



Unité HBAN
Parc de Tourvoie - BP 44 - 92163 Antony
Cedex



Université de Tours
UFR Sciences et Techniques
Parc de Grandmont
37200 Tours

Septembre 2008

PREVISION D'ETIAGE PAR MODELISATION HYDROLOGIQUE : MISE AU POINT D'UNE METHODE D'EVALUATION



Pierre-François STAUB

RAPPORT de Stage rédigé en vue de l'obtention du
MASTER 2 PROFESSIONNEL « Géo-Hydrosystèmes Continentaux en Europe »
SPECIALITE « Tranferts »

Année 2007-2008

*Sous la direction de Charles Perrin
Unité Hydrosystèmes et Bioprocédés, Centre Cemagref d'Antony*

Remerciements:

Ma gratitude va d'abord à Charles Perrin, qui (avec insistance !) a bien voulu m'accorder sa confiance et m'a permis de bénéficier d'une période riche en enseignements scientifiques et en bonne humeur au sein de l'unité Hydrosystèmes et Bioprocédés du CEMAGREF d'Antony. Sa promptitude légendaire à rendre service avec efficacité et générosité ne s'est jamais démentie, ainsi que j'ai pu l'expérimenter personnellement à bien des reprises.

J'exprime aussi ma reconnaissance envers Vazken Andréassian et Cécile Loumagne qui ont bien voulu m'accueillir dans le cadre de leur équipe de chercheurs. Mon souhait est que le travail ici présenté puisse à l'avenir d'une façon ou d'une autre bénéficier à cette équipe.

Reconnaissance aussi à Julien Lerat, Florent Lobligeois et Maria-Hélène Ramos qui dans une bonne humeur naturelle et permanente m'ont assisté avec pertinence, en particulier dans mes errements de programmeur Scilab débutant.

Un merci spécial à Audrey Valéry, infortunée compagne de bureau, qui n'a jamais bronché en dépit du crissement continu de ma souris sur mon bureau, et qui est toujours restée fairplay en dépit de mes victoires réitérées au badminton (je plaisante, on sait qu'elle gagne à tout, tout le temps !).

Ma sympathie va aussi à Pierre Javelle, Lionel Berthet, Nicolas Le Moine, Marine Riffard, Mamoutou Tangara, Jean-Louis Rosique et Adnan Tahir, pour leur amicale présence et les conseils qu'ils ont pu me dispenser au cours de ces mois de stage.

Quelques mots aussi pour Paula qu'il m'aura fallu délaissier pendant une bonne part de ces heures consacrées à la rédaction de cet opuscule...

SOMMAIRE :

RESUME:	4
ABSTRACT:	4
RESUMO:	4
I. INTRODUCTION	5
A. CONTEXTE DU TRAVAIL EFFECTUE	5
B. OBJECTIFS SPECIFIQUES DU TRAVAIL EFFECTUE	6
C. ORGANISATION ET CHAPITRAGE DU PRESENT DOCUMENT	7
II. LE MODELE HYDROLOGIQUE GR4J	8
III. PRESENTATION DES DEUX BASSINS VERSANTS AYANT SERVIS A L'ELABORATION DE LA VERSION PROTOTYPE DE L'OUTIL TAGE	10
IV. PRESENTATION DE LA METHODE APPLIQUEE POUR LA GENERATION DES PREVISIONS : L'OUTIL TAGE-GR4J	10
A. PRINCIPE GENERAL DE FONCTIONNEMENT DE L'OUTIL NUMERIQUE DE PREVISION D'ETIAGE TAGE	10
B. EVALUATION DES PERFORMANCES DE PREVISION	11
C. CONFRONTATION AVEC LE MODELE INTERANNUEL (MODELE « H »)	13
D. DEFINITION ADOPTEE POUR L'EVENEMENT D'ETIAGE : QUE SOUHAITE-T-ON PREVOIR ?	13
E. SEUIL D'ETIAGE :	14
F. SCENARIOS CONDITIONNANT LES PREVISIONS PROBABILISTES.	14
G. CRITERE DE DECLENCHEMENT DE LA PREVISION	15
H. CRITERE PROBABILISTE DE FRANCHISSEMENT DE SEUIL (« CRITERE D'ALERTE »)	16
I. INITIALISATION DE LA PREVISION PAR LA SIMULATION	16
J. MISE EN ŒUVRE DE LA SIMULATION GR4J : CALAGE CONTROLE PREALABLE	16
V. LES DIVERS TYPES D'INDICATEURS DE PERFORMANCE DE LA PREVISION	18
A. LES SCORES D'ERREUR.	19
1. L'ECART MOYEN MENSUEL.	19
2. LA RMSE TOTALE	20
3. LA RMSE SUR LA MEDIANE q_{50}^p DES PREVISIONS	20
4. ECART TYPE DES PREVISIONS :	20
5. ERREURS RELATIVES :	21
B. LES SCORES D'ALERTES	21
1. LA PROBABILITE DE DETECTION (PROBABILITY OF DETECTION) :	21
2. LE TAUX DE FAUSSES ALERTES (FALSE ALARMS RATE) :	21
3. LE BIAIS A GRANDE ECHELLE :	21
4. L'INDICE DE SUCCES CRITIQUE :	22
5. LE SCORE DE MENACES EQUITABLE (EQUITABLE THREATS SCORE) :	22
6. L'INDICE DE ROUSSEAU :	22

C.	LES SCORES PROBABILISTES :	22
D.	SCORES DE COMPETENCE :	23
<u>VI.</u>	<u>LES DEUX VOIES EXPLOREES EN MATIERE DE DEFINITION DU SEUIL D'ETIAGE</u>	<u>24</u>
<u>VII.</u>	<u>RESULTATS DE LA PREVISION D'ETIAGES DEFINIS PAR UN SEUIL FIXE</u>	<u>26</u>
A.	CHOIX DU SEUIL D'ETIAGE	26
B.	JOURS D'ETIAGES OBSERVES	26
C.	APPLICATION DES INDICATEURS A L'EVALUATION DE LA PERFORMANCE DU MODELE « HISTORIQUE »	27
1.	JOURS D'ETIAGES PREVUS PAR LE MODELE H.	27
2.	SCORES DETERMINISTES CONTINUS : ECARTS ENTRE DEBITS OBSERVES ET PREVUS.	27
3.	SCORE PROBABILISTE : RPS.	29
4.	SCORES DETERMINISTES PAR CATEGORIES : PERFORMANCES DE LA PREVISION EN TERMES D'ALERTE D'ETIAGE.	29
D.	PERFORMANCES DE TAGE ASSOCIE A GR4J	30
1.	PARAMETRISATION	30
2.	RESULTATS BRUTS ET INTERPRETATIONS.	31
a.	Jours d'étiages prévus par le modèle TAGE-GR4J.	31
b.	Scores déterministes continus : écarts entre débits observés et prévus.	32
c.	Score probabiliste : RPS.	34
d.	Scores déterministes par catégories : performances de la prévision en termes d'alertes d'étiage.	35
e.	Vérification des propriétés de l'indice de Rousseau.	39
3.	SCORES DE COMPETENCE DE TAGE-GR4J VS. H.	41
<u>VIII.</u>	<u>DETERMINATION DES HORIZONS MAXIMAUX DE PREVISION UTILE</u>	<u>43</u>
<u>IX.</u>	<u>APPLICATION DE TAGE-GR4J A 261 BASSINS FRANCAIS</u>	<u>47</u>
<u>X.</u>	<u>CONCLUSIONS</u>	<u>52</u>
<u>XI.</u>	<u>ANNEXE A: RESULTATS DES SIMULATIONS GR4J SUR LES DEUX BASSINS TESTS.</u>	<u>55</u>
<u>XII.</u>	<u>ANNEXE B: DETERMINATION DU CRITERE D'ALERTE OPTIMAL</u>	<u>58</u>
<u>XIII.</u>	<u>ANNEXE C: SCORES DE PERFORMANCE OBTENUS AVEC TAGE-GR4J POUR LE BASSIN "A" POUR LE CAS "SEUIL FIXE"</u>	<u>60</u>
A.	SCORES DETERMINISTES CONTINUS : ECARTS ENTRE DEBITS OBSERVES ET PREVUS	60
B.	SCORE PROBABILISTE : RPS.	62
C.	SCORES DETERMINISTES PAR CATEGORIES : PERFORMANCES DE LA PREVISION EN TERMES D'ALERTE D'ETIAGE	63
<u>XIV.</u>	<u>ANNEXE D: SCORES DE COMPETENCES AVEC TAGE-GR4J POUR LE CAS "SEUIL FIXE"</u>	<u>66</u>

**XV. ANNEXE E: RESULTATS DE LA PREVISION PAR TAGE-GR4J D'ETIAGES
DEFINIS COMME EXTREMES SAISONNIERS ("SEUIL VARIABLE")** **70**

A. CHOIX DU SEUIL D'ETIAGE	70
B. JOURS D'ETIAGES PREVUS PAR TAGE-GR4J.	70
C. SCORES DETERMINISTES CONTINUS : ECARTS ENTRE DEBITS OBSERVES ET PREVUS.	71
D. SCORES RPS.	73
E. SCORES DETERMINISTES PAR CATEGORIES : PERFORMANCES DE LA PREVISION EN TERMES D'ALERTE D'ETIAGE.	74

XVI. ANNEXE H : ETUDE BIBLIOGRAPHIQUE SUR LES GENERATEURS DE PLUIE **80**

XVII. BIBLIOGRAPHIE **83**

INDEX DES FIGURES **85**

INDEX DES TABLEAUX **87**

RESUME:

Nous présentons l'élaboration d'un outil numérique dédié à la prévision probabiliste des étiages en rivière à des échéances de quelques semaines ou quelques mois. Cet outil, baptisé TAGE (Technique d'Anticipation pour la Gestion des Etiages), repose sur la prise en compte de scénarios de pluie P et d'évapotranspiration potentielle E générés pour la période de prévision. Ces scénarios, extraits d'archives climatiques journalières, sont représentatifs des P et E moyennes sur la surface d'un bassin versant et sont injectés dans un modèle hydrologique « pluie-débit » global, lequel restitue des gammes de débits prévus à l'exutoire du bassin. Les résultats concernant deux bassins versants tests sont présentés, et sont utilisés pour éprouver divers indicateurs numériques de performances de prévision. Les performances sont déclinées en fonction des mois de l'année, afin de faire ressortir la dimension saisonnière des potentialités de prévision. La performance de l'outil est évaluée en considérant la durée maximale d'anticipation au-delà de laquelle les prévisions n'apportent plus d'informations significativement plus riches que celles issues des archives historiques de débits.

Mots clefs: Etiage, Prévision Probabiliste, Modèle hydrologique, Modèle global Pluie-Débit

ABSTRACT:

We introduce a new computational tool dedicated to the probabilistic forecast of river low flows, focusing on lead times ranging from a few weeks to a few months. This tool, called « TAGE », is based on rainfall R and potential evapotranspiration E scenarios generated for the time interval between the forecast date and the day for which the forecasted flows are requested. These scenarios, extracted from daily climate data archives, represent the R and E averaged over the catchment area, and are injected in a « rainfall-runoff » lumped hydrologic model, which in turn delivers the predicted ranges of river flow at the catchment outlet. Results regarding two test basins are presented, that are used to assess various forecast performance scoring procedures. The performance is broken down into monthly scores, in order to disclose the seasonal aspect of the forecasting potentialities. The tool performance is assessed with respect to the maximum lead-time beyond which the forecasts no longer bring significantly improved information as compared to that information directly available from the river flow measurement archives.

Key words: Low flow, Probabilistic forecast, Hydrological Model, Rainfall-Runoff lumped model

RESUMO:

Apresentamos neste trabalho a concepção de um programa numérico dedicado à previsão probabilística de secas nos rios num prazo de algumas semanas ou alguns meses. Este programa, baptizado TAGE (Técnica de Antecipação para a Gestão das sEcas), tem em conta cenários de chuva P e de evapotranspiração potencial E gerados para o período da previsão. Estes cenários, extraídos de arquivos climáticos diários, são representativos das P e E médias sobre a superfície de uma bacia, e são introduzidos num modelo hidrologico “chuva-débito” global, o qual restitui gamas de débitos previstos à saída da bacia. Os resultados relativos às duas bacias teste são apresentados, e são utilizados para provar diversos indicadores numéricos de desempenho da previsão. Os desempenhos são apresentados em função dos meses do ano, afim de mostrar a dimensão sazonal das potencialidades de previsão. O desempenho do programa é avaliado considerando a duração máxima de antecipação para além da qual as previsões não fornecem informações significativamente mais ricas do que aquelas originárias dos arquivos históricos dos débitos.

Palavra-chave: Seca, Previsão probabilística, Modelo hidrologico, Modelo chuva-débito global

I. INTRODUCTION

A. CONTEXTE DU TRAVAIL EFFECTUE

Le CEMAGREF d'Antony développe depuis les années 1980 une expertise en hydrologie quantitative fondée sur des modèles conceptuels « pluie-débit » globaux construits pour reproduire les débits aux exutoires de bassins versants, à partir de données de précipitation et d'évapotranspiration moyennées sur l'ensemble de la surface de ces bassins.

Ces modèles, qui ne reposent pas sur la description de processus physiques élémentaires, sont essentiellement développés dans une perspective de gestion quantitative des ressources en eau. L'une des forces de ces modèles provient du fait qu'ils rendent compte de façon détaillée et continue dans le temps des quantités d'eau qui circulent entre les divers compartiments hydrologiques du bassin. Ils ont donc une « mémoire » des conditions hydriques passées, conditions qui jouent un rôle clé dans la détermination des débits futurs.

Cette caractéristique les rend naturellement aptes à produire des prévisions. De nombreuses applications sont développées au CEMAGREF en matière de prévisions à court terme, en particulier pour améliorer la prévision des crues.

Pour les prévisions à plus long terme, telles celles des étiages, les travaux de la communauté scientifique sont moins nombreux, en dépit des forts enjeux sociétaux d'une gestion des basses eaux. Ces enjeux sont à la fois :

- *sociaux* (l'approvisionnement en eau potable représente des enjeux vitaux),
- *économiques* (le manque d'eau est préjudiciable à l'industrie comme à l'agriculture et au tourisme),
- *écologiques* (impact sur le milieu aquatique),
- *énergétiques* (notre plus gros préleveur en France est EDF, pour ses centrales thermiques)
- *réglementaires* (l'Etat doit appuyer ses décisions de restriction sur des éléments hydrologiques incontestables).

Pour les étiages, la perspective d'un changement climatique est un facteur d'autant plus aggravant que des températures de l'eau plus élevées en été (même si le débit restait inchangé) aggraveraient les restrictions de prélèvement ou de rejets de refroidissement des centrales EDF.

Dans tous les cas, l'anticipation des situations de crise favorise un meilleur arbitrage entre les usagers et permet de limiter les impacts sur l'environnement. Les outils de modélisation actuels ne permettant pas d'estimer avec suffisamment d'anticipation les débits, dans le cadre d'une prévision à long terme. Ce stage s'inscrit dans une démarche entreprise au Département Ressources en Eau, Usages et Risques du CEMAGREF d'Antony afin de proposer des solutions pour y remédier.

Le projet vise à construire une méthode générale permettant la prévision à moyen terme (plusieurs semaines ou mois) du débit en rivière.

Cette méthode se donne pour objectifs d'être...

- suffisamment *générale* pour pouvoir être appliquée à une très grande part du territoire métropolitain. Dans ce sens il est prévu de mettre en œuvre l'outil de prévision sur un millier de bassins versants français jaugés.

- suffisamment *frugale* pour pouvoir être alimentée par les réseaux de mesure existants,
- suffisamment *robuste* pour être transférable à courte échéance aux opérationnels.

La prévision proposée devra être nécessairement probabiliste, étant donné l'incertitude importante liée à notre incapacité de prévoir le temps (pluviométrie surtout) à de telles échéances.

L'importance de l'horizon de prévision fait jouer aux pluies futures un rôle important sur les débits à venir. De nombreux auteurs ont choisi d'adopter la position de générer un grand nombre de scénarios de pluie probables et de simuler les séries correspondantes de débits futurs. A partir de ces nombreuses séries de débit générées, on peut calculer des intervalles de confiance pour la période ciblée, ce qui permet au gestionnaire de prendre des décisions de gestion. Cette méthodologie de prévision des étiages a été détaillée (Michel, 1989) puis appliquée (Perrin et al., 2001) sur des bassins versants de Bretagne et de Moselle.

S'agissant de la construction des scénarios de pluies futures, plusieurs approches sont proposées dans la littérature (nous livrons à ce propos en Annexe de ce document, §XVI, l'analyse bibliographique que nous avons pu réaliser de l'état de l'art en la matière):

- Exploitation d'archives climatiques : préférentiellement utilisées lorsque de longues séries pluviométriques sont disponibles, cette approche consiste simplement à rejouer la climatologie passée observée ; c'est celle que nous exploiterons dans ce travail.
- Générateur stochastique de pluie : l'approche, paramétrique, permet de caractériser les propriétés statistiques des pluies et leur fonctionnement temporel. Il s'appuie sur des lois courantes (par ex. exponentielle, Weibul, etc.) et des processus de chaînes de Markov ; c'est cette approche qui avait été notamment adoptée par Perrin et al. (2001).
- Générateur basé sur le ré-échantillonnage des séries observées : par rapport à l'approche exploitant des séries historiques au sens strict, ceci permet de générer un plus grand nombre de scénarios. Des travaux ont été menés récemment au Cemagref qui utilisent cette démarche, notamment ceux de J.Lerat dédiés à l'étude des étiages sur le bassin du Rhin dans le cadre du projet européen NeWater (Sauquet et al., 2007).

Ces différentes approches ne sont, ni les unes ni les autres, parfaites. Elles vont trouver des limitations liées à la disponibilité des données, la variabilité des événements observés, l'adéquation des lois utilisées. La démarche entreprise vise à explorer les complémentarités entre ces approches pour produire des scénarios les plus « représentatifs » possible.

B. OBJECTIFS SPECIFIQUES DU TRAVAIL EFFECTUE

Le travail accompli dans ce stage se situe en aval de cette démarche de réflexion entreprise de longue date au CEMAGREF sur une méthode de prévision probabiliste des étiages. Il s'est agi pour nous de contribuer à la matérialisation de ces idées par le développement d'un outil capable de produire des prévisions de débits en périodes de basses eaux, et apte à évaluer à la fois la qualité et l'utilité de ces prévisions par le biais d'indicateurs adaptés.

Ce développement a néanmoins nécessité de notre part des choix et des réflexions complémentaires :

- Afin de figer une stratégie de prévision pertinente: choix des moments de déclenchement des prévisions, choix d'une définition de l'évènement d'étiage, mode de prise en compte de la saisonnalité, ...

- Afin de définir une stratégie concise et fiable d'évaluation de la performance des prévisions : mise en place d'indicateurs internes pour le contrôle des données intermédiaires, choix des indicateurs finaux.
- Afin de concevoir un code modulaire, permettant des évolutions ultérieures simples.

Pour élaborer la version initiale de l'outil de prévision présenté ici, nous avons choisi de mettre en œuvre les prévisions à partir des débits générés par **le modèle hydrologique GR4J** (Perrin, 2002 ; voir description au §II), et ce lorsqu'on lui soumet un ensemble de scénarios de pluies P et d'évapotranspiration E journalières **issus des archives de mesures passées**.

C. ORGANISATION ET CHAPITRAGE DU PRESENT DOCUMENT

Dans un premier temps (§II) nous présentons les grandes lignes du fonctionnement du modèle hydrologique GR4J, qui est au cœur du dispositif de prévision.

Le chapitre III est dédié à la présentation des deux bassins versants tests qui nous ont servi à éprouver l'outil développé au cours de sa phase d'élaboration.

Au chapitre suivant (§IV) nous présentons le principe de fonctionnement de l'outil de prévision « TAGE » que nous avons développé. Nous en profitons pour introduire tous les concepts manipulés et sollicités au sein de l'outil. Nous y exposons les motifs de nos choix quant à la définition de la stratégie de prévision adoptée.

Le chapitre V explicite les divers indicateurs de performance retenus pour l'évaluation des prévisions.

Le chapitre VI précise les deux types de définition de l'évènement « étiage » que nous avons testées au cours de ce travail, et exprime les bénéfices et désavantages respectifs des ces deux alternatives.

Le chapitre VII présente les résultats des prévisions obtenues pour le premier type d'étiage (« seuil fixe »). Y sont comparées les performances respectives du modèle interannuel et de l'outil TAGE-GR4J.

Le chapitre VIII propose une méthode d'évaluation de l'horizon de prévision utile maximal à partir de certains des indicateurs utilisés, et expose les résultats correspondant pour les deux types d'étiage considérés.

Le chapitre IX illustre l'utilisation de TAGE à 261 bassins français, et présente de façon spatialisée les potentialités de prévisions utiles atteignables avec le modèle hydrologique GR4J.

Le chapitre X conclut ce rapport en discutant les résultats obtenus et en dessinant les perspectives d'évolution de l'outil TAGE.

Un certain nombre d'annexes accompagnent par ailleurs ce rapport, dans lesquelles le lecteur pourra trouver des informations complémentaires, traitant respectivement les thèmes suivants :

- les résultats de la simulation de GR4J appliqué à ces bassins tests, en amont et en dehors de toute prévision ; nous y déduisons des éléments d'information qui seront utiles à l'interprétation des résultats de prévision ultérieurs (§XI).
- la façon dont nous avons déterminé un critère optimal de déclenchement d'alerte (§XII)

- le détail des scores de performance de prévision obtenus pour le second des deux bassins tests utilisés (§XIII).
- le détails des scores de compétence obtenu par l'outil TAGE par comparaison avec l'utilisation des archives historiques de débit (§XIV)
- le détail des résultats obtenus pour les deux bassins avec le deuxième type de définition d'étiage considéré (§XV).
- Une étude bibliographique personnelle sur l'état de l'art en matière de générateurs de scénarios climatiques (§XVI).

Les références bibliographiques citées dans ce rapport sont restituées dans le dernier chapitre (§XVII).

II. LE MODELE HYDROLOGIQUE GR4J

Le modèle hydrologique GR4J (Perrin, 2002) a été choisi ici du fait de l'expertise sur ce modèle au sein de l'équipe du Cemagref. Un autre modèle hydrologique du même type aurait cependant pu être utilisé à la place.

Le schéma structurel de ce modèle est présenté en Figure 1.

