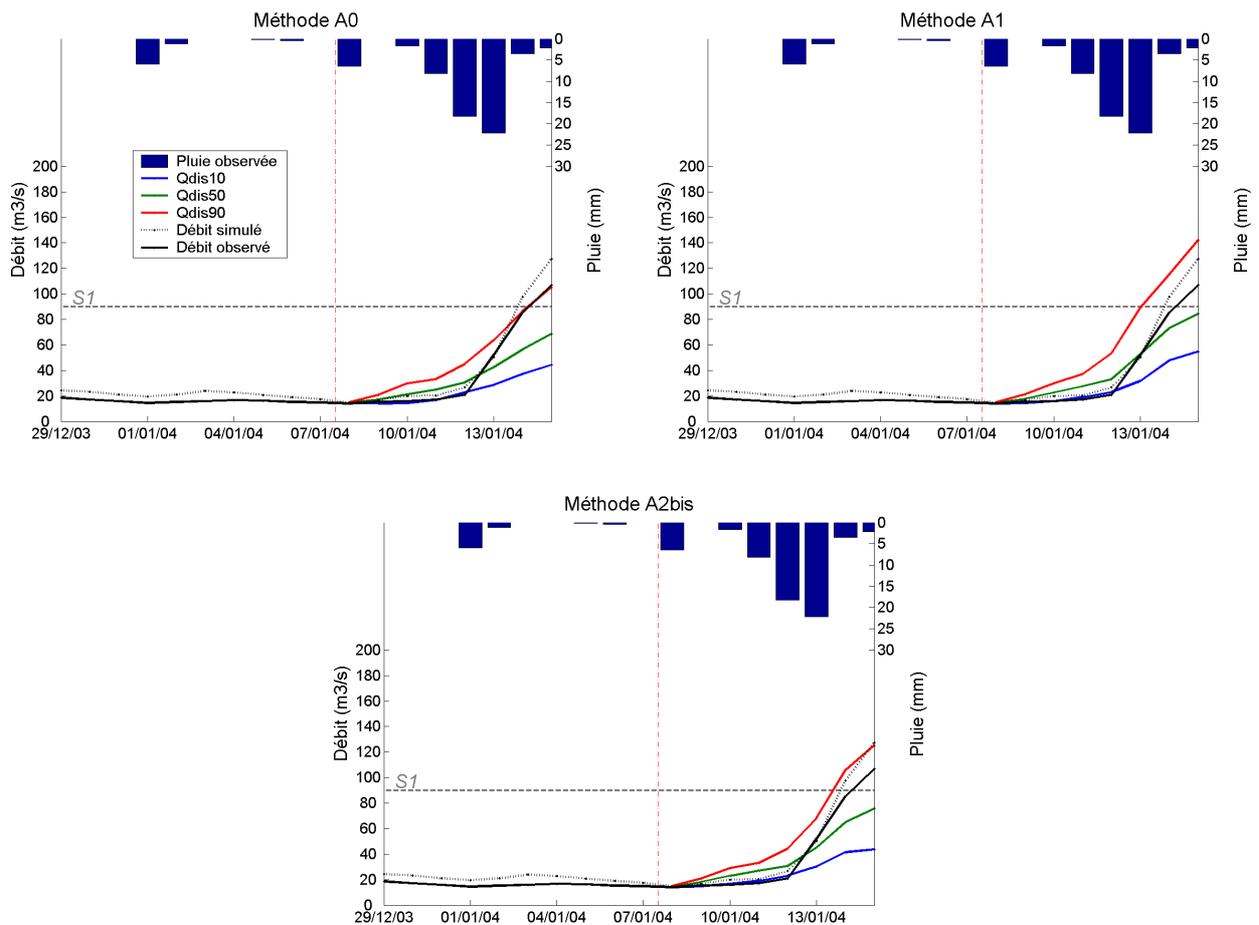


Figure 11.22 : Prévisions probabilistes de débit du Loing à Episy, émises le 8 janvier 2004 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus des méthodes A0, A1 et A2bis.

Nous proposons d'étudier un autre exemple de bonne alerte pour le bassin du Doubs à Neublans. Une crue du Doubs se produisit au début du mois de janvier 2003 avec une pointe de crue observée le 6 janvier. L'algorithme A2bis a été appliqué sur la période précédant la pointe de crue. La Figure 11.23 indique l'évolution des scénarios de débit prévu entre le 27 décembre 2002 et le 5 janvier 2003.

Dès le 27 décembre, un signe avant-coureur apparaît puisque la borne supérieure de la courbe enveloppe (Q_{dis90}) dépasse le seuil $S1$ à partir du 31/12. A échéance $J+7$ (3 janvier), le quantile médian Q_{dis50} atteint également le seuil $S1$, ce qui correspond très bien au débit simulé par GR3P. En outre, la courbe du débit simulé se situe à l'intérieur de l'enveloppe formée par les scénarios Q_{dis10} et Q_{dis90} . Ce résultat indique une certaine justesse des prévisions hydrologiques.

Le lendemain 28 décembre, la hausse du niveau du Doubs est confirmée avec une hausse des débits prévus, les scénarios Q_{dis50} et Q_{dis90} dépassant le seuil $S1$ le 3 janvier, ce qui est conforme à ce qui est simulé par le modèle. Cependant on note une légère défaillance de la prévision du modèle CEP puisqu'une hausse plus sensible est déjà prévue pour les premières échéances, un maximum local étant atteint le 31/12. Cette erreur ne se retrouve plus le jour

de prévision suivant. En effet le 29 décembre, les débits associés aux trois scénarios, bien que légèrement sur-estimés, augmentent progressivement jusqu'à la dernière échéance. Par ailleurs, l'enveloppe formée par les deux scénarios *Qdis10* et *Qdis90* piège également la valeur du débit simulé pour les huit échéances, traduisant une assez bonne fiabilité de la prévision.

Le 30 décembre, l'épisode pluvieux conséquent observé début janvier est anticipé un peu trop tôt par la méthode d'analogie, ce qui se traduit par un pic de crue précoce. Ce petit décalage s'efface un peu le lendemain 31 décembre, le scénario *Qdis90* atteignant son maximum le 5 janvier.

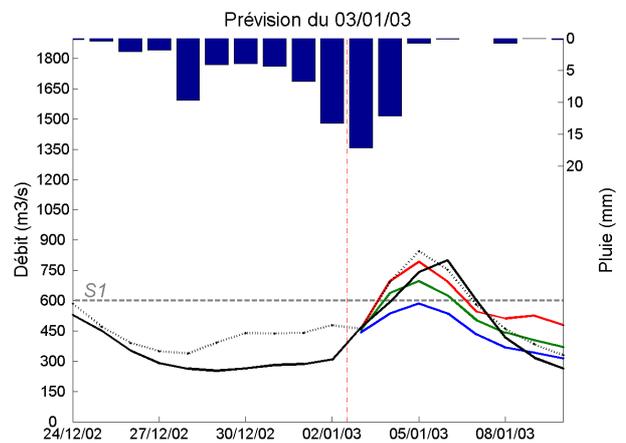
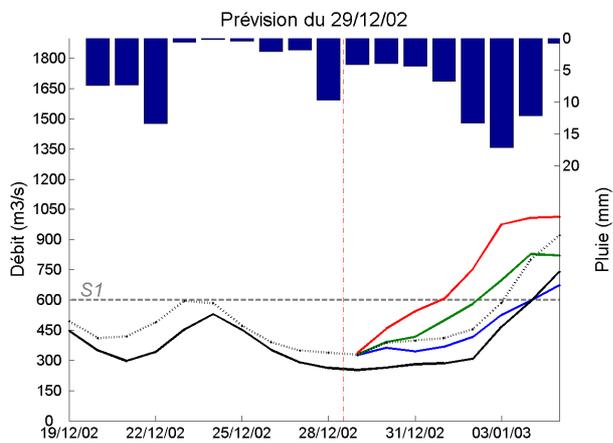
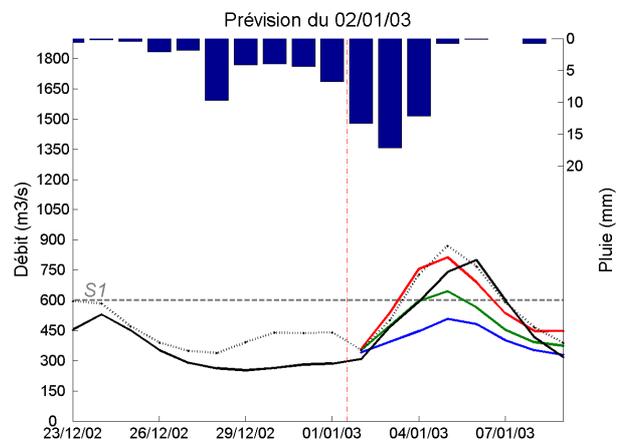
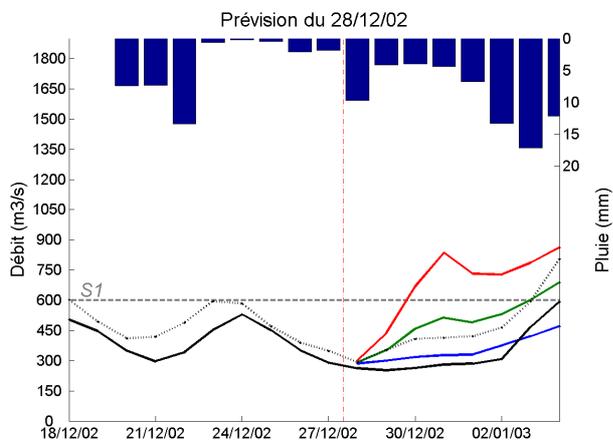
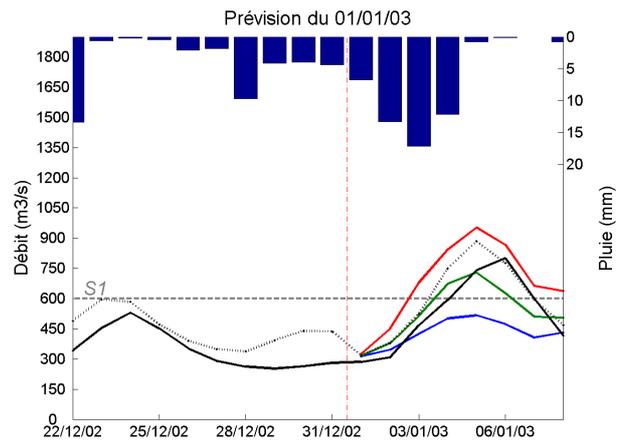
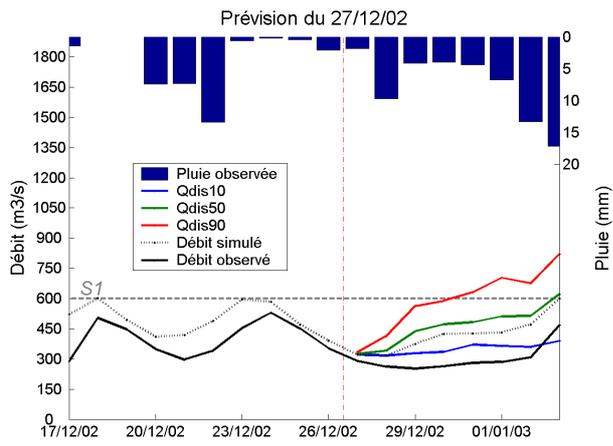
Le lendemain 1er janvier, les prévisions de précipitations semblent cohérentes dans le temps puisque le pic de crue prévu se produit le même jour que celui simulé par le modèle (le 5 janvier). Cependant la pointe n'est observée que le lendemain 6 janvier : ce décalage temporel d'une journée semblerait donc lié à une défaillance de GR3P. Par ailleurs, d'une part la courbe du débit simulé est bien incluse dans l'enveloppe *Qdis10-Qdis90*, et d'autre part près de trois quarts des débits prévus pour le 5 janvier dépassent le débit seuil *S1*. La chaîne de prévision permet ainsi d'anticiper aisément l'épisode de crue, et ce près de cinq jours à l'avance.

Le 2 janvier, les débits prévus pour le 5 sont un peu plus faibles que ceux prévus la veille. Les scénarios de précipitations sont donc probablement légèrement inférieurs à ceux générés la veille pour le 5. Cela n'empêche pas d'obtenir plus de 50 % des débits dépassant le seuil *S1*. En revanche le débit simulé est supérieur au quantile *Qdis90*, ce qui traduit un biais assez prononcé de la prévision hydrologique, probablement du fait d'une prévision des précipitations sous-estimée.

Nous retrouvons à peu près les mêmes résultats le lendemain 3 janvier. La seule différence réside dans les incertitudes qui s'avèrent réduites par rapport à la veille. Ainsi l'écart entre les quantiles *Qdis10* et *Qdis90* ne s'élève plus qu'à $200 \text{ m}^3/\text{s}$, contre $320 \text{ m}^3/\text{s}$ la veille.

Le 4 janvier, la crue est déjà avérée, mais la prévision probabiliste pour les huit jours suivants est très correcte, la distribution des débits étant centrée sur la valeur de débit simulé par le modèle. En outre les incertitudes sont faibles, l'enveloppe restant étroite même jusqu'à *J+7*.

Enfin le 5 janvier, jour où la pointe de crue est simulée par le modèle, la chaîne permet de reproduire presque parfaitement la décrue.



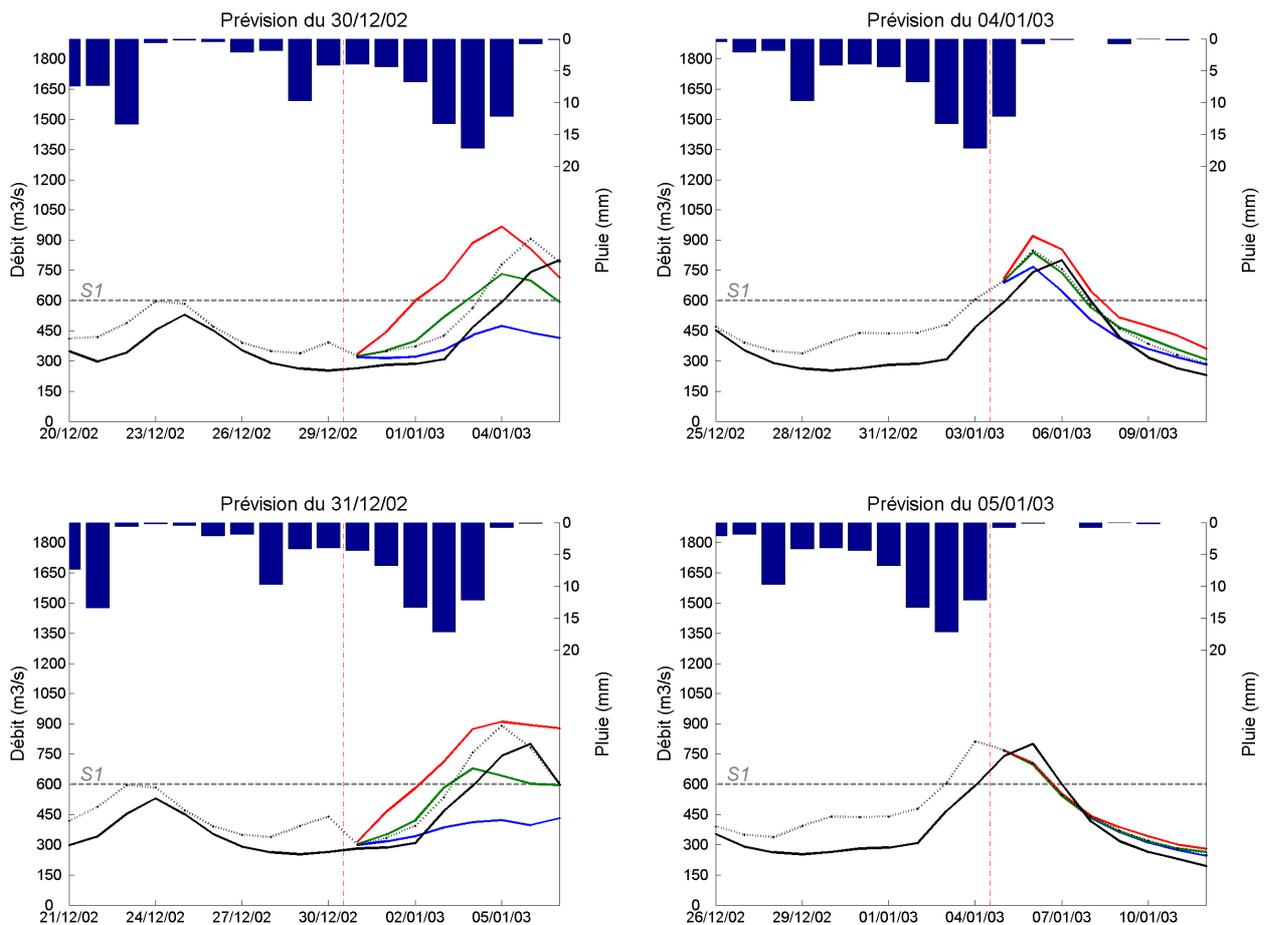


Figure 11.23 : Evolution des prévisions probabilistes de débit du Doubs à Neublans, émises entre le 27 décembre 2002 et le 5 janvier 2003 pour les huit jours suivants, lorsque le modèle est alimenté par les scénarios de précipitations issus de la méthode A2bis.

11.4.5 Comparaison à une prévision déterministe expertisée

Le SPC Seine moyenne-Yonne-Loing établit des prévisions de cote et de débit de certains affluents de la Seine à partir de prévisions de pluie émises par Météo-France (MF). L'objectif de cette section est de comparer, à titre d'illustration, les prévisions issues des méthodes d'analogie à des prévisions déterministes expertisées par MF que le SPC nous a transmises. Les zones sur lesquelles les prévisions sont fournies ont beaucoup évolué dans le temps, si bien qu'une comparaison des méthodes sur une longue série de prévisions est difficile à établir. De plus, les sous-bassins de la Seine que nous avons définis dans cette thèse ne sont pas tous identiques à ceux que le SPC Seine moyenne-Yonne-Loing a en charge de surveiller. Parmi les sous-bassins communs figure celui de l'Yonne à Courlon, retenu pour cette comparaison.

Nous avons évalué les différentes méthodes sur une série de 18 prévisions (15/01/2004 au 01/02/2004). L'échéance maximale couverte par les prévisions de MF sur cette période est de quatre jours ($J+3$) : par conséquent, les valeurs de pluie prévues ont été comparées sur le même échantillon de journées pour les quatre échéances, soit 15 journées (18 janvier au 1er

février 2004). Durant cette période, plusieurs épisodes de précipitations ont été observés sur le bassin de l'Yonne à Courlon (Figure 11.24).

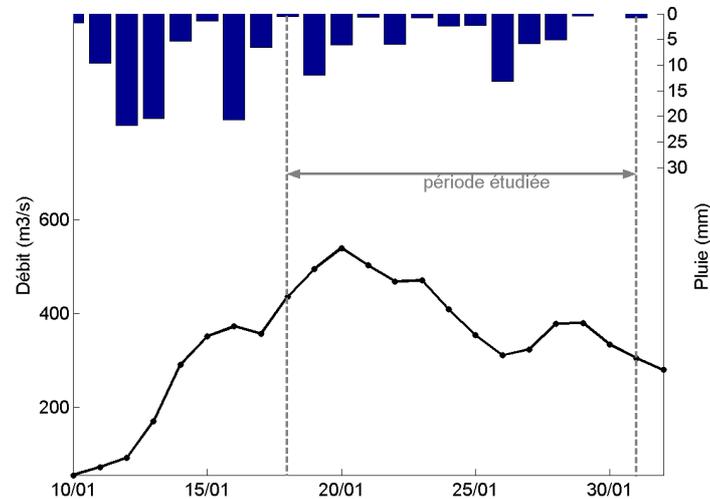


Figure 11.24 : Hydrogramme et précipitations observés entre le 10 janvier et le 2 février 2004 sur le bassin de l'Yonne à Courlon.

La hausse du niveau observée entre le 10 et le 20 janvier est due à de fortes précipitations qui se sont produites avant la période d'étude. Il aurait été intéressant d'évaluer les différentes prévisions sur ces événements, mais les bulletins de MF n'ont pas été disponibles.

Les prévisions de précipitations que MF transmet chaque jour au SPC sont constituées de trois valeurs de pluie pour chaque échéance : un minimum, une valeur moyenne et un maximum. Nous avons choisi de retenir la valeur moyenne pour cette comparaison. La valeur retenue pour chacune des méthodes d'analogie (A0 et A2bis) est celle du quantile $Pdis60$ de la distribution empirique des précipitations.

Les prévisions ont été comparées en évaluant sur la période de 15 jours l'écart absolu moyen (MAE) entre pluie observée et pluie prévue pour chaque échéance (Figure 11.25). Sur cette période, la prévision de MF est meilleure pour les deux premières échéances (J et $J+1$) avec un écart absolu moyen plus faible qu'avec le quantile $Pdis60$ des méthodes d'analogie. En revanche pour les échéances suivantes ($J+2$ et $J+3$), la tendance est inversée. L'écart est plus faible lorsque la prévision est définie par le quantile $Pdis60$ obtenu par les méthodes A0 et A2bis. L'erreur entre la pluie observée et la valeur moyenne définie par MF augmente sensiblement entre $J+1$ et $J+3$. Par ailleurs, la valeur de l'écart lorsque la prévision est obtenue par les méthodes d'analogie s'avère presque constante sur les quatre premières échéances. La prévision est même plus fiable à $J+1$ et $J+2$, ce qui montre à nouveau que les prévisions adaptées par analogie peuvent être instables. Il est donc nécessaire de comparer, pour un jour donné, les prévisions émises les jours précédents afin d'avoir une idée de la fiabilité de la prévision par une méthode d'analogie. Enfin, nous remarquons finalement que l'erreur de prévision, sur cette courte période, est équivalente entre les deux méthodes A0 et

A2bis, cette dernière étant légèrement avantagée, notamment pour les trois premières échéances.

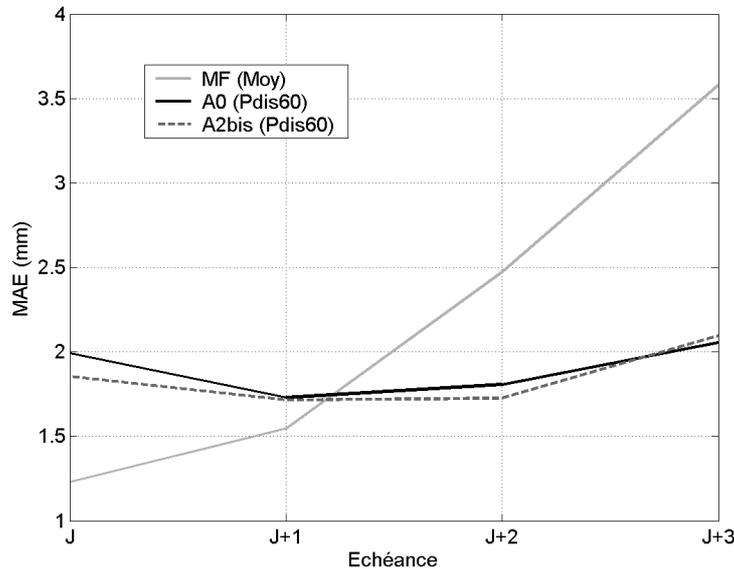


Figure 11.25 : Evolution de l'erreur absolue moyenne entre pluie observée et pluie prévue en fonction de l'échéance de prévision, lorsque la pluie prévue est constituée de la valeur moyenne de MF, des valeurs du quantile Pdis60 obtenues par les méthodes A0 et A2bis.

11.5 Conclusions

Les différents algorithmes de sélection d'analogues, développés dans la seconde partie de ce manuscrit, ont été appliqués dans ce chapitre dans un contexte de prévision opérationnelle. Une période d'une durée de trois ans a été retenue (01/10/2001-01/10/2004), sur laquelle nous disposons de prévisions archivées élaborées par le CEPMMT. Cela nous a permis d'évaluer, sur un échantillon de plus d'un millier de journées, les performances des prévisions de précipitations adaptées par analogie. Des scores identiques à ceux employés lors de l'évaluation des performances dans un contexte de prévision parfaite ont été appliqués. Les résultats ont révélé que :

- les méthodes A2 et A2bis développées dans cette thèse s'avèrent globalement plus performantes que les autres méthodes pour les trois ou quatre premières échéances (J à $J+2$ ou $J+3$), les méthodes devenant équivalentes pour les échéances suivantes ;
- l'intérêt d'appliquer une prévision adaptée par analogie par rapport à une prévision naïve telle que la prévision climatologique devient limité à partir de l'échéance $J+6$;
- la qualité de la prévision des variables d'analogie fournie par le MNP a un impact sur les performances dès l'échéance $J+1$, mais la baisse de la fiabilité de la prévision du MNP n'influence que la justesse de la prévision probabiliste de pluie, la finesse de la prévision n'étant liée intrinsèquement qu'à l'algorithme ;

- la vitesse verticale est une variable moins robuste que les géopotentiels ou les variables d'humidité, la justesse de prévision de l'algorithme A2bis devenant plus faible que celle de l'algorithme A0 à partir de l'échéance $J+2$.