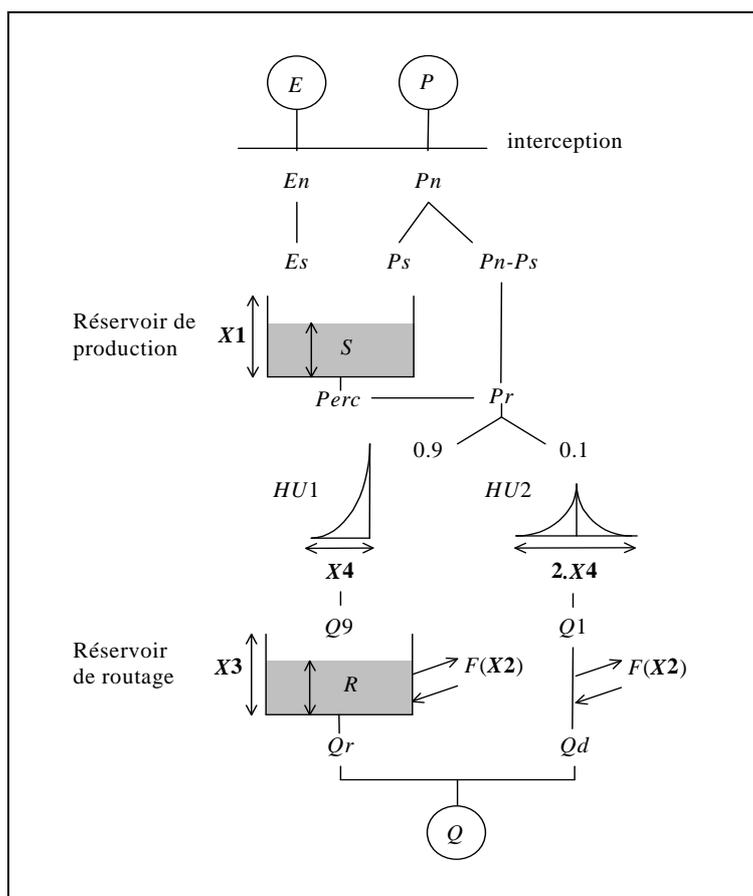


Figure 1: schéma du modèle GR4J décrivant les entrées, les variables internes, les réservoirs et paramètres associés.

GR4J reçoit le qualificatif de modèle "global", par opposition aux modèles hydrologiques dits "distribués" qui reposent sur un maillage spatial du bassin versant. Au contraire, les modèles globaux tels GR4J sont alimentés par des données d'entrée (P, E) moyennées sur l'ensemble du bassin versant, et ne restituent les débits Q qu'à l'exutoire du bassin.

GR4J ne prétend pas à une description des processus physiques individuels intervenant au sein du bassin versant. Il simule la circulation des flux à partir de compartiments virtuels, appelés "réservoirs"; les capacités de ces compartiments ainsi que les lois numériques régissant les échanges entre ces réservoirs sont paramétrées au cours d'une phase de calage du modèle dont nous avons traité au §IV.J. Il s'agit d'un modèle « continu », c'est-à-dire que le débit à une date t dépend de l'état des compartiments du modèle aux dates antérieures.

GR4J est un modèle fonctionnant au pas de temps journalier. Cela signifie qu'il est nourri par des données de pluie et d'évaporation représentant des variables cumulées sur 24 heures, et qu'il produit en retour des débits moyens sur 24 heures. Son application est donc pertinente pour la modélisation et la prévision de phénomènes se déployant avec des temps caractéristiques supérieurs ou égaux à une journée, tels les variations de ressources en eau, les étiages, et les crues pour les grands bassins versants.

GR4J intègre deux réservoirs distincts:

- I. un réservoir associé à une fonction dite de "Production", dont le niveau de remplissage est noté "S", et qui correspond à la fraction de l'eau précipitée qui sera rendue difficilement disponible pour le réseau hydrographique (disponible seulement par "percolation"). L'état de saturation de ce réservoir, couplé aux conditions d'évapotranspiration, détermine la part complémentaire de flux qui sera elle rendue plus aisément disponible pour un transfert vers le cours d'eau (pluie efficace).
- II. un réservoir associée à une fonction dite de "Routage" (ou Transfert), dont le niveau de remplissage est noté "R", et qui recueille 90% de la pluie efficace (les 10% restant échappent à l'effet tampon occasionné par le réservoir, et servent à rendre compte des petits pics à haute fréquence observés sur les hydrogrammes). Ce réservoir est constitué d'eau facilement disponible pour le cours d'eau. L'eau est restituée au cours d'eau en une quantité calculée par une loi de vidange, une partie pouvant en outre être échangée vers l'extérieur des limites topographiques du bassin (échanges souterrains).

La phase d'initialisation des états du modèle décrite au §IV.I consiste à déterminer les valeurs de S et R à l'issue de la première année de la période de calage.

Les flux transférés entre compartiments sont redistribués temporellement à l'aide de fonction appelées "hydrogrammes unitaires" HU, dont l'étalement temporel (temps de base) est paramétré et traduit la cinétique de processus propres au bassin.

GR4J repose au total sur le calage des 4 paramètres définis comme suit:

X1: capacité du réservoir de production	(mm)
X2: coefficient d'échanges souterrains	(mm)
X3: capacité à un jour du réservoir de routage	(mm)
X4: temps de base de l'hydrogramme unitaire HU1	(jour)

III. PRESENTATION DES DEUX BASSINS VERSANTS AYANT SERVIS A L'ELABORATION DE LA VERSION PROTOTYPE DE L'OUTIL *TAGE*

Deux bassins tests ont été utilisés pour l'élaboration de l'outil. Il s'agit des mêmes bassins qui avaient servi à illustrer l'application de la méthode de prévision d'étiage par Perrin et al (2001). Ces auteurs avaient conclu que les caractéristiques très contrastées de ces bassins se traduisent par des potentialités très différentes en matière de prévision d'étiage. C'est pourquoi nous avons fait le choix de reprendre ces bassins pour développer et tester notre outil de prévision, afin d'exposer nos indicateurs de performance à des conditions suffisamment variées.

Le premier bassin, baptisé ici "A" (code Banque Hydro "A7016010") est situé sur la rivière Moselle. Sa taille à l'exutoire à la station de jaugeage de Cutsines est de 6830 km². Au total 21 pluviomètres répartis sur le bassin sont utilisés pour calculer la précipitation moyenne. L'évapotranspiration potentielle moyenne annuelle est de 660mm, et la moyenne annuelle des précipitations est de 1030mm.

Le second bassin baptisé ici "J" (code Banque Hydro "J3205710") est situé en Bretagne, et la station de jaugeage localisée sur la rivière Aber Oest. Le bassin est petit (24 km²) et essentiellement agricole. Deux pluviomètres ont été utilisés pour évaluer les précipitations locales. L'évapotranspiration potentielle moyenne annuelle est de 670 mm, et la moyenne annuelle des précipitations est de 1180 mm.

Les chroniques de pluie et d'évapotranspiration potentielle, ainsi que de débits journaliers observés (q^{obs}) associés, étaient pré-conditionnées dans des fichiers dédiés à chaque bassins versant que nous avons utilisés (l'évapotranspiration est établie à partir de la formule de Oudin et al (2005); nous n'avons donc pas eu de travail spécifique à effectuer pour la mise en forme des données d'entrée.

IV. PRESENTATION DE LA METHODE APPLIQUEE POUR LA GENERATION DES PREVISIONS : L'OUTIL *TAGE-GR4J*

A. PRINCIPE GENERAL DE FONCTIONNEMENT DE L'OUTIL NUMERIQUE DE PREVISION D'ETIAGE *TAGE*

Le schéma général de la technique de prévision d'étiage mise en œuvre dans ce travail est résumé par le diagramme de la Figure 2 où sont illustrés les chemins et étapes que parcourent les flux de données.

Nous donnons dans les paragraphes qui suivent une description de chacune des variables et des étapes impliquées dans ce diagramme. Toutefois, pour plus de clarté et afin de mieux introduire les objectifs associés à chaque étape, nous avons fait le choix de présenter cette description à rebours le diagramme : nous partons des résultats finaux et des notions les plus accessibles, pour remonter ensuite progressivement le chemin qui a permis de générer ces résultats, à travers des concepts plus spécifiques.

Nous avons baptisé cet outil numérique « TAGE » (acronyme pour *Technique d'Anticipation pour la Gestion des Étiages*). L'outil développé et présenté ici dans sa version initiale est conçu pour pouvoir s'adapter à diverses façons de générer des scénarios climatiques (prévisions d'ensemble, basées sur les scénarios historiques ou stochastiques, paramétriques ou non) ainsi qu'à divers modèles hydrologiques globaux aptes à convertir ces scénarios en prévisions de débit. A terme, nous visons à ce que TAGE constitue un cadre au sein duquel la comparaison de diverses combinaisons {générateur climatique/modèle hydrologique} soit rendue aisée et fiable.

Précisons que dans la version ici présentée le système TAGE n'a pas vocation à être un outil opérationnel mais un outil d'évaluation rétrospective des capacités de prévision d'étiage des modèles hydrologiques.

Une partie importante de notre travail a consisté dans le codage de l'outil. Le langage adopté est Scilab (<http://www.scilab.org/>), environnement de développement gratuit et bien adapté au développement d'algorithmes de calcul. Au cours du codage nous avons largement bénéficié des bibliothèques HYDROGR de codes Scilab et C dédiés à l'hydrologie et développées au Cemagref par Julien Lerat (<http://lerat.julien.free.fr/>), qui contiennent en particulier des fonctions "clefs en main" assurant l'exécution de simulations hydrologiques ou de traitements de données.

Le code TAGE se décompose sommairement dans les modules suivants:

- ❖ un programme principal générant en sortie les jours d'étiages observés et prévus par le modèle hydrologique et d'après les archives de débits, ainsi que les débits associés. Cela à partir des entrées suivantes:
 - le nom du fichier (P, E, q) du bassin versant jaugé sur lequel travailler
 - la période cible choisie pour la prévision (cf.§ IV.D)
 - le type d'étiage considéré (cf.§ IV.E)
 - l'étendue temporelle h_{\max} des scénarios (cf. § IV.F)
 - le critère d'alerte choisi (cf. §IV.H)

- ❖ des modules de calcul de scores de performance de la prévision exploitant les sorties du programmes principal. Trois modules dédiés sont créés:
 - un module de calcul des scores d'alertes d'étiage (cf.§V.B)
 - un module de calcul des scores d'erreur sur les débits (cf. §V.A)
 - un module de calcul de score probabiliste (cf. §V.C)

- ❖ un programme "hôte" de lancement des programmes précédents, qui permet un enchaînement de séquences pour une série de bassins versants, et qui restitue sous forme d'un fichier de synthèse l'ensemble des résultats.

B. EVALUATION DES PERFORMANCES DE PREVISION

Nous cherchons à évaluer la qualité de la prévision des étiages, en prenant en compte la justesse avec laquelle des événements d'étiage sont anticipés, la justesse et la précision avec laquelle les débits observés sont anticipés, ainsi que l'ampleur des capacités d'anticipation.

Les résultats d'évaluation que nous donnerons ici pour chaque bassin seront déclinés en scores de performance mensuels (cf. §V), afin de tenir compte des modifications de performances induites par les variations du climat et de l'hydrométrie au cours de l'année.

Ces scores mensuels seront l'agrégation de résultats journaliers enregistrés pour les jours d'étiage observés ou prévus durant ces mois au cours de toutes les années de la chronique associée au bassin. Le §V donne une description détaillée des scores utilisés.

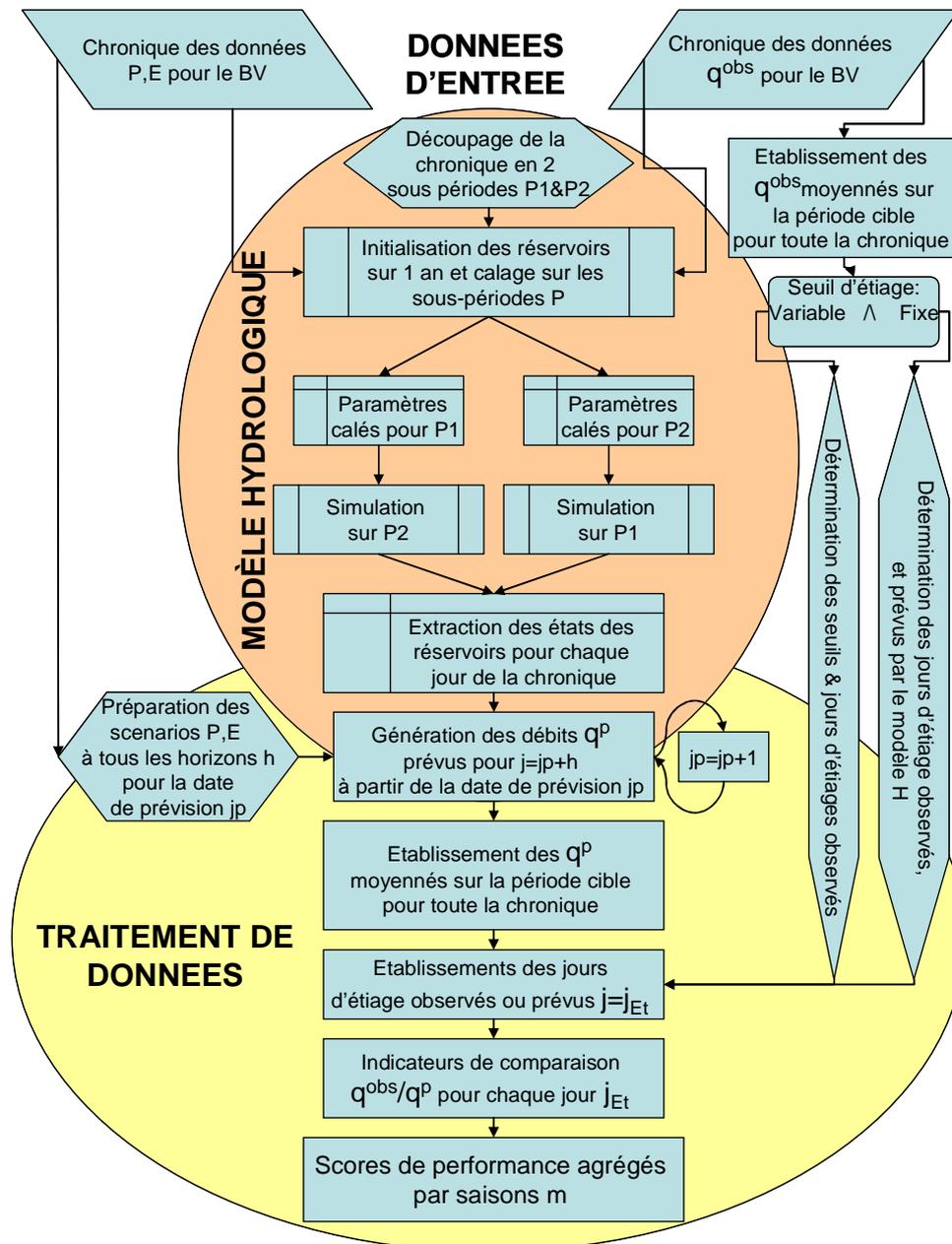


Figure 2 : diagramme générique de fonctionnement de l'outil de prévision TAGE.

Ces scores de performance restent néanmoins des entités abstraites du point de vue d'un utilisateur potentiel. Il était donc important de les traduire en une quantité en rapport direct avec les besoins opérationnels de prévision. Il nous a paru naturel de choisir comme indicateur ultime de la qualité de la prévision d'un modèle le délai maximal au-delà duquel les prévisions ne sont plus intéressantes.

Nous désignerons dans la suite ce délai par l'expression "**Horizon de Prévision Utile Maximal**" (HPUM).

Le HPUM sera donc l'horizon de prévision au-delà duquel les scores de performance ne seront plus assez bons pour que la prévision soit jugée "utile". Nous verrons par la suite quel critère nous avons choisi pour décider de cette utilité.

C. CONFRONTATION AVEC LE MODELE INTERANNUEL (MODELE « H »)

En l'absence de toute modélisation des processus hydrologiques, et pour les bassins jaugés (c'est-à-dire ceux pour lesquels il existe un suivi des débits à l'exutoire), les gestionnaires de ressource en eau peuvent tout de même émettre des prévisions quant aux débits à venir, en se basant sur les débits enregistrés historiquement. Si les débits sont suivis journalièrement, le gestionnaire dispose pour chaque jour calendaire de l'ensemble des débits mesurés ce même jour aux années précédentes. De la sorte il lui est possible de faire une analyse fréquentielle des écoulements et d'émettre des hypothèses de nature statistique sur la probabilité de voir le débit passer sous un certain seuil, qualifié de seuil d'étiage.

Pour simplifier la lecture de ce rapport, ce modèle basé sur les seules données de débits historiques sera désigné par l'expression « **modèle H** ».

Les prévisions basées sur un modèle hydrologique ne seront intéressantes qu'à la condition que la qualité d'information qu'elles apportent à un gestionnaire soit significativement supérieure à celle qui est obtenue avec le modèle H. Nous nous attacherons donc ici à intégrer ce critère pour évaluer la performance du modèle GR4J.

D. DEFINITION ADOPTEE POUR L'EVENEMENT D'ETIAGE : QUE SOUHAITE-T-ON PREVOIR ?

Diverses grandeurs caractéristiques des faibles débits, que l'on peut extraire des hydrogrammes, ont été définies dans la littérature (voir les compilations de Smakthin, 1998 et 2001 ou Abi-Zeid, 1999), rapportées et recommandées comme bons indicateurs des étiages. Mais il n'existe pas de manière standardisée de définir l'étiage.

Deux aspects rentrent néanmoins systématiquement en jeu :

- la définition d'un seuil de débit, en deçà duquel les cours d'eau est susceptible d'être qualifié comme étant en étiage. Ce seuil est typiquement basé sur des considérations liées à la fréquence d'apparition des débits : on prend alors pour seuil la valeur d'un quantile donné des débits, ou bien une valeur moyenne ayant une fréquence de retour donnée, tel le QMN_x (débit mensuel minimal annuel de période de retour x ans)
- la prise en compte d'une dimension temporelle : les faibles débits doivent avoir perdurés plus d'un certain temps pour qu'on puisse parler d'étiage, un faible débit journalier isolé étant souvent de conséquence moins significative qu'un déficit enregistré sur plusieurs jours. En France il est courant de caractériser la sévérité des étiages à partir du débit moyen minimal observé sur une période de 7 jours consécutifs (VCN7).

Dans le cadre de ce travail, nous considérerons qu'il y a étiage à une date j seulement si *la moyenne des débits sur une période cible « p » de plusieurs jours centrée sur la date j* prend une valeur inférieure à une valeur seuil donnée. Cela permet de lisser les éventuelles fluctuations liées à

l'imprécision de la débitmètrie, et par ailleurs de prendre en compte la dimension de durée de l'événement. Dans ce travail, pour illustrer la méthode mise au point, nous avons choisi une période cible p de une semaine, ce qui signifie que nous travaillerons systématiquement sur des chroniques de débits (observés ou prévus) lissés sur une semaine. Le fait d'utiliser des modèles journaliers, tels GR4J, rend la période cible paramétrable à des multiples de la journée, cette période pouvant même être réduite à $p=1$ jour au cas où les besoins opérationnels le requerraient.

E. SEUIL D'ÉTIAGE :

Pour ce qui concerne le seuil d'étiage, c'est-à-dire la valeur du débit lissé en deçà de laquelle on enregistrera un « jour d'étiage », nous avons dans ce travail exploré deux voies mutuellement exclusives : soit le seuil a une valeur fixe, indépendamment du jour calendaire considéré, soit cette valeur dépend du jour calendaire et est calculée en fonction des débits enregistrés historiquement ce même jour. La première méthode permet la comparaison entre GR4J et le modèle H, la seconde permet la prise en compte du caractère inhabituellement faible des débits à une saison donnée. Plus de détails sur ces deux types de définition de seuil d'étiage sont donnés au §VI.

F. SCENARIOS CONDITIONNANT LES PREVISIONS PROBABILISTES.

Dans la première version de l'outil TAGE présentée ici, la génération des débits journaliers prévus pour un jour j d'une année donnée s'effectue, pour chaque date de prévision j_p , en injectant successivement dans GR4J l'ensemble des n_s scénarios de pluie P et d'évapotranspiration E issus de l'ensemble des autres années de la chronique historique de ces variables. Si l'on note n_{An} le nombre d'années de données modélisables, on aura donc $n_s = n_{An} - 1$ scénarios générés pour chaque prévision. L'intervalle de temps entre la date de prévision j_p et le jour central de la période cible j est appelé « **horizon de prévision** », noté h . Chaque scénario s'étale donc sur une durée h . TAGE génère pour chaque date de prévision j_p un nombre n_{max} de prévisions, correspondant à autant de périodes cibles d'anticipation. Autrement dit les scénarios sont initialement provisionnés sur un horizon total $h_{max} = n_{max} \cdot p$, et chaque horizon h correspond à un nombre entier « n » de périodes cibles p compris entre 1 et n_{max} ($h = n \cdot p$). La Figure 3 illustre la manière dont les horizons de prévision sont définis dans cette méthode pour une date de prévision donnée.

Pour la validation de la méthode présentée dans ce travail, nous avons utilisé $n_{max} = 20$, soit $h_{max} = 20$ semaines. Les bassins versants utilisés pour tester notre outil étaient jaugés sur $n_{An} = 11$ années (entre 1995 et 2005). Au final cela signifie que, pour chaque date de prévision j_p , des séries de $n_{An} - 1 = 10$ débits sont modélisés pour 20 dates cibles j différentes, chacune espacée de la précédente d'une période cible.