Nous avons ensuite analysé le comportement du modèle hydrologique lorsqu'il est alimenté par des prévisions opérationnelles de précipitations. L'évaluation des performances a été basée sur les mêmes scores que ceux utilisés en contexte de prévision parfaite des précipitations. Les performances obtenues ne sont pas identiques sur les bassins tests. De manière générale, le modèle hydrologique filtre les erreurs issues des méthodes d'analogie, si bien que les différences entre les méthodes en termes de débit prévu deviennent moins perceptibles. La comparaison des débits prévus lorsque le modèle est alimenté en prévisions de pluie obtenues dans les contextes de prévision opérationnelle et parfaite a permis de détecter le biais dû au MNP à partir de l'échéance $J+3$. Nous avons également montré que l'erreur se propage plus fortement dans la chaîne de prévision lorsque la méthode appliquée utilise les champs de vitesse verticale pour sélectionner les analogues. Ainsi, lorsque la vitesse verticale est utilisée, les erreurs de prévision du MNP se répercutent plus sensiblement et plus rapidement sur les prévisions de débit. En revanche, nous avons montré que l'évaluation des prévisions dépend du score utilisé. Ainsi, les méthodes A0 et A1 s'avèrent plus performantes que les autres si l'évaluation est basée sur le score *CRPSS*, qui donne un poids équivalent à toute la gamme de débits possibles ; tandis qu'elles sont moins performantes que les autres si la prévision est évaluée par le score *RMSE*, qui donne plus de poids aux écarts importants. Nous avons déduit de ces résultats que les débits sont plus souvent sur-estimés par les méthodes A0 et A1 que par les autres méthodes.

La part respective dans l'erreur de prévision des différentes mailles de la chaîne (MNP, méthode d'analogie, modèle GR3P) a également été évaluée. L'erreur due au modèle GR3P s'avère prépondérante, suivie de près par l'erreur due à la méthode d'analogie. En revanche, la part de l'erreur due à la qualité de prévision du MNP est négligeable devant les deux autres sources.

Nous nous sommes ensuite intéressés aux scores *POD* et *FAR* en fonction du seuil. Pour cela, nous avons considéré des seuils de débit proches de ceux utilisés par les services de prévision des crues sur les bassins tests. L'analyse s'est focalisée sur les événements supérieurs au premier seuil (*S1*), qui sont en nombre suffisants pour autoriser un traitement statistique. Les résultats ont montré que le taux de détection est assez élevé, même lorsqu'un faible quantile (tel que le *Qdis10*) est utilisé comme prévision. Par ailleurs, la capacité à détecter les événements baisse plus rapidement sur les premières échéances, puis elle semble se stabiliser sur les dernières échéances, notamment lorsque le quantile de prévision est élevé. Le taux de fausses alertes est en revanche plutôt stable jusqu'à l'échéance $J+5$, puis il augmente les échéances suivantes. Cette augmentation s'est avérée moins sensible lorsque l'analogie sur la vitesse verticale est recherchée, ce qui confirme les résultats obtenus précédemment en prévision parfaite en termes de *FAR*.

Enfin, nous avons étudié quelques cas afin d'illustrer la capacité de la chaîne de prévision à

bien prévoir des événements de crue. Nous avons considéré des cas de défaut d'alerte, de fausse alerte et de bonne alerte. Nous pouvons retenir de cette étude les éléments suivants :

- la méthode A2bis est à privilégier en raison du taux de fausses alertes plus faible, et ce en dépit d'un taux de détection un peu plus faible ;
- les prévisions obtenues par une méthode d'analogie en contexte opérationnel doivent être croisées jour après jour avec les prévisions des jours précédents afin de se donner une idée du potentiel d'un épisode de crue. En effet, nous avons noté des irrégularités dans les prévisions, si bien que l'on ne peut pas se fier à la prévision du jour J sans la comparer aux prévisions récentes émises pour le même jour ;
- une méthode d'analogie étant biaisée pour les événements forts, les prévisions sont notamment utiles lorsque plusieurs épisodes pluvieux se succèdent, contribuant ainsi à une saturation progressive du bassin et conduisant ensuite à une crue ; les événements ponctuels forts peuvent être anticipés en débaisant éventuellement la prévision, ce qui n'a pas été fait ici.

Nous avons également comparé, sur une courte période à titre d'illustration, les prévisions par analogie à des prévisions expertisées par Météo-France, sur le bassin de l'Yonne à Courlon pour les quatre premières échéances (J à $J+3$). Les prévisions expertisées se sont avérées plus performantes pour les deux premières échéances (J et $J+1$), puis ce sont les méthodes d'analogie qui ont conduit aux erreurs les plus faibles pour les deux dernières échéances ($J+2$ et $J+3$).

Conclusions & perspectives

Ces travaux de thèse ont visé à adapter une méthode de prévision probabiliste des précipitations à des grands bassins versants français, ici la Saône à Lyon et la Seine à Paris, dans une perspective d'application à la prévision des débits. Cette méthode est fondée sur le principe d'analogie : en considérant une journée cible pour laquelle on souhaite prévoir les précipitations associées, l'objectif est de rechercher dans une archive les situations météorologiques passées les plus similaires à celle du jour cible. Cette identification repose sur un choix judicieux de variables les plus informatives et sur des critères numériques pré-définis. Les cumuls pluviométriques des journées analogues sont ensuite exploités pour construire la distribution empirique des pluies, et estimer ainsi de manière probabiliste la quantité de précipitations attendue pour la journée cible. Nous avons repris la méthode développée et appliquée par Bontron (2004) sur des bassins réactifs du Sud-Est de la France, et qui a formé notre méthode de référence nommée A0.

La méthode de référence A0 est constituée de deux niveaux de sélection d'analogues : le premier niveau (niveau 1) consiste à rechercher l'analogie en terme de circulation générale, et le second (niveau 2) permet de ne retenir, parmi les situations sélectionnées au niveau 1, que les situations présentant les mêmes champs d'humidité.

Nous disposons pour cette thèse de plusieurs bases de données que nous avons comparées. Ceci nous a permis de choisir l'archive pluviométrique et l'archive météorologique qui ont été considérées pour la suite du travail. Les différents critères numériques permettant d'évaluer les performances d'une méthode de prévision probabiliste ont également été présentés.

Nous nous sommes attelés au développement de la méthode de référence en appliquant différents algorithmes au bassin de la Saône. Cette tâche a été effectuée dans un contexte de prévision parfaite, c'est-à-dire que les situations cibles n'étaient pas prévues par un modèle de prévision, mais décrites par des champs météorologiques réanalysés, directement extraits de l'archive ERA-40.

La pré-sélection initiale des situations candidates de l'algorithme A0 dans une fenêtre temporelle de quatre mois centrées sur le jour cible permet de prendre en compte explicitement les effets saisonniers, et donc implicitement la température. Or, ce critère

calendaire figé s'avère trop rigide et semble dans certains cas écarter des situations intéressantes, si bien qu'une situation analogue potentiellement informative du cumul de précipitation prévu, située à l'extérieur de la fenêtre temporelle, ne peut être considérée. Afin de supprimer ce critère calendaire, la température a été introduite dans un niveau préliminaire de sélection des situations candidates. Après optimisation des paramètres, le gain de performance obtenu s'est avéré substantiel, que ce soit sur le bassin de la Saône ou sur celui de la Seine. Ce nouvel algorithme, que nous avons nommé A1 par la suite, a été ensuite validé sur une période cible indépendante de celle sur laquelle l'optimisation avait été effectuée.

Dans un second temps, nous avons cherché à introduire une variable d'analogie qui permette de caractériser et de localiser les fronts, générateurs de l'essentiel des précipitations sur nos bassins tests. Plusieurs variables ont été envisagées, mais c'est finalement la vitesse verticale qui s'est révélée la plus intéressante pour améliorer la sélection d'analogues. Désormais, trois niveaux de sélection d'analogues sont appliqués après la sélection préliminaire sur l'état thermique : un premier niveau (niveau 1) en termes de circulation générale, un second (niveau 2) en termes de mouvement vertical, et un dernier (niveau 3) en termes d'humidité. Cet algorithme, baptisé A2, a été validé avec succès, et est apparu plus performant que l'algorithme A1 sur les deux bassins tests.

Enfin, nous avons cherché à optimiser les procédures d'identification des analogues selon les variables de pression et d'humidité de la méthode de référence A0. Pour ce faire, la méthode A2, constituée d'un niveau pré-sélectif suivi de trois niveaux de sélection (niveaux 1, 2 et 3), a été appliquée. Cette optimisation a montré, sur les deux bassins tests, l'intérêt d'une mesure de l'humidité à deux niveaux de pression plutôt qu'un : l'un des niveaux se situe dans les très basses couches, et l'autre dans les couches moyennes de la troposphère. Cet algorithme amélioré, baptisé A2bis, a ensuite été validé avec succès. La même variante à deux niveaux d'humidité (A1bis) a été proposée pour l'algorithme A1, et les résultats se sont également révélés concluants.

A l'issue de cette phase de développement, nous avons finalement conservé les quatre algorithmes (A1, A1bis, A2, A2bis) pour l'application hydrologique, qui constitue l'essentiel de la troisième partie de ce document. L'objectif de cette partie était d'alimenter un modèle pluie-débit en scénarios de prévision de pluie générés par les méthodes d'analogie dans deux contextes : en prévision parfaite (situations cibles extraites des réanalyses) et en prévision opérationnelle (situations cibles prévues par un modèle numérique de prévision). Ainsi, nous avons pu étudier le comportement de la chaîne de prévision dans ces deux contextes.

Le modèle hydrologique choisi est GR3P, qui est un modèle global conceptuel à réservoirs, spécifiquement adapté à la prévision des crues. Il a d'abord été calé sur chacun des sous-bassins de la Saône et de la Seine sur la plus longue période possible, puis nous avons vérifié qu'il était capable de reproduire correctement le débit lorsqu'il est alimenté par des pluies observées. Par ailleurs, une méthodologie a dû être développée pour la génération des scénarios probabilistes de débits à partir des prévisions issues d'une méthode d'analogie par

adaptation statistique. La principale difficulté réside dans la construction de scénarios cohérents dans le temps pour les échéances désirées. En effet, en adaptation statistique, les situations analogues retenues chaque jour sont indépendantes dans le temps, puisque les sorties d'un MNP sont adaptées indépendamment pour chaque échéance. Pour y remédier, une approche innovante a été adoptée : nous avons construit des séquences de pluie en tirant aléatoirement, à chaque échéance, une valeur dans la distribution empirique définie par les situations analogues retenues. Les scénarios sont suffisamment nombreux pour conserver le caractère probabiliste aux prévisions de débit.

Nous avons ensuite procédé à une analyse approfondie des performances de prévision de pluie dans les contextes de prévision parfaite et opérationnelle. Pour ce dernier cas, une période d'une durée de trois ans a été retenue (01/10/2001-01/10/2004), sur laquelle nous disposons de prévisions archivées élaborées par le centre européen CEPMMT. Cela nous a permis d'établir un traitement statistique pertinent sur les résultats. Les prévisions issues des méthodes développées s'avèrent beaucoup plus fines que celles issues de la méthode de référence A0, notamment lorsque la vitesse verticale est introduite (algorithme A2). En revanche, les gains en justesse de prévision sont moindres. En contexte opérationnel, les méthodes qui intègrent une analogie sur les mouvements verticaux se sont révélées globalement plus performantes que les autres méthodes pour les trois ou quatre premières échéances (J à $J+2$ ou $J+3$). Par ailleurs, la qualité de la prévision des variables d'analogie fournie par le modèle numérique a un impact sur les performances dès l'échéance $J+1$, mais la baisse de la fiabilité n'influence que la justesse de la prévision probabiliste de pluie, la finesse de la prévision n'étant liée intrinsèquement qu'à l'algorithme. Nous avons également déduit que la vitesse verticale est une variable moins robuste que les géopotentiels ou les variables d'humidité, la justesse de prévision de l'algorithme A2bis devenant plus faible que celle de l'algorithme A0 à partir de l'échéance $J+2$.

L'évaluation des performances de prévision en fonction du seuil de précipitations a permis de mettre en évidence plusieurs éléments : nous avons remarqué, à l'aide de scores globaux, que les méthodes A2 et A2bis sont généralement les plus performantes, notamment pour des événements pluvieux modérés ou d'intensité supérieure. Cependant, l'analyse du taux de détection et du taux de fausses alertes a permis de montrer que les méthodes A2/A2bis sont moins aptes que les autres à détecter des événements pluvieux conséquents, tandis que la méthode A1bis semble plus adaptée à la détection d'événements, notamment d'intensité faible à modérée. En revanche, l'introduction de la vitesse verticale dans l'algorithme permet de diminuer assez nettement le taux de fausses alertes. Le gain de performance apporté par les méthodes A2 et A2bis, révélé par les scores globaux, est donc principalement dû à la diminution du taux de fausses alertes. Par ailleurs, la prévision des événements forts est très souvent sous-estimée par les méthodes d'analogie, du fait d'une longueur limitée de l'archive de situations passées. Cette sous-estimation s'est révélée accrue lorsque la vitesse verticale est introduite, *i.e.* lorsque les méthodes A2/A2bis sont appliquées.

Nous avons ensuite alimenté le modèle GR3P en scénarios de précipitations issus des

méthodes d'analogie développées dans cette thèse. La chaîne de prévision ainsi formée a été évaluée dans les contextes de prévision parfaite et opérationnelle des précipitations.

L'analyse des prévisions de débit obtenues dans un contexte de prévision parfaite des précipitations a montré que de manière générale, les différences de performance observées entre les méthodes sont faibles. Le modèle hydrologique semble ainsi filtrer les différences de prévision de pluie sur les prévisions de débit, même pour une échéance de l'ordre de la semaine. En nous intéressant à des événements supérieurs à un seuil de débit, nous avons retrouvé des résultats identiques à ceux obtenus sur les pluies : le taux de détection d'un événement est systématiquement plus faible mais les fausses alertes sont plus rares lorsque les scénarios de pluie injectés dans le modèle sont générés par une méthode intégrant l'analogie sur la vitesse verticale. Ce fut aussi l'occasion d'étudier la sensibilité des performances de prévision au choix d'un quantile de débit particulier.

Nous avons également comparé les débits prévus en contexte opérationnel avec ceux prévus lorsque le modèle est alimenté en prévisions de pluie obtenues en contexte de prévision parfaite. Le biais dû à la prévision des champs des variables d'analogie par le modèle numérique a ainsi été détecté à partir de l'échéance $J+3$. En particulier, l'erreur semble se propager plus fortement dans la chaîne de prévision lorsque la méthode appliquée utilise les champs de vitesse verticale pour sélectionner les analogues. L'évaluation de la part respective des différents maillons de la chaîne (modèle météorologique, méthode d'analogie, modèle GR3P) dans l'erreur de prévision a montré que celle due au modèle GR3P s'avère prépondérante, suivie de près par l'erreur due à la méthode d'analogie. En revanche, la part de l'erreur due à la qualité de prévision du modèle numérique est finalement négligeable devant les deux autres sources.

Nous nous sommes ensuite intéressés aux événements supérieurs à un seuil de débit. Afin d'être au plus près des attentes des services de prévision des crues, nous avons considéré des seuils de débit proches des seuils de vigilance crues. Les résultats ont montré que le taux de détection est assez élevé, même lorsqu'un faible quantile est utilisé comme prévision. Par ailleurs, la capacité à détecter les événements baisse plus rapidement sur les premières échéances, puis elle semble se stabiliser sur les dernières échéances, notamment lorsque le quantile de prévision est élevé. Le taux de fausses alertes est en revanche plutôt stable jusqu'à l'échéance $J+5$, puis il augmente les échéances suivantes. Cette augmentation s'est avérée moins sensible lorsque l'analogie sur la vitesse verticale est recherchée, ce qui confirme les résultats obtenus en prévision parfaite.

Cette analyse en contexte opérationnel s'est achevée par une étude de cas de crue. Il en ressort que les prévisions obtenues par une méthode d'analogie en contexte opérationnel doivent être croisées jour après jour avec les prévisions des jours précédents afin de se donner une idée du potentiel d'un épisode de crue. En effet, nous avons noté des irrégularités dans les prévisions, si bien que l'on ne peut pas se fier à la prévision du jour J sans la comparer aux prévisions récentes émises pour le même jour. Il faut également garder à l'esprit qu'une méthode d'analogie est affectée d'un biais, d'une sous-estimation, pour les

événements forts. De ce fait, les prévisions sont notamment utiles lorsque plusieurs épisodes pluvieux se succèdent, contribuant ainsi à une saturation progressive du bassin et conduisant ensuite à une crue. En revanche, les événements ponctuels forts pourraient être anticipés en débiaisant éventuellement la prévision. Ceci n'a pas été fait ici mais mériterait d'être exploité¹.

Par ailleurs, pour une prévision probabiliste de débits obtenue à partir d'une prévision de pluie adaptée par analogie, nous pouvons donner quelques recommandations, sur la base de l'expérience acquise dans ces travaux, afin d'exploiter les prévisions. Si l'on se réfère aux quantiles de référence tels que les quantiles au non-dépassement d'ordre 10 %, 50 % et 90 % ($Qdis10$, $Qdis50$ et $Qdis90$) :

- il est possible de déclencher une alerte lorsque le quantile $Qdis50$ dépasse le premier seuil de vigilance ($S1$), à condition que les incertitudes ne soient pas trop fortes, *i.e.* que le quantile $Qdis10$ ne soit pas trop éloigné du seuil. Pour des seuils plus élevés, en raison du biais dans la prévision de pluie, il est préférable de s'appuyer sur un quantile plus élevé tel que le quantile $Qdis90$;
- la vigilance peut être renforcée lorsque la valeur de ce quantile devient élevée avec l'augmentation de l'échéance. Dans ce cas, cela renseigne sur le risque d'une montée significative du niveau du cours d'eau. Il est également utile, dans ce cas, de s'intéresser aux prévisions de pluie adaptées par analogie : on peut alors se fier aux pluies prévues les plus fortes de la distribution des analogues, voire directement aux champs de pression associés à ces situations analogues.

Enfin, nous avons également comparé, sur une courte période à titre d'illustration et pour un des sous-bassins tests, les prévisions par analogie à des prévisions expertisées par Météo-France, pour les quatre premières échéances (J à $J+3$), seules échéances accessibles par les services de prévision des crues. Nous retiendrons de cette étude comparative que des prévisions expertisées peuvent apporter une information plus utile et plus juste pour les deux premières échéances (J et $J+1$). En revanche, pour des échéances supérieures, les services de prévision des crues peuvent s'appuyer sur une méthode de prévision de pluie adaptée par analogie, d'une part car l'information peut s'avérer complémentaire de l'information apportée par une prévision expertisée, et d'autre part car cette approche propose des prévisions à horizon plus lointain. Par ailleurs, cette approche permet de mieux gérer les incertitudes puisqu'elle fournit une prévision probabiliste. Il aurait été intéressant d'effectuer cette comparaison sur une plus longue série, mais la disponibilité des données, très limitée, ne nous a pas permis de le faire.

En conclusion, la méthode des analogues a été sensiblement améliorée au cours de ce travail. Ainsi, le gain de performance final obtenu par l'algorithme développé dans cette thèse s'élève à plus de 10 % par rapport à l'algorithme de référence mis au point par Bontron

¹ Notons que des techniques permettant de débiaiser les prévisions sont employées par des utilisateurs opérationnels de la prévision par analogie.

(2004). Ce gain est par ailleurs du même ordre de grandeur que celui que Bontron avait obtenu par rapport à l'algorithme développé par Guilbaud (1997). Pour une application opérationnelle, nous recommanderions d'appliquer la version la plus développée (méthode A2bis), étant donné que le gain de performance moyenne est le plus élevé. Pour une utilisation spécifique à la détection d'événements (dépassement de seuils), il serait souhaitable d'appliquer en plus un algorithme incluant la pré-sélection selon la température mais qui n'intègre pas d'analogie sur les mouvements verticaux. En effet, la méthode A1 (ou A1bis) permettrait de mieux détecter des dépassements de seuil ; tandis que l'algorithme incluant l'analogie sur la vitesse verticale permettrait de diminuer les fausses alertes dans certains cas (prévisions à inter-comparer).

Au-delà du travail d'amélioration de la méthode de prévision de pluie, ce travail de thèse a permis de synthétiser toute une série de critères d'évaluation de la performance d'outils de prévisions probabilistes, qui peuvent être utilisés à profit par les SPC et le SCHAPI :

- décomposition du score *CRPS* en finesse et justesse de prévision ;
- probabilité de détection, taux de fausses détection, taux de fausses alertes et score d'Hansen-Kuipers ;
- diagramme ROC et score AROC ;
- diagramme de Talagrand ;
- diagramme de fiabilité.

En perspectives...

Outre l'utilisation justifiée d'une méthode de prévision de pluie par recherche d'analogues, il est encore possible d'améliorer ce type d'approche. En effet, plusieurs hypothèses fortes pourraient être levées :

- en contexte opérationnel, nous avons conservé un nombre d'analogues retenues identique à chaque échéance de prévision. Or on a déjà démontré par le passé qu'il est préférable d'augmenter la taille de l'échantillon d'analogues avec l'échéance ;
- les pluies observées lors des situations analogues, qui forment la distribution empirique prévue pour une journée cible, ont un poids identique quel que soit le degré d'analogie de la situation météorologique associée. Il nous semble fort intéressant, bien que des essais aient déjà été faits (Plasse, 2008), de lever cette hypothèse d'équiprobabilité en pondérant les situations analogues en fonction des valeurs des critères d'analogie associées (*i.e.* du degré de similitude) ;
- les prévisions par analogie, comme nous avons pu le remarquer dans ces travaux, sont biaisées pour les événements pluvieux importants. Une étude approfondie du biais de la prévision pourrait permettre, à terme, de débiaiser la prévision afin que chaque quantile d'une distribution ait une probabilité d'être dépassé identique à sa fréquence théorique ;

- lorsque plusieurs variables sont considérées à un niveau de sélection d'analogues, les valeurs des critères d'analogie sont généralement évaluées par une simple moyenne ou par une somme. Or il est envisageable de pondérer les valeurs en fonction de la pertinence des variables, ou d'accorder un décalage temporel dans la mesure des variables.

En plus des pistes que nous venons d'évoquer, les performances de prévision pourraient encore être améliorées par l'utilisation des réanalyses ERA-40 à haute résolution (résolution initiale de 1,125°) pour certaines variables. En effet, tout au long du développement de l'algorithme, seule la résolution 2,5° a été exploitée pour les champs météorologiques. Or, il serait intéressant d'utiliser les réanalyses ERA-40 à 1° pour décrire les champs de vitesse verticale ou d'humidité, même si les tests qui ont été faits sur les variables d'humidité n'ont pas été très concluants pour le bassin de la Saône. Il pourrait être notamment bénéfique d'effectuer ce test sur des bassins de plus petite taille pour décrire plus finement les mouvements verticaux.

La question de l'applicabilité des méthodes développées dans cette thèse à d'autres bassins versants mérite également d'être posée. Les variables principalement introduites (température, vitesse verticale) ne sont pas nécessairement spécifiques aux grands bassins versants français tels que ceux de la Saône et de la Seine. En effet, tous les bassins français sont notamment soumis aux perturbations frontales. Il semble donc peu probable que, sur d'autres bassins versants, la méthode développée ici soit moins performante que la méthode de référence, par exemple sur les bassins à crues rapides du Sud-Est de la France.

Pour une application à la prévision hydrologique opérationnelle, la prévision des pluies sur de grands bassins et à des échéances de plusieurs jours, voire de la semaine, pose encore des questions dans la recherche d'analogues :

- comment combiner les pluies des situations analogues de plusieurs sous-bassins de façon à respecter la cohérence spatiale du champ de pluie ?
- comment combiner les situations analogues de plusieurs journées successives, tout en gardant une cohérence temporelle ?

Ces questions sont ouvertes et méritent donc d'être étudiées au travers d'une analyse approfondie.

Le mot de la fin

Il est nécessaire de garder à l'esprit que les performances d'une méthode d'adaptation des sorties de modèles météorologiques par analogie dépendent de celles du modèle météorologique lui-même. On peut raisonnablement penser qu'avec l'amélioration constante des modèles, les prévisions par analogie seront également de plus en plus performantes. Ce fait assure sans doute la pérennité de ce type de méthode.

Enfin, il serait erroné de considérer la méthode de prévision de pluie adaptée par analogie comme une alternative aux autres approches. Cette méthode est plutôt à considérer comme un complément des autres méthodes de prévisions, qu'elles soient expertisées ou directement produites par un modèle numérique. De façon générale, la prise en compte des incertitudes joue un rôle fondamental dans la prise de décisions, c'est pourquoi les méthodes de prévision probabiliste, telles que la prévision d'ensemble ou la prévision par analogie, constituent de ce point de vue un outil très utile voire indispensable pour les services opérationnels.

Références bibliographiques

- Altava-Ortiz V., A. Barrera, M.C. Llasat, M.A. Prat, J. Gibergans-Baguena and M. Barnolas**, 2006. Application of the MM5 and the analogous method to heavy rainfall event, the case of 16-18 October 2003 in Catalonia. *Advances in Geosciences*, 7, 313-319.
- Anderson J.**, 1996. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *Journal of Climate*, 9, 1518-1530.
- Bannayan M. and G. Hoogenboom**, 2008. Weather analogue: A tool for real-time prediction of daily weather data realizations based on a modified k-nearest neighbor approach. *Environmental Modelling and Software*, 23, 703-713.
- Barnett T.P. and R.W. Preisendorfer**, 1978. Multifield analog prediction of short-term fluctuations using a climate state vector. *Journal of the Atmospheric Sciences*, 35, 1771-1787.
- Ben Daoud A., E. Sauquet, Ch. Obled, M. Lang and G. Bontron**, 2008a. Meteorological variables used in an analog method: comparison between ERA-40 and NCEP/NCAR re-analyses. *Geophysical Research Abstracts*, 10, EGU2008-A-04137.
- Ben Daoud A., E. Sauquet, Ch. Obled, M. Lang and G. Bontron**, 2008b. A flood forecasting system based on probabilistic precipitation scenarios: performance evaluation and sensitivity analysis to the re-analysis archive. *Proceedings of the International Interdisciplinary Conference HydroPredict'2008 on Predictions for Hydrology, Ecology, and Water Resources Management: Using Data and Models to Benefit Society, Prague, Czech Republic, 15-18 Sept.*, 207-210.
- Ben Daoud A., E. Sauquet, M. Lang, Ch. Obled and G. Bontron**, 2009a. Comparison of 850-hPa relative humidity between ERA-40 and NCEP/NCAR re-analyses: detection of suspicious data in ERA-40. *Atmospheric Science Letters*, 10, 43-47.
- Ben Daoud A., E. Sauquet et Ch. Obled**, 2009b. Reconstitution du risque nivo-pluviométrique durant le mois d'octobre 1859 par une technique d'analogie. *Colloque Isère 1859-2009, 150ème anniversaire de la crue de référence. 5 Nov., Grenoble*.
- Ben Daoud A., G. Bontron, E. Sauquet, Ch. Obled, et M. Lang**, 2009c. La prévision des précipitations par recherche d'analogues : état de l'art et perspectives. *La Houille Blanche*, 6, 1-6.
- Berrisford P., D. Dee, K. Fielding, M. Fuentes, P. Källberg, S. Kobayashi and S. Uppala**, 2009. The ERA-Interim archive. *ERA Report Series, ECMWF*.
- Berther L., V. Andréassian, C. Perrin and P. Javelle**, . How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*, 13, 819-831.
- Bindon H.H.**, 1940. Relation between equivalent potential temperature and wet-bulb potential temperature. *Monthly Weather Review*, 68, 243-245.
- Blifernicht J. and A. Bardossy**, 2008. Probabilistic forecast of daily areal precipitation focusing on extreme events. *Natural Hazards and Earth System Sciences*, 7, 263-269.
- Bois P. et Ch. Obled**, 1976. Prévision des avalanches par des méthodes statistiques : aspects méthodologiques et opérationnels. *La Houille Blanche*, 6/7, 509-531.
- Bompert P., G. Bontron, S. Celie et M. Haond**, 2009. Une chaîne opérationnelle de prévision hydrométéorologique pour les besoins de la production hydroélectrique de la CNR. *La Houille Blanche*, 5, 54-60.
- Bontron G.**, 2004. Prévision quantitative des précipitations : adaptation probabiliste par recherche d'analogues. Utilisation des réanalyses NCEP-NCAR et application aux précipitations du Sud-Est de la France. *Thèse de Doctorat, Institut National Polytechnique de Grenoble*.

- Brown T.A.**, 1974. Admissible scoring systems for continuous distributions. *Manuscript P-5235, The Rank Corporation, Santa Monica, CA.*
- Dal Piva E., M. A. Gan, and V. Brahmananda Rao**, 2008. An objective study of 500-hPa moving troughs in the southern hemisphere. *Monthly Weather Review*, 136, 2186-2200.
- Dawson J., D. Scott and G. McBoyle**, 2009. Climate change analogue analysis of ski tourism in the northeastern USA. *Climate Research*, 39, 1-9.
- Déqué M.**, 2003. La prévision numérique à l'échelle saisonnière: que sait-on faire et que peut-on espérer?. *La Météorologie*, 41, 20-29.
- Déqué M., C. Drevet, A. Braun and D. Cariolle**, 1994. The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling. *Climate Dynamics*, 10, 249-266.
- Diomede T., F. Nerozzi, T. Paccagnella and E. Todini**, 2006. The use of meteorological analogues to account for LAM QPF uncertainty. *Natural Hazards and Earth System Sciences*, 12, 141-157.
- Djeboua A. and Ch. Obled**, 2003. QPF's up to 6 days lead-time: analog based versus deterministic mesoscale models during MAP. *Proceedings of the 5th International EGS Plinius Conference on Mediterranean Storms, Ajaccio, 1-3 Oct..*
- Duband D.**, 1970. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. *Thèse de 3ème cycle de la Faculté des Sciences de Paris.*
- Duband D.**, 1974. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. *Congrès de la Société Hydrotechnique de France, XIIIèmes journées de l'Hydraulique, Paris, France.*
- Duband D.**, 1980. Dynamic selection of analogue flow patterns to produce quantitative precipitation forecasts. *WMO Symposium on probabilistic and statistical method in weather forecasting, Nice, France, 8-12 septembre.*
- Durand Y., E. Brun, L. Merindol, G. Guyomarc'h, B. Lesaffre and E. Martin**, 1993. A meteorological estimation of relevant parameters for snow models. *Annals of Glaciology*, 18, 65-71.
- Durand Y., G. Giraud, E. Brun, L. Merindol and E. Martin**, 1999. A computer based system simulating snowpack structures as a tool for regional avalanche forecast. *Journal of Glaciology*, 45, 469-485.
- Durand Y., M. Laternser, G. Giraud, P. Etchevers, B. Lesaffre and L. Mérindol**, 2009. Reanalysis of 44 years of climate in the French Alps (1958-2002): methodology, model validation, climatology and trends for air temperature and precipitation. *Journal of Applied Meteorology and Climatology*, 48, 429-449.
- Gibergans-Baguena J. and M. C. Llasat**, 2007. Improvement of the analog forecasting method by using local thermodynamic data. Application to autumn precipitation in Catalonia. *Atmospheric Research*, 86, 173-193.
- Gibson J.K., P. Kallberg, S. Uppala, A. Nomura, A. Hernandez, and E. Serrano**, 1997. ERA Description. *ECMWF ERA-15 Project Report Series*, 1.
- Gleisner H., P. Thejll, M. Stendel, E. Kaas et B. Machenhauer**, 2005. Solar signals in tropospheric re-analysis data: Comparing NCEP/NCAR and ERA40. *Journal of Atmospheric and Solar-Terrestrial Physics*, 67, 785-791.
- Golding B.W.**, 2000. Quantitative precipitation forecasting in the UK. *Journal of Hydrology*, 239, 286-305.
- Guilbaud S.**, 1997. Prévision quantitative des précipitations journalières par une méthode statistico-dynamique de recherche d'analogues. *Thèse de Doctorat, Institut National Polytechnique de Grenoble.*
- Habets F., A. Boone, L. Champeaux, P. Etchevers, L. Franchisteguy, E. Leblois, E. Ledoux, P. Le Moigne, E. Martin, S. Morel, J. Noilhan, P. Quintana-Segui, F. Rousset-Regimbeau et P. Viennot**, 2008. The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. *Journal of Geophysical Research*, 113, doi:10.1029/2007JD008548.
- Hamill T. M.**, 2001. Interpretation of rank histogram for verifying ensemble forecasts. *Monthly Weather Review*, 129, 550-560.
- Hamill T. M., J. S. Whitaker, and S. L. Mullen**, 2006. Reforecasts: an important dataset for improving weather predictions. *Bulletin of the American Meteorological Society*, 87, 33-46.
- Hamill T.M.**, 1997. Reliability diagrams for multicategory probabilistic forecasts. *Weather and*

Forecasting, 12, 736-741.

- Hamill T.M. and J.S. Whitaker**, 2006. Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application. *Monthly Weather Review*, 134, 3209-3229.
- Hamill T.M. and J.S. Whitaker**, 2007. Ensemble Calibration of 500-hPa Geopotential height and 850-hPa and 2-m temperatures using reforecasts. *Monthly Weather Review*, 135, 3273-3280.
- Hanson C.E., J.P. Palutikof and T.D. Davies**, 2004. Objective cyclone climatologies of the North Atlantic - a comparison between the ECMWF and NCEP reanalyses. *Climate Dynamics*, 22, 757-769.
- Hersbach, H.**, 2000. Decomposition of the ranked probability score for ensemble prediction systems. *Weather Forecasting*, 15, 559-570.
- Houtemaker P.L., L. Lefaivre and J. Derome**, 1996. The RPN ensemble prediction system. *Proceedings of ECMWF Seminar on Predictability, Vol. II, Reading, United Kingdom, ECMWF*, , 121-146.
- HYDRATEC**, 1998. Inondations en Ile-de-France - Nouvelles études, nouvelles connaissances. *Rapport de Synthèse, IIBRBS*.
- Ito A. and T. Sasai**, 2006. A comparison of simulation results from two terrestrial carbon cycle models using three climate data sets. *Tellus*, 58b, 513-522.
- Jolliffe I.T. and D.B. Stephenson**, 2003. Forecast verification. A practitioner's guide in atmospheric science. *John Wiley & Sons Edition*.
- Jung T.**, 2005. Systematic errors of the atmospheric circulation in the ECMWF forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 131, 1045-1073.
- Kalkstein S., J. Scott Greene, D.M. Mills, A.D. Perrin, J.P. Samenow and J.-C. Cohen**, 2008. Analog european heat waves for U.S. cities to analyze impacts on heat-related mortality. *Bulletin of the American Meteorological Society*, 89, 75-85.
- Kallberg P., A. Simmons, S. Uppala and M. Fuentes**, 2004. The ERA-40 archive. *ERA-40 Project Report Series*, 17.
- Kalnay E., M. Kanamitsu, R. Kistler, W. Collins, D. deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, Roy Jenne and D. Joseph**, 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77, 437-471.
- Kanamitsu M., W. Ebisuzaki, J. Woollen, S-K. Yang, J. J. Hnilo, M. Fiorino and G. L. Potter**, 2002. NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83, 1631-1643.
- Kistler R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne and M. Fiorino**, . The NCEP-NCAR 50-year reanalysis: Monthly Means CD-Rom and documentation. *Bulletin of the American Meteorological Society*, 82, 247-267.
- Klemes V.**, 1986. Operational testing of hydrological simulation models. *Hydrological Sciences*, 31, 13-24.
- Kopf S., M. Ha-Duong and S. Hallegatte**, 2008. Using maps of city analogues to display and interpret climate change scenarios and their uncertainty. *Natural Hazards and Earth System Sciences*, 8, 905-918.
- Lacaz Y., P. Javelle, D. Goutx et L. Berthet**, 2008. Influence des pluies futures sur la prévision des crues : Evaluation opérationnelle pour le cas du SPC "Seine moyenne - Yonne - Loing". *Colloque SHF Prévisions Hydrométéorologiques, 18-19 Nov., Lyon*.
- Lang M., J. Lavabre, E. Sauquet et B. Renard**, 2006. Estimation de la crue centennale pour les plans de prévention des risques d'inondations. *Collection Update Sciences & Technologies, Editions QUAE*, 232 pages.
- Lefevre R.J. and J.W. Nielsen-Gammon**, 1995. An objective climatology of mobile troughs in the Northern Hemisphere. *Tellus*, 47A, 638-655.
- Leroy A. et J.-P. Céron**, 2007. Un défi de la prévision saisonnière: la descente d'échelle. *La Météorologie*, 58, 36-44.
- Llasat M.C., J. Gibergans, S. Guilbaud and J.Y. Rodriguez**, 1997. The use of the vertical thermodynamic data in the daily rain classification. Application to the improvement of an analog technique. *INM/WMO International symposium on cyclones and hazardous weather in the*

- Mediterranean regions*, 395-402.
- Lorenz E.N.**, 1956. Empirical orthogonal functions and statistical weather prediction. *Science Report N°1, Statistical Forecasting Project, M.I.T., Cambridge*, 48 pp.
- Lorenz E.N.**, 1963. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20, 130-141.
- Lorenz E.N.**, 1969. Atmospheric predictability as revealed by natural occurring analogues. *Journal of the Atmospheric Sciences*, 26, 636-646.
- Malardel S.**, 2005. Fondamentaux de Météorologie à l'école du temps. *Météo-France, Cépaduès-Editions*, 709 pages.
- Martin E.**, 1995. Modélisation de la climatologie nivale des Alpes françaises, application des techniques de régionalisation à l'étude de l'impact d'un changement climatique sur l'enneigement. *Thèse de Doctorat, Université Paul Sabatier, Toulouse*.
- Martin E., G. Thirel, F. Rousset-Regimbeau, J. Noilhan, L. Franchisteguy, J.-M. Soubeyrou et F. Habets**, 2009. Vers une prévision d'ensemble des débits à l'échelle des grands bassins français. *La Houille Blanche*, 5, 88-94.
- Marty R.**, 2010. Désagrégation de prévisions quantitatives et probabilistes de précipitations et utilisation dans une chaîne de prévision hydrologique pour les bassins à crues éclair. *Thèse de Doctorat, Université de Grenoble*.
- Marty R., I. Zin and Ch. Obled**, 2008. On adapting PQPFs to fit hydrological needs: the case of flash flood forecasting. *Atmospheric Science Letters*, 9, 73-79.
- Marty R., I. Zin et Ch. Obled**, 2009. Assimilation de prévisions probabilistes de précipitations dans une chaîne de prévision hydrologique pour les bassins à crues rapides. *Vème rencontre Annuelle Georisque, GESTER, Montpellier*.
- Matheson J.E. and R.L. Winkler**, 1976. Scoring rules for continuous probability distributions. *Management Science*, 22, 1087-1096.
- Mathevet T.**, 2005. Quels modèles pluie-débit globaux au pas de temps horaire ? Développements empiriques et comparaison de modèles sur un large échantillon de bassins versants. *Thèse de Doctorat, Ecole Nationale du Génie Rural des Eaux et Forêts, Paris*.
- Molteni F., R. Buizza, T.N. Palmer and T. Petroliaigis**, 1996. The ECMWF ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, 122, 73-119.
- Nash J.E. and J.V. Sutcliffe**, 1970. River flow forecasting through conceptual models. Part 1: a discussion of principles. *Journal of Hydrology*, 10, 282-290.
- Obled Ch. and W. Good**, 1980. Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland). *Journal of Glaciology*, 25, 315-345.
- Obled Ch., G. Bontron and R. Garçon**, 2002. Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach. *Atmospheric Research*, 63, 303-324.
- Pailleux J., J.-F. Geleyn et E. Legrand**, 2000. La prévision numérique du temps avec les modèles Arpège et Aladin. *La Météorologie*, 8, 32-60.
- Paquet E., J. Gailhard et R. Garçon**, 2006. Evolution de la méthode du GRADEX : approche par type de temps et modélisation hydrologique. *La Houille Blanche*, 5, 80-90.
- Pardé M.**, 1933. Fleuves et rivières. 5ème édition, *Librairie Armand COLIN, Paris*, 241 pages.
- Pelletier R.**, 2004. Reconstitution de 10 ans de lames d'eau sur 46 bassins versants du Rhône. *Météo-France Centre-Est, Division Développements-Etudes-Climatologie*.
- Perrin C., C. Michel and V. Andreassian**, 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, 242, 275-301.
- Perrin C., C. Michel and V. Andreassian**, 2003. Improvement of parsimonious model for streamflow simulation. *Journal of Hydrology*, 279, 275-289.
- Perrin C., J. Lerat, L. Oudin et V. Andréassian**, 2005. Inventaire et diagnostic des modèles simples de prévision des crues déjà existants sur le bassin de la Seine. *Rapport final, Cemagref, Groupement d'Antony*.
- Pettitt A.N.**, 1979. A non-parametric approach to the change-point problem. *Applied Statistics*, 28, 126-

135.

- Plasse J.**, 2008. Application d'une méthode de prévision des pluies par analogie au bassin de la Seine à Paris. *Rapport de stage de fin d'étude, Cemagref*.
- Quintana-Segui P., P. Le Moigne, Y. Durand, E. Martin, F. Habets, M. Baillon, C. Canellas, L. Franchisteguy, and S. Morel**, 2008. Analysis of near surface atmospheric variables: Validation of the SAFRAN analysis over France. *Journal of Applied Meteorology and Climatology*, 47, 92-107.
- Roebbert P.J. and G.W. Reuter**, 2002. The sensitivity of precipitation to circulation details. Part II: mesoscale modeling. *Monthly Weather Review*, 130, 3-23.
- Roebbert P.J. and L.F. Bosart**, 1998. The sensitivity of precipitation to circulation details. Part I: an analysis of regional analogs. *Monthly Weather Review*, 126, 437-455.
- Rossby C.G.**, 1932. Thermodynamics applied to air mass analysis. *Massachusetts Institute of Technology, Papers in Physical Oceanography and Meteorology*, 1, n°3.
- Rousset-Regimbeau F.**, 2007. Modélisation des bilans de surface et des débits sur la France ; application à la prévision d'ensemble des débits. *Thèse de Doctorat de l'Université Paul Sabatier, Toulouse*.
- Rousset-Regimbeau F., F. Habets, E. Martin and J. Noilhan**, 2007. Ensemble streamflow forecasts over France. *ECMWF Newsletter*, 111, 21-27.
- Ruosteenoja K.**, 1988. Factors affecting the occurrence and lifetime of 500 mb height analogues: a case study based on large amount of data. *Monthly Weather Review*, 116, 368-376.
- Sauquet E.**, 2006. Mapping mean annual river discharges: geostatistical developments for incorporating river network dependencies. *Journal of Hydrology*, 331, 300-314.
- Simmons A., S. Uppala, D. Dee and S. Kobayashi**, 2006. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter*, 110, 25-35.
- Simmons A.J., P.D. Jones, V. da Costa Bechtold, A.C.M. Beljaars, P.W. Kallgerg, S. Saarinen, S.M. Uppala, P. Viterbo and N. Wedi**, 2004. Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *Journal of Geophysical Research*, 109.
- Stanski H.R., L.J. Wilson and W.R. Burrows**, 1989. Survey of common verification methods. *Meteorology research report 89-5, Service de l'environnement atmosphérique, Downsview, Canada*.
- Sterl A.**, 2004. On the (In)Homogeneity of Reanalysis Products. *Journal of Climate*, 17, 3866-3873.
- Talagrand O., R. Vautard and B. Strauss**, 1997. Evaluation of probabilistic prediction systems. *Proceedings of ECMWF workshop on predictability, Reading, England, october, , 1-25*.
- Tangara M.**, 2005. Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global. *Thèse de Doctorat, Ecole Pratique des Hautes Etudes de Paris*.
- Tartaglione N., M. Maugeri, F. Dalan, M. Brunetti, T. Nanni and A. Speranza**, 2009. Searching for resemblance between large-scale sea level pressure patterns leading to "intense" precipitation events over Italy. *Theoretical and Applied Climatology*, 95, 183-196.
- Teweles J. and H. Wobus**, 1954. Verification of prognosis charts. *Bulletin of the American Meteorological Society*, 35, 455-463.
- Thévenot N.**, 2004. Prévision quantitative des précipitations : adaptation par une méthode d'analogie de la prévision d'ensemble du CEPMMT - Aspects opérationnels. *Mémoire de Master 2 TUE, Grenoble*.
- Thirel G., F. Rousset-Regimbeau, E. Martin and F. Habets**, 2008. On the impact of short-range meteorological forecasts for ensemble streamflow prediction. *Journal of Hydrometeorology*, 9, 1301-1317.
- Toth Z. and E. Kalnay**, 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, 74, 2317-2330.
- Toth Z. and E. Kalnay**, 1997. Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, 125, 3297-3319.
- Trigo I.F.**, 2005. Climatology and interannual variability of storm-tracks in the Euro-Atlantic sector: a comparison between ERA-40 and NCEP/NCAR reanalyses. *Climate Dynamics*, 26, 127-143.
- Uppala S.**, 2006. From ERA-15 to ERA-40 and ERA-Interim. *ECMWF/GEO Workshop on Atmospheric Reanalysis*, 17-21.

- Uppala S.M., P.W. Kallberg, A.J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J.K. Gibson, J. Haseler, A. Hernandez, G.A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R.P. Allan, E. Andersson, K. Arpe, M.A. Balmaseda, A.C.M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. H¹ Im, B.J. Hoskins, L. Isaksen, P.A.E.M. Janssen, R. Jenne, A.P. McNally, J.-F. Mahfouf, J.-J. Morcrette, N.A. Rayner, R.W. Saunders, P. Simon, A. Sterl, K.E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo and J. Woollen**, 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131, 2961-3012.
- Van Den Dool H.M.**, 1994. Searching for analogues, how long must we wait?. *Tellus*, 46A, 314-324.
- Vidal J.-P., E. Martin, L. Franchistéguy, M. Baillon and J.-M. Soubeyrou**, 2009. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, DOI: 10.2002/joc.
- von Storch H.**, 1995. Inconsistencies as the interface between climate research and climate impact studies. *Meteorologische Zeitschrift*, 4, 72-80.
- Vrac M., K. Hayhoe and M. Stein**, 2007. Identification and intermodel comparison of seasonal circulation patterns over North America. *International Journal of Climatology*, 27, 603-620.
- Wallis J.R. and E. Todini**, 1975. Comment upon the residual mass curve coefficient. *Journal of Hydrology*, 24, 201-205.
- Wetterhall F., A. Bardossy, D. Chen, S. Halldin and C.-Y. Xu**, 2006. Daily precipitation-downscaling techniques in three Chinese regions. *Water Resources Research*, 42, W11423.
- Wetterhall F., S. Halldin and C.-Y. Xu**, 2005. Statistical precipitation downscaling in central Sweden with the analogue method. *Journal of Hydrology*, 306, 174-190.
- Wetterhall F., S. Halldin and C.-Y. Xu**, 2007. Seasonality properties of four statistical-downscaling methods in central Sweden. *Theoretical and Applied Climatology*, 87, 123-137.
- Wilby R.L. and T.M.L Wigley**, 1997. Downscaling général circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21, 530-548.
- Wilks D.S.**, 1995. Statistical methods in the atmospheric sciences: an introduction. *San Diego, CA, USA: Academic Press*, 467 pp.
- Zorita E. and H. von Storch**, 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of Climate*, 12, 2474-2488.
- Zorita E., J. Hughes, D. Lettenmaier and H. von Storch**, 1995. Stochastic downscaling of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *Journal of Climate*, 8, 1023-1042.

Annexe 1

Type de communication	Communication orale
Titre	Reconstitution du risque nivo-pluviométrique durant le mois d'octobre 1859 par une technique d'analogie
Auteurs	Ben Daoud A., Sauquet E., Obled Ch.
Année	2009
Source	Colloque scientifique Isère 1859-2009, 150ème anniversaire de la crue de référence

RECONSTITUTION DU RISQUE NIVO-PLUVIOMETRIQUE DURANT LE MOIS D'OCTOBRE 1859 PAR UNE TECHNIQUE D'ANALOGIE

Aurélien Ben Daoud, Eric Sauquet

Cemagref – Unité de Recherche Hydrologie-Hydraulique
3 bis Quai Chauveau, CP 220, 69336 Lyon, Cedex 9, France

Tél: +33 (0)4 72 20 87 64, Fax: +33 (0)4 78 47 78 75, e-mail: aurelien.bendaoud@cemagref.fr

Charles Obléd

Laboratoire d'étude des Transferts en Hydrologie et Environnement
Institut National Polytechnique de Grenoble, B.P. 53, 38041 Grenoble, Cedex 9, France
Tél: +33 (0)4 76 82 50 53, Fax: +33 (0)4 76 82 50 14, e-mail: charles.obled@hmg.inpg.fr

I INTRODUCTION

L'objectif principal de ce travail est de reconstituer la séquence pluviométrique sur la période précédant la crue majeure de l'Isère qui s'est produite les 31 octobre et 1er novembre 1859. Afin d'inclure les différents épisodes ayant contribué à l'événement durant le mois d'octobre 1859, nous retenons une période large de 36 jours, qui s'étire du 1er octobre au 5 novembre 1859. Or nous ne disposons que de très peu d'information sur les précipitations relevées sur le bassin de l'Isère au cours de cette période. Pour y remédier, nous appliquerons une technique d'analogie permettant d'identifier des situations passées considérées comme analogues, d'un point de vue météorologique, aux situations observées en 1859.

Cette méthode doit d'abord être adaptée au bassin versant de l'Isère en amont de Grenoble. Une fois optimisée, la technique peut être appliquée et les dates analogues identifiées. A partir de la distribution des précipitations des situations analogues, nous pourrions reconstituer un scénario probable d'évolution de la pluviométrie sur la période d'intérêt. Par l'intermédiaire de la technique d'analogie, d'autres paramètres météorologiques seront également étudiés, tels que la température ou le géopotential, mais aussi les régimes de temps les plus vraisemblablement rencontrés durant la période. Cette étude s'inscrit dans la continuité des travaux menés par Jourdain et al. (2009).