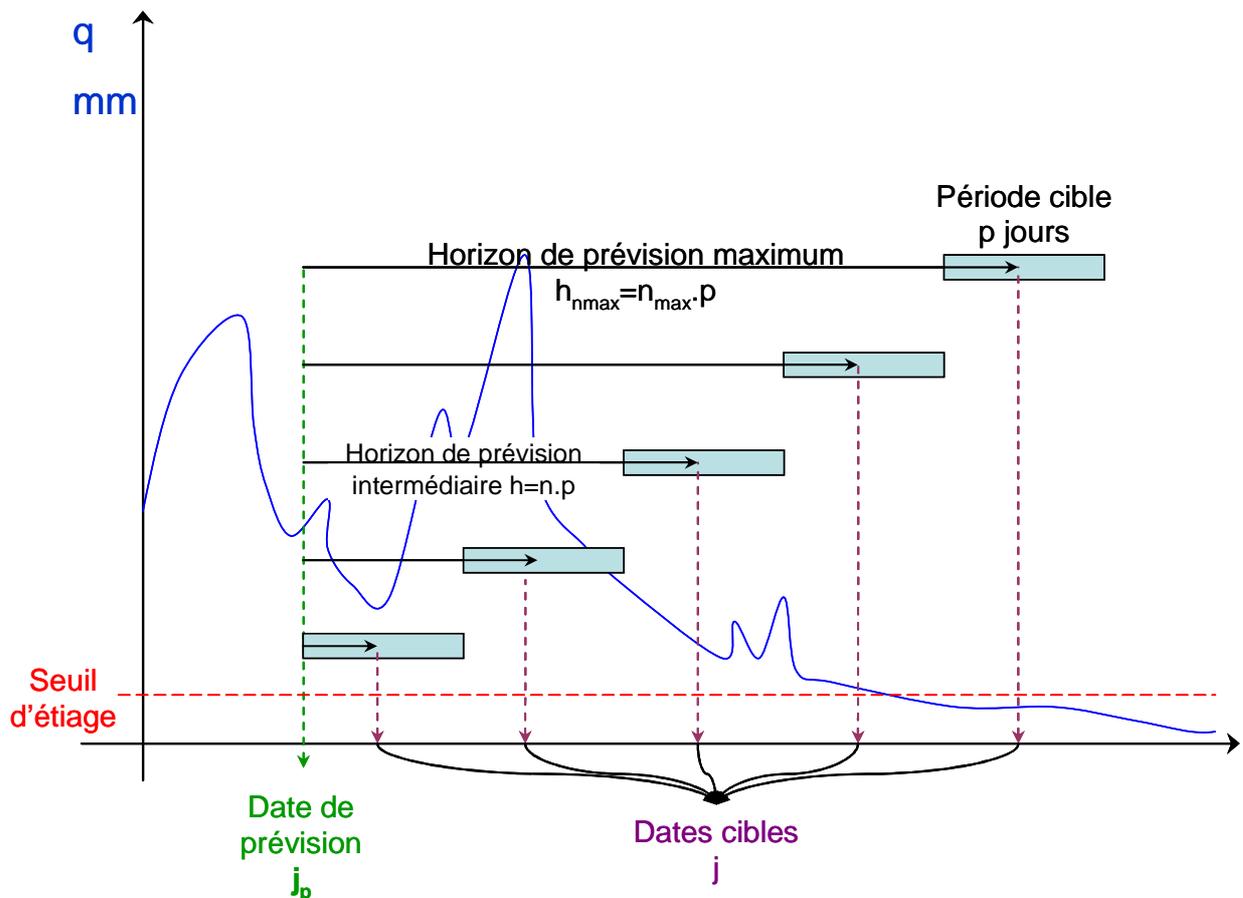


Figure 3: représentation schématique de la stratégie de prévision adoptée dans TAGE. La courbe bleue symbolise un hydrogramme. Pour chaque date de prévision j_p une série de n_{max} prévisions sont émises concernant autant de dates cibles j , chacune d'elle étant associée au débit lissé sur la période cible correspondante.

G. CRITERE DE DECLENCHEMENT DE LA PREVISION

Quand faire la prévision ? A une date donnée ? Dès que le débit passe sous un seuil donné ? Lorsque les précipitations cumulées passent sous un seuil donné ?...

D'un point de vue opérationnel, les critères qui amèneront un gestionnaire à émettre une prévision seront certainement très variables selon les enjeux et les caractéristiques du système qu'il doit gérer.

Afin d'éviter de produire des évaluations de performance dans des conditions qui seraient biaisées par le choix arbitraire d'un critère de déclenchement de la prévision, nous avons fait le choix d'émettre des prévisions en continue. Cela signifie que les opérations décrites ci-dessus et illustrées par la Figure 3 sont répétées pour tous les jours j_p de la chronique. Cela se fait au prix de temps de calculs plus importants, mais permet d'obtenir des évaluations de performance valides et représentatives quel que soit le jour où la saison à laquelle un utilisateur serait amené à déclencher ses prévisions. De ce point de vue le système proposé se distingue de travaux précédents (Sauquet et al., 2007) qui font reposer le déclenchement de la prévision sur une connaissance *a priori* des périodes d'étiage dans l'année pour le bassin considéré.

H. CRITERE PROBABILISTE DE FRANCHISSEMENT DE SEUIL (« CRITERE D'ALERTE »)

Le fait que TAGE génère, pour une date de prévision donnée et un horizon de prévision donné, une série de prévisions pour chaque date cible j signe son caractère de « modèle de prévision probabiliste ». Le modèle naïf « historique » H, qui associe à chaque jour j une série de débits historique est dans le même cas. Partant de là, il existe une marge dans la façon d'utiliser ces modèles pour la détection des événements d'étiage : pour chaque série de débits prévus, on doit définir quelle proportion de ces derniers doivent être inférieurs au seuil d'étiage pour qu'on puisse effectivement parler d'étiage. Cette proportion constitue le « critère d'alerte ». Pour le travail présenté ici d'élaboration de la méthode, nous avons opté pour une proportion de 10%. Les raisons qui nous ont poussé à faire ce choix sont exposées en Annexe (§ XII).

I. INITIALISATION DE LA PREVISION PAR LA SIMULATION

Dans la version initiale de TAGE présentée ici, la valeur de débit journalier prévu $q^p(j)$ pour chaque scénario $\{P(j_p, h), E(j_p, h)\}$ est obtenue à l'aide du modèle GR4J. Cela n'est toutefois possible que parce que les états des réservoirs de stockage S et de routage R de ce modèle sont connus au jour j_p de la prévision (NB : on considèrera ici en première approximation que l'état des hydrogrammes unitaires n'influe pas sur la prévision, compte tenu que les temps de redistribution impliqués [au maximum quelques jours] sont très inférieurs aux délais de prévision qui nous intéressent [plusieurs semaines]). Dans l'outil TAGE, les valeurs de S et R sont déterminées à l'issue d'une simulation préalable de l'ensemble des débits de la chronique, simulation qui s'effectue en injectant dans GR4J, pour chaque jour j de la chronique, les valeurs de P(j) et E(j) enregistrées ce même jour j. Cette simulation restitue donc les débits simulés $q^{sim}(j)$, et les états de S(j) et R(j) correspondants.

J. MISE EN ŒUVRE DE LA SIMULATION GR4J : CALAGE CONTROLE PREALABLE

La phase de simulation décrite ci-dessus fait elle-même partie d'une procédure d'évaluation du modèle, dite de 'split-sample test', c'est-à-dire des calages-contrôle croisés sur des périodes indépendantes. L'objet de cette procédure est double :

- 1- Le premier objectifs est de caler les 4 paramètres du modèle GR4J en minimisant, via un algorithme numérique itératif dédié, l'erreur entre les débits simulés $q^{sim}(j)$ et les débits observés $q^{obs}(j)$. L'erreur à minimiser est quantifiée sous la forme d'un critère numérique également appelé « fonction objectif », qui peut être défini de différentes façons. La fonction objectif utilisée ici est le critère dit de Nash-Sutcliffe (Nash, 1970), qui s'exprime comme suit :

$$NS(\%) = 1 - \frac{\sum_{j=1}^{j_{Max}} (q^{obs}(j) - q^{sim}(j))^2}{\sum_{j=1}^{j_{Max}} (q^{obs}(j) - \bar{q}^{obs})^2} \quad \text{Eq.1}$$

où \bar{q}^{obs} est la moyenne des débits observés sur toute la chronique.

Cette fonction est donc égale à 1 dans le cas d'un modèle parfait, et peut aller jusqu'à prendre des valeurs négatives si le modèle évalué présente des performances inférieures à celle du modèle naïf qui prédit un débit constant \bar{q}^{obs} .

Notons qu'il existe des variantes pour le choix de la fonction objectif, en particulier le critère NS peut être calculé non plus à partir des débits bruts q , mais à partir de leurs racines carrées \sqrt{q} , voire de leurs logarithmes $\log(q)$. Ces variantes permettent de favoriser davantage la contribution au critère d'ensemble NS des jours présentant des débits faibles, ce qui revêt un intérêt pour la prévision d'étiage. Nous n'avons pas utilisé ces variantes pour la phase d'élaboration de l'outil, mais il sera intéressant de les tester dans le travail qui suivra.

D'une façon générale, l'opération d'étalonnage des paramètres du modèle est absolument nécessaire dans la mesure où GR4J est un modèle conceptuel qui ne repose pas sur une description physique des processus.

- 2- Le second objectif de la procédure de calage-contrôle est de vérifier la pertinence des paramètres calés sur une période de validation indépendante. En effet le calage des paramètres sur une période limitée ne garantit jamais de façon certaine que ces valeurs de paramètres soient extrapolables sur d'autres périodes qui peuvent mettre en jeu des circonstances physiques différentes. Une façon d'évaluer la robustesse du calage effectué consiste à éprouver les débits simulés avec les paramètres calés sur une période de validation indépendante de la période ayant servi au calage. Cette opération prend le nom de « calage/contrôle ». Elle s'effectue en divisant la chronique en deux sous-périodes distinctes P1 et P2 : dans un premier temps les paramètres sont calés sur la période P1 et validés sur P2, puis un deuxième jeu de paramètres sont calés sur P2 et validés sur P1. Le contrôle consiste à vérifier que la valeur de la fonction objectif obtenue en validation sur P2 n'est pas significativement inférieure à celle obtenue en calage sur P1, et réciproquement. Ce procédé a été désigné sous le nom de « split sample test » (Klemes, 1986).

Notons que la simulation en mode continu requiert une période d'initialisation des états des réservoirs de stockage et de routage. Durant cette période initiale, les états de S et R évoluent à partir d'un état initial arbitraire avec les forçages observés, de sorte qu'au bout d'un certain temps il n'y a plus d'impact des conditions initiales sur les simulations. Dans la procédure utilisée, la période d'initialisation est prise égale à un an. Cela signifie que les années n°1 des périodes P1 et P2 sont inexploitable pour les prévisions, la simulation n'étant pas fiable sur ces années. Afin d'éviter de perdre ainsi deux années de la chronique, la période P2 comprend une année supplémentaire par rapport à P1 (la dernière année de P1 est prise en entier comme première année de P2). Ainsi, en dehors de la première année, les simulations sont exploitables pour toutes les années de la chronique, y compris l'année de P2 qui suit la dernière année de P1. Cette situation est illustrée par la Figure 4.

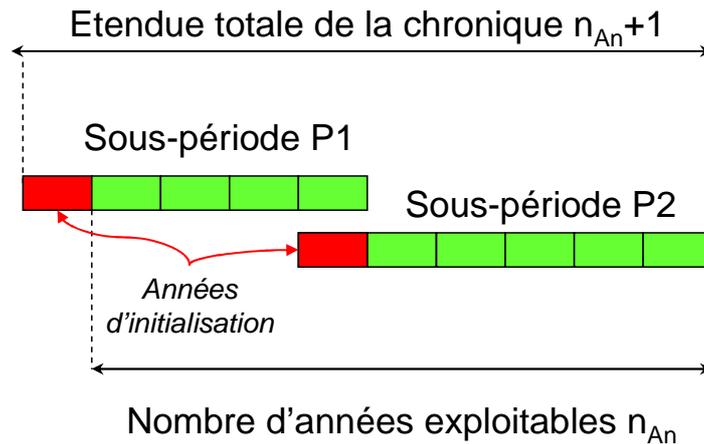


Figure 4: diagramme représentant la stratégie du « split-sample test » utilisé lors de la phase de calage-contrôle du modèle hydrologique sur une chronique de $n_{An}+1$ années de mesures (P,E,q).

Les résultats que nous avons obtenus en simulation sur nos deux bassins versants tests ainsi que les interprétations que nous en avons faites sont rapportés en Annexe de ce document, §XI. Cette étude préalable des résultats de simulation nous a notamment permis d'expliquer ce qui différencie fondamentalement les deux bassins du point de vue des potentialités de prévisions des débits futurs. L'analyse des valeurs des paramètres du modèle issues de l'optimisation montre en particulier que la capacité du réservoir de routage pour le bassin J excède d'un facteur 15 celle observée pour le bassin A. Cela rend les hydrogrammes du bassin J moins sujets aux variations pluviométriques et davantage déterminés par l'état des réservoirs des jours qui précèdent. D'un point de vue de prévisionniste, cela se traduit pour J par de des potentialités de prévision des débits sur des échelles de temps plus importantes. Dans la suite de ce travail nous qualifierons par l'expression "**effet mémoire**" cette propriété du bassin à stocker des réserves d'eau suffisantes pour alimenter de manière prédictible le débit à l'exutoire du bassin.

V. LES DIVERS TYPES D'INDICATEURS DE PERFORMANCE DE LA PREVISION

D'une façon générale nous nous sommes inspirés des méthodes de prévision basées sur le calcul de scores de performance qui largement utilisés en météorologie, science où la prévision a atteint un stade de maturité élevé. L'ensemble des indicateurs utilisés ici sont extraits de la thèse de Bontron (2004).

Nous avons dans ce travail exploré trois types d'indicateurs susceptibles de fournir des informations complémentaires sur les performances de la prévision :

- les indicateurs basés sur **l'erreur de la prévision**. Cette erreur est un écart algébrique ou quadratique entre débits observés et prévus. Ces indicateurs peuvent donc être dimensionnés (en mm ou en m³/s) pour évaluer l'erreur de débit associée à la prévision. Ils peuvent également être rapportés aux débits moyens, afin de fournir des indicateurs relatifs sans unité, permettant une meilleure comparaison de performance entre des saisons ou des bassins différents. Toutefois, même sous forme d'erreur relative la comparaison et l'agrégation des scores n'est pas forcément pertinente: une erreur de 10% pour un bassin où il coule beaucoup d'eau n'a pas la même

signification pour un bassin où il en coule peu. En fait pour pouvoir comparer des bassins différents il faudrait introduire une notion de tolérance sur l'erreur, propre à chaque bassin, ce qui est hors de notre propos.

- **Les scores de contingences**, ou d'alertes. Ces scores expriment la capacité d'une prévision à détecter et discriminer des événements associés au dépassement d'un seuil donné de débit. Ce sont donc des indicateurs d'occurrences, traitant d'observations de types « oui/non ». Leur avantage est d'être directement exploitables opérationnellement, mais ils n'offrent pas de capacité d'analyse des performances en termes de justesse ou de précision des débits prévus. Pour notre cas, ils permettront de comparer la pertinence de divers modèles quant à la prévision de l'occurrence des étiages, mais ils ne pourront par exemple pas être utilisés pour évaluer la capacité des modèles à prévoir la sévérité de ces étiages.
- **Les scores probabilistes par catégories**, qui opèrent une quantification de l'adéquation entre des gammes (ou « catégories ») de débits observés et des gammes de débits prévus. Ces scores sont bornés, typiquement entre 0 et 1, et permettent une comparaison directe des performances vis-à-vis de cette adéquation entre différents modèles, ou bien pour un même modèle entre des périodes ou des bassins différents. Ils ne fournissent pas d'information directe sur l'erreur absolue. De plus, une même erreur relative sur les débits pour deux bassins différents se traduira par des scores différents, dans la mesure où la définition des catégories dépend de la fonction de répartition des débits observés, elle-même dépendante du bassin.

Par ailleurs et de manière générale, nous avons cherché à évaluer la performance en tenant compte de la **saisonnalité** ; nous avons donc élaboré des indicateurs pour chaque mois « m » de l'année, en regroupant tous les événements d'étiages ayant eu lieu durant ce mois pour toutes les années de la chronique. Ces événements correspondent aux jours d'étiage observés ou prévus, notés j_{Et} .

A. LES SCORES D'ERREUR.

1. L'ÉCART MOYEN MENSUEL.

Nous définissons l'écart comme la différence entre le débit observé q^{obs} et le débit q^p issu d'une prévision k individuelle, pour un jour j_{Et} donné.

Compte tenu...

- du nombre d'événements $n_{j_{Et}(m, iAn)}$ enregistrés pour le mois m de l'année iAn ,
- des n_{An} années exploitables de la chronique
- des $n_{An}-1$ prévisions disponibles pour chaque jour,

... l'écart moyen mensuel s'écrit :

$$Ecart(m) = \frac{\sum_{iAn=1}^{n_{An}} \sum_{i=1}^{n_{j_{Et}(m, iAn)}} \sum_{k=1}^{n_{An}-1} (q_k^p(j_{Et}[i]) - q^{obs}(j_{Et}[i]))}{(n_{An} - 1) \sum_{iAn=1}^{n_{An}} n_{j_{Et}(m, iAn)}} \quad Eq.2$$

Cet écart algébrique sera positif si le modèle surestime globalement les débits des jours j_{Et} et négatif dans le cas contraire.

2. LA RMSE TOTALE

Ou « Root Mean Square Error », s'appuie sur les carrés des écarts individuels précédents, et s'exprime comme suit :

$$RMSE(m) = \sqrt{\frac{\sum_{iAn=1}^{n_{An}} \sum_{i=1}^{nj_{Et}(m,iAn)} \sum_{k=1}^{n_{An}-1} (q_k^p(j_{Et}[i]) - q^{obs}(j_{Et}[i]))^2}{(n_{An} - 1) \sum_{iAn=1}^{n_{An}} nj_{Et}(m,iAn)}} \quad \text{Eq.3}$$

Parce qu'elle tient compte de chaque prévision individuelle, la RMSE prend en compte la dispersion de ces prévisions : plus les prévisions seront dispersées, plus la RMSE croîtra. Dans l'idéal la RMSE est nulle, ce qui correspond à des prévisions de biais nul (moyenne des prévisions $\bar{q}^p = \bar{q}^{obs}$) et de dispersion nulle.

3. LA RMSE SUR LA MEDIANE q_{50}^p DES PREVISIONS

Elle s'exprime comme suit :

$$RMSE_{Med}(m) = \sqrt{\frac{\sum_{iAn=1}^{n_{An}} \sum_{i=1}^{nj_{Et}(m,iAn)} (q_{50}^p(j_{Et}[i]) - q^{obs}(j_{Et}[i]))^2}{\sum_{iAn=1}^{n_{An}} nj_{Et}(m,iAn)}} \quad \text{Eq.4}$$

Cette quantité informe sur le centrage de la distribution des prévisions autour du débit observé, ce qui s'apparente d'une mesure de justesse. Contrairement à la RMSE totale, elle n'est pas sensible à la dispersion des prévisions.

4. ECART TYPE DES PREVISIONS :

Il s'exprime comme suit

$$EcartType(m) = \sqrt{\frac{\sum_{iAn=1}^{n_{An}} \sum_{i=1}^{nj_{Et}(m,iAn)} \sum_{k=1}^{n_{An}-1} (q_k^p(j_{Et}[i]) - \bar{q}^p(m))^2}{(n_{An} - 1) \sum_{iAn=1}^{n_{An}} nj_{Et}(m,iAn)}} \quad \text{Eq.5}$$

Cette quantité, qui n'est pas à proprement parler un critère d'erreur, informe sur la dispersion des prévisions autour de leur valeur moyenne. Cette quantité est indépendante de la justesse des prévisions.

5. ERREURS RELATIVES :

Afin de pouvoir comparer entre eux des bassins présentant des gammes de débit différentes, il est utile de rapporter les erreurs précédemment décrites à la moyenne des débits \bar{q}^{obs} sur la période considérée, comme suit:

$$RMSE\%(m) = 100 \frac{RMSE(m)}{\bar{q}^{obs}(m)} \quad \text{Eq.6}$$

L'usage de quantités relatives peut être intéressant également pour comparer les variations de l'erreur avec celle d'autres scores sans unités, tels les scores probabilistes.

B. LES SCORES D'ALERTES

En tout et pour tout on peut considérer quatre types d'alertes, dont les fréquences qualifient les performances du système de prévision :

- Les alertes prévues et effectivement observées, ou « **alertes vraies** », dont le nombre sera noté « **a** ».
- Les alertes prévues et non observées, ou « **fausses alertes** », désignées par « **b** »
- Les alertes observées non prévues, ou « **alertes manquées** », désignées par « **c** ».
- Les jours sans alertes prévues ni observées, désignés par « **d** ».

Le nombre total de prévisions émises **a+b+c+d** est noté **N**.

A partir de ces quantités, les prévisionnistes ont défini un certain nombre d'indicateurs simples ou plus complexes, dont nous décrivons ici quelques exemples qui seront utilisés dans le cadre de ce travail.

1. LA PROBABILITE DE DETECTION (PROBABILITY OF DETECTION) :

$$POD = \frac{a}{a+c} \quad \text{Eq.7}$$

C'est le rapport entre le nombre d'alertes vraies et le total des alertes observées.

2. LE TAUX DE FAUSSES ALERTES (FALSE ALARMS RATE) :

$$FAR = \frac{b}{a+b} \quad \text{Eq.8}$$

C'est le rapport entre le nombre de fausses alertes et le total des alertes prévues.

3. LE BIAIS A GRANDE ECHELLE :

$$Biais = \frac{a+b}{a+c} \quad \text{Eq.9}$$

C'est le rapport entre nombres d'alertes prévues et d'alertes observées.

4. L'INDICE DE SUCCES CRITIQUE :

$$CSI = \frac{a}{a + b + c} \quad \text{Eq.10}$$

C'est le rapport entre alertes vraies et total des alertes, prévues ou observées. Ce score à l'avantage d'être synthétique et aisément interprétable pour rendre compte de l'efficacité de la prévision. Il est utilisable si la fréquence des évènements donnant lieu à des alertes est faible devant celle des évènements autres. Cela ne sera pas forcément le cas pour notre étude, les seuils de définition d'étiage pouvant dans certains cas sélectionner un nombre important d'évènements au sein d'une période d'intérêt.

5. LE SCORE DE MENACES EQUITABLE (EQUITABLE THREATS SCORE) :

$$ETS = \frac{a - R}{a + b + c - R} \quad \text{avec} \quad R = \frac{(a + b)(a + c)}{N} \quad \text{Eq.11}$$

Ce score s'apparente dans sa formulation au CSI, mais il est retranché une composante R correspondant à l'espérance du nombre d'alertes vraies qui pourraient être générées par un système qui générerait aléatoirement les (a+b) alertes émises, la fréquence relative des évènements d'étiage étant égale à (a+c)/N.

En omettant ainsi les prévisions faites « par chance », on peut comparer entre elles des séries de prévisions associées à des régimes climatiques différents, par exemple des prévisions effectuées à des saisons différentes, ou bien sur des bassins différents.

6. L'INDICE DE ROUSSEAU :

(Rousseau, 1980) :

$$IR = \frac{a.d - E^2}{(a + E)(d + E)} \quad \text{avec} \quad E = \frac{b + c}{2} \quad \text{Eq.12}$$

Sans entrer dans les détails de l'établissement de cette formulation (voir pour cela la thèse de Bontron), retenons que cet indice prend la valeur 1 pour une prévision idéale, et pénalise à la fois les prévisions dont la composante « chance » est importante (comme l'ETS) ainsi que les prévisions qui présentent un biais à grande échelle important. C'est donc un bon candidat pour juger de manière synthétique et exigeante de la qualité des prévisions.

C. LES SCORES PROBABILISTES :

Guidés par les recommandations de Bontron, 2004, nous avons choisi pour ce travail d'éprouver le « Ranked Probability Score » ou "RPS" (Epstein, 1969). Ce score s'applique à des prévisions probabilistes regroupées par « catégories » ordonnées entre elles, telles des gammes de débit en rivières. Il est défini par:

$$rps(j) = \frac{1}{G-1} \sum_{K=1}^G \left(\sum_{L=1}^K p_L(j) - \sum_{L=1}^K \delta_L(j) \right)^2 \quad \text{Eq.13}$$

G est le nombre de catégories, $p_L(j)$ est la probabilité associée à la catégorie L par la prévision pour le jour j. $\sum_{L=1}^K p_L(j)$ représente la probabilité cumulée que la prévision se trouve dans une catégorie $\leq K$ le jour j. δ_L est une fonction qui prend la valeur 1 si l'observation est dans la catégorie L, et 0 sinon.

Le principe général du RPS est donc de comparer les probabilités au non-dépassement prévues et observées.

Ce score prend la valeur zéro pour une prévision parfaite, et 1 pour la plus mauvaise des prévisions (cas où toutes les prévisions sont dans une gamme extrême opposée à celle où se situent l'observation). Dans ce travail, compte tenu du nombre limité ($=n_{An}-1$) de prévisions émises chaque jour ($n_{An}-1=10$ pour les tests du prototype et 35 pour les bassins sur lesquels il sera appliqué) nous sommes limités à trois catégories, chacune représentant un tercile des débits historiques enregistrés pour le jour j. Ainsi la catégorie la plus basse correspond aux débits compris entre zéro et le quantile 33% « Q33 » des débits historiques du jour j, la catégorie moyenne correspond à l'intervalle Q33(j)-Q66(j), et la dernière aux débits $> Q66(j)$.

Pour caractériser la performance RPS pour la prévision d'étiage sur une période « m » de l'année, on calcule la moyenne sur l'ensemble de la chronique de débits du rps en ne considérant que les jours d'étiage J_{Et} prévus ou observés durant cette période, comme suit :

$$RPS(m) = \frac{\sum_{iAn=1}^{n_{An}} \sum_{i=1}^{nj_{Et}(m, iAn)} rps(j_{Et}[i])}{\sum_{iAn=1}^{n_{An}} nj_{Et}(m, iAn)} \quad \text{Eq.14}$$

D. SCORES DE COMPETENCE :

Afin de comparer les performances respectives, évaluées par le score S, d'un modèle de prévision M avec celles d'un modèle de référence M_{ref} , on calcule le score de compétence suivant :

$$Comp_S(M) = \frac{S(M) - S(M_{ref})}{S_{Parfait} - S(M_{ref})} \quad \text{Eq.15}$$

$S_{parfait}$ correspond à la valeur optimale prise par S dans le cas hypothétique d'une prévision parfaite. $Comp_S$ vaudra 1 dans l'idéal ($S(M)=S_{parfait}$), et deviendra négatif si le modèle M est plus mauvais que le modèle de référence.

Dans le cadre de ce travail, nous nous servirons de la compétence pour comparer le modèle TAGE avec le modèle naïf « H » basé sur la simple connaissance des débits historiques, que nous choisirons comme référence.

VI. LES DEUX VOIES EXPLORÉES EN MATIÈRE DE DÉFINITION DU SEUIL D'ÉTIAGE

Dans une première voie, le seuil est choisi comme fixe quel que soit le jour considéré. Cette méthode simpliste sélectionnera donc préférentiellement les jours d' « étiage » parmi les saisons les plus sèches. Ce type de définition fixe du seuil permet de comparer directement les performances de prévision d'étiage obtenues par TAGE-GR4J d'une part et le modèle H d'autre part. Son désavantage principal réside dans le fait qu'elle sélectionnera comme « étiages » des débits certes faibles mais tout à fait usuels pour certaines périodes de l'année (en été notamment), et que réciproquement des débits exceptionnellement bas à des périodes de hauts débits pourront ne pas être détectés, car supérieurs au seuil choisi.

La valeur du seuil que nous avons retenue est le QMNA2 (valeur moyenne mensuelle minimale de fréquence de retour 2 ans) ; nous avons fait ce choix car le QMNA est un débit caractéristique couramment employé pour caractériser les régimes d'étiage en hydrologie. De plus la valeur du QMNA2 s'est avérée sélectionner un nombre idéal d'événements sur les deux bassins prototypes éprouvés, à savoir environ 30 événements par an, soit environ 300 sur chacune des deux chroniques de 10 ans utilisées pour l'élaboration de la méthode. Ce nombre n'est ni trop grand (le risque d'un seuil trop élevé étant de « banaliser » la notion d'étiage et de perdre en finesse dans l'analyse des performances spécifiques des modèles hydrologiques pour les faibles débits), ni trop petit (il permet de conserver un nombre de plusieurs dizaines d'événements pour les mois d'été, ce qui crédibilise les interprétations statistiques qui seront faites pour ces périodes).

Dans une seconde voie, la valeur absolue du seuil d'étiage est adaptée pour chaque jour de l'année. Pour chaque jour calendaire, le seuil correspond à un certain percentile des débits enregistrés historiquement ce jour là. Pour le prototype nous avons opté pour le percentile 20% des débits journaliers, c'est-à-dire le second le débit le plus bas sur les 10 disponibles chaque jour calendaire. Ce type de définition du seuil d'étiage permet d'éviter de sélectionner comme étiage des débits certes faibles mais tout à fait habituels pour leur saison, et permet la détection de flux inhabituellement bas même à des saisons où le débit est en moyenne relativement élevé. Le désavantage de cette méthode du seuil variable est que le modèle H n'est pas applicable. En effet puisque le seuil est défini à l'aide des chroniques historiques de débit, tous les jours calendaires sont situés sur un pied d'égalité vis-à-vis du modèle H : aucun jour calendaire ne présente des caractéristiques qui puissent le porter plus qu'un autre à être considéré comme un jour d'étiage. Pour ce type de définition de seuil, le modèle H est donc intrinsèquement incompétent, et on ne pourra donc pas comparer ses performances avec TAGE-GR4J pour la détection d'étiages.

Par ailleurs, afin d'éviter de sélectionner en tant qu'étiages des débits trop élevés (en hiver par exemple), nous avons doublé ce seuil variable d'un second seuil, fixe, lequel permet de filtrer les débits hauts. Cette situation est illustrée par la Figure 5 pour le bassin J, année 1997. Les débits observés et simulés sont affichés respectivement en bleu et vert, et les seuils fixe et variable sont quant à eux représentés respectivement en traits rouges interrompus et continus. Les triangles noirs représentent les débits prévus à une semaine d'avance lorsqu'un étiage est détecté par le modèle, c'est-à-dire quand le débit prévu est inférieur au seuil. On voit que ce système ne déclenche que marginalement des alarmes hors de la saison estivale (par sous-passement du seuil fixe) et que seuls une partie des jours d'été sont considérés comme des étiages du fait de leur caractère inhabituellement bas. La Figure 6 présente plus en détails les hydrogrammes obtenus cette année là durant la période estivale. Dans le cadre de ce travail nous avons fixé ce seuil haut au quantile 25% de l'ensemble des débits de la chronique. Ce choix s'est imposé parce qu'il permet de sélectionner un

nombre d'étiages observés similaire à celui obtenu avec la méthode du seul seuil fixe décrit plus haut. Cette similitude nous permettra de faire des comparaisons entre les résultats obtenus par les deux méthodes de définition du seuil d'étiage.

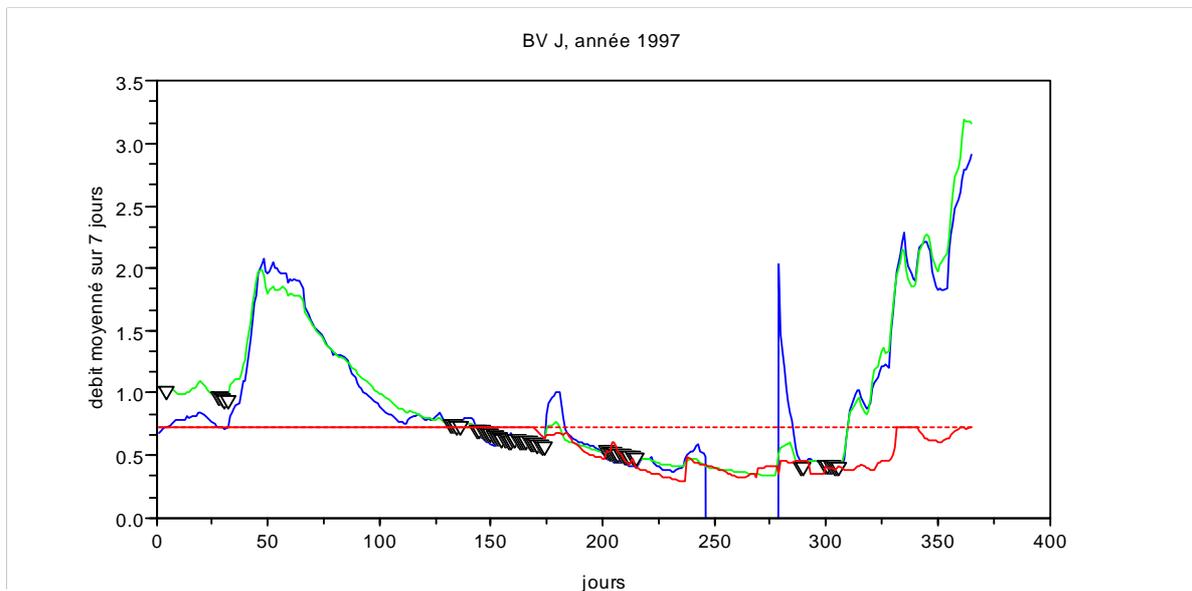


Figure 5: illustration pour toute l'année 1997 sur le bassin J des seuils d'étiages variable (trait rouge continu) et fixe (rouge interrompu) superposés sur les hydrogrammes observé (bleu) et simulé (vert). Les triangles noirs matérialisent les jours associés à des alertes d'étiage.

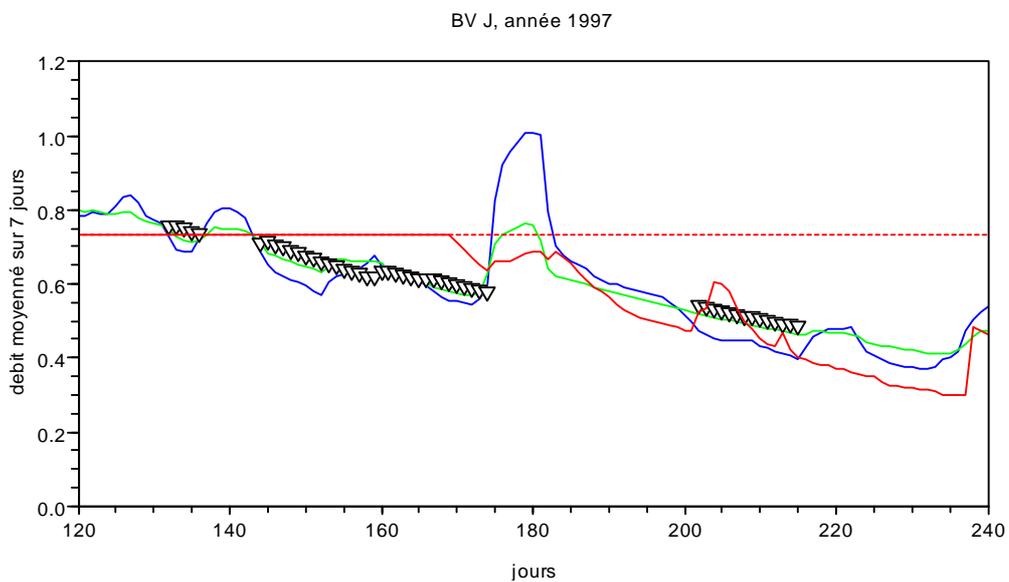


Figure 6: zoom de la figure précédente sur la période mai-août 1997.

VII. RESULTATS DE LA PREVISION D'ETIAGES DEFINIS PAR UN SEUIL FIXE

Le présent chapitre traite uniquement des résultats de prévision obtenus avec une définition de l'étiage basée sur un seuil fixe. Un travail équivalent a toutefois été effectué à partir d'un seuil saisonnier; les résultats correspondants sont rapportés en Annexe de ce document, au §XV.

A. CHOIX DU SEUIL D'ETIAGE

Pour la phase prototype du modèle nous avons travaillé sur des chroniques courtes (11 années) ; il nous fallait donc choisir une définition d'étiage qui ne soit pas trop extrême afin de conserver un nombre suffisant d'évènements sur lesquels travailler. Notre choix s'est porté sur la valeur du QMNA2 (valeur du module mensuel sous-passé un an sur deux), qui nous a permis de constituer des échantillons d'étiages supérieurs à 50 événements pour les mois d'été.

Les valeurs du QMNA2 pour les deux bassins prototypes sont les suivantes :

- QMNA2=0.378 mm/j pour BV A
- QMNA2=0.498 mm/j pour BV J

B. JOURS D'ETIAGES OBSERVES

Compte tenu de ces seuils, nous avons enregistré les quantités suivantes d'évènements d'étiages pour les deux bassins (*Tableau 1*).

Tableau 1 : nombre de jours d'étiages à seuil fixe observés chaque mois pour les deux bassins

Etiages observés	Jan	Fev	Mar	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Dec	Total
BV A	0	0	0	12	5	22	87	81	92	32	7	0	338
BV J	0	0	0	0	0	0	32	75	98	71	65	0	341

Les nombres totaux d'évènements ainsi sélectionnés sont semblables pour les deux bassins. Les étiages interviennent plus tôt dans l'année pour le bassin A, mais disparaissent plus tôt en automne.

Les débits moyens durant ces évènements d'étiages sont reportés pour chaque mois et chacun des bassins dans le *Tableau 2* :

Tableau 2 : débits moyens des jours correspondants à des étiages à seuil fixe observés pour les deux bassins

Débits moyens observés pendant les étiages , mm	Jan	Fev	Mar	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Dec
BV A				0.35	0.36	0.34	0.33	0.31	0.29	0.28	0.34	
BV J							0.46	0.41	0.42	0.43	0.43	

C. APPLICATION DES INDICATEURS A L'EVALUATION DE LA PERFORMANCE DU MODELE « HISTORIQUE »

1. JOURS D'ETIAGES PREVUS PAR LE MODELE H.

Les *Tableau 3* et *Tableau 4* confrontent les nombres d'étiages observés (déjà rapportés au § VII.B) et prévus par le modèle H, respectivement pour les bassins J et A. Globalement H prévoit 3 fois plus de jours d'étiages qu'il n'y en a effectivement d'observés.

Tableau 3 : comparaison des nombres mensuels de jours d'étiage observés et prévus par H pour le bassin J.

BV J	Jan	Fev	Mar	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Dec	Total
Etiages observés	0	0	0	0	0	0	32	75	98	71	65	0	341
Etiages prévus	0	0	0	0	0	0	149	270	225	240	259	0	1143

Tableau 4 : comparaison des nombres mensuels de jours d'étiage observés et prévus par H pour le bassin A

BV A	Jan	Fev	Mar	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Dec	Total
Etiages observés	0	0	0	12	5	22	87	81	92	32	7	0	338
Etiages prévus	0	0	0	76	40	85	264	229	208	150	56	0	1108

2. SCORES DETERMINISTES CONTINUS : ECARTS ENTRE DEBITS OBSERVES ET PREVUS.

La Figure 7 expose, pour le bassin A et pour chaque mois de l'année présentant des étiages prévus ou observés, la distribution des écarts bruts (en mm) entre débits prévus par H et débits observés. Les écarts absolus sont minimaux en été et plus importants au printemps et en automne, essentiellement du fait que les jeux de prévisions impliquent des débits plus élevés. L'absence totale de biais est attendue puisque pour le modèle H les prévisions correspondent aux données observées !

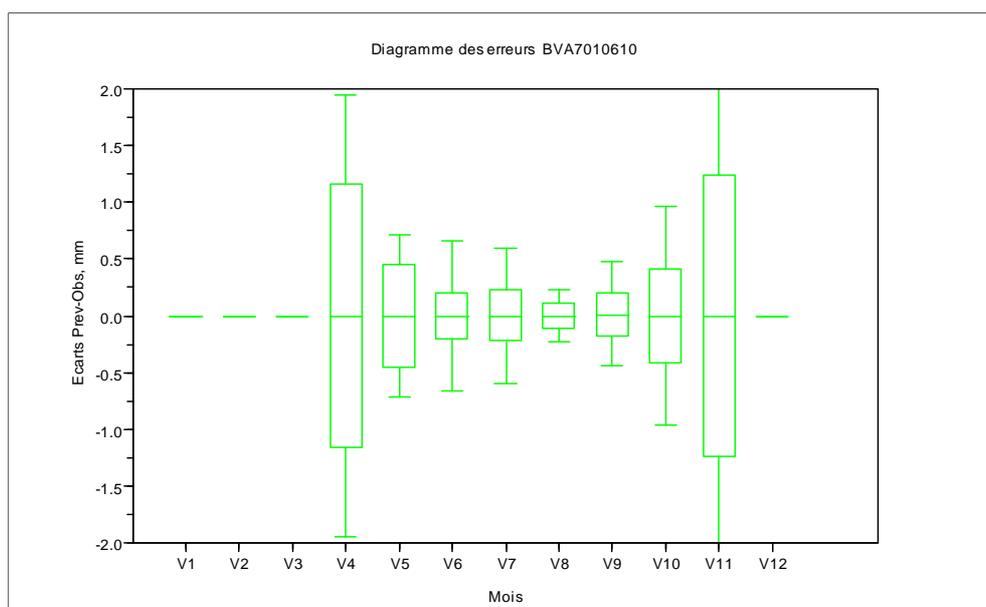


Figure 7 : distribution des écarts entre débits prévus par le modèle H et observés pour les mois à étiage pour le bassin A. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.

Le *Tableau 5* restitue les valeurs relatives de la RMSE totale, de la RMSE sur la médiane des prévisions et de l'écart type sur les prévisions. Il apparaît de ce tableau que la RMSE totale mensuelle s'explique à parts sensiblement égales par l'écart de la médiane et par la dispersion autour de celle-ci.

Tableau 5 : scores d'erreur relatives mensuelles obtenus avec le modèle H pour le bassin A.

BV A	avril	mai	juin	juillet	août	septembre	octobre	novembre
RMSE tot%	112	50	102	90	62	119	96	143
RMSE Médiane %	99	35	74	66	45	85	74	108
E.Type%	76	34	68	62	42	83	70	96

Les mêmes informations pour le bassin J sont rapportées dans la Figure 8 pour les écarts bruts, et dans le *Tableau 6* pour les erreurs relatives.

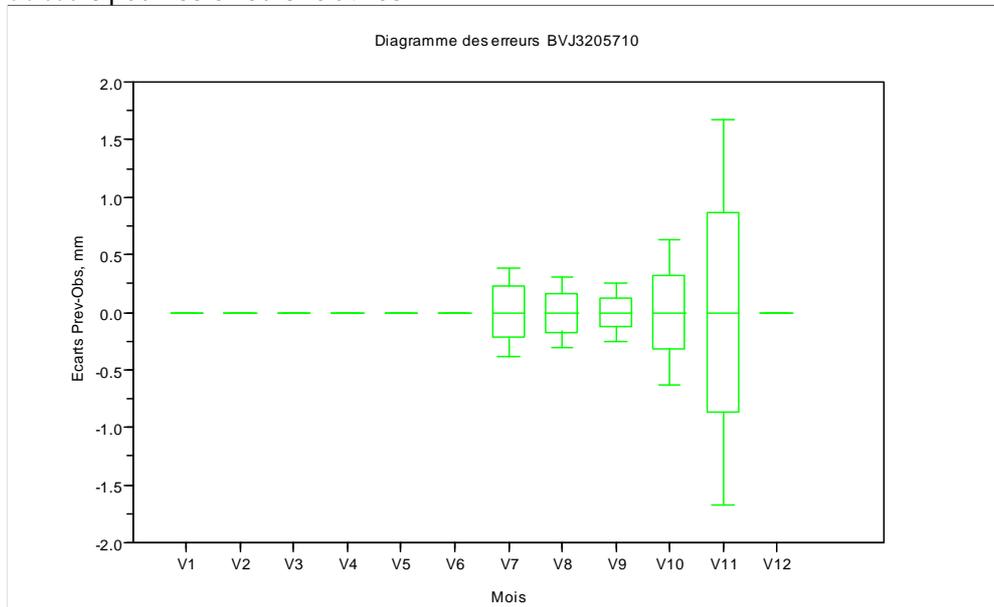


Figure 8 : distribution des écarts entre débits prévus par le modèle H et observés pour les mois à étiage pour le bassin J. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.

Tableau 6 : scores d'erreur relatives mensuelles obtenus avec le modèle H pour le bassin J.

BV J	juillet	août	septembre	octobre	novembre
RMSE tot%	40	39	44	76	103
RMSE Médiane %	30	29	32	60	78
E.Type%	28	27	29	56	71

Les erreurs sont sensiblement moindres pour le bassin J que pour le bassin A, ce qui s'explique par des variations interannuelles de débit plus faibles.