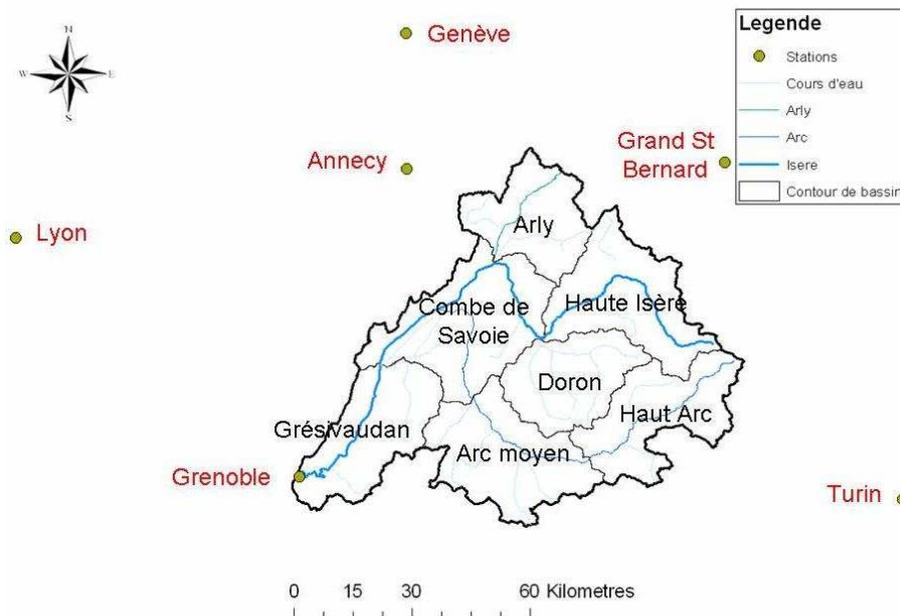


Figure 1. Bassin de l'Isère en amont de Grenoble et sous-bassins considérés pour l'étude.

Dans le cadre de ce travail, le bassin de l'Isère en amont de Grenoble a été divisé en sept sous-bassins (Figure 1), ceci afin d'évaluer la contribution de chacun d'entre eux à la crue de l'Isère (Panthou, 2009 ; Obled et al., 2009). Dans la première section de cet article, nous décrivons la méthodologie employée et présenterons les bases de données disponibles dans le cadre de cette étude. La seconde section vise à présenter les résultats de l'optimisation de la technique d'analogie utilisée. Enfin, nous reconstituerons la séquence pluviométrique de la période d'étude dans la dernière partie et nous caractériserons les régimes de temps probables auxquels le bassin de l'Isère a été soumis avant l'événement de crue.

II MÉTHODOLOGIE

Dans un premier temps, nous allons décrire brièvement la technique d'analogie employée puis nous présenterons les données disponibles nécessaires à son application.

II.1 Technique d'analogie

La technique d'analogie utilisée dérive d'une méthode de prévision des précipitations développée en France au cours des années 1970 (Duband, 1970). Cette méthode repose sur l'hypothèse suivant laquelle deux situations météorologiques similaires doivent engendrer des effets locaux similaires (Lorenz, 1969). Ainsi pour un jour cible dont on désire prévoir le cumul de précipitations, on recherche dans une archive les journées passées pour lesquelles la situation météorologique observée est proche de la situation prévue pour le jour cible. Les situations passées les plus similaires sont alors retenues, puis les pluies observées au cours de ces journées sont extraites d'une archive pluviométrique. Il est alors possible de construire la distribution empirique des pluies observées, puis par exemple d'en déduire les valeurs associées à des quantiles au non-dépassement. Cette méthode permet ainsi de prévoir les précipitations de manière probabiliste.

Outre l'usage courant de cette technique pour la prévision opérationnelle des précipitations, il est également possible de l'employer pour reconstituer des événements passés. C'est ainsi dans ce but que nous avons appliqué cette technique aux situations météorologiques précédant l'événement de la crue de l'Isère de 1859.

Pour caractériser une situation météorologique, les champs de pressions sont usuellement exploités pour mesurer le degré de similitude (cf. Ben Daoud et al., 2009 pour un état de l'art de la méthode employée en France). Dans notre cas, seuls les champs de pression au niveau de la mer (P_{mer}) observés en 1859 sont disponibles. Le critère permettant de mesurer la ressemblance entre deux situations est le critère TWS de Teweles and Wobus (1954), qui permet de comparer la forme de deux champs de P_{mer} . La valeur du critère TWS est de 0 pour deux formes strictement identiques, de 200 pour deux formes totalement opposées, et de 100 pour deux formes indépendantes.

L'objectif est donc de rechercher dans une période plus récente les situations les plus analogues à celles observées en 1859 en termes de champ de P_{mer} . Une telle technique ne peut être appliquée que si l'on dispose de deux archives dans lesquelles il est possible d'extraire, pour un jour passé, d'une part les champs de P_{mer} observés, et d'autre part les cumuls pluviométriques observés.

II.2 Données disponibles

Pour appliquer la technique d'analogie sur la période qui nous intéresse, nous devons disposer de trois archives : une archive des champs de P_{mer} observés en 1859, une archive des champs de P_{mer} observés dans une période plus récente, et une archive des cumuls de précipitation observés dans cette même période récente.

Les champs de pression sont généralement extraits d'une archive de réanalyses météorologiques. Celles-ci sont constituées d'observations interpolées sur une grille régulière, dont les données assimilées sont contraintes par un modèle numérique atmosphérique, ce qui permet de garantir une cohérence dans le temps et dans l'espace.

II.2.1 Champs de pression observés en 1859

Nous disposons de deux archives contenant les champs de Pmer observés entre 1er octobre et le 5 novembre 1859 :

- les réanalyses météorologiques EMULATE à résolution 5°, au pas de temps 24h ;
- l'archive reconstituée par Météo-France à résolution 2,5°, au pas de temps 24h (cf. Jourdain et al., 2009).

II.2.2 Champs de pression observés dans une période récente

Plusieurs archives de réanalyses sont à notre disposition :

- les réanalyses météorologiques EMULATE à résolution 5°, au pas de temps 24h, qui couvrent la période 1850-2003 ;
- les réanalyses météorologiques ERA-40 (Uppala et al., 2005) à résolution 2,5°, au pas de temps 6h, qui couvrent la période 1957-2002 ;
- les réanalyses météorologiques NCEP/NCAR (Kalnay et al., 1996) à résolution 2,5°, au pas de temps 6h, qui couvrent la période 1948-2007.

Ces réanalyses ne couvrent donc pas la même période, et les champs qu'elles contiennent n'ont pas la même résolution, ce qui nous incitera plus tard à faire un choix.

II.2.3 Précipitations observées dans une période récente

Dans le cadre de ce travail, nous disposons des réanalyses pluviométriques SAFRAN (Vidal et al., 2009). Les précipitations sont disponibles la période 1970-2006, au pas de temps horaire, sur une grille de résolution 8 km. Nous en avons déduit les lames d'eau journalières (6h TU le jour J à 6h TU le jour J+1) observées sur la période. Les pluies de bassin ont été construites, après ré-aggrégation, à partir des lames d'eau produites sur les mailles SAFRAN. Le contour des bassins versants a été obtenu à partir d'un modèle numérique de terrain à maille kilométrique (Global 30 Arc Second Elevation Data Set, U.S. Geological Survey (Sauquet, 2006)).

III OPTIMISATION DE LA METHODE DE RECONSTITUTION DES PRECIPITATIONS

Dans un premier temps, nous devons adapter la méthode à la zone d'étude. Les paramètres à optimiser sont le domaine spatial sur lequel les champs de Pmer sont comparés, fonction de la localisation géographique, et le nombre de situations analogues retenues, fonction de la taille de l'archive. Le protocole de calage des paramètres est celui retenu lorsque l'analogie est exploitée pour la prévision probabiliste de pluie.

III.1 Choix préliminaires

L'optimisation s'appuie sur un score d'évaluation : le CRPS (Hersbach, 2000). La performance est finalement exprimée en termes de gain relatif (CRPSS) de CRPS par rapport à une prévision de référence (ici la prévision climatologique). Le score CRPSS vaut 1 pour une prévision probabiliste parfaite (distribution parfaitement fine et centrée sur l'observation), il vaut 0 (ou est négatif) pour une prévision équivalente à (ou moins bonne que) la prévision de référence.

Etant donnée que l'archive NCEP/NCAR (61 ans) est plus longue que l'archive ERA-40 (45 ans), nous allons retenir l'archive NCEP/NCAR pour la recherche de situations analogues. Nous disposons ainsi de 61 années de situations candidates pour rechercher des situations analogues à celles observées en 1859. En revanche, si nous désirons reproduire la distribution empirique des pluies associée aux analogues, il est nécessaire de se restreindre à une période plus courte, qui recouvre à la fois la période de disponibilité des champs de Pmer et celle des précipitations. Ainsi, la période commune aux archives NCEP/NCAR et SAFRAN est celle du 1/8/1970 au 30/7/2006, soit 36 années.

Deux possibilités s'offrent à nous pour les situations observées en 1859 : soit on travaille sur les champs de Pmer extraits de l'archive EMULATE (à résolution 5°), soit on travaille sur les champs de Pmer fournis par Météo-France (à résolution $2,5^\circ$). Ces dernières sont *a priori* plus fiables que celles d'EMULATE, puisque les données françaises assimilées pour les réanalyses EMULATE ne sont constituées que de celles de Paris (Jourdain et al., 2009). Le champ de Pmer n'est donc pas forcément très fiable à l'échelle de la France. Néanmoins, les deux possibilités seront explorées dans cette section.

Dans un premier temps, compte tenu des degrés de liberté, il convient de fixer le nombre d'analogues à retenir puisque celui-ci dépend *a priori* de la longueur de la période de situations candidates. Dans notre cas, avec une période longue de 36 ans, nous savons par expérience qu'un nombre de 30 doit être proche du nombre optimal. Nous fixons donc à présent le nombre d'analogues à retenir à 30. En outre, une fenêtre temporelle de plus ou moins deux mois autour d'un jour cible a été appliquée pour l'extraction des analogues, afin de tenir compte de l'effet saisonnier sur les précipitations (Lorenz, 1969). Ainsi pour une situation du mois d'octobre 1859, nous ne rechercherons l'analogie que sur les journées passées des mois d'août à décembre.

III.2 Domaine spatial d'analogie

Le domaine est optimisé par exploration spatiale à partir d'un domaine élémentaire de départ qu'il faut préalablement choisir. La fenêtre d'analogie est agrandie d'un pas de grille dans chacune des quatre directions (nord, sud, ouest, est). Pour chaque jour, on identifie les 30 situations analogues et on construit une distribution empirique constituée des 30 valeurs de pluie correspondantes. Le score CRPSS est ainsi calculé sur toute la chronique sur les sept sous-bassins de l'Isère, et les performances obtenues pour chaque domaine sont comparées afin de retenir le domaine pour lequel les performances sont les plus élevées. On réitère le processus jusqu'à une taille de fenêtre raisonnable. C'est finalement le domaine pour lequel le score CRPSS est maximal qui sera considéré comme domaine optimal.

Les deux archives de Pmer observée sont testées pour identifier leur pouvoir prédictif respectif. Nous avons optimisé le domaine d'une part sur une grille de résolution 5° pour pouvoir chercher dans l'application les analogues aux situations extraites d'EMULATE ; et d'autre part sur une grille de résolution $2,5^\circ$ pour pouvoir chercher les analogues aux situations observées, reconstituées par Météo-France.

Pour exploiter les réanalyses EMULATE, il faut dégrader les champs de pression NCEP/NCAR afin d'obtenir des données compatibles sur une grille de $5^\circ \times 5^\circ$. Deux méthodes ont été testées : a) la valeur en un point est égale à celle donnée à $2,5^\circ$ en ce même point, b) la valeur en un point est égale à la moyenne des valeurs aux points qui entourent celui-ci à $2,5^\circ$ (huit points ainsi que le point central). En théorie, puisque la valeur associée à une maille est représentative de tous les points situés dans cette maille, c'est la méthode b) qu'il faudrait retenir.

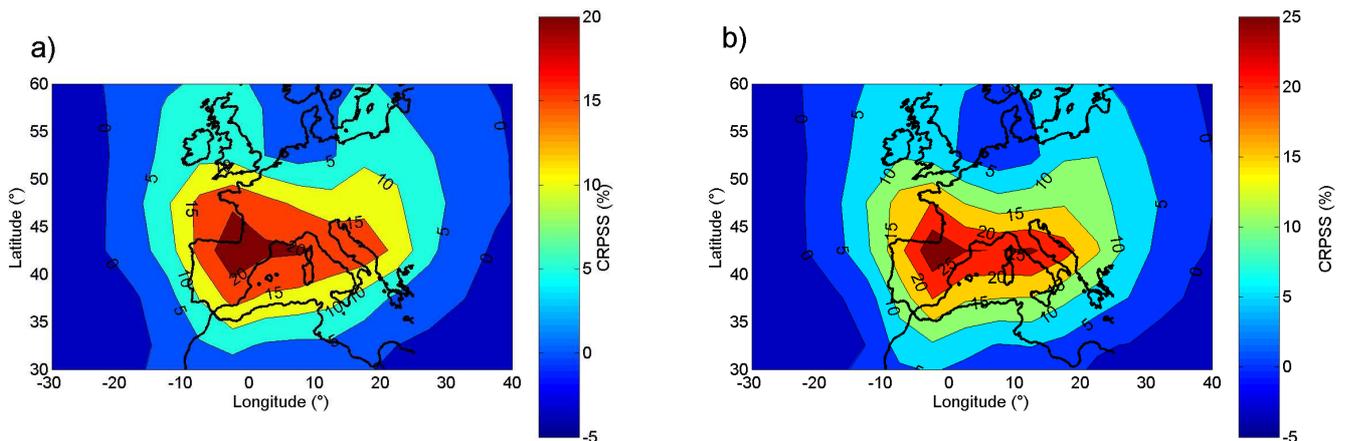


Figure 2. Performance (CRPSS) obtenue sur tous les domaines élémentaires de $5^\circ \times 5^\circ$ de côté, pour l'ensemble du bassin de l'Isère en amont de Grenoble, sur la période 1970-2006, en retenant 30 analogues. a) lorsque la valeur à 5° est égale à celle à $2,5^\circ$ au même point, b) lorsque la valeur à 5° est égale à la moyenne des valeurs environnantes à $2,5^\circ$.

Afin de choisir le domaine élémentaire de départ, nous avons cartographié les valeurs de CRPSS de tous les domaines élémentaires de 2×2 points de grille, pour chacune des deux méthodes (Figure 2). La valeur

maximale de CRPSS est obtenue pour le même domaine élémentaire ($-5^{\circ}\text{E}/0^{\circ}\text{E}$; $40^{\circ}\text{N}/45^{\circ}\text{N}$) : celle-ci vaut 24,2% avec la méthode a), et 28,8% avec la méthode b). C'est donc la seconde méthode que l'on va retenir.

La Figure 3 est identique à la Figure 2 avec des domaines élémentaires de résolution de $2,5^{\circ}$. La zone la plus pertinente est relativement similaire à celle obtenue à 5° . En revanche, les valeurs de CRPSS sont plus faibles à $2,5^{\circ}$ qu'à 5° , avec une valeur maximale de 25,4%.

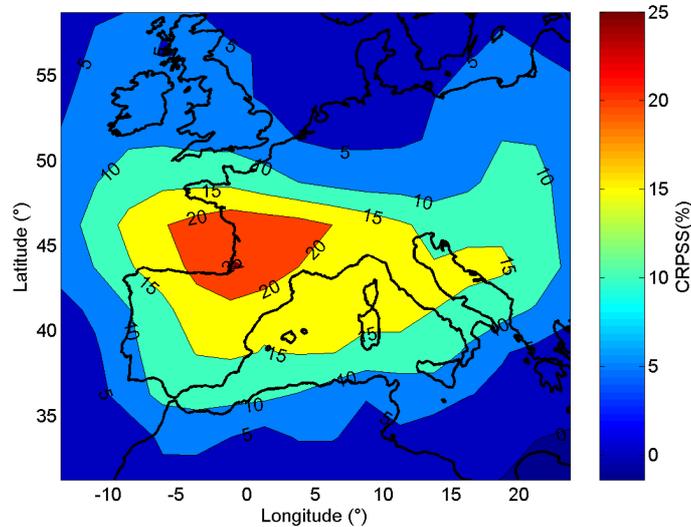


Figure 3. Performance (CRPSS) obtenue sur tous les domaines élémentaires de $2,5^{\circ}\times 2,5^{\circ}$ de côté, pour l'ensemble du bassin de l'Isère en amont de Grenoble, sur la période 1970-2006, en retenant 30 analogues.

Nous avons pu ensuite optimiser le domaine d'analogie, sur la période commune aux archives NCEP/NCAR et SAFRAN, avec les deux résolutions (Figure 4). Le domaine de départ correspond au domaine élémentaire le plus pertinent : ($-5^{\circ}\text{E}/0^{\circ}\text{E}$; $40^{\circ}\text{N}/45^{\circ}\text{N}$) à résolution 5° , ($-2,5^{\circ}\text{E}/0^{\circ}\text{E}$; $42,5^{\circ}\text{N}/45^{\circ}\text{N}$) à résolution $2,5^{\circ}$. Le domaine n'est donc pas identique si on travaille à résolution $2,5^{\circ}$ ($-7,5^{\circ}\text{E}/20^{\circ}\text{E}$; $37,5^{\circ}\text{N}/52,5^{\circ}\text{N}$) ou à 5° ($-5^{\circ}\text{E}/20^{\circ}\text{E}$; $40^{\circ}\text{N}/45^{\circ}\text{N}$), et apparaît même très grand à $2,5^{\circ}$.

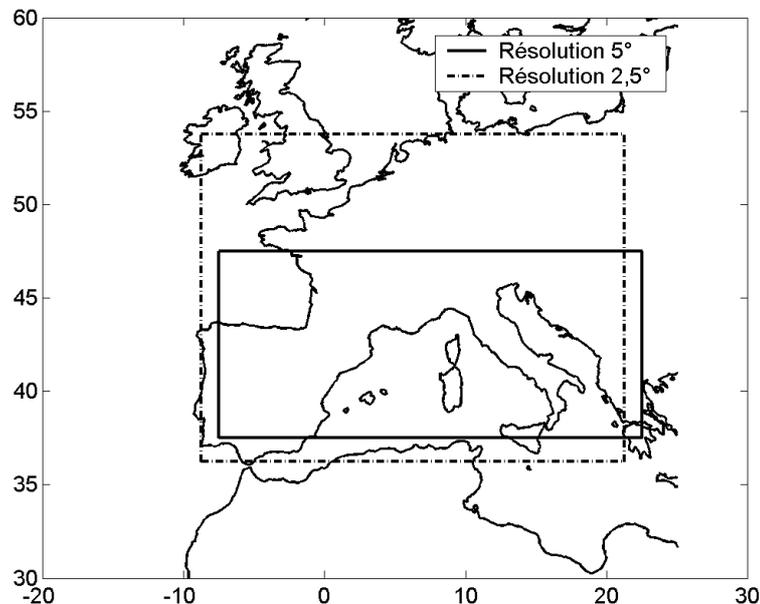


Figure 4. Domaines d'analogie optimisés avec l'archive NCEP/NCAR, pour le bassin de l'Isère, sur la période 1970-2006, avec les résolutions 5° et $2,5^{\circ}$.

III.3 Nombre de situations analogues à sélectionner

Le domaine ayant été optimisé pour le bassin de l'Isère, nous pouvons à présent optimiser le nombre d'analogues à sélectionner, qui était jusqu'ici fixé à 30, avant de lancer la recherche de situations analogues aux situations observées en 1859.

Ce nombre a été optimisé avec chacune des résolutions. La Figure 5 montre l'évolution des performances (CRPSS) en fonction de ce nombre, obtenues par application du domaine optimisé précédemment pour les résolutions 2,5° et 5°.

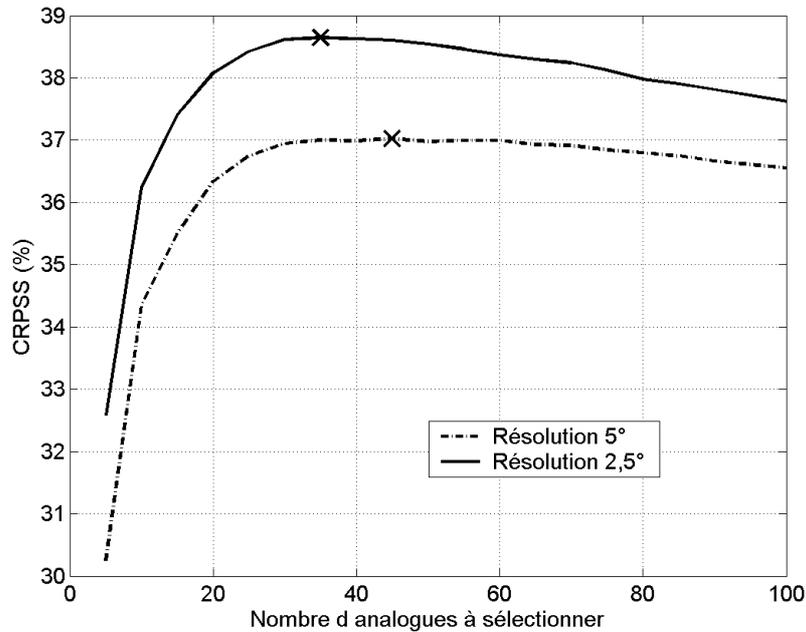


Figure 5. Evolution des performances (CRPSS) en fonction du nombre d'analogues sélectionnées, lorsque les résolutions 2,5° et 5° sont utilisées.

Finalement, le nombre optimal est de 35 si la résolution des champs de Pmer est de 2,5°, et de 45 si celle-ci est de 5°, même si l'optimum est atteint pour un nombre compris entre 35 et 60. En termes de performance, celle-ci est améliorée en appliquant la résolution 2,5° (CRPSS de 38,6%) par rapport à la résolution de 5° (CRPSS de 37%). Ce résultat peut paraître contradictoire avec celui obtenu sur un domaine élémentaire, mais il peut s'expliquer de la manière suivante. Théoriquement, la taille du domaine d'analogie doit être cohérent avec l'échelle des phénomènes météorologiques que l'on veut détecter. Ainsi lorsque le domaine est élémentaire, une résolution de 2,5° permet de capturer moins d'éléments qu'une résolution de 5°. En revanche après optimisation, il est normal d'obtenir de meilleures performances à 2,5° puisque le domaine optimisé est nettement plus grand que celui obtenu à 5.

Notre choix est donc fait : nous allons rechercher les situations analogues à celles de 1859 avec les champs de Pmer définis à 2,5°. Nous retiendrons donc 35 situations par la suite. Nous pouvons à présent rechercher les situations analogues aux situations du 1 octobre au 5 novembre 1859, et ainsi décrire plus précisément les événements météorologiques responsables de la crue de l'Isère des 31 octobre et 1^{er} novembre 1859.

IV RECONSTITUTION DU RISQUE NIVO-PLUVIOMETRIQUE DURANT LE MOIS

D'OCTOBRE 1859

La recherche de situations analogues permet, lorsque les précipitations observées dans le passé sont disponibles, de reconstituer l'évolution des précipitations probables. Ceci constitue notre premier objectif. De plus, à l'aide des dates des situations analogues, nous pouvons également extraire d'autres variables descriptives des situations météorologiques. Ceci nous permettra, dans le second volet de cette section, d'analyser l'évolution de la masse d'air et des types de temps rencontrés sur la période d'étude.

IV.1 Reconstitution de l'évolution de la pluviométrie sur les sous-bassins de l'Isère

Les 35 situations retenues, les plus analogues aux situations observées en 1859, ont été extraites sur la période allant du 1/8/1970 au 30/7/2006. Les situations de 1859, en termes de champ de Pmer, sont décrites à résolution 2,5°. Les situations analogues sont également décrites à résolution 2,5°, et ont été extraites des réanalyses NCEP/NCAR. Les cumuls observés lors des situations analogues ont été classés et des quantiles au non-dépassement ont été déduits (quantiles d'ordre 20, 60 et 90 %). La Figure 6 montre les courbes reliant ces quantiles, pour trois sous-bassins versants de l'Isère, du 01/10 au 05/11 1859. Les valeurs maximales de chaque distribution sont également indiquées afin d'avoir une idée des plus fortes précipitations observées lors des situations analogues.

Le début du mois d'octobre 1859 fut relativement sec, comme le montrent ces graphiques avec des quantiles très faibles du 1^{er} au 5 octobre. En revanche, le reste du mois semble agité avec une alternance d'épisodes pluvieux et de courtes périodes sèches. Un premier épisode pluvieux se produit entre le 9 et le 16, puis on observe un autre épisode, plus conséquent, du 20 au 24 octobre. Un troisième épisode se produit à la fin du mois d'octobre et au tout début du mois de novembre, moment où la crue s'est produite. Néanmoins, les valeurs des quantiles obtenues sont dans l'ensemble bien inférieures à la valeur du cumul décennal, indiqué sur les graphiques. On peut supposer ainsi que les cumuls de pluie, considérés à l'échelle de la journée, ne sont pas excessifs.

Ces résultats ne représentent que le potentiel des épisodes pluvieux qui ont eu lieu durant cette période. Il faudra se reporter à l'étude de Obled et al. (2009) pour étudier plus précisément la répartition des précipitations sur les différents sous-bassins versants retenus.

IV.2 Etude de l'évolution de la situation météorologique sur la période

Il est également possible d'extraire, pour chacune des dates analogues, d'autres variables météorologiques afin de caractériser les situations météorologiques de 1859. En effet, le champ de Pmer seul ne peut expliquer complètement l'événement.