3. SCORE PROBABILISTE : RPS.

Les scores RPS obtenus par le modèle H, rapportés dans les *Tableau 7* et *Tableau 8* pour les 2 bassins, sont quasi identiques (tous les scores se situent entre 0.22 et 0.23), quel que soit le mois ou le bassin. Cette similitude n'est pas fortuite, mais tient au fait que, pour chaque mois et chaque bassin, les 3 catégories de débits supportant le score RPS utilisé sont définies à partir des mêmes données que le modèle H, à savoir la chronique des débits historiques.

Tableau 7 : scores RPS mensuels obtenus par le modèle H pour le bassin A

BV A	avril	mai	juin	juillet	août	septembre	octobre	novembre
RPS	0.22	0.23	0.22	0.23	0.22	0.23	0.22	0.23

Tableau 8 : scores RPS mensuels obtenus par le modèle H pour le bassin J

BV J	juillet	août	septembre	octobre	novembre
RPS	0.22	0.22	0.22	0.22	0.23

Cette constance du RPS pour H invite à considérer que ce score est un outil simple pour juger de la capacité de TAGE-GR4J à produire des prévisions meilleures que H.

4. SCORES DETERMINISTES PAR CATEGORIES : PERFORMANCES DE LA PREVISION EN TERMES D'ALERTE D'ETIAGE.

Le *Tableau 9* rapporte les valeurs des scores d'alertes obtenus par H pour le bassin A. Les taux de fausses alertes (FAR) sont élevés et varient entre 60-70% en été, jusque 100% en mai et novembre, mois pour lesquels les biais sont très forts. Les valeurs de CSI sont acceptables en été mais fortement abaissées en ETS, ainsi qu'en indice de Rousseau ; le mois de juin montre les meilleurs scores pour l'ETS et l'IR.

Tableau 9 : scores mensuels d'alertes pour le bassin A obtenus avec le modèle H

BV A	avril	mai	juin	juillet	août	septembre	octobre	novembre
FAR%	89	100	76	68	66	62	82	100
POD%	67	0	91	98	98	87	84	0
CSI%	10.0	0.0	23.0	32.0	34.2	36.4	17.4	0.0
ETS%	6.4	-1.5	17.1	4.5	10.2	10.4	7.9	-2.1
Biais	6.33	8.00	3.86	3.03	2.83	2.26	4.69	8.00
IR	0.04		0.24	-0.24	-0.01	0.07	-0.01	

Les mêmes scores pour le bassin J sont présentés dans le Tableau 10. Les taux de fausses alertes sont un peu moins mauvais que pour le bassin A, les étiages étant par ailleurs quasiment tous détectés. L'ETS abat fortement les scores honorables obtenus en CSI. Septembre montre les meilleurs scores ETS et IR.

Tableau 10 : :scores mensuels d'alertes pour le bassin J obtenus avec le modèle H

BV J	juillet	août	septembre	octobre	novembre
FAR%	79	73	56	73	78
POD%	97	99	100	90	89
CSI%	20.7	27.3	43.6	25.9	21.8
ETS%	11.3	3.2	16.2	3.8	0.9
Biais	1.31	2.87	2.23	2.31	2.05
IR		-0.34	0.15	-0.22	-0.39

Le Figure 9 illustre pour le bassin A les nombres jours d'étiages effectivement observés avec étiages vrais détectés par le modèle H, ainsi que les fausses alertes et les étiages non détectés. On voit que le prix à payer d'une détection de quasi tous les événements d'étiage est un nombre élevé de fausses alertes. La proportion de fausses alertes est globalement d'autant plus élevée que le nombre d'étiages vrais est faible.

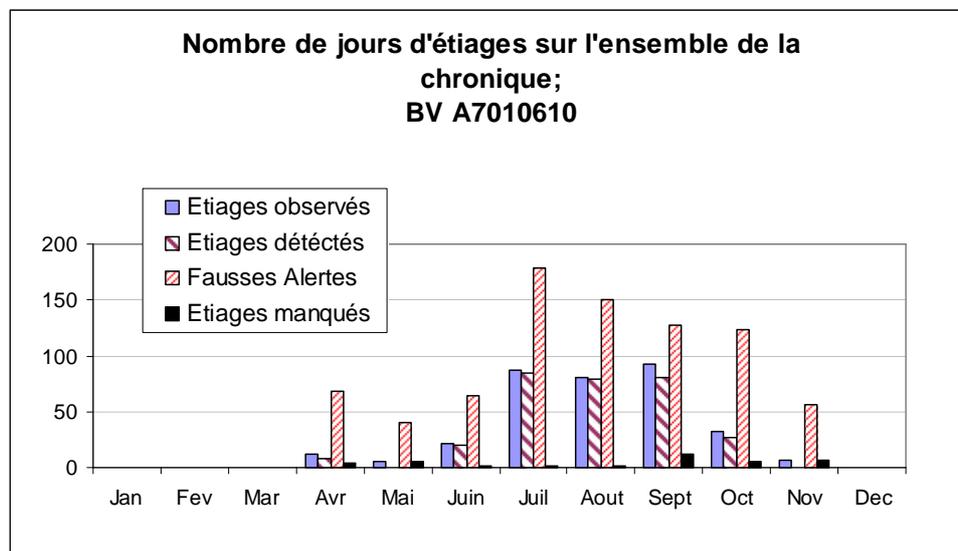


Figure 9: analyse des étiages prévus par le modèle H par comparaison avec les étiages observés pour le bassin A.

D. PERFORMANCES DE TAGE ASSOCIE A GR4J

1. PARAMETRISATION

Ainsi nous l'avons déjà exprimé au § IV.D, la période cible expérimentée ici est de 7 jours, et la prévision s'étend entre 1 et 20 périodes cibles, soit au maximum 5 mois. Pour ce qui concerne le

critère d'alerte, nous avons choisi de prendre le même que celui adopté pour le modèle H (cf Annexe §XII), ce qui permettra une comparaison directe des performances de prévision.

2. RESULTATS BRUTS ET INTERPRETATIONS.

Dans cette section nous présenterons en détails les résultats obtenus pour le bassin J. Nous y renseignons aussi les résultats principaux obtenus pour le bassin A; les détails des scores obtenus pour ce bassin A sont consignés en Annexe de ce document, §XIII.

a. **Jours d'étiages prévus par le modèle TAGE-GR4J.**

La Figure 10 illustre les nombres totaux de jours d'étiage prévus par TAGE-GR4J en fonction de l'horizon de prévision exprimé en nombre « n » de semaines, pour les deux bassins. Globalement, ce nombre d'étiages prévus croît avec n. Cela s'interprète par le fait que la dispersion des débits prévus pour un jour cible j s'accroît quand la prévision est faite avec une anticipation de plus en plus grande. La gamme des débits prévus étant plus élevée, il y a plus de chance que l'un de ces débits passe sous le seuil d'étiage, d'où le déclenchement d'une alerte supplémentaire pour ce jour cible.

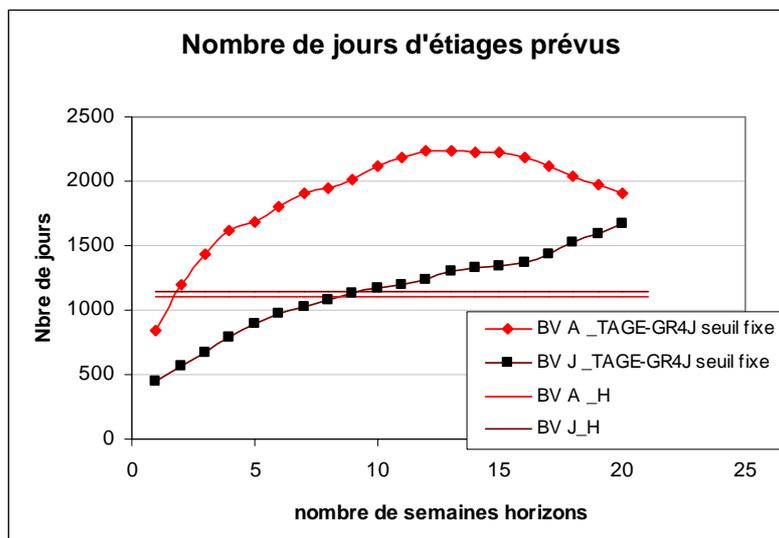


Figure 10: Nombre de jours d'étiages prévus par TAGE-GR4J pour les deux bassins en fonction de l'horizon de prévision. Comparaison avec le modèle H.

L'augmentation de la dispersion des débits prévus traduit en fait la perte de l'effet mémoire du bassin : pour des prévisions à courte échéance, la gamme des débits prévisibles est fortement influencée par l'état initial des réservoirs au moment de la prévision. Dans ce cas l'étendue de la gamme des précipitations prévues par les scénarios entre la date de prévision et la date cible influe de manière limitée sur les débits observés au jour cible. Par contre lorsque le délai de prévision s'accroît l'influence de l'état des réservoirs au moment de la prévision devient insignifiante, et les précipitations intervenues depuis la prévision deviennent le déterminant principal pour les débits cibles.

De ce point de vue il vaut de s'intéresser à la différence de comportement entre les deux bassins dans la Figure 10. Pour le bassin A, le nombre d'étiages prévus s'accroît très rapidement, ce qui laisse imaginer un effet mémoire très limité, c'est-à-dire des temps de décharge des réservoirs très courts. Au contraire pour le bassin J, la montée est nettement plus lente, d'où la présomption d'un effet mémoire plus important. La Figure 10 rapporte aussi le nombre de jours d'étiage prévus par le modèle

H. Si le nombre d'étiages prévus était entièrement déterminé par la dispersion des débits prévus, on s'attendrait dans l'idéal à ce que ce nombre constitue une limite supérieure pour le nombre d'étiages prévus par les modèles hydrologiques : le modèle H correspond au cas d'un effet mémoire totalement annulé. Toutefois d'autres facteurs entrent en jeu, comme par exemple le biais systématique du modèle : plus le modèle hydrologique sous-estime les débits en période d'étiage, plus le nombre de fausses alertes croîtra, et avec elles le nombre d'étiages prévus. Ce biais interne du modèle (nous ne parlons pas ici du biais à grande échelle), dont nous verrons qu'il est plus important pour le bassin A que pour le bassin J, explique certainement pourquoi le dépassement de la limite fixée par le modèle H est plus important pour le bassin A que pour le bassin J.

La Figure 11 décompose le nombre d'étiages prévus par mois de l'année en fonction de l'horizon de prévision, pour le bassin J. Les étiages effectivement observés sont également représentés. On voit qu'à mesure que l'horizon de prévision s'accroît, non seulement le nombre d'alertes croît pour les mois à étiage (de juillet à novembre), mais des fausses alertes sont créées également pour des mois sans étiage, tels décembre, puis juin et enfin mai. Pour les horizons les plus élevés, le nombre d'alertes approche les 300, ce qui correspond à environ 30 alertes par mois sur les 10 années de chronique considérées : le niveau de discrimination des situations d'étiage est alors clairement nul.

Jours d'étiages prévus par TAGE. BV J

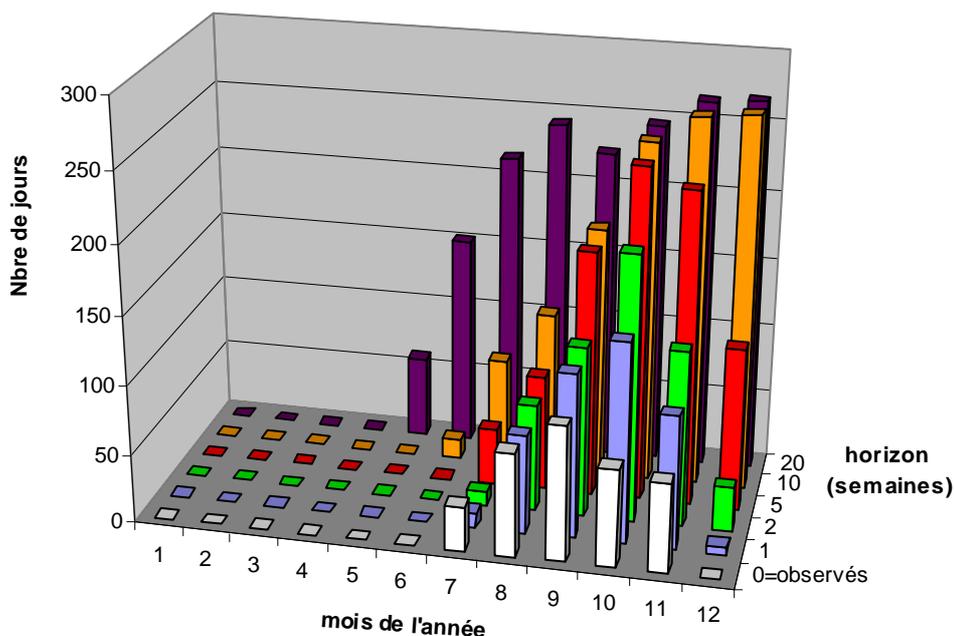


Figure 11: Nombres mensuels de jours d'étiage observés et prévus par TAGE-GR4J en fonction de l'horizon de prévision, pour le bassin J.

b. Scores déterministes continus : écarts entre débits observés et prévus.

Les Figure 12: et Figure 13 représentent respectivement les RMSE relatives mensuelles (cf. définition § V.A.5) et les écarts relatifs moyens mensuels (cf. définition § V.A.1) pour le bassin J. Les RMSE sont nettement inférieures à celles obtenues ci-dessus avec le modèle H. A l'exception de septembre, la RMSE croît lentement mais indéfiniment à mesure que l'horizon croît, ce qui supporte l'idée que l'effet mémoire contenu dans le modèle hydrologique aide à la description anticipée des débits.

Par comparaison avec le bassin A (cf. Annexe de ce document, §XIII), les écarts sont nettement moindres en valeurs absolues (<30%) et leur signe est différent selon les mois et les horizons. Les RMSE sont aussi inférieures à celles du bassin A, notamment en été.

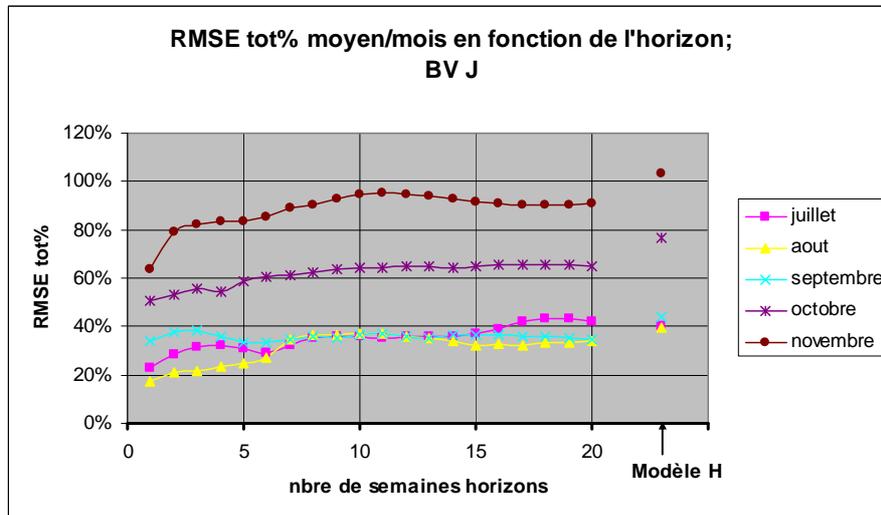


Figure 12: RMSE totales mensuelles relatives pour les mois à étiages du bassin J selon l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

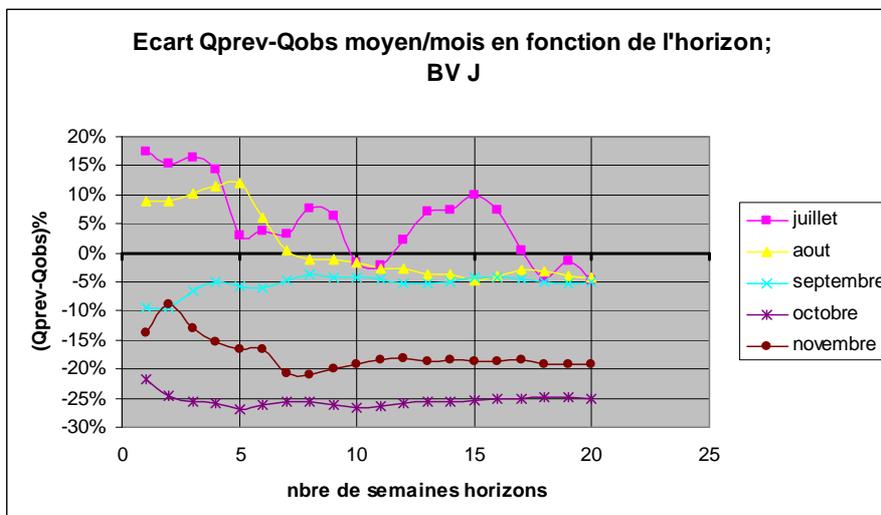


Figure 13: Ecart mensuels relatifs entre débits prévus et observés pour les mois à étiages du bassin J selon l'horizon de prévision

Les Figure 14 et Figure 15 restituent les distributions des écarts pour le bassin J, selon le modèle H (en vert), et selon TAGE-GR4J (en rouge) respectivement pour des horizons d'une semaine et deux mois. On constate que les écarts TAGE-GR4J et leurs distributions restent significativement moindres que celles données par le modèle H, et ce jusqu'à deux mois de délais de prévision (n=8 semaines) dans le cas des mois d'été, qui sont par ailleurs les plus significatifs en matière d'étiage. Ces figures illustrent qualitativement l'intérêt de la modélisation GR4J pour les prévisions d'étiage à long terme dans le cas de ce bassin.

Diagramme des erreurs BVJ3205710

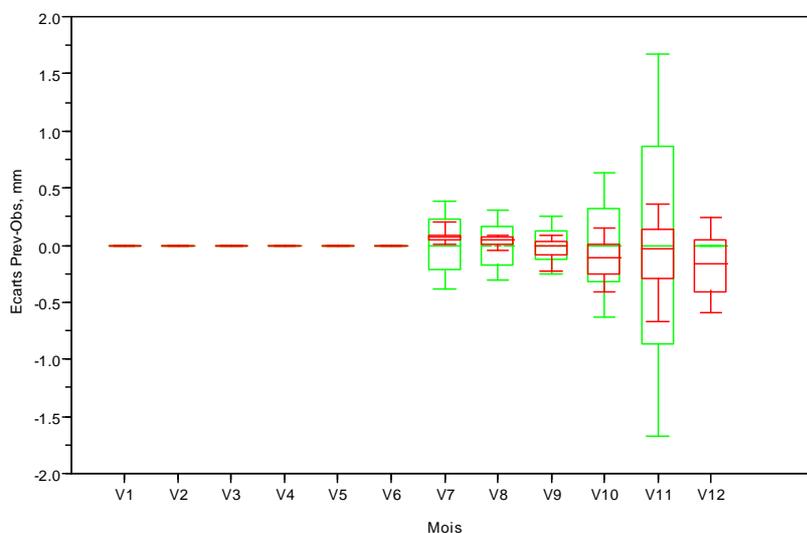


Figure 14: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la **première** semaine d'horizon, bassin J. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.

Diagramme des erreurs BVJ3205710

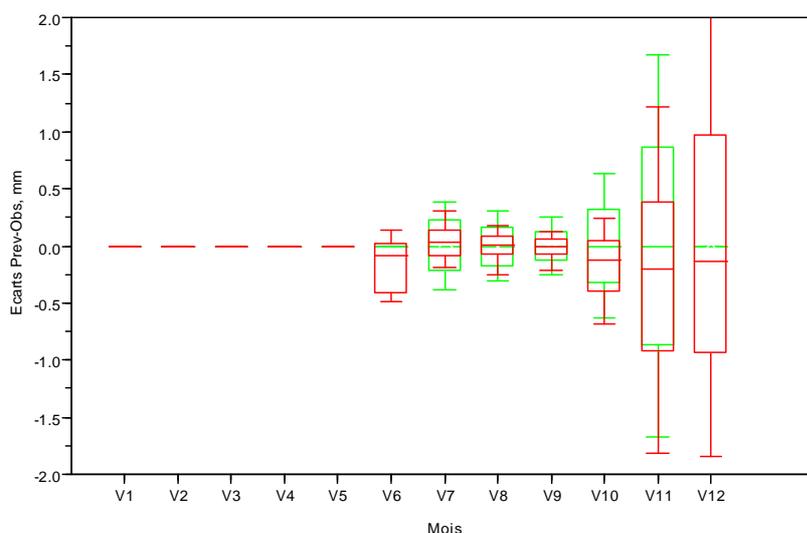


Figure 15: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la **huitième** semaine d'horizon, bassin J. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.

c. Score probabiliste : RPS.

Les scores RPS mensuels obtenus pour le bassin J sont données dans la Figure 16. On observe une dégradation progressive du score pour les mois d'été et pour novembre. Toutes les courbes semblent tendre vers une même valeur limite aux alentours de 0.20-0.22. Pour le mois d'octobre, l'écart moyen important enregistré dès les premiers horizons (cf. Figure 13) est responsable d'un score RPS élevé. Le même phénomène ne se produit pas en novembre du fait de l'amplitude importante de la catégorie RPS centrale.

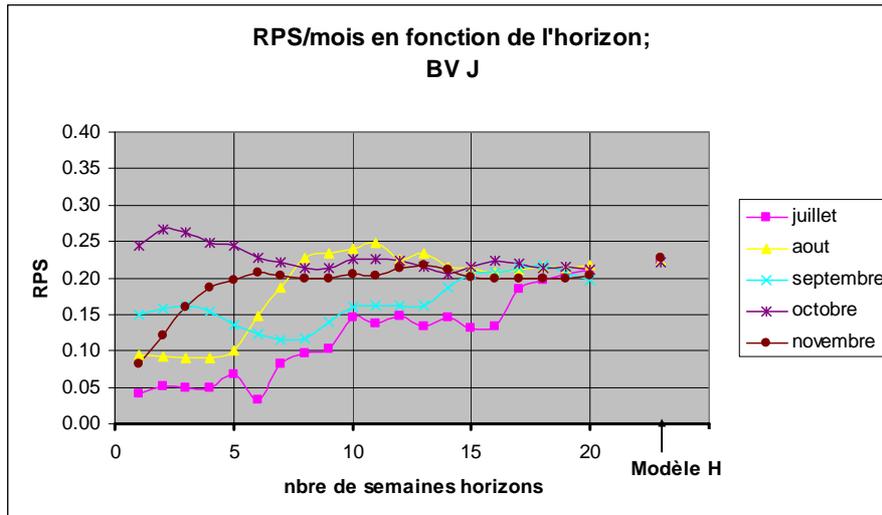


Figure 16: scores RPS mensuels pour le bassin J en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

En se souvenant que les valeurs du RPS obtenues pour les deux bassins avec le modèle H étaient systématiquement de 0.22-0.23, on peut traduire la tendance des courbes RPS de TAGE-GR4J comme une dégradation des prévisions vers une performance aussi médiocre que H. Cette dégradation est rapide dans le cas du bassin A, et plus lente pour J.

Par comparaison, les scores RPS obtenus pour le bassin A (cf. Annexe de ce document, §XIII.B) atteignent leurs maximum pour des horizons de zéro à deux semaines, avec des valeurs de RPS particulièrement élevées en été, indiquant que le score RPS constitue un indicateur particulièrement sévère et sensible aux biais des modèles pour les périodes sèches.

d. Scores déterministes par catégories : performances de la prévision en termes d'alertes d'étiage.

Pour le bassin J, les Figure 17 à Figure 22 tracent respectivement les évolutions des indicateurs mensuels suivants, définis au § V.B :

- taux de fausses alertes FAR
- probabilité de détection POD
- indice de succès critique CSI
- score de menace équitable ETS
- biais à grande échelle B.
- indice de Rousseau IR

Le taux de fausses alertes FAR croît progressivement avec l'horizon h pour les mois d'été, ce qui traduit une dégradation lente des performances du modèle. Les valeurs maximales atteintes sont similaires à celles obtenus avec le modèle H (cf. Tableau 10).

La probabilité de détection des étiages POD croît elle aussi progressivement; elle part néanmoins de valeurs élevées, sauf pour le mois de juillet. Ce fait peut être mis en regard de la Figure 13 qui montrait que TAGE GR4J surestime sensiblement les débits d'étiage en juillet, ce qui explique sa difficulté relative à les détecter en mode de prévision probabiliste. Aux horizons lointains, les POD ont des valeurs semblables à celles du modèle H.

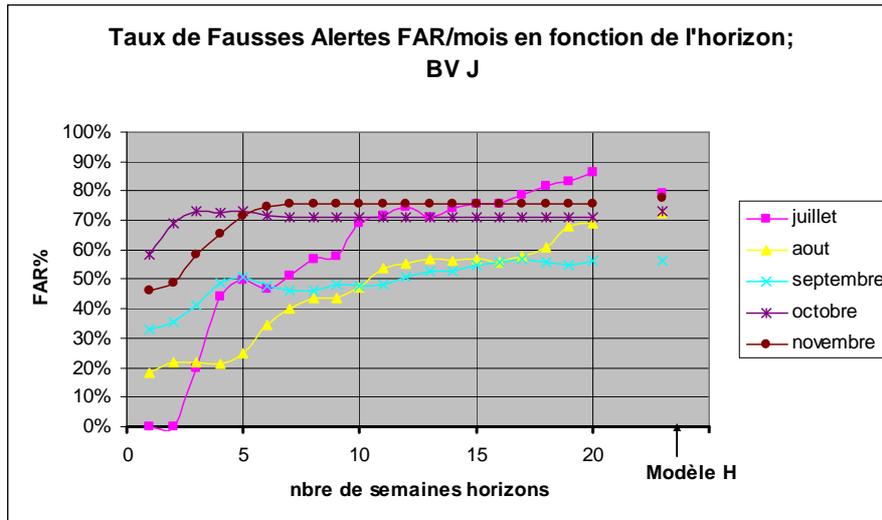


Figure 17 : taux mensuels de fausses alertes pour le bassin J en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

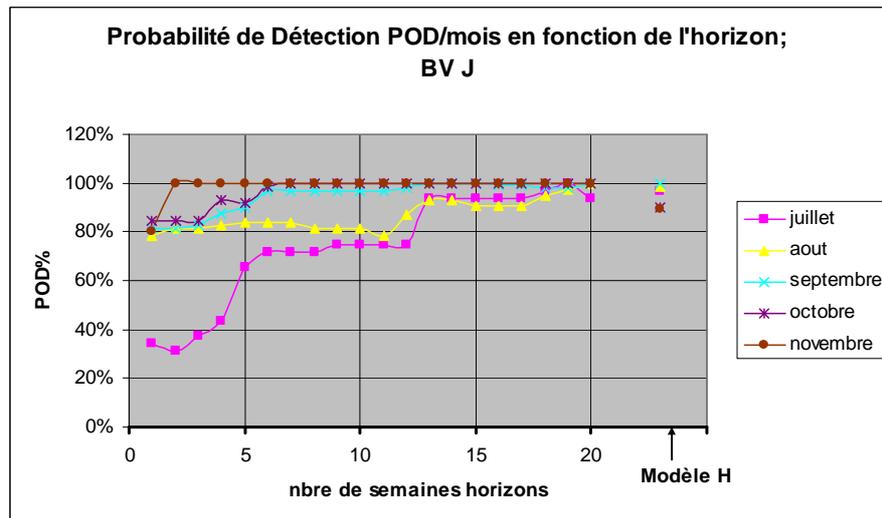


Figure 18 : probabilité de détection des étiages pour le bassin J en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

Les valeurs de CSI décroissent progressivement. Elles sont très supérieures à celles du modèle H, sauf pour les derniers horizons où les valeurs sont semblables. Il convient toutefois de considérer les valeurs de CSI pour les mois à fortes densités d'étiages (c-à-d tous les mois ici représentés, sauf juillet) avec précaution, le CSI n'étant pas adapté à ce type de situations.

L'ETS reste élevé pour les mois d'été, mais abat fortement les scores des autres mois comparativement au CSI. Cela traduit le fait que les performances intrinsèques de TAGE-GR4J, indépendamment des prévisions justes dues au hasard, sont meilleurs en été pour ce bassin. Les valeurs obtenues aux derniers horizons ressemblent de près à celles obtenues avec le modèle H.

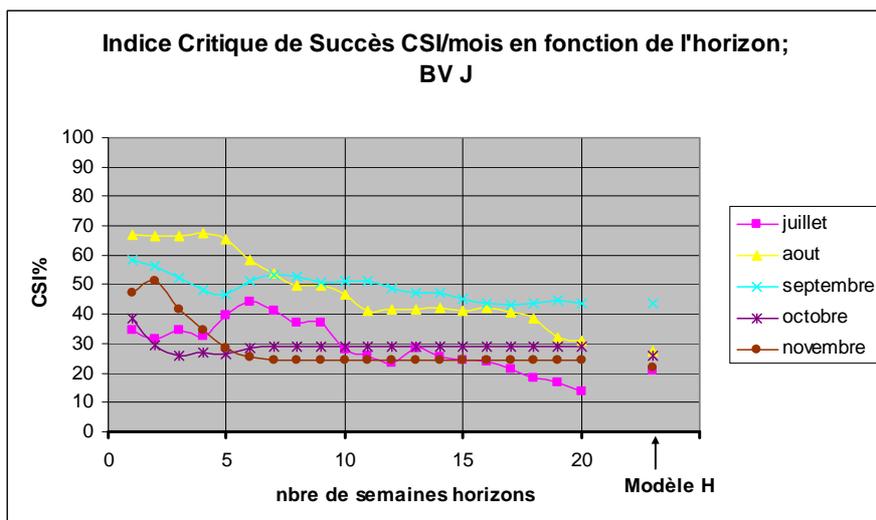


Figure 19 : Indices mensuels de Succès Critique pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.

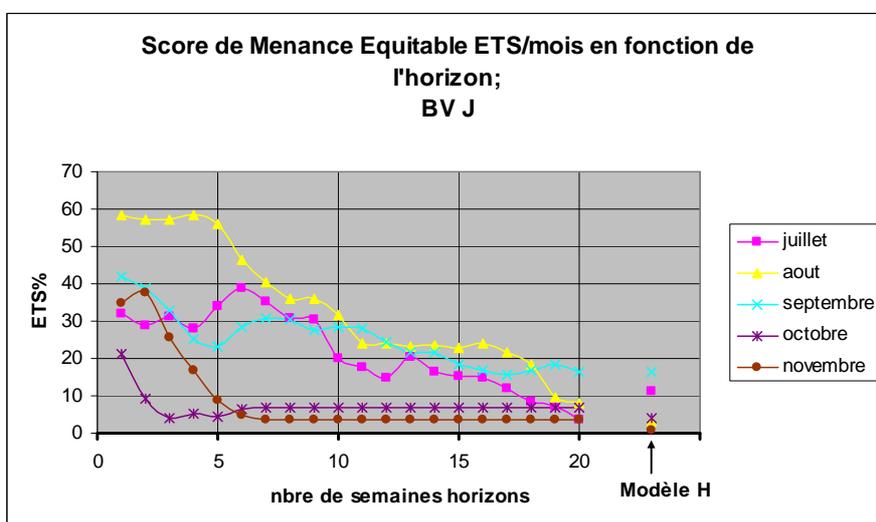


Figure 20 : Scores mensuels de Menaces Equitables pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.

Le biais à grande échelle B est une fonction croissante de l'horizon h , de manière monotone pour les mois d'été, et avec une saturation au-delà de 6 semaines pour l'automne. En juillet, pour les premiers horizons, le biais est nettement inférieur à l'unité, ce qui signifie qu'une bonne part des étiages ne sont pas détectés. Ce fait peut être mis en regard de la Figure 13 qui montrait que TAGE GR4J surestime sensiblement les débits d'étiage en juillet, ce qui explique sa difficulté relative à les détecter en mode de prévision probabiliste. Pour des horizons plus élevés la dispersion des prévisions prend le pas sur l'écart moyen, et le biais croît. B atteint des valeurs semblables à celles vues avec le modèle H pour août et septembre, et des valeurs quelques peu supérieures pour les mois d'automne, constat à relier au fait que TAGE-GR4J sous-estime sensiblement les débits d'étiage à ces périodes (cf. Figure 13).

Les indices de Rousseau obtenus pour les mois d'été restent positifs même pour des horizons de 4 à 5 mois. Pour les mois d'automne, l'indice tombe vite sous zéros, du fait du biais à grande échelle important.

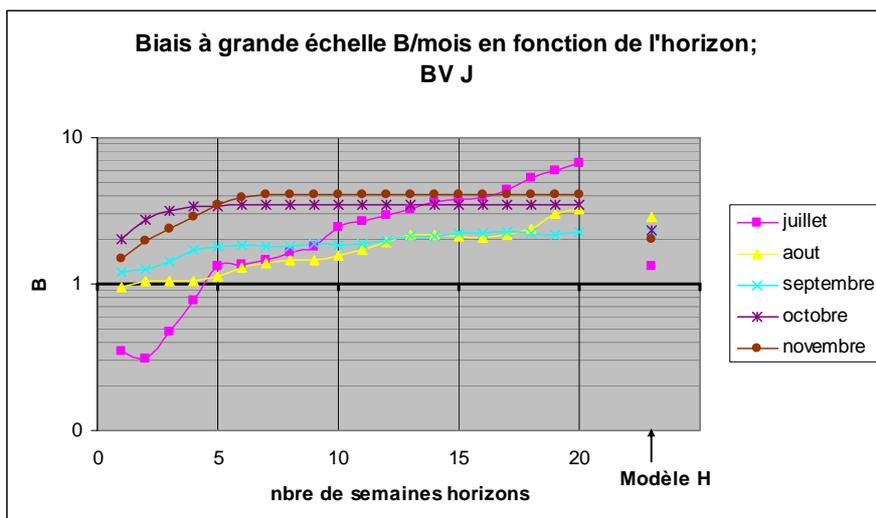


Figure 21 : Biais à grande échelle mensuels pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.

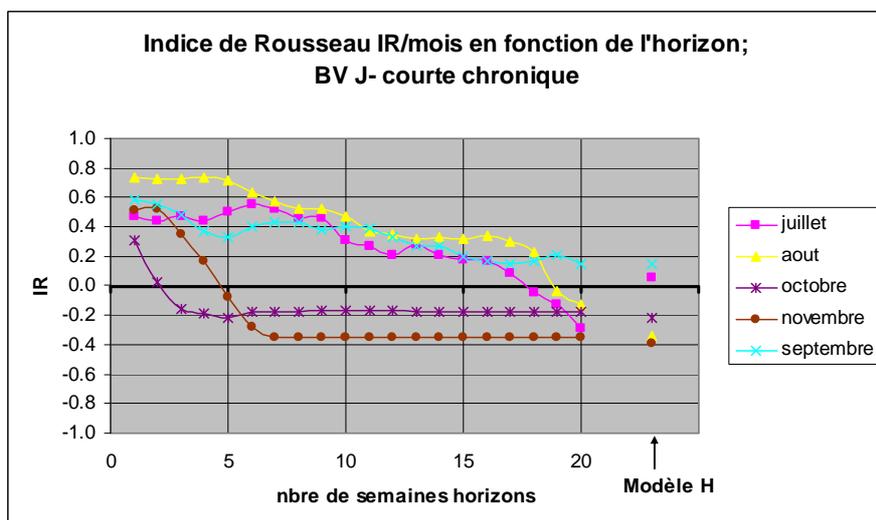


Figure 22 : Indices mensuels de Rousseau pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.

Par comparaison le bassin A a les caractéristiques suivantes (cf. détails en Annexe de ce document, §XIII.C):

- les taux de fausses alertes sont nettement supérieurs, surtout aux horizons de prévision les moins lointains
- les valeurs de CSI sont plus faibles.
- les biais sont néanmoins très nettement supérieurs à ceux vus pour le bassin J, en dépit de densités d'étiages comparables en été.
- les Indices de Rousseau sont très inférieurs à ceux obtenus pour le bassin J. Pour comparaison avec la Figure 22, la Figure 23 montre les résultats obtenus pour l'indice IR avec le bassin A. La différence de comportement est flagrante, l'indice s'écroulant très rapidement au-dessous des valeurs similaires ou inférieures à celles obtenus avec le modèle H, ce qui est symptomatique d'un effet mémoire très limité.

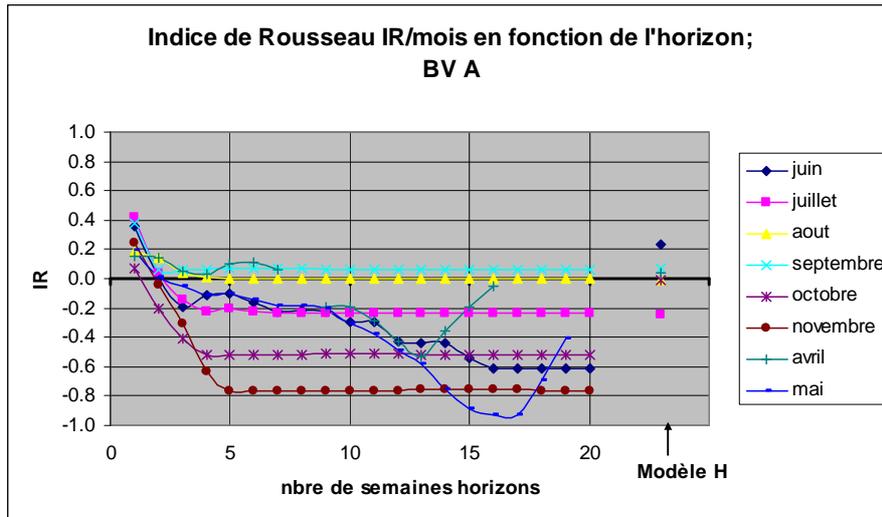


Figure 23 : Indices mensuels de Rousseau pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin A. Le dernier point de chaque mois correspond à la performance du modèle H.

D'une manière générale, pour le bassin A, les scores sont aussi mauvais ou pires que ceux obtenus avec le modèle H dès que l'on dépasse le deuxième horizon.

e. Vérification des propriétés de l'indice de Rousseau.

Ainsi que nous l'avons précisé lors de la définition de l'indice de Rousseau, ce dernier pénalise à la fois les prévisions dont la composante « chance » est importante (comme le fait l'ETS) ainsi que les prévisions qui présentent un biais à grande échelle important. La Figure 24 restitue les valeurs de IR obtenues pour tous les mois, tous les horizons de prévisions, et pour les deux bassins versants A et J, en fonction de la valeur du biais à grande échelle. Comme attendu, les meilleures valeurs de l'indice sont obtenues pour des biais proches de 1, c'est-à-dire pour des prévisions non biaisées à grande échelle, telles celles obtenues en août pour le bassin J aux petits horizons de prévision. De part et d'autre de la valeur $B=1$, on s'attend à une diminution de l'indice, ce que la figure confirme. Toutefois on observe que l'indice diminue plus ou moins vite selon les séries de données considérées. Cela est à mettre au crédit de la propriété qu'à l'indice de pénaliser davantage les prévisions pour lesquelles la composante « chance » est importante. Or la probabilité de produire des alertes justes « par chance » est accrue si le nombre d'étiages vrais est important. Les nombres en encadrés superposés à certaines courbes dans la Figure 24 correspondent aux nombres d'étiages observés pour les mois en question. On observe que plus ce nombre est élevé, c'est-à-dire plus le facteur « chance » est grand, plus IR décroît rapidement avec le biais.

Pour résumer, on peut dire que l'indice de Rousseau pénalise les prévisions biaisées, et ce d'autant plus sévèrement que la période considérée est significative en termes de fréquence des étiages. En soi cela pourrait s'avérer problématique puisque les performances des prévisions en période « humides » sont ainsi évaluées de manière plus optimiste que celles en périodes sèches. Toutefois il apparaît qu'en pratique le problème ne pose pas, puisque la valeur du biais pour les périodes humides s'avère nettement plus élevée qu'en périodes sèches, de sorte que le facteur « chance » ne compense pas cette différence.

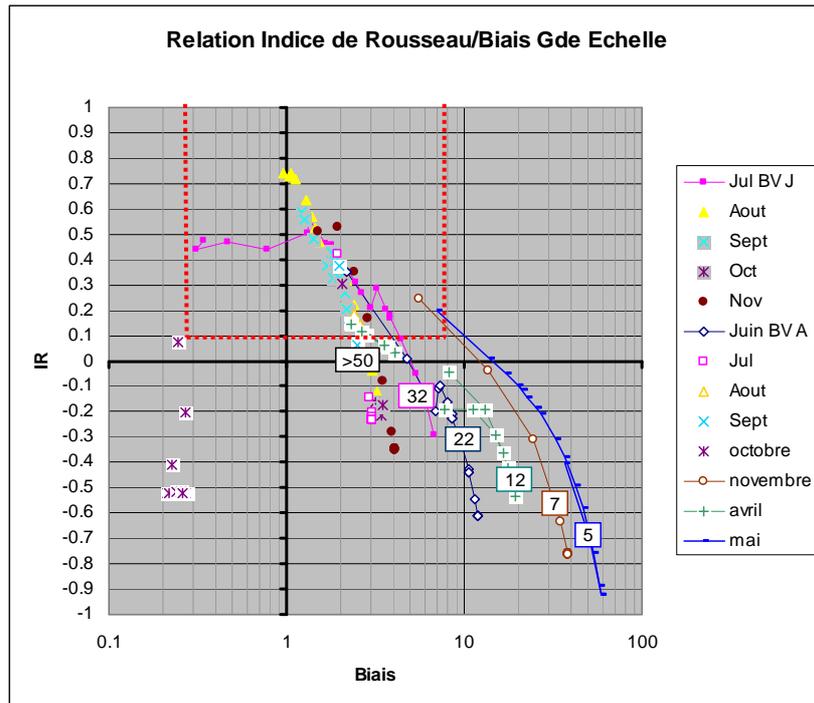


Figure 24: diagramme de corrélation entre l'indice de Rousseau et le biais à grande échelle pour les deux bassins A et J (seuil fixe). Les chiffres en encadré représente le nombre d'étiages observés pour les mois correspondants. La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau > 0.1

A partir de la Figure 24 on peut établir une valeur limite empirique de IR en-deçà de laquelle la prévision peut être jugée inutile, parce que trop biaisée. A titre expérimental, nous avons dans ce travail opté pour la limite $IR=0.1$. Les traits à tirets rouges dans la Figure 24 mettent en lumière le domaine des biais associés à cette limite inférieure. Pour nos deux bassins, 90% des points présents dans ce domaine correspondent à des biais inférieurs à 4, cette valeur $B=4$ correspondant par ailleurs à la plus mauvaise valeur obtenue pour B avec le modèle H pour une période comprenant un nombre d'étiage significatifs (mois de juin pour le BV A). Choisir la limite $IR=0.1$ revient donc à considérer la performance $B < 4$ comme un pré-requis pour que 90% des prévisions soient « utiles ».

La dimension d'utilité est plus facile à appréhender avec l'indice de succès critique CSI, qui comme nous l'avons vu s'exprime simplement comme le ratio entre alertes vraies et nombre total d'alertes. La Figure 25 rapporte le diagramme de corrélation entre IR et CSI, pour les mêmes données que celles de la figure précédente. On y voit que la limite $IR > 0.1$ se traduit empiriquement par des valeurs $CSI > 10\%$ si l'on considère toutes les données, et $CSI > 20\%$ pour 90% d'entre elles. (NB : la valeur $CSI > 20\%$ n'est pas un critère de tri : beaucoup de séries de données de la Figure 25 montrent un $CSI > 20\%$ mais des $IR < 0.1$). La valeur $CSI > 20\%$ nous paraît correspondre à une condition nécessaire « d'utilité » raisonnable.

En tout état de cause, des critères plus fondés et plus exigeants pourront être définis par la suite sans que l'outil TAGE n'ait à en être modifié : ce critère n'est que de l'un de ses paramètres.