La période sur laquelle les données pluviométriques sont disponibles s'étend du mois d'août 1970 à juillet 2006, tandis que la période couverte par les réanalyses NCEP/NCAR est plus large. C'est la période limitée de l'archive pluviométrique qui a restreint jusqu'à présent notre recherche de situations analogues. Or si nous ne nous intéressons plus aux précipitations, mais aux données supplémentaires disponibles dans les réanalyses, nous pouvons étendre la période candidate dans laquelle nous recherchons les situations analogues. Puisque les réanalyses NCEP/NCAR couvrent la période allant de 1948 à 2007, nous pouvons maintenant rechercher des situations analogues, en termes de champ de Pmer, à celles de 1859, jusqu'en 1948. Pour une raison de simplicité, nous avons retenu le même nombre de situations que précédemment (égal à 35).

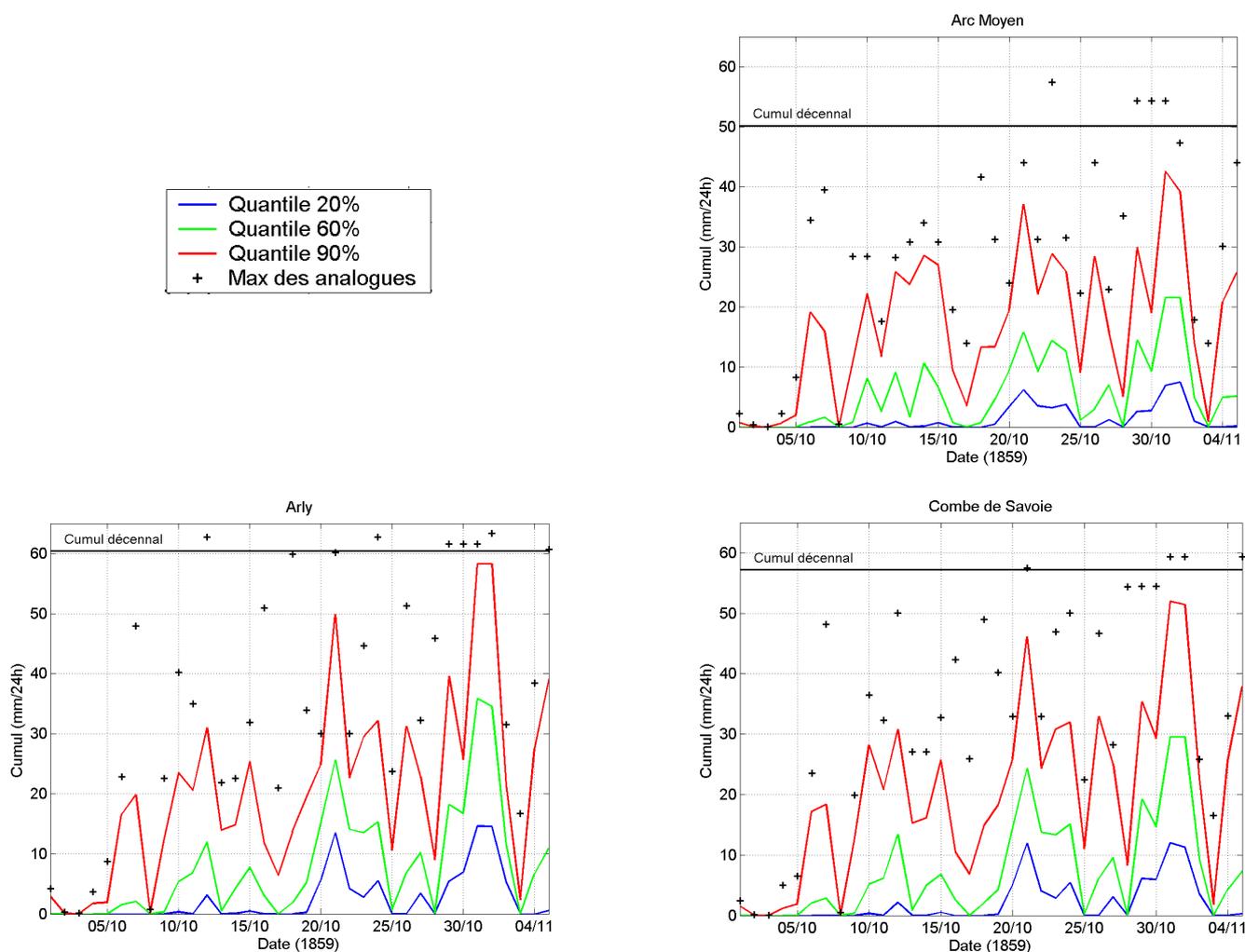


Figure 6. Evolution des quantiles au non-dépassement prévus sur la période du 01/10 au 05/11/1859, sur les sous-bassins Arc Moyen, Arly et Combe de Savoie.

IV.2.1 Température

Nous nous sommes intéressés à la température qui aurait pu être observée à basse altitude. Notre choix s'est orienté vers la température au niveau de pression 850 hPa (soit aux environs 1500 m). En complément, nous avons étudié le géopotentiel à 850 hPa, qui nous renseigne sur les altitudes associées à ces températures. Nous avons extrait ces variables au point de grille le plus proche des sous-bassins de l'Isère en amont de Grenoble (5°E ; 45°N).

Avec la fenêtre temporelle de plus ou moins deux mois autour de chaque jour cible, il est donc possible, pour les jours d'octobre 1859, que des analogues soient situées en plein mois d'août. Or c'est à l'automne que la température diminue rapidement dans les basses couches. Il serait donc plus intéressant de sélectionner les analogues dans une fenêtre temporelle plus restreinte. Nous avons donc appliqué deux fenêtres temporelles pour respecter la saisonnalité : l'une de plus ou moins un mois et l'autre de plus ou moins deux mois.

Pour chaque journée de la période considérée en 1859, 35 situations analogues ont été sélectionnées sur la période couverte par les réanalyses NCEP/NCAR (1948-2007). La température à 850 hPa a été extraite au point de grille défini précédemment. Nous avons alors construit, pour chaque jour cible de 1859, la distribution des températures, à partir de laquelle nous avons déduit les quantiles d'ordre 20, 60 et 90%. En outre, nous avons retenu la température observée lors de la situation la plus analogue à chaque situation cible (Analogue n°1). La Figure 7 présente l'évolution des statistiques de température observée lors des situations analogues, après avoir appliqué une fenêtre de plus ou moins deux mois (gauche) et de plus ou moins un mois (droite).

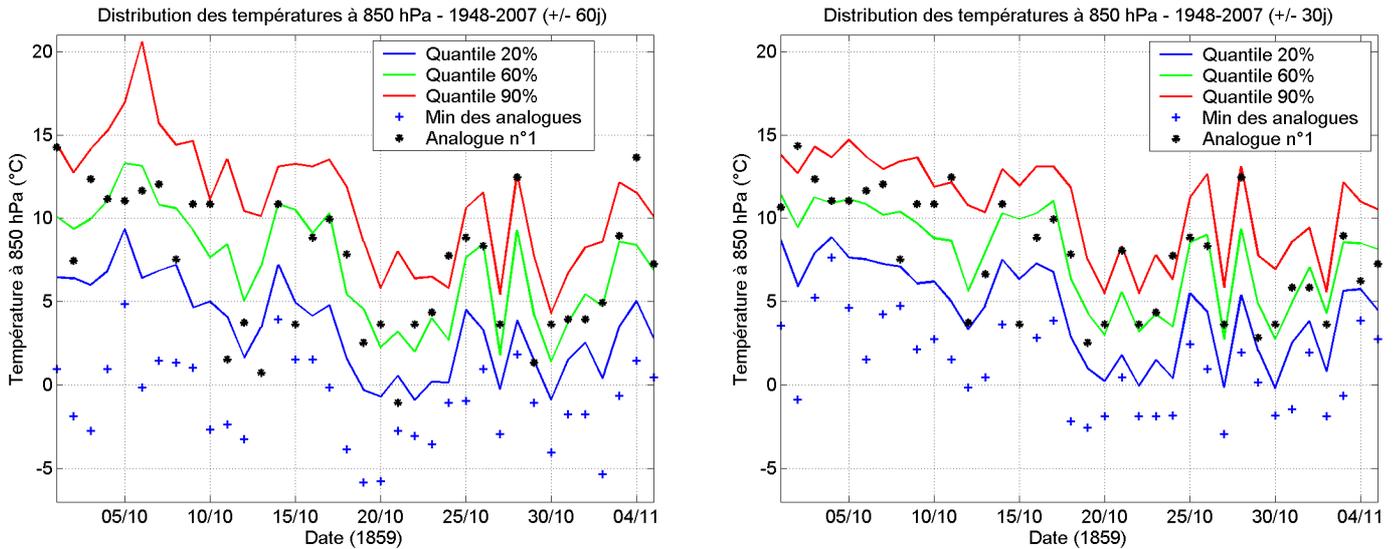


Figure 7. Evolution de la distribution des températures à 850 hPa observées lors des journées analogues à celles de 1859, sur la période 1948-2007 : en appliquant une fenêtre de +/- 2 mois (gauche) et une fenêtre de +/- 1 mois (droite).

De la même manière, la distribution des géopotentiels a pu être établie pour chaque jour cible de 1859. Le géopotentiel moyen de chaque distribution et le géopotentiel de la première analogue en ont été déduits (Figure 8). Afin de respecter la saisonnalité du géopotentiel, nous avons ici appliqué une fenêtre temporelle de plus ou moins deux mois autour de chaque jour cible.

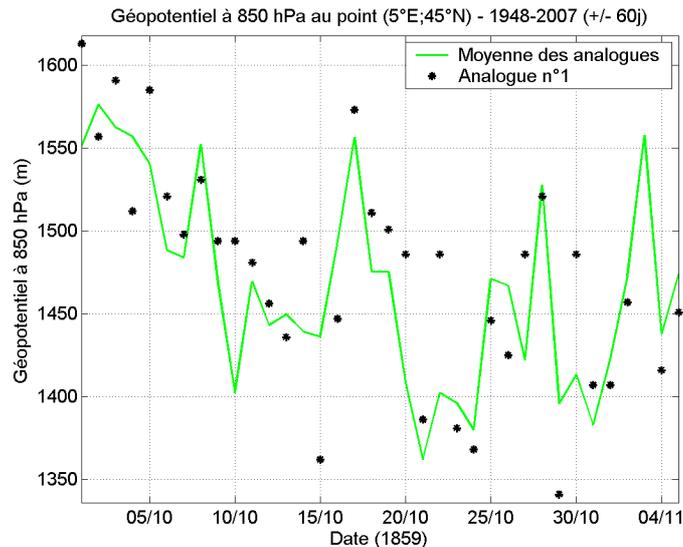


Figure 8. Evolution du géopotentiel à 850 hPa établie à partir des situations analogues aux situations cibles de 1859 : moyenne de chaque distribution et valeur du géopotentiel associé à la situation la plus analogue à chaque situation cible. La fenêtre temporelle dans laquelle les analogues ont été sélectionnées est de large de quatre mois.

L'examen de la Figure 7 montre que la fenêtre temporelle de quatre mois apparaît finalement plus pertinente. En effet en réalité, elle permet, pour les journées d'octobre 1859, de sélectionner des situations analogues du mois de décembre. Ainsi les distributions tiennent compte de ces situations, avec notamment des températures extrêmes basses plus faibles.

Nous pouvons déduire de la Figure 7 cinq périodes distinctes :

- période douce du 1/10 au 17/10 ;
- période froide du 19 au 24/10 ;
- période assez douce du 25 au 28/10 ;
- période froide les 29 et 30/10 ;
- période douce à partir du 31/10.

En outre, la Figure 8 indique une fluctuation assez faible du géopotential. Sur notre période, l'altitude à laquelle on aurait observé la température à 850 hPa est d'environ 1450 m. Nous retrouvons néanmoins des altitudes plus basses du 21 au 24 et du 29 au 31 octobre, ce qui indique le passage probable d'une dépression ou d'un talweg à proximité durant ces journées. En revanche, le géopotential remonte fortement du 1^{er} au 3 novembre, ce qui traduit le radoucissement observé à l'aide des températures. Une hausse du champ de pression pourrait également être à l'origine de cette brusque remontée du géopotential.

Il y aurait donc eu des épisodes de temps perturbé combinés à des changements de température. On remarque notamment, en analysant conjointement la Figure 6 et la Figure 7, que des précipitations se sont probablement produites au cours de l'épisode plus froid entre le 20 et le 24 octobre. Un manteau neigeux significatif se serait alors constitué en altitude. Ensuite, un radoucissement de la masse d'air aurait accompagné l'épisode pluvieux du 31 octobre au 2 novembre. Ceci aurait eu pour conséquence une fonte partielle des stocks de neige constitués en altitude.

Enfin, pour vérifier la pertinence des graphiques précédents, nous pouvons les comparer avec l'évolution des températures relevées par exemple à Grenoble (Figure 9).

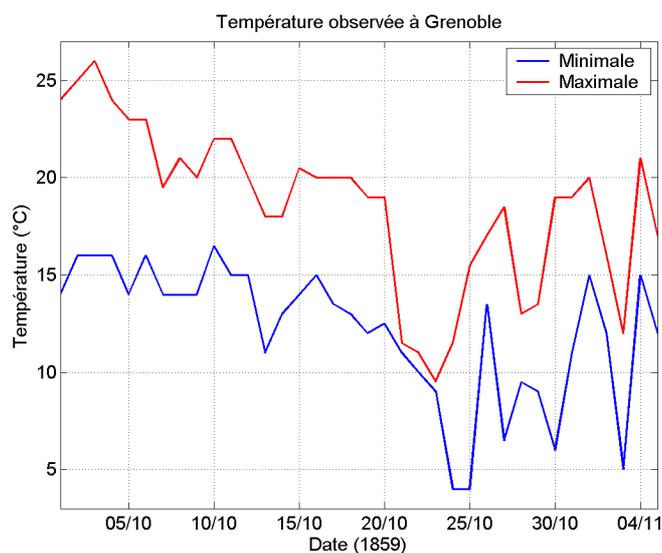


Figure 9. Evolution des températures minimales et maximales relevées à Grenoble sur la période du 1 octobre au 5 novembre 1859.

Nous retrouvons ici globalement une évolution de température à Grenoble similaire à l'évolution de la température à 850 hPa reconstituée par les analogues. Les différences, telles que celles observées les 28 et 29 octobre (températures limitées à Grenoble mais assez élevées en altitude), pourraient s'expliquer notamment par la présence de couverture nuageuse ou par l'occurrence de précipitations.

IV.2.2 Régimes de temps

Il est également possible, à l'aide de la technique d'analogie, d'analyser la situation synoptique de la période de 1859. Nous avons d'abord cherché à décrire l'évolution du régime de temps qui a pré-dominé durant cette période. Pour ce faire, nous avons utilisé la classification par types de temps (TDT) développée

par EDF (Paquet et al., 2006). Les champs moyens associés aux TDT sont représentés sur la Figure 10. La période totale sur laquelle les TDT sont disponibles couvrent les années 1953-2006. Dans cette partie, nous n'avons donc recherché les situations analogues à celles de 1859 en termes de Pmer que dans cette période. Nous en avons déduit, pour chaque jour cible de 1859, la proportion de situations analogues associées à chaque TDT (Figure 11).

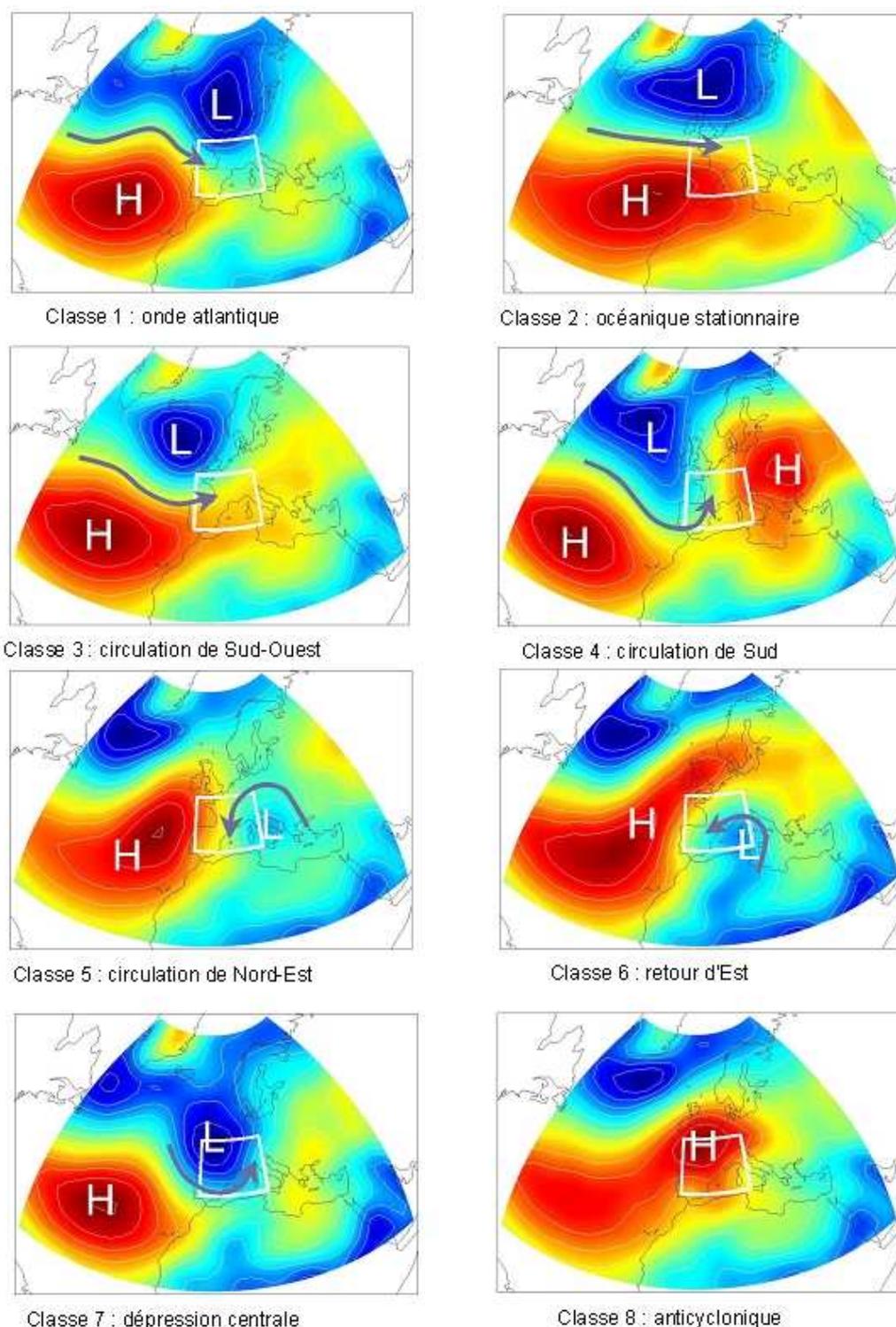


Figure 10. Classification par TDT établie par EDF.

Les résultats obtenus confirment que le début du mois d'octobre fut marqué par des conditions anticycloniques.

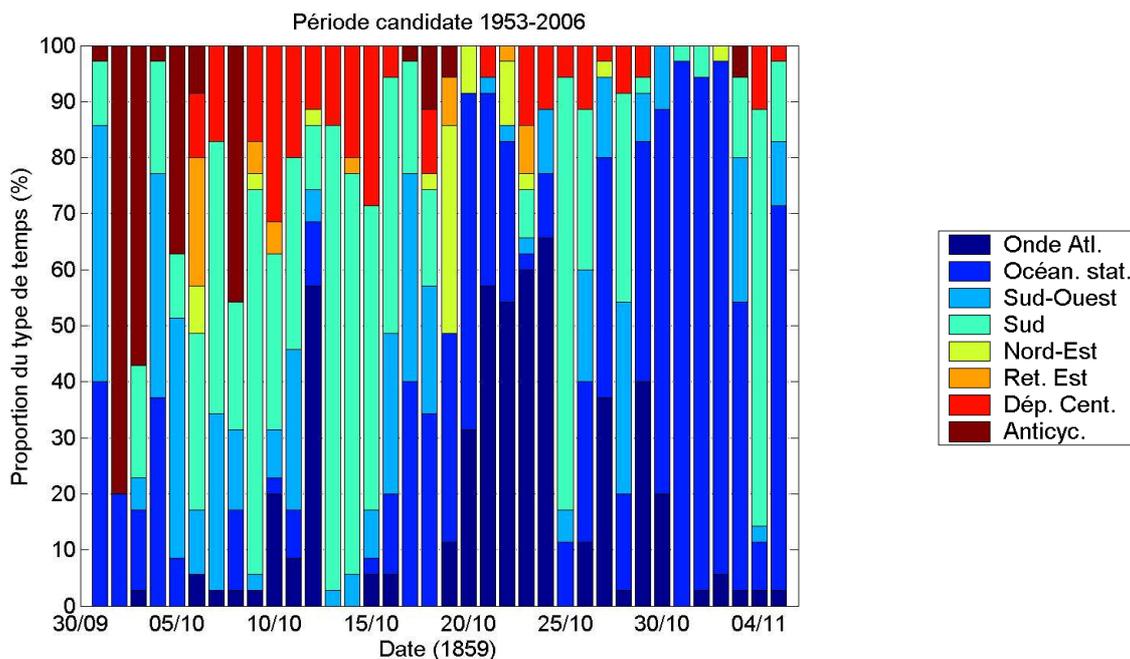


Figure 11. Répartition des TDT EDF parmi les situations analogues à chaque situation cible de la période de 1859.

Entre le 6 et le 16 octobre, le TDT le plus représenté par les analogues retenues est le type Sud (classe 4). En revanche, aucun TDT ne prédomine sur la période du 17 au 20 octobre. Nous avons vu, lors de l'analyse des précipitations obtenues par les analogues, que la période du 21 au 24 octobre était humide avec une forte probabilité de précipitations assez conséquentes. Le TDT qui prévaut durant cette période est le type Onde Atlantique (classe 1), ce qui confirme les résultats obtenus sur les précipitations. En effet, ce type de temps est associé à une alternance d'ondes baroclines génératrices de précipitations. Par ailleurs nous remarquons que pour la journée du 25 octobre, le TDT largement représenté est le type Sud (classe 4), ce qui confirme bien le radoucissement qui s'est produit à cette date là (cf. températures à 850 hPa, Figure 7). La fin du mois reste dominée par un flux d'ouest avec la récurrence du type de temps Océanique Stationnaire (classe 2). Ainsi entre le 30 octobre et le 2 novembre, pas moins de 70% des situations analogues retenues appartiennent à cette classe. Ce résultat confirme l'occurrence de précipitations sur la zone d'intérêt durant cette période, mais cela ne permet pas d'affirmer que ces précipitations furent accompagnées d'un radoucissement. En revanche, dans ce type de configuration synoptique, le risque de vents forts de secteur sud-ouest est réel. Cet élément est peut-être également à l'origine de la fonte nivale qui s'est probablement déclenchée en altitude à cette période.

Pour achever cette étude, nous avons étudié la circulation atmosphérique à partir des champs de géopotential à 500 hPa (Z500). Puisque nous disposons des dates des situations analogues à chaque situation cible de la période de 1859 en termes de Pmer, les champs de Z500 ont été extraits des réanalyses NCEP/NCAR pour chacune de ces dates. Finalement pour chaque jour cible entre le 1/10 et le 5/11/1859, le champ moyen de Z500 obtenu sur les champs des analogues a été calculé. Pour que l'analogie soit de meilleure qualité, nous avons recherché les dates analogues dans la période la plus longue possible (1948-2007). Nous avons cartographié ces champs pour chaque jour cible, et pour chacun de ces champs, nous avons comparé le champ de Pmer reconstitué par Météo-France (Jourdain et al., 2009). La Figure 12 illustre les situations rencontrées lors de deux journées contrastées : le 23 octobre et le 1^{er} novembre. Les cartes restantes sont reportées en annexe.

Le 23 octobre, journée incluse dans une période pluvieuse, nous avons vu précédemment qu'un refroidissement s'était probablement produit. Nous retrouvons ce résultat grâce à ces cartes. En effet, le champ de Z500 indique un talweg prononcé axé des îles britanniques à la Méditerranée occidentale, synonyme d'arrivée d'air froid en altitude et donc d'un changement de masse d'air. Le champ de Pmer est particulièrement cohérent avec le champ d'altitude. Une dépression se creuse ainsi en Méditerranée, entraînant le basculement du vent au secteur nord sur le sud-est de la France. Le 1^{er} novembre, on remarque que les champs d'altitude et de surface sont fort similaires, témoignant d'un flux zonal très marqué (cf.

analyse de Jourdain et al., 2009). L'air advecté sur la région d'étude est par conséquent très humide et relativement doux.

De manière générale, les résultats obtenus sur l'ensemble de la période rejoignent nos conclusions de l'analyse des TDT.