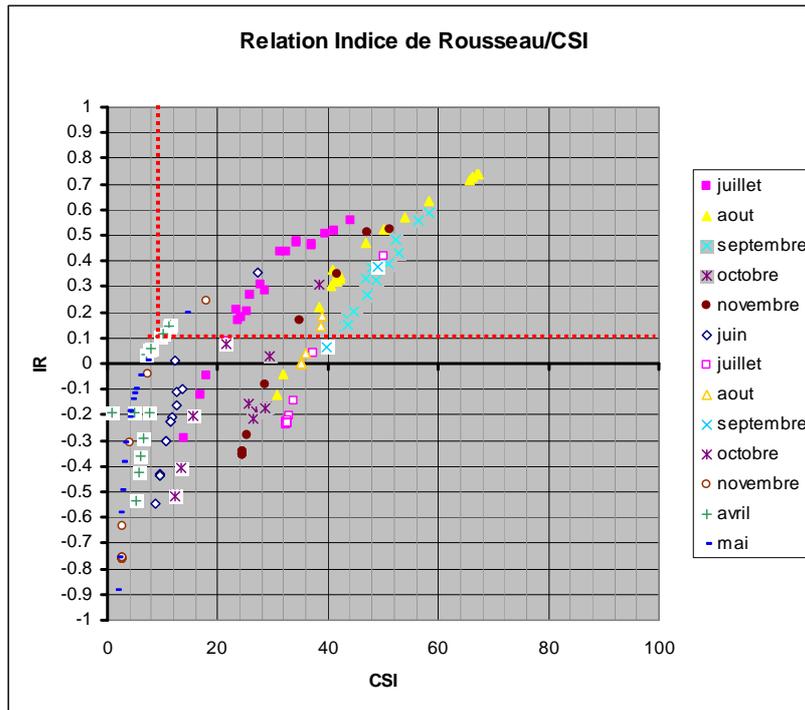


Figure 25: diagramme de corrélation entre l'indice de Rousseau et l'indice de succès critique pour les deux bassins A et J (seuil fixe). La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau > 0.1

3. SCORES DE COMPETENCE DE TAGE-GR4J VS. H.

Pour les deux bassins A et J, les scores de compétences relatifs aux scores probabilistes RPS du modèle TAGE-GR4J en référence au modèle H sont rapportés dans les Figure 26 et Figure 27.

Pour le bassin A, on observe pour les mois de printemps et juillet que le score passe rapidement sous zéro et subit des fluctuations ultérieures, sans jamais toutefois remonter à la valeur du premier horizon. Pour les autres mois, aucune tendance n'est observée, la compétence étant nulle ou négative dès le premier horizon. Pour les mois d'août/septembre, la compétence est particulièrement négative, indiquant une performance du modèle TAGE-GR4J fortement inférieure au modèle H. L'explication tient aux valeurs de RPS pour TAGE-GR4J qui sont plus élevées que pour H durant ces mois, sujet déjà traité au § VII.D.2.c

Pour le bassin J, la compétence $Comp_{RPS}$ de TAGE-GR4J décroît progressivement pour rejoindre une valeur nulle à des horizons allant de 2 mois (août) jusque 4 mois (juillet). Le score d'octobre est toujours négatifs ou nul, du fait de la valeur élevée du RPS correspondant pour TAGE-GR4J, comme discuté précédemment (cf. § VII.D.2.c).

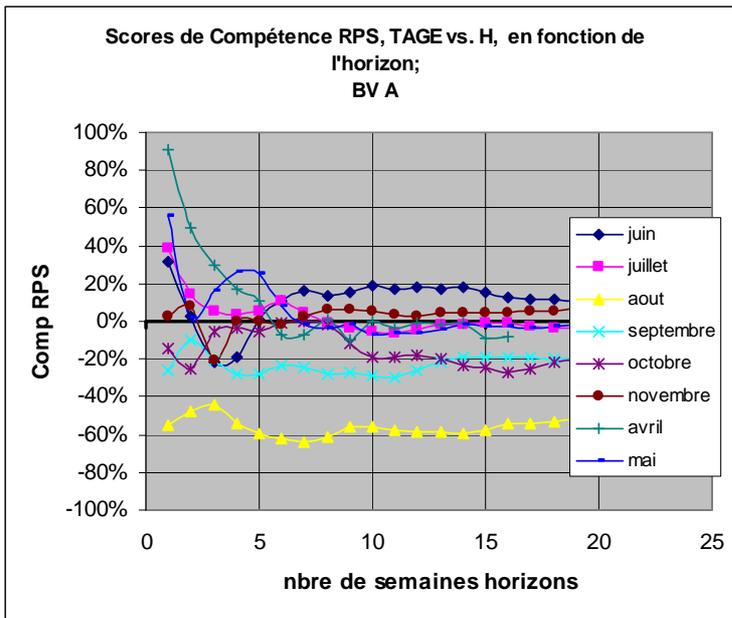


Figure 26: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score RPS pour le bassin A.

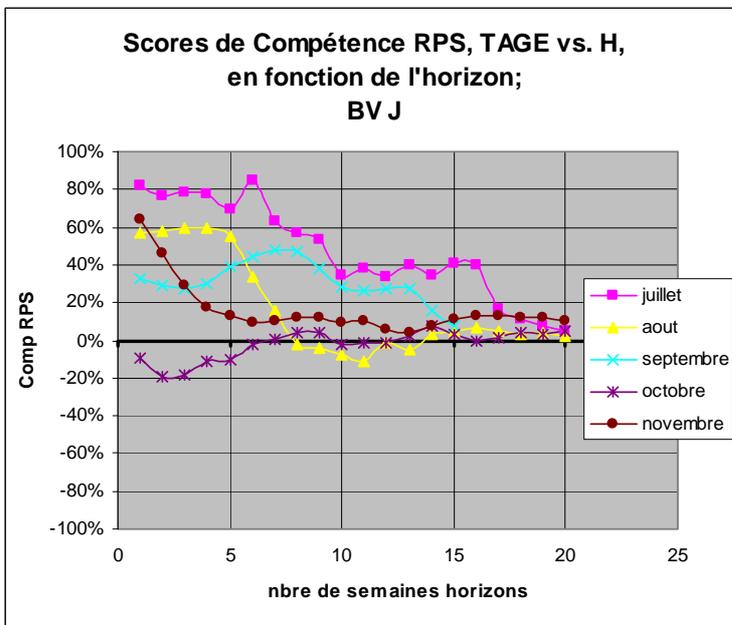


Figure 27: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score RPS pour le bassin J.

Les scores de compétences calculés à partir des scores RMSE et ETS sont présentés et commentés en Annexe de ce document, §XIV.

En synthèse, on constate une grande diversité des performances selon les mois, visible surtout pour le bassin J. TAGE-GR4J est nettement plus performant que H sur le bassin J, notamment en été. La compétence RPS semble exacerber davantage les différences que les autres scores (sensibilité).

La compétence sur RMSE a parfois un comportement asymptotique et reste >0.

La compétence ETS est moins sévère que celle basée sur RPS, probablement car moins sensible aux écarts entre débits prévus et observés.

VIII.DETERMINATION DES HORIZONS MAXIMAUX DE PREVISION UTILE

Dans le cadre de ce travail, nous avons fait le choix de définir l'utilité de l'outil de prévision selon les deux modalités suivantes :

- 1- Si les archives historiques sont utilisables pour effectuer des prévisions (cas des étiages définis par un seuil fixe), TAGE doit être meilleur que H. Cela signifie qu'il doit démontrer un score de compétence significativement positif pour l'indicateur de qualité retenu. Nous avons vu d'après les exemples des bassins A et J que, en dehors de toute tendance, la compétence RPS fluctuait avec une amplitude de l'ordre de 0.1 point. Nous imposerons donc ici une « positivité sécurisée » avec le critère **Score de Compétence >0.1**.
- 2- Le taux de succès du dispositif d'alerte d'étiage, représenté par l'indice CSI, ne doit jamais être inférieur à une valeur donnée « raisonnable ». Dans le cadre de ce travail d'approche, nous situerons cette limite minimale à CSI=20%, ce qui correspond à la moyenne des CSI obtenus par le modèle H sur nos deux bassins de test. Par ailleurs, il est important que le biais à grande échelle ne soit pas trop élevé (notamment afin de ne pas multiplier le nombre de fausses alertes). On peut par exemple se fixer la limite $B < 4$, 4 étant le biais maximal observé pour le modèle H pour des prévisions remplissant la condition CSI>20%. En observant la Figure 24 et la Figure 25, on s'aperçoit que ces deux conditions CSI>20% et Biais<4 sont remplies à plus de 95% dès que l'on choisit le critère IR>0.1. Nous choisirons donc ce dernier critère **IR>0.1** pour sécuriser la qualité de la prévision du modèle TAGE-GR4J, indépendamment de sa compétence vis-à-vis de H.

L'Horizon de prévision utile maximal, noté **HPUM**, sera donc celui qui remplira **conjointement** les deux conditions :

- **Comp_{RPS}>0.1**
- **IR>0.1**

Nous donnons dans les *Tableau 11* et *Tableau 12*, pour le cas des étiages définis avec un seuil fixe, les résultats concernant les horizons maximaux de prévision pour les deux bassins. A titre indicatif, nous avons également affiché les scores de CSI et de Biais à grande échelle obtenus pour l'horizon HPUM, ainsi que les mêmes quantités obtenues avec le modèle H.

Tableau 11 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin A à **seuil fixe**, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM et avec le modèle H.

		avril	mai	juin	juillet	août	septembre	octobre	novembre
BV A	Horizon max issu de Compétence RPS>0.1	5	6	2	7	0	0	0	2
	Horizon max issu de IR>0.1	2	1	1	1	2	1	1	1
	HPUM (semaines)	2	1	1	1	0	0	0	1
	CSI% (HPUM)	11	14	27	50				18
	CSI%(H)	10	0	23	32	34	36	17	0
	Biais (HPUM)	2.3	7	2.2	2				5.5
	Biais (H)	6.3	8.0	3.9	3.0	2.8	2.3	4.7	8.0

Tableau 12 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin J à **seuil fixe**, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM et avec le modèle H.

		juillet	août	septembre	octobre	novembre
BV J	Horizon max issu de Compétence RPS>0.1	18	7	14	0	5
	Horizon max issu de IR>0.1	16	18	20	1	4
	HPUM (semaines)	16	7	14	0	4
	CSI% (HPUM)	24	54	47		35
	CSI% (H)	21	27	44	26	22
	Biais (HPUM)	3.8	1.4	2.1		2.9
	Biais (H)	1.3	2.9	2.2	2.3	2.0

On constate que c'est le critère $Comp_{RPS>0.1}$ qui prime en août/septembre/octobre ; pour les autres saisons c'est l'indice de Rousseau qui détermine le HPUM.

Les HPUM du bassin J sont très élevés en été (de 7 semaines à 4 mois); ils sont systématiquement très faibles pour le bassin A, voire nuls en été.

Les scores de biais et de CSI obtenus aux horizons maximums restent raisonnables et compétitifs avec ceux obtenus avec le modèle H, ce qui indique l'utilité potentielle de la prévision faite avec TAGE-GR4J à ces horizons.

S'agissant du cas des étiages définis comme des extrêmes saisonniers (seuil variable) le modèle H n'est pas apte à prévoir les jours d'étiage, ainsi qu'expliqué préalablement (cf.§ VI). Pour cette raison nous n'avons pas calculé de score de compétence de TAGE-GR4J en référence à H à seuil variable. Nous aurions bien entendu pu appliquer le modèle H a posteriori sur les jours d'étiage prévus par TAGE-GR4J, mais cela n'aurait à notre sens pas présenté d'intérêt pratique, puisque cette comparaison n'aurait correspondu à aucune situation opérationnelle potentielle.

Toutefois, nous avons noté que pour les deux bassins tous les scores RPS mensuels obtenus à seuil variable tendent, pour les horizons lointains, vers des valeurs identiques à celles observées précédemment à seuil fixe. Cette reproductibilité incite à choisir une valeur RPS de référence identique à celle déterminée à seuil fixe, à savoir $RPS=0.22$. Puisque le critère pour déterminer le HPUM à seuil fixe était que le score de compétence RPS soit supérieur à 10%, cela revient à choisir une valeur limite égale à 90% de la valeur de référence, **soit $RPS \leq 0.20$** .

Par ailleurs, afin de bénéficier de l'expérience acquise par les prévisions faites à seuil fixe, nous avons conservé le critère sur l'indice de Rousseau $IR > 0.10$ comme déterminant possible pour extraire le HPUM à seuil variable. Les comportements de cet indice vis-à-vis du biais à grande échelle et du CSI (Figure 62 et Figure 63), qui ont été vérifiés comme conformes à ceux déjà observés à seuil fixe, nous ont confortés dans cette voie.

Les *Tableau 13* et *Tableau 14* restituent pour les bassins A et J les valeurs du HPUM défini comme le minimum des deux horizons de prévisions issus respectivement du critère $RPS \leq 0.20$ et du critère $IR > 0.10$.

Nous avons également retranscrit dans ces tableaux les valeurs du CSI et du biais à grande échelle correspondant au HPUM.

Tableau 13 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin A à seuil variable, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM.

		avril	mai	juin	juillet	août	septembre	octobre	novembre	décembre
BV A	Horizon max issu de $RPS \leq 0.20$	4	5	2	2	0	0	0	1	2
	Horizon max issu de $IR > 0.1$	3	2	2	1	0	0	1	1	2
	HPUM (semaines)	3	2	2	1	0	0	0	1	2
	CSI (HPUM)	17	29	22	28				24	21
	Biais (HPUM)	2.3	3.4	3.2	3.4				2.6	2.1

Tableau 14 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin J à seuil variable, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM.

		mai	juin	juillet	août	septembre	octobre	novembre	décembre
BV J	Horizon max issu de $RPS \leq 0.20$	18	13	16	20	13	0	6	8
	Horizon max issu de $IR > 0.1$	9	11	16	20	20	2	3	4
	HPUM (semaines)	9	11	16	20	13	0	3	4
	CSI (HPUM)	17	29	32	31	32		30	18
	Biais (HPUM)	4.7	3.4	3.1	1.2	1.9		3.8	5.6

La Figure 28 compare les HPUM mensuels issus des deux types de seuillage d'étiage, pour les deux bassins. Les résultats sont très similaires, à l'exception du mois d'août pour le bassin J.

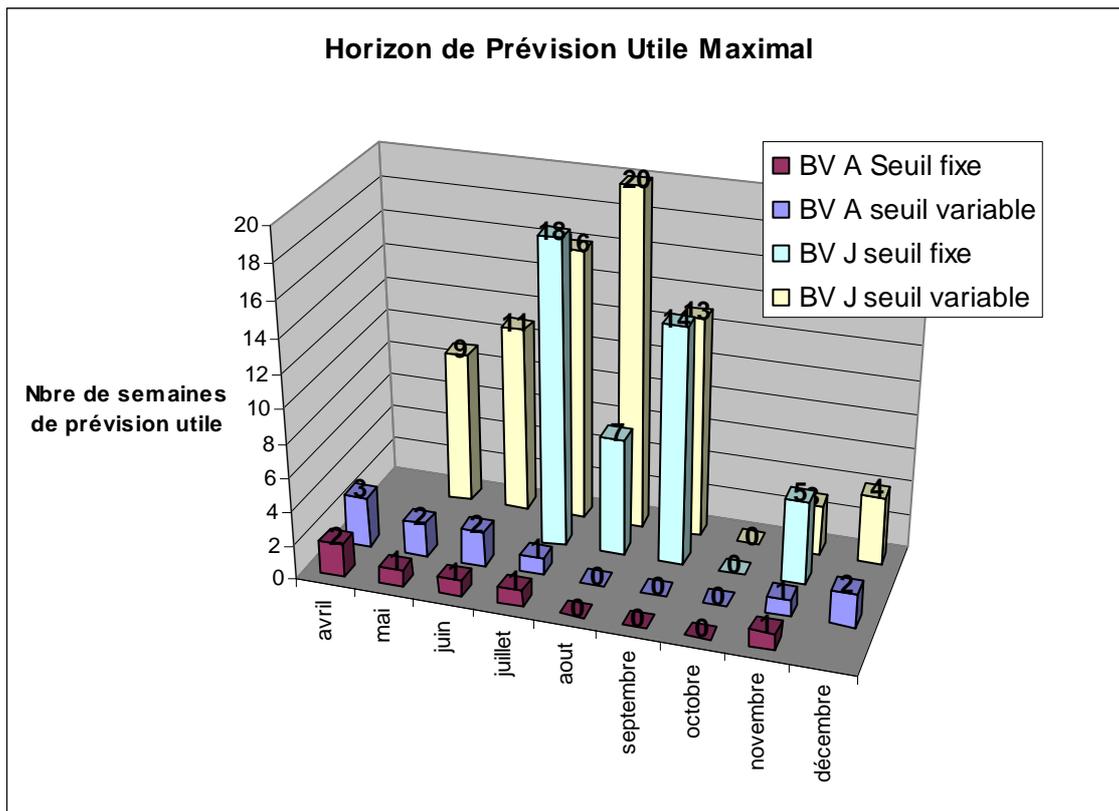


Figure 28: Comparaison des Horizons de Prédiction Utile Maximums mensuels obtenus avec TAGE-GR4J pour les deux bassins et les deux types d'étiages considérés dans ce travail.

On distingue donc bien deux types de bassins pour lesquels les potentiels de prévision hydrologique pour les étiages sont très différents. Pour le bassin A, aucune prévision n'est utile en été, et l'horizon de prévision utile reste très limité pour les autres mois (typiquement 1 semaine, et en tous cas inférieur à 3 semaines). Le fait que les mois d'été soient encore moins bien lotis que les autres tient essentiellement à ce que le modèle TAGE-GR4J sous-estime fortement les débits observés durant ces périodes (en moyenne -30%).

Pour le bassin J au contraire, les mois d'été ont des étiages dont la prévision avec TAGE-GR4J s'avère utile entre 2 et 5 mois à l'avance. Les étiages saisonniers de printemps (qui pour ce cas n'existent qu'à seuil variable) ou d'automne peuvent eux aussi faire l'objet de prévisions utiles avec TAGE-GR4J 1 à 3 mois à l'avance. Le bassin J se caractérise donc par une forte influence de l'état des réservoirs sur les débits futurs, ce qui autorise des prévisions à long termes si la dynamique de décharges des réservoirs est bien modélisée. Cette fois les mois de printemps et d'automne sont désavantagés par rapport aux mois d'été, car les débits observés y sont sensiblement sous-estimés par les prévisions ; ces dernières sont par ailleurs imprécises du fait de la forte variation interannuelle des précipitations durant ces mois. Ces effets couplés génèrent des taux de fausses alertes et des biais à grande échelle importants, qui pénalisent la prévision à long terme.

IX. APPLICATION DE TAGE-GR4J A 261 BASSINS FRANCAIS

Nous avons pour finir appliqué le système TAGE-GR4J sur un vaste jeu de 261 bassins français pour lesquels des archives de pluie, évapotranspiration potentielle et débit existent sur la période 1970-2006.

Les conditions suivantes ont été adoptées pour la génération des prévisions d'étiage:

- le calage de la simulation s'est effectué à partir d'une fonction objectif de Nash-Suttcliffe basée sur le logarithme des débits, notée ci-après *Nash LogQ*. Cette option accentue le poids des débits faibles dans l'optimisation et permet en principe une meilleure simulation des étiages.
- la période cible choisie pour lisser les débits était de 11 jours
- l'horizon de prévision maximal h_{max} était de 4 mois et demi (12 périodes).
- le seuil d'étiage a été pris variable, égal pour chaque jour calendaire au quantile 10% des débits journaliers historiques, avec pour valeur plafond le quantile 20% de l'ensemble des débits de la chronique.
- le critère de franchissement de seuil a été pris égal à 20% des prévisions.
- les HPUM mensuels ont été déterminés sous la double condition que l'indice de Rousseau soit supérieur à 0.3 (après avoir démontré que cela ne sélectionne que des prévisions avec $CSI > 20\%$) et que le score de compétence RPS soit supérieur de 10% à celui obtenu par le modèle H lorsqu'il est appliqué aux jours d'étiages observés ou prévus avec TAGE-GR4J.

Par ailleurs, pour chaque bassin BV nous avons défini un horizon de prévision utile maximal caractéristique, établi à partir des HPUM mensuels pondérés par le nombre de jours d'étiage observé mensuel $N_{j_{Et}}(m)$, comme suit:

$$HPUM(BV) = \frac{\sum_{m=1}^{12} N_{j_{Et}}(m) \cdot HPUM(m)}{\sum_{m=1}^{12} N_{j_{Et}}(m)}$$

La Figure 29 donne la répartition des valeurs du HPUM caractéristique pour les 261 bassins étudiés. On constate que pour 50% d'entre eux l'horizon utile typique est inférieur à 2 semaines (ce qui masque toutefois le fait que pour certains mois, l'HPUM est plus fort), et que dans 10% des cas il excède 6 semaines.

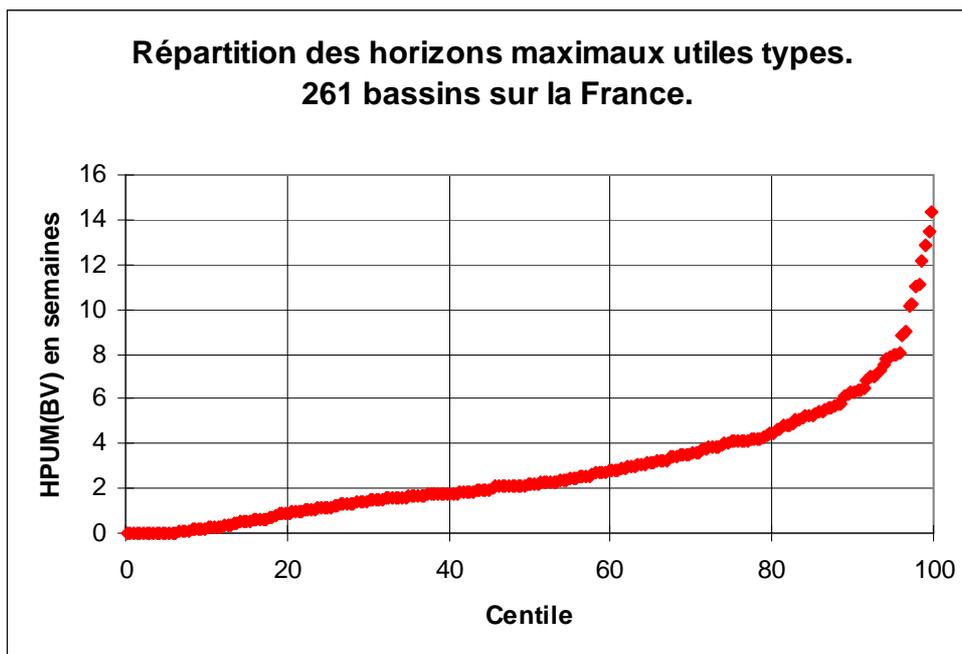


Figure 29: Diagramme de répartition des horizons de prévision utiles maximaux caractéristiques pour les 261 bassins étudiés.