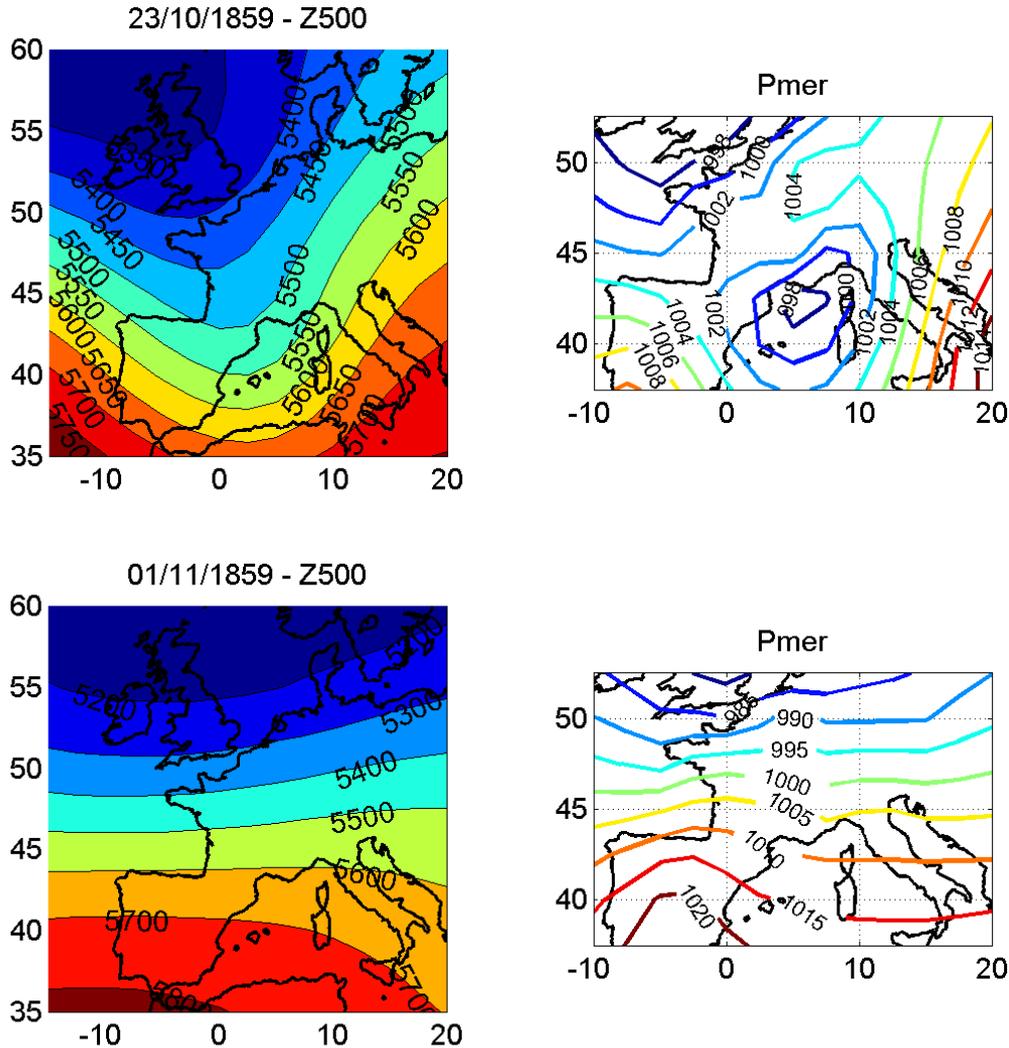


Figure 12. Situations météorologiques reconstituées pour les journées du 23 octobre et du 1er novembre 1859. Champ moyen de géopotential à 500 hPa obtenu à l'aide des analogues (gauche) et champ de Pmer reconstitué par Météo-France (droite).

V CONCLUSION

L'objectif de cette étude était de reconstituer la séquence pluviométrique de la période précédant la crue de l'Isère de 1859. Pour y répondre, une technique d'analogie a été employée. Dans un premier temps, cette technique a été adaptée au bassin de l'Isère. Les paramètres (domaine d'analogie, nombre d'analogues retenues) ont ainsi été optimisés sur le bassin.

Pour chaque journée entre le 1^{er} octobre et le 5 novembre 1859, nous avons ensuite sélectionné 35 situations analogues en termes de champ de pression au niveau de la mer, dans une période plus récente (1970-2006). Grâce aux précipitations observées au cours de ces journées, nous avons pu construire une distribution empirique des pluies pour chaque journée de 1859, et ce sur les sept sous-bassins de l'Isère définis. Nous avons déduit des résultats plusieurs périodes pluvieuses entrecoupées de périodes sèches. Nous avons notamment détecté deux séquences probablement pluvieuses : la première du 20 au 24 octobre et la seconde du 29 octobre au 2 novembre. Néanmoins, les précipitations journalières telles qu'elles ont été quantifiées ne semblent pas exceptionnelles.

Nous avons également étudié, sur la même période, les distributions des températures mesurées au niveau de pression 850 hPa, observées lors des journées analogues d'une période récente. Cette analyse a permis de mettre en évidence une alternance de périodes froides et chaudes. Notamment la période pluvieuse du début de la seconde décennie d'octobre 1859 aurait été associée à un refroidissement de la masse d'air, ce qui aurait permis à un stock de neige de se constituer en altitude. Par la suite, un radoucissement se serait produit à la toute fin du mois d'octobre, moment où un nouvel épisode de précipitations aurait probablement eu lieu. Ces éléments combinés auraient alors engendré une fonte partielle du manteau neigeux en altitude et ainsi contribué à la crue de l'Isère.

Enfin, nous avons analysé les types de temps prédominants durant la période précédant la crue. En outre, nous avons étudié la circulation atmosphérique à échelle synoptique à l'aide du champ moyen de géopotential à 500 hPa construit grâce aux dates analogues retenues. Il s'avère que la période froide et humide du début de la seconde décennie d'octobre aurait été associée à une ondulation atlantique, avec une alternance de talwegs et de dorsales. En revanche, c'est probablement une circulation d'ouest marquée qui a prévalu durant la fin du mois d'octobre 1859. Des perturbations atlantiques accompagnées d'un radoucissement auraient alors été à l'origine de précipitations modérées mais aussi de la probable fonte partielle du manteau neigeux en altitude.

Nos travaux s'inscrivent dans la continuité et la complémentarité de ceux menés par Jourdain et al. (2009). Les résultats semblent bien s'accorder avec leurs conclusions. Obled et al. (2009) ont également complété notre étude par l'utilisation des dates analogues que nous avons générées, ce qui leur a permis de reconstituer plus précisément un scénario de précipitations et de températures sur la même période.

Ce travail révèle également le fort intérêt d'appliquer une technique d'analogie à un autre contexte que celui de la prévision, tel que celui de la reconstitution de données historiques assorties des incertitudes.

Références bibliographiques

Ben Daoud A., 2009. **La prévision des précipitations par recherche d'analogues : état de l'art et perspectives**. La Houille Blanche (sous presse).

Duband D., 1970. **Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations**. Thèse de 3^{ème} cycle de la Faculté des Sciences de Paris.

Hersbach H., 2000. **Decomposition of the ranked probability score for ensemble prediction systems**. Weather Forecasting, 15, 559-570.

Jourdain S., Clavel A. et A. Auffray, 2009. **Contexte météorologique des mois d'octobre et novembre 1859**. Colloque Isère 1859 -2009, 150^{ème} anniversaire de la crue de référence, 5 nov. 2009.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D., 1996. **The NCEP/NCAR 40-Year Reanalysis Project**. Bulletin of the American Meteorological Society, 77, 437-471.

Lorenz E.N., 1969. **Atmospheric predictability as revealed by natural occurring analogues**. Journal of the Atmospheric Sciences, 26, 636-646.

Obled Ch., Panthou G., Garçon R. et F. Gottardi, 2009. **Crue de l'Isère 1859 : Reconstitution d'un scénario de précipitations et de températures pour la période 1er Octobre- 5 Novembre 1859**. Colloque Isère 1859 -2009, 150^{ème} anniversaire de la crue de référence, 5 nov. 2009.

Panthou G., 2009. **Etude hydrométéorologique d'un événement nivo-pluvial : La crue de l'Isère 1859**. Mémoire de Master 2P ES Université J. Fourier. 70 p.

Paquet, E., Gailhard, J. and R. Garçon, 2006. **Evolution of GRADEX method: improvement by atmospheric circulation classification and hydrological modelling**. La Houille Blanche, 5, 80-90.

Sauquet E., 2006. **Mapping mean annual river discharges : geostatistical developments for incorporating river network dependencies**. Journal of Hydrology, 331, 300-314.

Teweles J. and H. Wobus, 1954. **Verification of prognosis charts**. Bulletin of the American Meteorological Society, 35, 455-463.

Uppala, S.M., Kallberg, P.W., Simmons, A.J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hølm, E., Hoskins, B.J., Isaksen, I., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J., 2005. **The ERA-40 re-analysis**. Quarterly Journal of the Royal Meteorological Society, 131, 2961-3012.

Vidal J.-P., Martin E., Franchistéguy L., Baillon M. and J.-M. Soubeyroux, 2009. **A 50-year high-resolution atmospheric reanalysis over France with the Safran system**. International Journal of Climatology, DOI: 10.1002/joc.2003.

Annexe 2

Type de communication	Article
Titre	Comparison of 850-hPa relative humidity between ERA-40 and NCEP/NCAR re-analyses: detection of suspicious data in ERA-40
Auteurs	Ben Daoud A., Sauquet E., Lang M., Obled Ch. And Bontron G.
Année	2009
Source	Atmospheric Science Letters

Comparison of 850-hPa relative humidity between ERA-40 and NCEP/NCAR re-analyses: detection of suspicious data in ERA-40

Aurélien Ben Daoud,^{1*} Eric Sauquet,¹ Michel Lang,¹ Charles Obled² and Guillaume Bontron³

¹Cemagref, UR HHLY, 3 bis Quai Chauveau – CP220, F-69336 Lyon, France

²Laboratoire des Transferts en Hydrologie et Environnement, Institut National Polytechnique de Grenoble, B.P. 53, 38041 Grenoble, Cedex 9, France

³Compagnie Nationale du Rhône, Direction de l'Energie – Centre de Gestion de Production du Rhône, 2 rue André Bonin, 69316 Lyon, Cedex 4, France

*Correspondence to:

Aurélien Ben Daoud, Cemagref,
Hydrology-Hydraulics Research
Unit, Lyon, France.

E-mail:

aurelien.bendaoud@cemagref.fr

Abstract

An exploratory study is performed on the 850-hPa relative humidity data (RH850) extracted from the Re-Analysis Project (ERA-40) and National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) re-analyses covering a domain centred on western Europe (60°W–60°E, 15°N–75°N). The largest deviations between the two data archives are observed over the North Atlantic Ocean. In addition, unrealistic values of RH850 are detected in the ERA-40 re-analysis at resolutions of both 2.5° and 1.125°. There is no strong correlation between RH850 provided by ERA-40 and observations from radio sounding stations, thus ruling out a straightforward correction of the detected anomalous values. Copyright © 2009 Royal Meteorological Society

Keywords: re-analyses; relative humidity; ERA-40; NCEP/NCAR; western Europe; radio sounding; anomalous values

Received: 9 September 2008
Revised: 20 November 2008
Accepted: 20 November 2008

1. Introduction

Ever since the first indications of a tendency to climate change in the second half of the 20th century, we have needed to acquire reliable and homogeneous meteorological data both in time and space. Consequently, in the 1990s, this led to a new kind of dataset known as re-analysis.

Among the currently available re-analysis archives, two are long enough for studying climate evolution; the first data archive is issued by the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) (NNR), covering the period from 1948 to the present, while the second data archive corresponds to the 45-year Re-Analysis Project (ERA) of the European Centre for Medium-Range Weather Forecasts (ECMWF), which covers the 1957–2002 period. The NNR dataset was provided by a global spectral model with T62 truncation (~2.5° resolution) and 28 vertical levels (Kalnay *et al.*, 1996). Initially, the ERA data were obtained through a global spectral model with T159 truncation (~1.12° resolution) and 60 vertical levels (Uppala *et al.*, 2005). ERA data are freely distributed to the scientific community, but after being downgraded to a lower resolution corresponding to a T95 truncation (~2.5° resolution); as a result, the ERA dataset is more often used for applications. We should note that ERA and NNR are not from the same generation. In addition, a different data assimilation system is used for each archive. The amount of data assimilated by the model is slightly

larger for ERA than for NNR, including new data provided by satellites. Even if ERA is more recent and provides a finer resolution for a restricted community, the choice between these two data archives is not necessarily evident in advance. The users of re-analysis data should firstly perform an exploratory investigation, including a comparison between the two archives over the study area at different scales. This approach is important consideration in our study, which presents the main findings of such a dataset comparison.

We focus our results on one variable that has a physical meaning for the prediction of precipitation: the 850-hPa relative humidity (RH850). This variable, which is used in an operational precipitation forecasting system based on meteorological analogues (Bontron and Obled, 2003), is currently extracted from NNR. Since this system is operating in south-eastern France, the present study is mainly concerned with western Europe.

The data are presented in Section 2, and the results are discussed in Section 3, first for the comparison of RH850 between ERA and NNR at 2.5° resolution, and then for the analysis of ERA RH850 at finer resolution (1.125°). ERA RH850 values are also compared with observations provided by radio sounding. Finally, the conclusions are given in Section 4.

2. Data and methods

In a first step, RH850 data were extracted from the ERA and the NNR with the same spatial resolution

(2.5°) and full time resolution (six hourly) over a common period from 1st September 1957 to 31st August 2002. The extraction area covers most of Europe and part of the North Atlantic Ocean (60°W–60°E, 15°N–75°N).

Secondly, RH850 data were extracted from ERA with the full spatial resolution (1.125°) over a small area (30.375°W–30.375°E, 33.750°N–64.125°N), for the 1958–2001 period. Finally, we considered an area taking in eastern France with a high density of radio sounding observations. Eight available radio sounding stations were selected: Trappes (Tr), Nancy (Na), Lyon (Ly) and Nîmes (Ni) in France; Stuttgart (St) and Idar-Oberstein (Id) in Germany; Payerne (Pa) in Switzerland; and Milano (Mi) in Italy. RH850 data were extracted from the free-access WYOMING soundings database (<http://weather.uwyo.edu/upperair/europe.html>).

3. Results

3.1. Comparison of RH850 between ERA and NNR at 2.5° resolution

The deviations between the NNR and ERA fields are calculated and averaged over the whole common period 1957–2002 (i.e. 65 744 values) for each grid point of the extraction area. Mean deviations (MDs) are displayed in Figure 1.

Globally, the two archives yield comparable values for points located in the continental area. However, strong differences between the two archives are detected over the North Atlantic Ocean. The MD averaged over the domain is equal to 4.6%, a value non-negligible in comparison with the measurements of RH850 which range from about 10% to 100% in this area. Since the data assimilated into the two re-analyses are similar, such deviations seem too high. Moreover, the maximum MD value, reached on a grid point located in Greenland (37.5°W, 70°N), is equal to 49.5%.

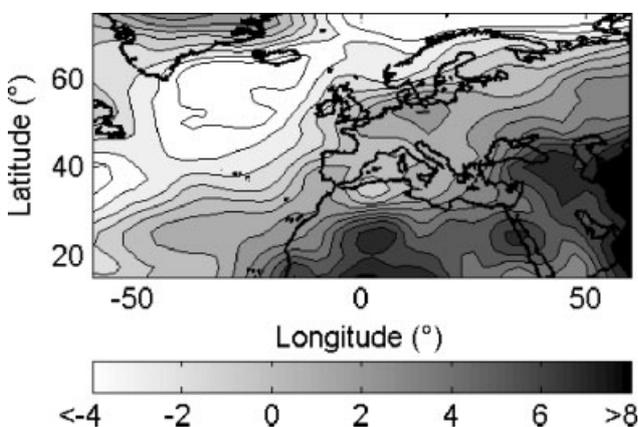


Figure 1. Mean deviation of RH850 (in %) between ERA and NNR, computed over the whole common period (1/9/1957–31/8/2002). The contour interval is 2%.

The two archives display totally different behaviours in terms of extreme values: unrealistic values of RH850 can be found in ERA, but not in NNR. Thus, relatively few values detected in ERA are lower than 0% (negative anomalous values), whereas values higher than 100% (positive anomalous values) are commonly observed at many grid points. Values of RH850 slightly above 100% can be measured in some cases (e.g. during intense convective phenomena) when the air is supersaturated. Nevertheless, values higher than 102 or 103% in ERA are clearly abnormal. On the other hand, negative values are totally unrealistic.

Since there are numerous values of RH850 higher than 100%, the following analysis concentrates on such cases. RH850 values from ERA above 100% are sorted using a bin width of 1% (i.e. class 116 contains all values between 115.5 and 116.5), except for the last class (>120) that pools values higher than 119.5. The corresponding NNR values are extracted, considered as a sample and finally summarised by box plots (Figure 2).

The NNR RH850 medians are high, ranging from 95% in the case of anomalous values of 101% up to more than 99% for those higher than 110%. This result shows a significant overall agreement between NNR RH850 values and the magnitude of the ERA anomalous values. However, for each positive anomalous value lower than 119%, at least one NNR value is equal to 0%. The results of this analysis suggest we should be prudent when using ERA RH850 at 2.5° resolution in view of the large number of positive anomalous values.

To supplement this analysis, we examine the ERA RH850 data with the full spatial resolution (1.125°) by searching for anomalous values and comparing them with radiosonde observation data. The main results are given in the following sub-section.

3.2. Further analysis of RH850 anomalous values in ERA at 1.125° resolution

In a second step, we examine the ERA RH850 data with the full resolution (1.125°) over the extraction area. Anomalous values are also detected, demonstrating that such data are not due to the interpolation procedure. Figure 3 shows the proportions of (Figure 3(a)) negative and (Figure 3(b)) positive anomalous values computed at each grid point, with the extreme values of (Figure 3(c)) negative and (Figure 3(d)) positive anomalous values attained at each grid point.

The proportions of negative anomalous values over the period are not so large (between 0 and 0.15% of cases), whereas positive anomalous values are much more common (between 0.1 and 13.3% of cases). These results are in accordance with the compilation obtained at 2.5° resolution.

Negative anomalous values of RH850 mainly appear in western Europe, with the highest proportion

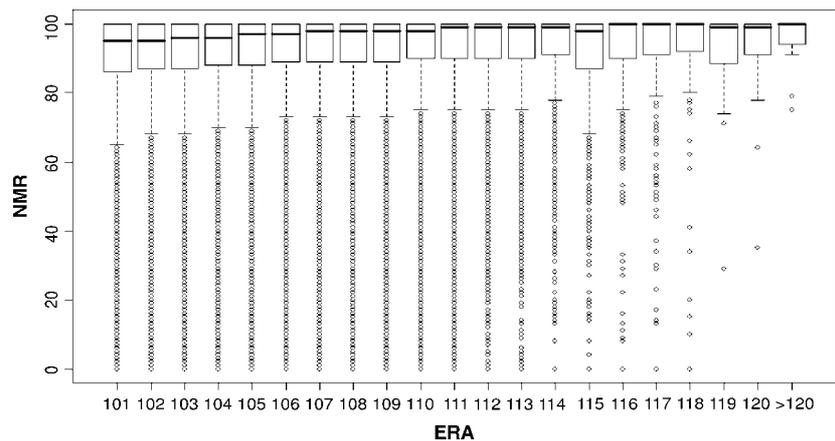


Figure 2. Relation between NNR and ERA RH850 values for each positive anomalous value detected at a grid point of the extraction domain and for the whole period. The box plots are defined by the first quartile Q25, the median and the third quartile Q75. The limits of the whiskers are given by the closest observations above $Q25 - 1.5(Q75 - Q25)$ and below $Q75 - 1.5(Q75 - Q25)$; open circles indicate outliers. Units are %.

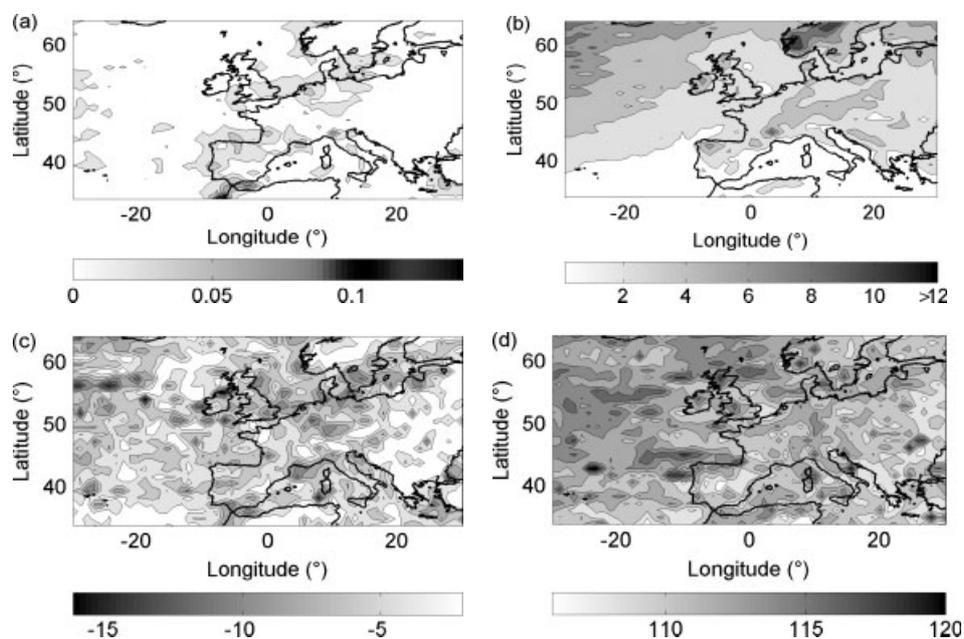


Figure 3. Proportions of (a) negative and (b) positive anomalous values, along with extreme values of (c) negative and (d) positive anomalous values at each grid point; computed for the period 1958–2001. Units are % and the contour interval is (a) 0.02% for (a) and 2% for (b), (c) and (d).

observed along the north Moroccan coast. Positive anomalous values of RH850 are more often observed above 30°N , especially in polar regions and over Scandinavia. In addition, a large proportion of positive anomalous values are detected in regions with contrasted topography in western Europe.

Extreme values for both anomalous value classes are particularly unrealistic. Indeed, the lowest negative anomalous values lie between -1 and -10% , with a minimum of -16.2% at a grid point located in the North Atlantic Ocean (19.125°W , 54°N). Nevertheless, there are few points at which no anomalous value can be detected. As shown in Figure 3(d), most of the highest positive anomalous values lie between 110% and 115% , with a maximum of 121.8% also located in the North Atlantic Ocean (23.625°W , 42.75°N).

We also compare the ERA RH850 data with the observed RH850 values provided by radiosonde, which are assumed to be assimilated into the model. Figure 4(a) shows the location of radio sounding stations with respect to the re-analysis grid points. Each of these stations is associated with the nearest grid point. The coefficients of determination (R^2) are computed by considering all concomitant RH850 values extracted from radio sounding and ERA. In Figure 4(b), we plot R^2 values against the Euclidean distance (D) between the grid point and the related station.

The R^2 coefficients range from 0.69 for Payerne to 0.87 for Nîmes, demonstrating a relatively strong correlation between the two datasets. The maximum R^2 corresponds to the shortest distance (in the case of

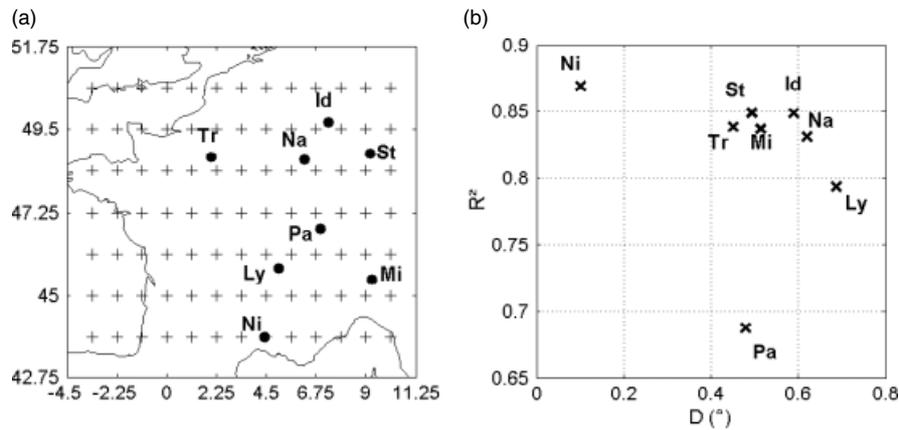


Figure 4. Location of radio sounding stations with respect to ERA-40 re-analysis grid points (a) and coefficient of determination computed for each station, as a function of the Euclidean distance D between the station and the closest grid point (b). Distance units are in decimal degrees ($^{\circ}$).

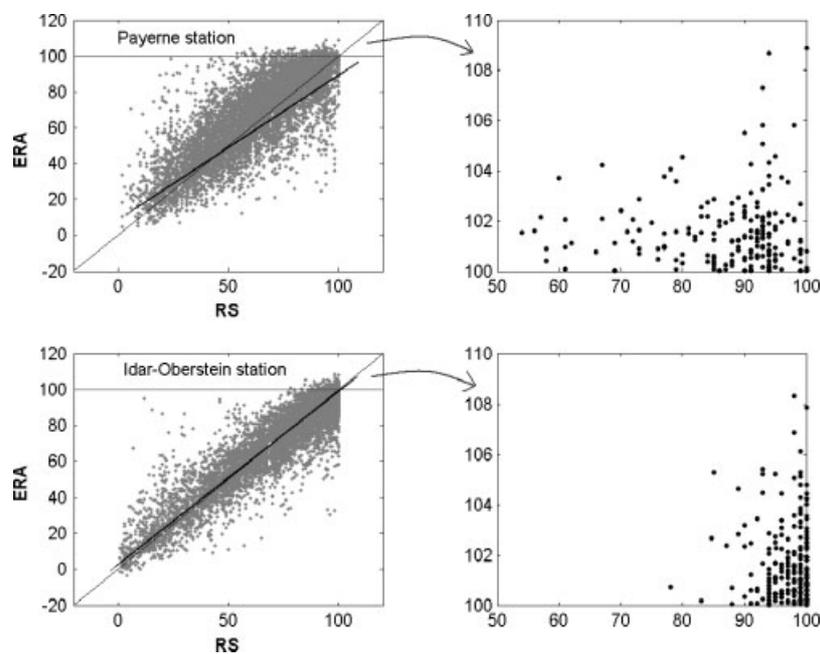


Figure 5. Consistency between radiosonde RH850 and ERA RH850 data for all available values (on left) and only considering ERA positive anomalous values (on right). The linear regression is plotted on the left (bold line). The graphs present data for the Payerne station (top) and Idar-Oberstein station (bottom). Units are in %.

Nîmes), while we detect a weak tendency for R^2 to be higher when stations and grid points are close together (Figure 4). The only station that seems atypical is Payerne, which is located in a mountainous area of Switzerland. The reason could be that the grid-cells in this region remain too coarse to describe the contrasted relief, in spite of the high ERA resolution of 1.125° .

Figure 5 illustrates the consistency between radiosonde and ERA data at nearby grid points for the stations of Payerne (top) and Idar-Oberstein (bottom), where we observe the highest proportions of ERA anomalous values. For each station, the relation between the two datasets is plotted on the left, whereas only points corresponding to positive anomalous values are taken into account on the right.

The radio sounding data appear to be correctly assimilated by the model, since the points are evenly

scattered around the one-to-one line and the linear regressions deviate only slightly from the bisector. If we focus on positive anomalies, all values measured by radiosonde are higher than 50%; nevertheless, the graphs exhibit two different behaviours. For ERA values lying between 100 and 102%, the Payerne station values are included within a large interval, between 55 and 100%. On the other hand, this is not the case for the Idar-Oberstein station, where all corresponding values are higher than 78%, with most values being higher than 90%.

The six other stations represent intermediate cases in terms of the relation between ERA RH850 anomalous values and radio sounding data. Globally, we fail to detect any strong correlation between the variables; hence, we cannot expect a straightforward

correction of the biases produced by ERA-40 from radio sounding values.

4. Discussion and conclusions

We compare the RH850 extracted from the ERA-40 and the NCEP/NCAR datasets at 2.5° resolution over a domain centred on western Europe (60°W–60°E, 15°N–75°N) by computing deviations averaged over the period from 1st September 1957 to 31st August 2002.

The two archives yield similar values for most of the continental parts of the study area. In these areas, both ERA and NNR are constrained by the same information derived from extended and dense observation networks. This is not the case for the oceanic domains, where values are more likely influenced by the physics introduced into the atmospheric model and the method used for extrapolation to unmonitored locations. We detect some strong discrepancies between the archives, even on the continent: in the ERA dataset, few values of RH850 are lower than 0%, but numerous values are higher than 100%.

No rule can be found to account for the location of these anomalous values, since the distribution is relatively homogeneous. Thus, even in well-monitored regions such as western Europe, the model still produces anomalous values. The results remain similar, irrespective of the resolution. Therefore, our analysis suggests that ERA RH850 data should be used with prudence in view of the large number of anomalous values.

To correct these anomalous values at 1.125° resolution, one solution would consist of using nearby radio sounding stations for grid points located on continental areas. Unfortunately, over the test region (eastern France), we are unable to detect any strong correlation between ERA RH850 and RH850 radiosonde data, which inhibits a possible correction using these sources. Another solution for handling normal RH850 data would involve truncating or compressing the time-series to take fully into account the natural range of relative humidity. In parallel with these approaches, we need to investigate the consequences of the corrections in terms of consistency with other variables.

The main remaining question is: what are the reasons for such unrealistic values? Are they due to the model physics? We require additional investigations, including, for example, the recalculation of RH850 using the 850-hPa specific humidity, or the analysis of relative humidity on other pressure levels. If we cannot find a way of correcting these data, the solution of this issue could be provided by the ECMWF through a new generation of re-analysis called ERA-Interim. This new system will be run for the data-rich 1990s and 2000s, and will be extended by the 2010s (Uppala, 2006).

Acknowledgements

The NCEP/NCAR re-analysis data were downloaded from the NOAA web site (<http://www.cdc.noaa.gov/>). The ECMWF ERA-40 re-analysis data were obtained from the ECMWF data server (http://data.ecmwf.int/data/d/era40_daily/). The authors are grateful to Peter Bechtold (ECMWF) for his helpful comments, Etienne Leblois (Cemagref) and Angélique Godart (LTHE) for the extraction of radio sounding data. M.S.N. Carpenter post-edited the English style.

References

- Bontron G, Obled CH. 2003. New developments in quantitative precipitation forecasts by analog sorting techniques. In *Proceedings of the 5th International EGS Plinius Conference on Mediterranean Storms*, Ajaccio 1–3 Oct. 2003, ed. by EGS.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne Roy, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
- Uppala S. 2006. From ERA-15 to ERA-40 and ERA-Interim. ECMWF/GEO Workshop on Atmospheric Reanalysis, 17–21.
- Uppala SM, Kallberg PW, Simmons AJ, Andrae U, Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hòlm E, Hoskins BJ, Isaksen I, Janssen PAEM, Jenne R, McNally AP, Mahfouf J-F, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131**: 2961–3012, DOI: 10.1256/qj.04.176.

Annexe 3

Type de communication	Article
Titre	La prévision des précipitations par recherche d'analogues : état de l'art et perspectives
Auteurs	Ben Daoud A., Sauquet E., Lang M., Obled Ch. et Bontron G.
Année	2009
Source	La Houille Blanche

La prévision des précipitations par recherche d'analogues : état de l'art et perspectives

*Precipitation forecasting through an analog sorting technique : state of the art
and further investigations*

AURÉLIEN BEN DAOUD, ERIC SAUQUET, MICHEL LANG

Cemagref – Unité de Recherche Hydrologie-Hydraulique
3 bis Quai Chauveau, CP 220, 69336 Lyon, Cedex 9, France
Tél : +33 (0)4 72 20 87 64, Fax : +33 (0)4 78 47 78 75, e-mail : aurelien.bendaoud@cemagref.fr

CHARLES OBLED

Laboratoire d'étude des Transferts en Hydrologie et Environnement
Institut National Polytechnique de Grenoble, B.P. 53, 38041 Grenoble, Cedex 9, France
Tél : +33 (0)4 76 82 50 53, Fax : +33 (0)4 76 82 50 14, e-mail : charles.obled@hmg.inpg.fr

GUILLAUME BONTRON

Compagnie Nationale du Rhône
Direction de l'Ingénierie – Eau et Environnement, 2 rue André Bonin, 69316 Lyon, Cedex 04, France
Tél : +33 (0)4 72 00 61 93, Fax : +33 (0)4 72 10 66 62, e-mail : G.Bontron@cnr.tm.fr

*L*a prévision des précipitations par analogie, adaptée des sorties de modèles numériques de prévision, s'est améliorée ces dix dernières années et est actuellement implémentée opérationnellement dans diverses infrastructures françaises. Le premier objectif de cette étude est de dresser l'état de l'art de cette approche. L'application de cette méthode nécessite une base de données contenant les champs des variables qui permettent de caractériser les situations météorologiques passées. Etant donné qu'il existe deux archives de génération différente (réanalyses ERA-40 et NCEP/NCAR), la sensibilité de cette méthode de prévision au choix de l'archive a été étudiée. Les résultats révèlent une faible sensibilité, même si de manière générale les performances sont légèrement supérieures avec l'archive ERA-40, notamment pour les événements pluvieux extrêmes. Enfin, des perspectives d'amélioration de la méthode susceptibles d'être exploitées à court terme sont évoquées.

*P*recipitation forecasting based on an adaptation of model outputs through an analog sorting technique has been improved for around ten years. The method runs operationally in several French institutions. First, this short paper presents the state of the art of this approach and the more recent developments. Second, a sensitivity analysis to the choice of the database from which the variables that characterise the past meteorological situations are extracted is performed. Two available archives are tested (ERA-40 and NCEP/NCAR re-analyses). The results show that despite the performances obtained with the ERA-40 database are slightly better, especially for heavy rainfall events, the sensitivity is weak. Finally, further ways for improvement that could be investigated are suggested.

I ■ INTRODUCTION

Malgré le progrès constant des modèles numériques de prévision du temps (MNP), la quantité de précipitation reste une variable extrêmement difficile à prévoir plusieurs jours à l'avance. En effet, la description exhaustive des processus à l'origine de la pluie est loin d'être connue, puisque de multiples paramètres entrent en jeu dans la formation des nuages et des précipitations (quantité de vapeur d'eau présente dans l'atmosphère, niveaux de saturation et de condensation, vents verticaux, relief, etc.).

Avec la constante augmentation de la résolution des MNP, la prévision des précipitations est généralement fiable jusqu'à 48 heures à l'avance. Néanmoins pour certaines situations météorologiques, il arrive encore que la prévision fournie par les MNP soit erronée moins de 24 heures avant l'occurrence d'un événement pluvieux.

En revanche, d'autres variables sont mieux prévues par les MNP puisque leur variabilité dans le temps et dans l'espace est décrite par des équations robustes. C'est le cas de la pression atmosphérique, variable clé pour réaliser des prévisions. Les champs de pression atmosphérique permet-

tent en effet de caractériser la circulation atmosphérique à grande échelle et leur analyse est sans doute la première étape du prévisionniste en météorologie. Avec l'expérience, le prévisionniste associe les champs de pression prévus pour les journées à venir à des situations qu'il a déjà rencontrées auparavant. S'il se souvient du temps qu'il a fait lors de ces journées passées au-dessus de la zone de prévision, il aura appliqué de manière instinctive le principe d'analogie entre situations météorologiques.

Plus formellement, le concept d'analogie en météorologie a été introduit par Lorenz [1]. Il repose sur l'hypothèse que deux situations synoptiques analogues doivent engendrer des effets locaux du même ordre. Il peut être utilisé pour corriger les biais des MNP ([2], [3]), étudier les caractéristiques météorologiques d'événements particuliers, ou encore émettre des prévisions probabilistes de précipitations. Cette méthode de prévision des pluies, assimilée à une technique d'adaptation statistique des sorties de MNP et communément appelée « méthode des analogues », s'appuie notamment sur une archive de variables d'analogie (prédicteurs), c'est-à-dire les variables qui décrivent la situation météorologique et à partir desquelles la similitude entre la situation prévue et une situation passée est évaluée. Depuis quelques années, les réanalyses météorologiques sont utilisées comme archives de variables d'analogie.

Nous dresserons tout d'abord l'état de l'art de la méthode des analogues en donnant l'historique puis en décrivant la dernière version développée par Bontron [4]. Nous évoquons quelques résultats d'un travail de thèse en cours portant sur la sensibilité de la méthode aux choix de l'archive de réanalyses météorologiques. Enfin, nous évoquerons les perspectives d'amélioration de la méthode qui s'offrent à la prévision opérationnelle des précipitations adaptée par analogie.

II ■ LA MÉTHODE DE PRÉVISION DES PRÉCIPITATIONS ADAPTÉE PAR ANALOGIE : ÉTAT DE L'ART

● II.1 BREF HISTORIQUE

Duband [5] a initié les travaux sur la prévision des précipitations par analogie en France et ses développements ont conduit Electricité De France (EDF) à implanter un système de prévision opérationnel au début des années 1970. Ce système utilisait les champs de pression observés le jour J et une extrapolation temporelle était effectuée à partir de ces champs prédicteurs pour les jours à venir. Ces champs n'étaient donc pas encore fournis par des MNP. Avec le progrès de ces derniers, la méthode a évolué dans les années 1980 [6] en devenant une méthode d'adaptation statistique des MNP, c'est-à-dire que les champs des variables sur lesquels l'analogie est recherchée étaient désormais prévus par des MNP. La méthode s'est développée ensuite jusque dans les années 1990, avec l'amélioration des prévisions de variables synoptiques. L'échéance de prévision a alors été repoussée à quatre jours. Guilbaud [7] a introduit un nouveau critère d'analogie pour la sélection d'analogues à l'aide

de variables définies sur une grille régulière [8]. A partir de ces travaux, Bontron [4] a amélioré les performances de la méthode en hiérarchisant la sélection des situations analogues sur deux niveaux et en utilisant une nouvelle archive de variables d'analogie : les réanalyses météorologiques. Cette mise à jour sera considérée comme un état de référence. Les différentes méthodes ont été implémentées opérationnellement à EDF, à la Compagnie Nationale du Rhône (CNR) en 2005 et dans plusieurs laboratoires. Parallèlement, le concept d'analogie pour la prévision des précipitations a été appliqué dans d'autres pays ([9], [10], [11], [12]).

● II.2 DESCRIPTION DE LA MÉTHODE DE RÉFÉRENCE

De manière générale, la recherche d'analogie repose sur le choix de plusieurs paramètres :

- la variable à prévoir et l'archive historique correspondante (pluviométrique dans notre cas) ;
- les variables d'analogie qui permettent de caractériser l'état de l'atmosphère et l'archive historique correspondante ;
- le critère qui permet d'identifier une situation analogue ;
- le domaine spatial sur lequel on recherche l'analogie ;
- le nombre de situations analogues à sélectionner.

La méthode dite « de référence », suggérée par Bontron [4], forme le point de départ de travaux de recherche, menés conjointement par la CNR et le Cemagref depuis novembre 2006, en concertation avec le Laboratoire d'étude des Transferts en Hydrologie et Environnement. Ces travaux visent à améliorer la prévision probabiliste de précipitations journalières. Dans la méthode de référence, la sélection de situations analogues à la situation du jour cible (jour C) s'effectue en deux étapes. Les variables, domaines, critères et nombre d'analogues à sélectionner pour chaque niveau d'analogie sont résumés dans le *Tableau 1*.

Tableau 1 : Paramètres de la sélection de situations analogues à chaque niveau d'analogie [4]

	Niveau 1	Niveau 2
Variables d'analogie	Géopotentiels à 1 000 hPa et 500 hPa	Humidité relative à 850 hPa et colonne totale d'eau
Domaine	Echelle synoptique à optimiser	Echelle locale à optimiser
Critère	Score de Teweles-Wobus (1954)	Erreur quadratique moyenne
Nombre d'analogues	A optimiser	A optimiser

La première sélection (niveau 1) permet de constituer un échantillon de situations similaires à la situation cible à échelle synoptique, en termes de circulation générale. A l'issue de cette étape, N1 situations sont alors retenues. La seconde sélection (niveau 2) permet de ne retenir, parmi les

situations de l'échantillon du niveau 1, que les situations analogues à la situation du jour C à une échelle plus locale, en termes d'humidité. N2 situations sont finalement retenues comme analogues. Les nombres N1 et N2, fonction notamment de la longueur de l'archive de variables d'analogie, ainsi que les domaines d'analogie doivent être optimisés pour la station ou le bassin versant d'intérêt. Finalement, l'échantillon de dates analogues permet, à l'aide des pluies observées au cours de ces journées, de construire la distribution empirique des pluies prévue pour le jour C.

III ■ SENSIBILITÉ DE LA MÉTHODE DE PRÉVISION AU CHOIX DE L'ARCHIVE DE VARIABLES D'ANALOGIE

L'archive de prédicteurs utilisée par Bontron [4], issue des réanalyses NCEP/NCAR [13] à résolution (2.5° x 2.5°), est apparue comme plus homogène que celles disponibles auparavant, lesquelles résultaient de l'archivage au fil du temps des analyses produites par les modèles opérationnels de prévision. Cette archive présente également l'intérêt d'inclure de nombreuses variables météorologiques, autres que le géopotential, telles que celles utilisées lors du second niveau de la sélection des situations analogues.

Aujourd'hui, deux archives sont disponibles sur des périodes suffisamment longues avec une résolution de 2,5° : les réanalyses ERA-40 (ERA) [14] et les réanalyses NCEP/NCAR (NNR) [13]. En préalable au travail sur la méthode en elle-même, les variables d'analogie extraites de ces bases de données ont été comparées afin d'apprécier les différences entre les deux réanalyses [15]. Des valeurs aberrantes d'humidité relative à 850 hPa ont été détectées dans l'archive ERA. Certaines valeurs sont inférieures à 0 % et une proportion non négligeable est supérieure à 100 %. Une tentative de correction de ces valeurs a ensuite été menée mais elle n'a pas été concluante.

Une analyse de la sensibilité de la méthode de référence au choix de l'archive de variables d'analogie a été effectuée sur le bassin versant de la Saône, dont les pluies de bassin ont été extraites des réanalyses SAFRAN [16]. Cette analyse a été réalisée dans un contexte de prévision parfaite, c'est-à-dire que les situations pour lesquelles on émet une prévision ne sont pas prévues par un MNP, mais tirées des archives de réanalyses. La période de situations cibles et candidates s'étend du 01/09/1972 au 30/08/2002 (soit 10 987 dates).

● III.1 OPTIMISATION DES NOMBRES D'ANALOGUES À SÉLECTIONNER À L'AIDE DU CRPS

La sensibilité de la méthode au choix de l'archive de variables d'analogie est d'abord étudiée en optimisant les nombres N1 et N2 d'analogues sélectionnées aux niveaux 1 et 2. Ils sont optimisés à l'aide du score Continuous Ranked Probability Score (CRPS) [17]. La performance de prévision est finalement exprimée en normalisant le CRPS par rapport à la climatologie. Le score résultant, nommé RSS (Relative

Skill Score) (1), a pour valeur 1 pour une prévision parfaite, 0 pour une prévision équivalente à la prévision climatologique, et est négatif pour une prévision de moins bonne qualité que la prévision climatologique :

$$RSS = \frac{CRPS_M - CRPS_{c\lim}}{CRPS_{parf} - CRPS_{c\lim}} \quad (1)$$

où $CRPS_M$ est le score calculé avec la méthode à évaluer, $CRPS_{c\lim}$ est le score calculé avec la prévision climatologique, et $CRPS_{parf}$ est le score parfait (valeur nulle).

Ce score a été calculé pour chaque couple (N1, N2). Le nombre N1 optimal est de 70 (resp. 80) lorsque l'archive ERA (resp. NNR) est utilisée. La Figure 1 montre l'évolution du RSS en fonction de N2 pour le nombre N1 optimal associé à chaque archive (ERA N2 et NNR N2). L'évolution du RSS en fonction de N1 (niveau 1 uniquement appliqué) a également été tracée pour chacune des archives (ERA N1 et NNR N1).

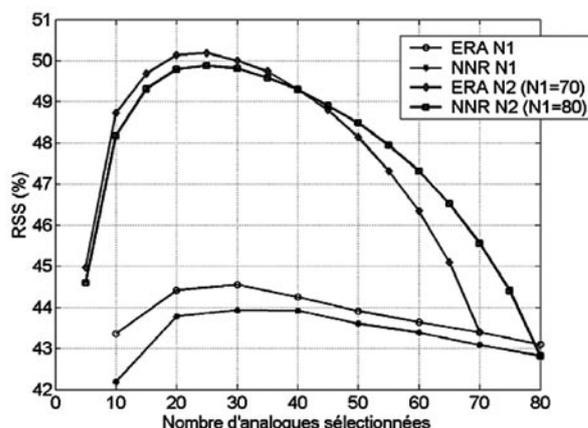


Figure 1 : Evolution du score RSS en fonction du nombre d'analogues sélectionnées : au niveau 1 seul (ERA N1 et NNR N1) et au niveau 2 lorsque N1 est égal au nombre N1 optimal associé à chaque archive (ERA N2 et NNR N2).

Bien que l'archive ERA contienne des valeurs irréalistes d'humidité, les performances de prévision semblent équivalentes avec les deux archives. Contrairement à ce que l'on aurait pu craindre, les performances au niveau 2 sont même légèrement supérieures avec l'archive ERA. Ce faible gain de performance avec ERA est également obtenu à l'issue du premier niveau d'analogie. Par ailleurs, le niveau 2 d'analogie est bien justifié du fait d'un gain de performances non négligeable apporté par son application (de l'ordre de 5 points de RSS). Ainsi, avec les archives utilisées et dans le cadre de l'application au bassin de la Saône, le nombre d'analogues à sélectionner au second niveau est de 25.

● III.2 EVALUATION DE LA QUALITÉ DE LA PRÉVISION À L'AIDE DU DIAGRAMME ROC

Le diagramme ROC (Receiver Operating Characteristic) est utilisé en médecine mais aussi depuis quelques années en météorologie [18]. Ce diagramme trace, pour un seuil S – ici de pluie – fixé, la probabilité de détection en fonction du taux de fausses détections. La sensibilité de la méthode au choix de l'archive de réanalyses a été étudiée à l'aide de ce diagramme. Pour un seuil donné, les courbes ROC associées aux archives ERA et NNR ont été tracées. En raison du très faible écart entre les deux courbes, nous avons décidé de comparer les aires estimées sous les courbes ROC (score AROC). Ce score vaut 1 pour une prévision parfaite, 0,5 pour une prévision équivalente à la simple prévision climatologique ; et s'il est inférieur à 0,5, la prévision s'avère de moins bonne qualité que la climatologie.

Cinq seuils de pluie ont été considérés : tous les cumuls de pluie supérieurs à 0 mm, et tous ceux supérieurs aux quantiles de pluie 50, 70, 90 et 99 %. Le *Tableau 2* présente les scores AROC obtenus pour chaque seuil lorsque chacune des archives est utilisée.

Tableau 2 : Valeurs du score AROC atteint pour différents seuils de pluie lorsque chacune des archives est utilisée

	0 mm	P50	P70	P90	P99
ERA	0,932	0,936	0,930	0,922	0,895
NNR	0,930	0,935	0,930	0,922	0,888

Aucune différence notable n'est observée entre les scores. Cependant, les performances sont légèrement supérieures lorsque l'archive ERA est utilisée, notamment pour les pluies extrêmes (supérieures au quantile P99). L'avantage de ce score est aussi de pouvoir discriminer la prévision entre deux événements : on remarque, lorsque les deux archives sont utilisées, que le score maximal est atteint pour le seuil P50, c'est-à-dire pour les situations dont les cumuls pluviométriques sont compris entre les quantiles P50 et P70 (*i.e.* pour les journées moyennement pluvieuses).

Supposons à présent qu'en prévision opérationnelle nous disposons des sorties d'un MNP différent du modèle utilisé pour les réanalyses. Pourrait-on chercher dans ces réanalyses des situations analogues à la situation prévue par le MNP sans que les performances ne soient diminuées ? Bien que nous soyons placés dans un contexte de prévision parfaite, nous nous sommes intéressés à l'impact sur les performances de l'utilisation d'une archive, dans laquelle on extrait les situations cibles, différente de celle dans laquelle on extrait les situations analogues. Nous avons ainsi envisagé les quatre cas suivants : situations cibles extraites d'ERA et analogues extraites d'ERA (EE), situations cibles extraites de NNR et analogues extraites de NNR (NN), situations cibles extraites d'ERA et analogues extraites de NNR (EN), situations cibles extraites de NNR et analogues extraites d'ERA

(NE). Le *Tableau 3* présente, en considérant chacun des cas, les scores AROC obtenus pour les seuils de pluie considérés précédemment.

Tableau 3 : Valeurs du score AROC atteint pour différents seuils de pluie, en considérant chacun des cas (XY : situations cibles extraites de X et analogues extraites de Y)

	0 mm	P50	P70	P90	P99
EE	0,932	0,936	0,930	0,922	0,895
NN	0,930	0,935	0,930	0,922	0,888
EN	0,929	0,933	0,927	0,922	0,891
NE	0,927	0,932	0,927	0,923	0,897

Les résultats montrent qu'en général les performances, bien qu'équivalentes, diminuent légèrement si deux archives différentes sont utilisées. Un gain de performance est cependant obtenu pour les pluies les plus importantes. Ce résultat doit être relativisé compte tenu du faible nombre de valeurs de l'échantillon (110 valeurs). Ainsi si l'on dispose des sorties de plusieurs modèles, cette perspective en mode opérationnel s'avérerait intéressante dans certaines conditions, par exemple si les sorties de l'un d'entre eux ne sont pas disponibles. Il reste tout de même à effectuer une série de tests via les sorties archivées des MNP opérationnels.

IV ■ PERSPECTIVES D'AMÉLIORATION

La partie précédente montre que l'utilisation des deux archives NNR et ERA conduit, dans un contexte de prévision parfaite, à des résultats équivalents pour une résolution spatiale de 2,5°. La prochaine étape est d'exploiter l'archive européenne ERA à sa résolution initiale de 1,125°, ce qui permet d'espérer une meilleure prise en compte des conditions locales lors de l'application du niveau 2. Néanmoins, les premiers résultats montrent que le gain de performance n'est pas significatif.

D'autre part, l'évaluation ne peut avoir un sens que si la prévision se fait dans un contexte opérationnel. Par conséquent, dans le but d'achever l'étude de la sensibilité de la méthode aux archives, des prévisions archivées produites par les modèles GFS et ECMWF seront utilisées pour vérifier l'efficacité de la méthode et pour évaluer la réelle aptitude des modèles à prévoir l'évolution des champs des variables d'analogie. Ce travail a déjà débuté dans le cadre d'une étude exploratoire appliquée à la Seine [19].

Parallèlement à ces tests, d'autres perspectives d'amélioration sont envisagées. Actuellement, pour la prévision d'un jour cible, les situations candidates au premier niveau de sélection sont automatiquement sélectionnées dans une fenêtre temporelle de plus ou moins deux mois autour du jour cible. Il n'y a donc aucune souplesse dans cette pré-sélection. Pour remédier à ce problème, il est envisagé de sélectionner les situations candidates selon leur température. Ainsi, un

certain nombre de situations, proches de la situation cible d'un point de vue thermique, pourrait être retenu. Une autre perspective envisagée est de détecter automatiquement les zones frontales, responsables des épisodes pluvieux sur des grands bassins peu montagneux comme les bassins tests de la Saône et de la Seine. Deux variables a priori pertinentes seront testées : la température potentielle équivalente, qui permet de détecter les changements de masse d'air et les zones d'instabilité atmosphérique ; et la vitesse verticale, qui permet de détecter les zones d'ascendance, responsables des formations nuageuses frontales. L'une de ces variables pourrait ainsi être introduite entre le premier niveau (circulation générale à échelle synoptique) et le second niveau (état hygrométrique de l'atmosphère à échelle locale).

Par ailleurs, la distribution des pluies fournies par les analogues est fondée sur l'équiprobabilité entre les situations. Or, une situation très analogue au niveau 1 ou au niveau 2 pourrait être affectée d'un poids plus important afin de tenir compte de la qualité d'analogie. Une méthode de pondération par les scores obtenus aux niveaux 1 ou 2 pourrait être ainsi envisagée.

Enfin, la méthode de prévision par analogie est utilisée en France principalement sur les bassins à crue rapide à horizon de trois à cinq jours. L'adaptation de la méthode aux bassins à crue lente, sur les bassins tests de la Saône et de la Seine, nécessite de reconsidérer le cumul pluviométrique à prévoir, ce dans le but d'anticiper les événements jusqu'à sept à dix jours à l'avance. Actuellement testée à un pas de temps journalier, il est également prévu de l'adapter à des pas de temps supérieurs pour ce type de bassin.

V ■ CONCLUSIONS

La méthode de prévision des précipitations adaptée par analogie s'est développée ces dernières années en France et a été implémentée dans plusieurs infrastructures. La dernière version, qui décompose la sélection d'analogues en deux niveaux, s'appuie sur une archive de variables d'analogie. Deux archives de réanalyses (ERA-40 et NCEP/NCAR) sont actuellement disponibles à résolution 2,5° sur des périodes suffisamment longues (de l'ordre d'une quarantaine d'années). La sensibilité de la méthode au choix de l'archive a été étudiée. Les résultats montrent que cette sensibilité est faible ; même si, pour les événements pluvieux extrêmes, les performances de la méthode sont supérieures lorsque l'archive ERA-40 est utilisée.

Plusieurs perspectives d'amélioration de la méthode ont été mentionnées et seront explorées à court terme.

Dans le cas de la prévision des précipitations sur des bassins à crue lente, il faudra adapter la méthode de prévision à l'échelle du temps de réaction de ces bassins, qui est de l'ordre de quelques jours.

VI ■ RÉFÉRENCES

[1] E. N. LORENZ (1969) — Atmospheric predictability as revealed by natural occurring analogues. *J. Atmos. Sci.* **26** 636-646

[2] G. M. CARTER, J. P. DALLAVALLE, H. R. GLAHN (1989) — Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Weather Forecasting*. **4** 401-412

[3] T. M. HAMILL, J. S. WHITAKER, S. L. MULLEN (2006) — Reforecasts : An important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.* **87** 33-46

[4] G. BONTRON (2004) — *Prévision quantitative des précipitations : adaptation probabiliste par recherche d'analogues. Utilisation des réanalyses NCEP-NCAR et application aux précipitations du Sud-Est de la France. Thèse de Doctorat, Institut National Polytechnique de Grenoble.*

[5] D. DUBAND (1970) — *Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. Thèse de 3^e cycle de la Faculté des Sciences de Paris.*

[6] D. DUBAND (1980) — Dynamic selection of analogue flow patterns to produce quantitative precipitation forecasts. *WMO Symposium on probabilistic and statistical method in weather forecasting, Nice, France, 8-12 septembre.*

[7] S. GUILBAUD (1997) — *Prévision quantitative des précipitations journalières par une méthode statistico-dynamique de recherche d'analogues. Thèse de Doctorat, Institut National Polytechnique de Grenoble.*

[8] CH. OBLED, G. BONTRON, R. GARCON (2002) — Quantitative precipitation forecasts : a statistical adaptation of model outputs through an analogues sorting approach. *J. of Atmos. Res.* **63** 303-324

[9] F. WETTERHALL, S. HALLDIN, C-Y. XU (2004) — Statistical precipitation downscaling in central Sweden with the analogue method. *J. Hydrol.* **306** 174-190

[10] C. MATULLA, X. ZHANG, X. WANG, J. WANG, E. ZORITA, S. WAGNER, H. VON STORCH (2007) — Influence of similarity measures on the performance of the analog method for downscaling daily precipitation. *Clim. Dyn.* **30** 133-144

[11] V. ALTAVA-ORTIZ, A. BARRERA, M. C. LLASAT, M. A. PRAT, J. GIBERGANS-BÀGUENA, M. BARNOLAS (2006) — Application of the MM5 and the analogous method to heavy rainfall event, the case of 16-18 October 2003 in Catalonia (Spain). *Adv. in Geosc.* **7** 313-319

[12] T. DIOMEDE, F. NEROZZI, T. PACCAGNELLA, E. TODINI (2008) — The use of meteorological analogues to account for LAM QPF uncertainty. *Hydrol. Earth Syst. Sci.* **12** 141-157

[13] E. KALNAY, M. KANAMITSU, R. KISTLER, W. COLLINS, D. DEAVEN, L. GANDIN, M. IREDELL, S. SAHA, G. WHITE, J. WOOLLEN, Y. ZHU, M. CHELLIAH, W. EBISUZAKI, W. HIGGINS, J. JANOWIAK, K.C. MO, C. ROPELEWSKI, J. WANG, A. LEETMAA, R. REYNOLDS, R. JENNE, D. JOSEPH (1996) — The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Meteorol. Soc.* **77** 437-471

[14] S.M. UPPALA, P.W. KALLBERG, A.J. SIMMONS, U. ANDRAE, V. DA COSTA BECHTOLD, M. FIORINO, J.K. GIBSON, J. HASELER, A. HERNANDEZ, GA. KELLY, X. LI, K. ONOGI, S. SAARINEN, N. SOKKA, R.P. ALLAN, E. ANDERSSON, K. ARPE, M.A. BALMASEDA, A.C.M. BELJAARS, L. VAN DE BERG, J. BIDLOT, N. BORMANN, S. CAIRES, F. CHEVALLIER, A. DETHOF, M. DRAGOSAVAC, M. FISHER, M. FUENTES, S. HAGEMANN, E. HÖLM, B.J. HOSKINS, L. ISAKSEN, P.A.E.M. JANSSEN, R. JENNE, A.P. MCNALLY, J-F. MAHFOUF, J-J. MORCRETTE, N.A. RAYNER, R.W. SAUNDERS, P. SIMON, A. STERL, K.E. TRENBERTH, A. UNTCH,

- D. VASILJEVIC, P. VITERBO, J. WOOLLEN (2005) — The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **131** 2961-3012
- [15] A. BEN DAOUD, CH. OBLED, M. LANG, E. SAUQUET, G. BONTRON (2008) — Comparison of 850-hPa relative humidity between ERA-40 and NCEP/NCAR re-analyses : detection of suspicious data in ERA. *Atmos. Sci. Letters (soumis pour publication)*. 40
- [16] P. QUINTANA-SEGUI, P. LE MOIGNE, Y. DURAND, E. MARTIN, F. HABETS, M. BAILLON, C. CANELLAS, L. FRANCHISTEGUY, S. MOREL (2007) — Analysis of near-surface atmospheric variables : Validation of the SAFRAN analysis over France. *J. Appl. Meteor. Climatol.* **47** 92-107
- [17] H. HERBACH (2000) — Decomposition of the ranked probability score for ensemble prediction systems. *Weather Forecasting.* **15** 559-570
- [18] I. T. JOLLIFFE, D. B. STEPHENSON (2003) — *Forecast verification. A practitioner's guide in atmospheric science.* John Wiley & Sons Ed.
- [19] J. PLASSE, A. BEN DAOUD (2008) — Application au bassin de la Seine amont d'une méthode de prévision des pluies adaptée par analogie : de la prévision de pluie à la prévision de débit. *Rapport technique Cemagref.*

Annexe 4

Type de communication	Communication orale / extrait d'actes
Titre	A flood forecasting system based on probabilistic precipitation scenarios: performance evaluation and sensitivity analysis to the re-analysis archive
Auteurs	Ben Daoud A., Sauquet E., Obled Ch., Lang M. and Bontron G.
Année	2008
Source	International Conference HydroPredict'2008

A flood forecasting system based on probabilistic precipitation scenarios: performance evaluation and sensitivity analysis to the re-analysis archive

A. BEN DAOUD¹, E. SAUQUET¹, C. OBLED², M. LANG¹ & G. BONTRON³

¹ Cemagref, Hydrology - Hydraulics Research Unit, 3 bis quai Chauveau CP220, F-69336 Lyon Cedex 09, France
aurelien.bendaoud@cemagref.fr

² Laboratoire d'Etude des Transferts en Hydrologie et Environnement, Institut National Polytechnique de Grenoble, BP 53, F-38041 Grenoble Cedex 9, France

³ Compagnie Nationale du Rhône, Direction de l'Ingénierie – Eau et Environnement, 2 rue André Bonin, 69316 Lyon Cedex 4

Abstract A probabilistic quantitative precipitation forecasting system is applied to the Saone river basin in order to predict ensemble discharge at one station. This method, strictly based on the meteorological analogy concept, depends on the input data. The sensitivity of the precipitation forecasting system to the input database is assessed using two different re-analysis archives. This sensitivity analysis shows in average, weak performance discrepancies.

Key words QPF; performance evaluation; ERA-40; NCEP/NCAR; flood forecasting

INTRODUCTION

Flood events may be anticipated by using quantitative precipitation forecasts (QPFs) introduced thereafter in a rainfall-runoff model to simulate discharges. The QPF method considered here is based on the analog sorting technique described by (Bontron and Obled, 2003). A first selection (level 1) of analogous situations is performed in terms of synoptic circulation is identified, and from this sample a subset of the most similar patterns in terms of hygrometric situation (level 2) is extracted. The predictors involved in the similarity criterion are extracted from a long meteorological archive and results may strongly depend on the choice of the meteorological database. The aim of this presentation is to assess the sensitivity of the QPF system to two different re-analysis archives: the first National Centers for Environmental Prediction/National Center for Atmospheric Research re-analysis (NRR) and the 45-Year European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA). The study area covers the Saone river basin in France. The predictands are the daily rainfall amounts for three sub-basins. Finally for illustration, a flood forecast is obtained from the GR4J rainfall-runoff model for a river basin within the study area.

SENSITIVITY ANALYSIS OF THE QPF SYSTEM TO THE RE-ANALYSIS ARCHIVE

This sensitivity analysis to the re-analysis archives is carried out in a context of perfect prognosis, that is the evaluation is based on a sequence of observations, not forecasted by a model in real time conditions. Numerous appropriate scores to evaluate probabilistic forecast performances can be found in the literature (for a complete list, see Jolliffe and Stefenson (2003)). Among them, the Continuous Ranked Probability Score (CRPS) (Hersbach, 2000) seems well suited for probabilistic forecast

verification. However, methods may lead to very similar CRPS values. Thus, the relative skill score (RSS) is used instead:

$$RSS(M) = \frac{CRPS(M) - CRPS(M_{ref})}{CRPS_{perf} - CRPS(M_{ref})} \quad (1)$$

where M is the method to evaluate, M_{ref} is a reference forecasting method, which prediction is known whatever the day of prediction, and $CRPS_{perf}$ is the $CRPS$ of a perfect forecast (*i.e.* a value of 0). Here, forecasts issued from M_{ref} are given by the climatological distribution of the predictand. The RSS score is a function of the numbers of selected situations at level 1 ($N1$) and 2 ($N2$). The performance is evaluated by testing combinations ($N1, N2$) with $N1$ between 10 and 100 and $N2$ ranging from 5 to $N1$.

The RSS score is plotted against $N2$ with fixed values for $N1$ on Fig. 1 (ERA L2 and NNR L2). Two additional curves are displayed, corresponding to the RSS score considering only level 1 as criterion of selection (ERA L1 and NNR L1).

Lastly, the optimum is reached at level 1 for $N1=70$ when using ERA database and $N1=80$ when using NNR, and for $N2 = 25$ with both archives. These values correspond to the maximums of ERA L2 and NNR L2 curves in Fig. 1. The analog method (AM) performs better when using ERA instead of NNR, since ERA provides the highest values for RSS . Small deviations between ERA L2 and NNR L2 curves suggest comparable performance around the optimal values. Lastly, the second step of the selection is well justified since the gain in RSS of about 5% measured between the maximum of the curves at level 1 and the curves at level 2 is not negligible.

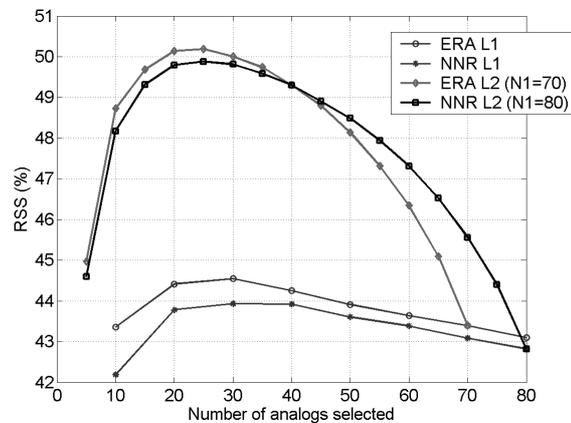


Fig. 1 Evolution of the RSS score with the number of analogs selected at level 1 (L1) and level 2 (L2), when ERA and NNR are used as predictors archive. Units are %.

The sensitivity analysis can be carried out using the receiver operating characteristics (ROC), described by Mason (1982) and Stanski *et al.* (1989), which quantifies both the false alarm rate and the hit rate. The ROCs are plotted for different thresholds of daily rainfall amount. Finally, the ROCs are summarised by the area value estimated under the curves ($AROC$). Thus the $AROC$ score lies between 0 and 1, a value of 0 corresponding to a bad prediction, a value of 1 to a perfect prediction and a value of 0.5 to a prediction equivalent to the climatological prediction. This forecasting procedure is applied over a period of 10987 days. Five thresholds are considered (Table 1): all precipitation amounts higher than 0 mm, and those higher than the quantiles 50% (P50), 70% (P70), 90% (P90) and 99% (P99) of the marginal daily rainfall distribution of the catchments.

Table 1 AROC values obtained for each threshold by using ERA and NNR.

	0mm	P50(mm)	P70(mm)	P90(mm)	P99(mm)
ERA	0.9321	0.9358	0.9301	0.9221	0.8948
NNR	0.9297	0.9348	0.9295	0.9223	0.8882

Differences between ERA and NNR in *AROC* scores are rather small. The most important deviation is detected for P99. Nevertheless, most of the highest scores are associated to ERA for all thresholds, except for P90. The AM seems to provide more reliable forecasts for heavy rainfall events when ERA is considered. Furthermore, the best scores for both archives corresponds to medium events, since the maximum of the *AROC* scores are detected for P50.

FLOOD CASE STUDY IN THE SAONE RIVER BASIN

The major event which occurred in January 1995 in the Doubs River sub-basin is one of the examples under study. The QPFs are introduced into the lumped GR4J rainfall-runoff model, running at a daily time step (Perrin et al., 2003). The GR4J model is calibrated using all the data available from 1973 to 2005. The deviation to the observations $err(t)$ is computed for each day, and linear relationships between $err(t)$ and $err(t+j)$, $j \geq 1$ day, are fitted to characterize the persistence of the model error: $err(t+j) = a_j err(t) + b_j$. For this study case, at the time of the forecast t_0 ,

- the AM is applied with data extracted from re-analyses, *i.e.* "perfect" meteorological fields to provide the empirical rainfall distribution for each day t_0+j , $j \geq 1$ day ; ten scenarios of rainfall forecasts are established by selecting randomly one value among the 25 daily rainfall amounts issued from the $N2=25$ analogous situations;
- the GR4J model is forced by the ten scenarios;

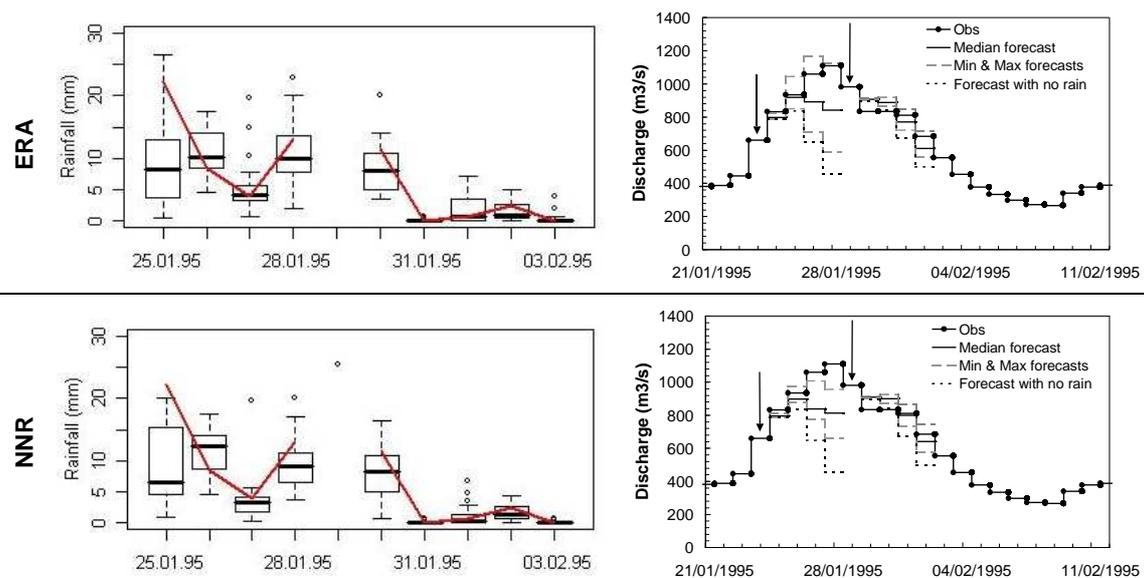


Fig. 2 Rainfall forecasts and related daily discharges for one of the major event in the Doubs river basin. Solid lines in the left panel represent observations. The empirical distribution of daily rainfall amounts provided by the AM are described by the box plots for two starting days. The arrows indicate the day of the forecast t_0 on the right

panel.

- a simple adaptive updating procedure to simulate the real-time reduction of forecast error is considered: the forecasted discharge at time t_0+j , $j \geq 1$ day is the sum of the value from GR4J run and the term $err(t_0+j) = a_j err(t_0) + b_j$.

Hindcasted discharges are compared with observed values on the right panel of Fig. 2. Here, we consider lead times below 4 days. In addition, as a reference, discharge forecasts with future rainfall equal to zero for time t_0+j , $j = 1 \dots 4$ are computed. Forecasts are established at time: $t_0 = 24/01/1995$ and $t_0 = 30/01/1995$.

On this specific event the AM fails to predict the major rainfall event of the 24/01/1995 in particular when NNR is used. ERA shows slightly better results since the observed values is within the range of the forecasted rainfall amounts for this day (Fig. 2, left). This behaviour observed on this specific event is of course consistent with results of Table 1 which demonstrate the better accuracy obtained for predicting high quantiles with ERA. On the other hand, both archives provide similar patterns for the recession curves.

CONCLUSIONS

The sensitivity of the AM to the re-analysis archive was examined in a perfect prognosis context with two different scores. Both indicators reveal small discrepancies between the use of ERA and NNR archives. Comparable forecasts are obtained for medium rainfall events but there is a slight increase in forecast skill from using ERA if heavy events are expected. We may suspect that the biases are similar for discharge forecasts. This fact has been observed in this paper on a case study.

Several ways for improving this analog method could be investigated. This approach is already attractive but new predictors and additional analogy criteria could be tested in order to increase more specifically the predictability of extreme events. This study provides preliminary results. It is not possible to conclude about the choice of the meteorological database to be used. To decide between the two re-analysis archives, the reliability of models outputs in an operational context has to be evaluated. This could be achieved applying this framework to re-forecasts.

Acknowledgements The NCEP/NCAR and ECMWF ERA-40 re-analysis data were obtained from the NOAA web site (<http://www.cdc.noaa.gov/>) and from the ECMWF data server (http://data.ecmwf.int/data/d/era40_daily/), respectively.

REFERENCES

- Bontron, G. & Obled, C. (2003) New developments in quantitative precipitation forecasts by analog sorting techniques. *Proceedings of the 5th International EGS Plinius Conference on Mediterranean Storms*. Ajaccio 1-3 Oct.
- Hersbach, H. (2000) Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather Forecasting*, **15**, 559-570.
- Jolliffe, I. T. & Stephenson D. B. (2003) Forecast verification: a Practitioner's Guide in Atmospheric Science. *Jon Wiley and sons Ed.* 239 p.
- Lorenz, E. N. (1969) Atmospheric predictability as revealed by natural occurring analogues. *J. Atmos. Sci.*, **26**, 636-646.
- Mason, I. B. (1982) A model for assessment of weather forecasts. *Austral. Met Mag.* **30**, 291-303.
- Perrin, C., Michel, C. & Andreassian, V. (2003) Improvement of a parsimonious model for streamflow simulation. *J. of Hydrology*, **279**, 275-289.
- Stanski, H. R., Wilson, L. J. & Burrows, W. R. (1989) Survey of common verification methods in meteorology. *World Weather Watch in Technical Report*, **8**, World Meteorological Organisation, Geneva.

Résumé

Cette thèse porte sur l'amélioration d'une méthode de prévision probabiliste des précipitations adaptée par analogie, appliquée aux grands bassins versants de la Saône et de la Seine. En considérant un jour cible dont on souhaite prévoir les précipitations, cette méthode permet de rechercher dans une archive les situations météorologiques passées les plus similaires. Les cumuls pluviométriques des journées analogues sont ensuite exploités pour estimer de manière probabiliste la quantité de précipitations attendue pour le jour cible. Une étape préliminaire consiste à construire et à choisir les archives nécessaires à l'application de ce type de méthode. A partir de l'algorithme de sélection d'analogues développé par Bontron (2004) sur des bassins du Sud-Est de la France, la méthode de prévision est ensuite améliorée en introduisant deux nouvelles variables : la température, qui permet de prendre en compte les effets saisonniers ; et la vitesse verticale, qui vise à mieux caractériser les mouvements atmosphériques verticaux. Finalement, la méthode est appliquée en contexte de prévision parfaite (situations cibles extraites de réanalyses) sur la plus longue période possible, puis en contexte opérationnel (situations cibles générées par un modèle de prévision) sur une durée de trois ans. Les scénarios de prévision de précipitations sont également exploités par le modèle pluie-débit GR3P pour la prévision des débits sur les bassins de la Saône et de la Seine. De manière générale, les résultats montrent qu'avec ce type d'approche, il est possible d'obtenir une prévision de débit utile, sur les grands bassins tests, pour une échéance de l'ordre de la semaine.

Mots clés : prévision probabiliste quantitative des précipitations, adaptation par analogie, réanalyses ERA-40, prévision hydrologique, bassin de la Saône, bassin de la Seine.

Abstract

This thesis aims at improving a probabilistic quantitative precipitation forecasting method based on analogs, formerly developed on small basins located in South-Eastern France (Bontron, 2004). The study area covers two French large basins of the Saône and the Seine. For a given target situation, this method consists in searching for the most similar situations observed in a historical meteorological archive. Precipitation amounts observed during analogous situations are then collected to derive the empirical predictive distribution function, i.e. the probabilistic estimation of the precipitation amount expected for the target day. A preliminary step is to build the databases necessary to apply such a method. The former forecasting method (Bontron, 2004) is improved by introducing two variables: temperature allows to take seasonal effects into account and vertical velocity enables a better characterization of the vertical atmospheric motion. The new algorithm is then applied in a perfect prognosis context (target situations come from a meteorological reanalysis) and then in an operational forecasting context (target situations come from weather forecasts) for a three years period. Precipitation scenarios are finally introduced in the rainfall-runoff model GR3P to provide discharge forecasts. Results for the two basins under study show that this approach yields useful forecasts up to one week ahead.

Key words : probabilistic quantitative precipitation forecasting, statistical adaptation, analogs, ERA-40 reanalyses, hydrological forecasting, Saône river basin, Seine river basin.