La Figure 30 présente qualitativement et de façon spatialisée les résultats de ces calculs. Au-delà de quelques fluctuations, on y repère des tendances régionales, les valeurs les plus élevées se trouvant au Nord (Somme [jusque 9 semaines], Seine aval, Meuse) et à l'Ouest (Bretagne ouest, basse Normandie), le Centre le Sud et l'Est présentant des potentialités de prévision d'étiages moins importantes. La carte ne fait pas apparaître de relation entre la taille des bassins et les potentialités de prévisions d'étiage.

HPUM

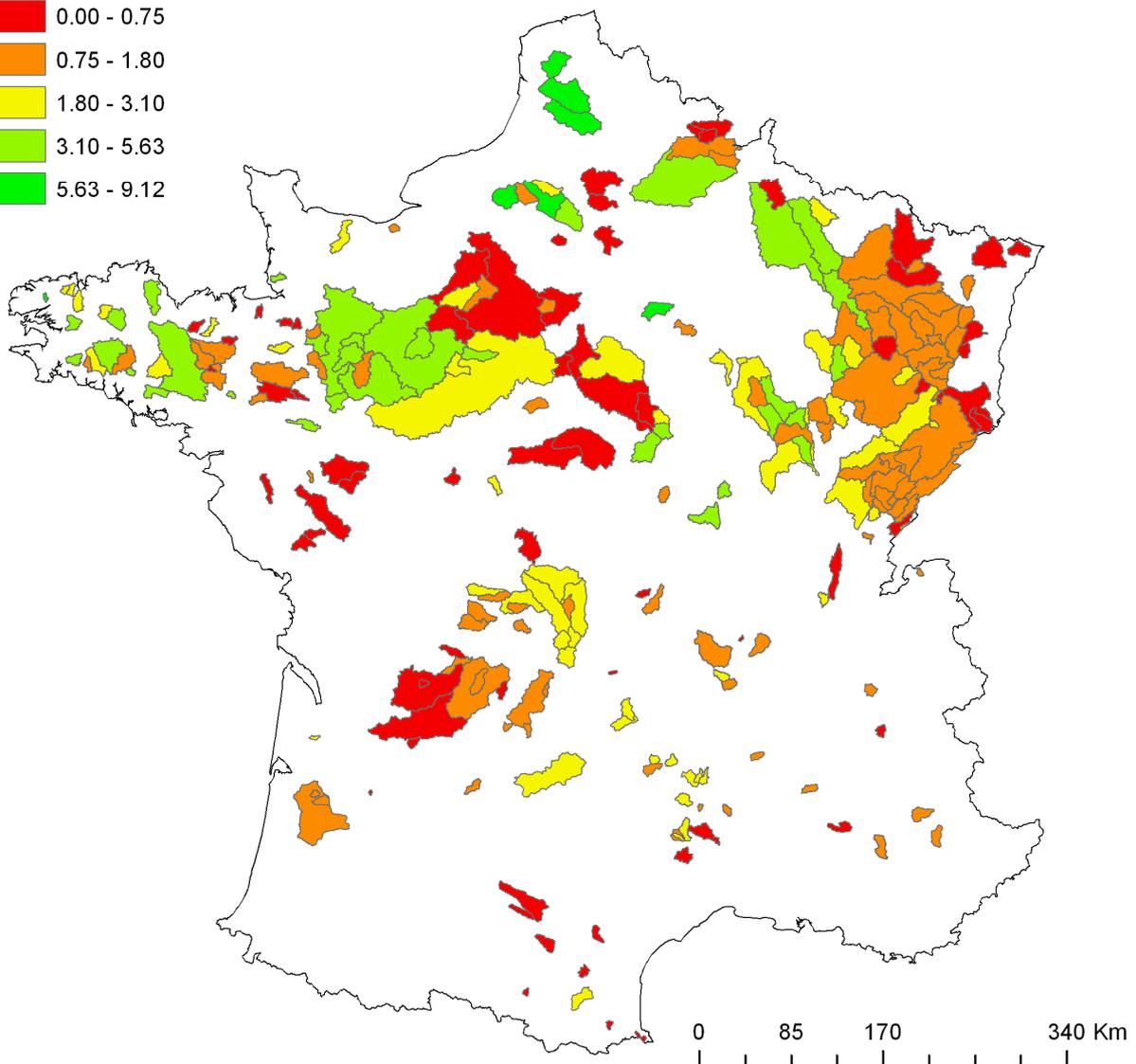
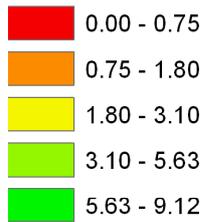


Figure 30: Carte des 261 bassins pour lesquels des prévisions ont été émises. La couleur des bassins est codée en référence à la valeur de HPUM(BV), lequel est exprimé en semaines.

Nous avons recherché si des variables explicatives des valeurs d'HPUM étaient disponibles parmi les 4 paramètres du modèle GR4J; nous avons également pris en compte la qualité de la simulation GR4J, avec l'intuition simple qu'une simulation de mauvaise qualité ne peut pas produire de bonne prévision. Le *Tableau 15* donne les valeurs des coefficients de détermination R entre ces divers paramètres pour l'ensemble des 261 bassins. La dernière ligne du tableau montre que les performances HPUM sont significativement corrélées aux seuls deux paramètres X3 et Nash LogQ, lesquels ne sont pas corrélés entre eux.

Tableau 15 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin A à **seuil variable**, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM.

	X1**	**X2**	**X3**	**X4**	Nash LogQ Val.	HPUM (BV)
X1**	1.00					
X2	-0.11	1.00				
X3	0.13	-0.41	1.00			
X4	-0.12	0.19	-0.20	1.00		
Nash LogQ Val.	-0.38	0.38	-0.13	0.33	1.00	
HPUM (BV)	-0.03	0.12	0.35	0.14	0.32	1.00

La Figure 31 montre la répartition spatiale des valeurs de Nash LogQ obtenues pour chaque bassin. Par comparaison avec la figure précédente, on observe effectivement que les bassins pour lesquels la simulation GR4J est médiocre (couleur Jaune, orange ou rouge) correspondent à des valeurs de HPUM faible. Dans ces cas (cas notamment de la Beauce, et de tout le quart sud est) les potentialités de prévision sont donc explicitement limitées par la qualité de la modélisation hydrologique. Toutefois on remarque que dans de nombreux cas, les HPUM sont faibles en dépit de simulations relativement bonnes (notamment au sud ouest et nord-est).

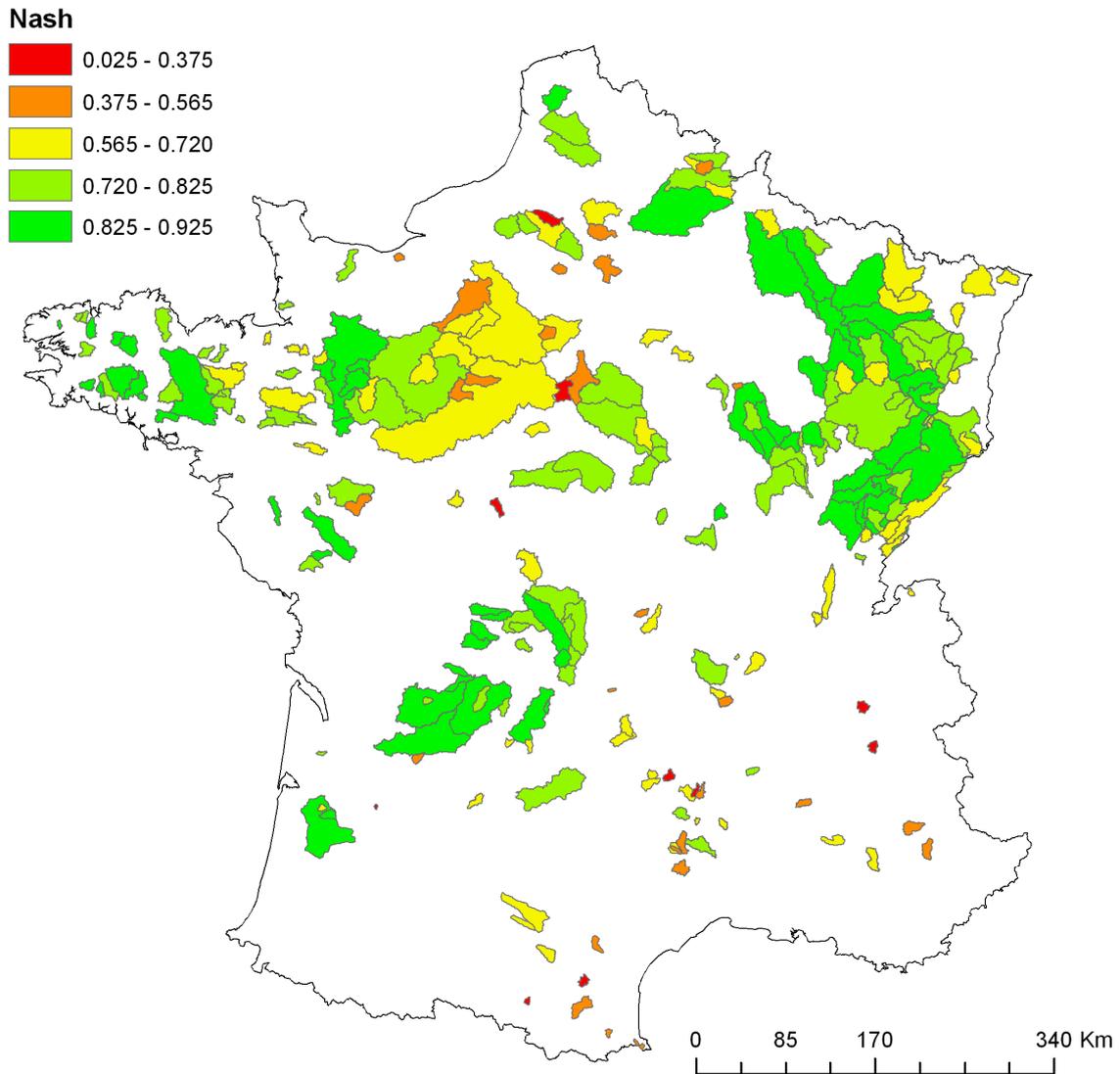


Figure 31: Carte des 261 bassins pour lesquels des prévisions ont été émises. La couleur des bassins est codée en référence à la valeur de Nash-LogQ (sans unité, variant entre 0 et 1).

La Figure 32 restitue la répartition spatiale du logarithme du paramètre X3, dont la valeur issue du calage de GR4J exprime la capacité du réservoir de routage. La corrélation avec la cartographie des HPUM (Figure 30) est globalement bonne, et en particulier les faibles valeurs d'HPUM enregistrées précédemment au sud-ouest et au nord-est, en dépit de simulations GR4J de bonnes qualités, s'expliquent bien par des capacités faibles du réservoir de routage.

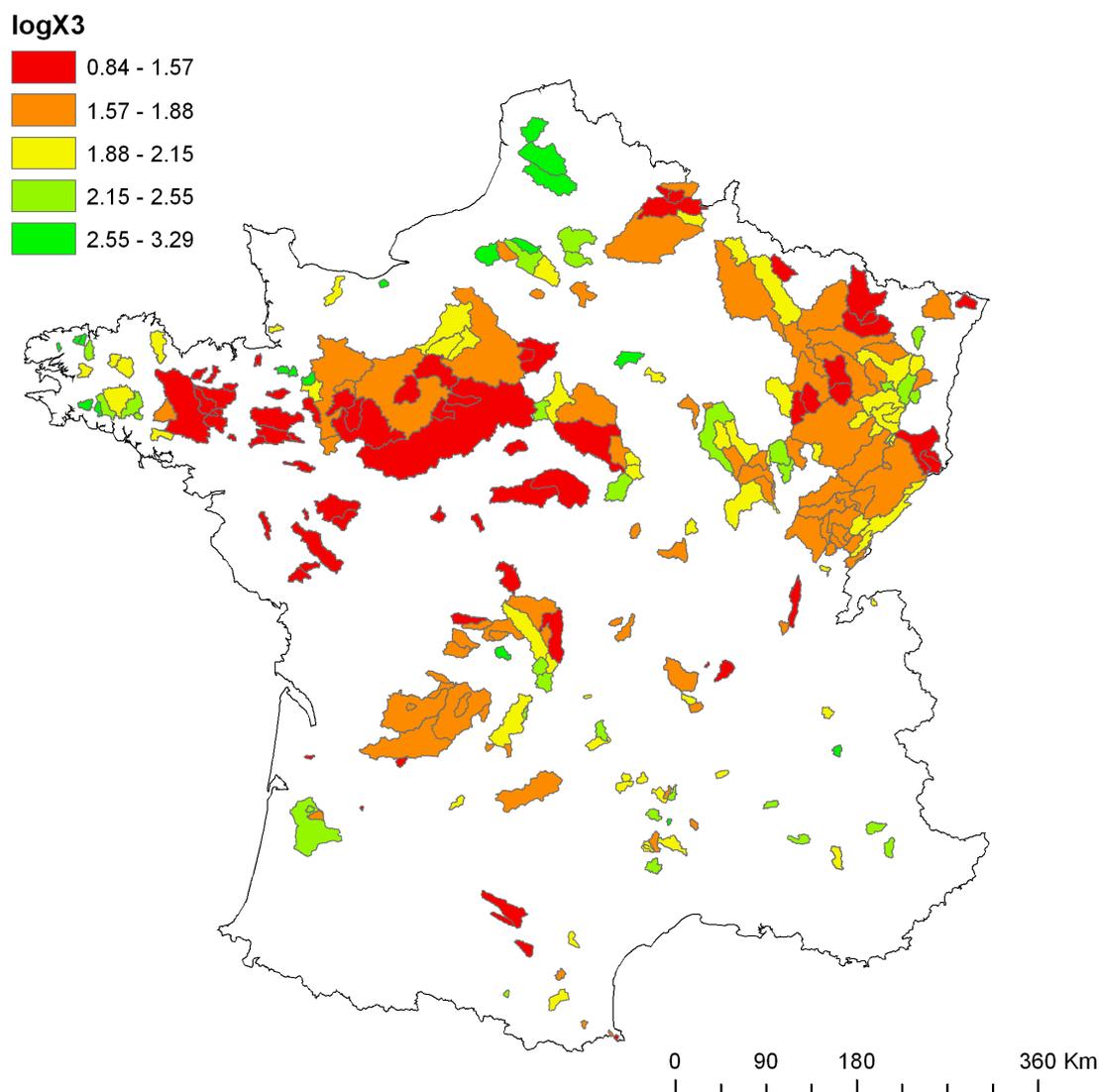


Figure 32: Carte des 261 bassins pour lesquels des prévisions ont été émises. La couleur des bassins est codée en référence à la valeur du logarithme du paramètre X3, lequel exprime la capacité du réservoir de routage.

X.CONCLUSIONS

Au cours de ce travail de stage nous avons pour objectif la mise au point d'une méthodologie probabiliste capable de produire et d'évaluer des prévisions à moyenne échéance (plusieurs mois) de débits d'étiage, à l'échelle d'un bassin versant, à partir de scénarios climatiques (pluie, évapotranspiration) et d'un modèle hydrologique global.

Notre travail a abouti à une version initiale d'un outil numérique original, baptisé TAGE, dédié à de telles prévisions. Cette expérience nous a placé dans une situation productive, dans laquelle il nous a fallu boucler des étapes de définition, de conception et de vérification d'un prototype.

La phase de définition nous a amené à acquérir des connaissances de bases à travers des échanges avec des chercheurs, et à nous familiariser avec la littérature scientifique traitant de divers domaines tels que la caractérisation ou la gestion des événements hydrologiques extrêmes (notamment les

étiages), la modélisation hydrologique, ou la prévision probabiliste. Dans un second temps, nous avons pu convertir ces informations pour conceptualiser l'outil que nous allons développer, en élaborant la stratégie de prévision et d'évaluation des performances qui sous-tend le code TAGE.

La phase de conception nous a amené à nous familiariser avec le langage informatique Scilab, pour produire un système original basé sur un code central générant les débits prévus, et des codes annexes dédiés au calcul des scores de performance. Le code ainsi produit est conçu pour pouvoir lire à la chaîne des fichiers, dans le but d'éprouver diverses combinaisons de modèles climatiques et hydrologiques sur des échantillons représentatifs de bassins versants. Le code TAGE est aisément paramétrable pour adapter sa résolution temporelle, ou bien son mode de définition des étiages.

La phase de vérification s'est déroulée en continu à travers le contrôle de la cohérence des résultats produits à tous les niveaux, depuis la prise en compte des données d'entrée jusqu'à la détermination des scores de performance.

A travers les prévisions émises pour deux bassins versants tests conférant des potentialités très distinctes vis-à-vis de la prévision des étiages, nous avons pu éprouver plusieurs méthodes de caractérisation des performances de l'outil. Nous avons cherché à "faire parler" divers indicateurs numériques et à comprendre leurs potentialités informatives vis-à-vis de la qualité des prévisions. Nous avons pu sur ces bassins discriminer les influences des variations saisonnières, éprouver l'influence du mode de définition du seuil d'étiage, et établir un cadre objectif pour mettre en compétition un modèle hydrologique (GR4J) avec un modèle naïf basé sur l'exploitation d'archives de débits.

A partir de ces réflexions nous avons pu proposer un critère synthétique d'évaluation en caractérisant la performance de la prévision par un horizon de prévision utile maximal.

Nous avons appliqué ce critère pour présenter une première cartographie des potentialités de prévision à l'aide de la combinaison TAGE-GR4J, en prenant en compte 261 bassins versants répartis sur le territoire français. Cet exercice aura montré que pour certaines régions des horizons de prévision utile compris entre 2 semaines et 3 mois sont atteignables dans 50% des bassins. Approfondir ce type d'approche devrait dans la suite permettre de révéler des déterminants des potentialités de prévision en termes de géologie des substrats et de climatologie régionale.

La phase suivante aura pour objectif d'éprouver les choix faits en termes de caractérisation de performance sur un jeu de bassins plus conséquents, impliquant des chroniques de données plus longues. Cela nous permettra de vérifier la robustesse du système d'indicateurs choisi, et de confirmer ou d'infirmer les interprétations que nous avons faites du comportement des indicateurs. Les critères seuils utilisés pour la détermination de l'horizon de prévision utile maximal devront probablement être révisés. Il sera aussi intéressant de vérifier la robustesse de l'outil en faisant varier ses paramètres d'analyse, tels que la résolution temporelle des prévisions (amplitude de la période cible), le critère probabiliste de dépassement de seuil, ou bien encore en modifiant la définition des seuils d'étiages.

La perspective est d'incorporer dans TAGE d'autres modèles de génération des pluies (modèles stochastiques) ou bien d'autres modèles hydrologiques globaux, afin de rechercher une combinaison optimale produisant les meilleures prévisions sur un échantillon important de bassins tests.

XI. ANNEXE A: RESULTATS DES SIMULATIONS GR4J SUR LES DEUX BASSINS TESTS.

Les Figure 33 et Figure 34 présentent en superposition les débits observés et simulés avec GR4J pour les deux bassins A et J. On y distingue nettement les cycles hydrologiques annuels. On observe sur la Figure 33 que pour le bassin A l'hydrogramme montre des fluctuations importantes à haute fréquence, nettement plus atténuées pour le bassin J.

Les valeurs du critère de Nash-Sutcliffe de calage et de validation pour les périodes P1 (août 1995-juillet 2000) et P2 (août 1999-juillet 2005) sont rapportés dans le *Tableau 16*.

Tableau 16 : critère de Nash-Sutcliffe obtenus pour les deux bassins pendant la phase de calage-contrôle.

	Calage P1	Calage P2	Contrôle P2	Contrôle P1
BV A	93.7%	92.7%	93.2%	92.2%
BV J	94.2%	94.9%	93.6%	94.3%

Les deux bassins présentent des valeurs de Nash-Sutcliffe supérieures à 90%, signant a priori une très bonne qualité de simulation, les valeurs de contrôle s'approchant par ailleurs de près de celles obtenues en calage.

Toutefois, au-delà de ces valeurs descriptives générales, on note que pour les faibles débits du bassin A, la simulation produit souvent des débits significativement inférieurs aux débits observés. A contrario, la simulation des débits du bassin versant J est nettement plus fidèle aux bas débits.

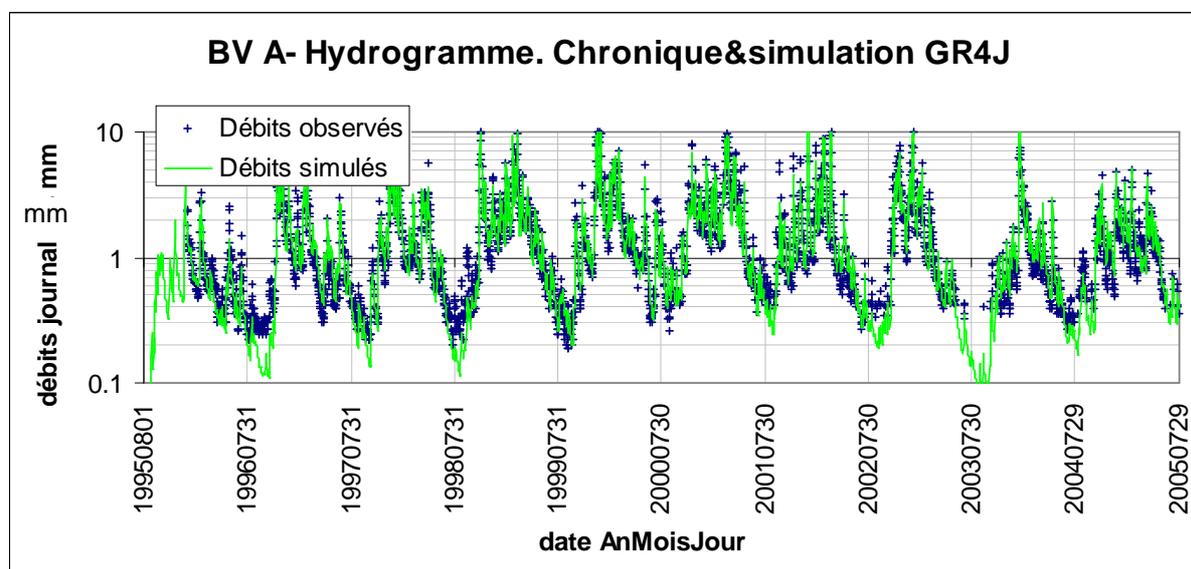


Figure 33: superposition des débits observés et simulés avec GR4J pour toute la chronique 1995-2005 du bassin A.

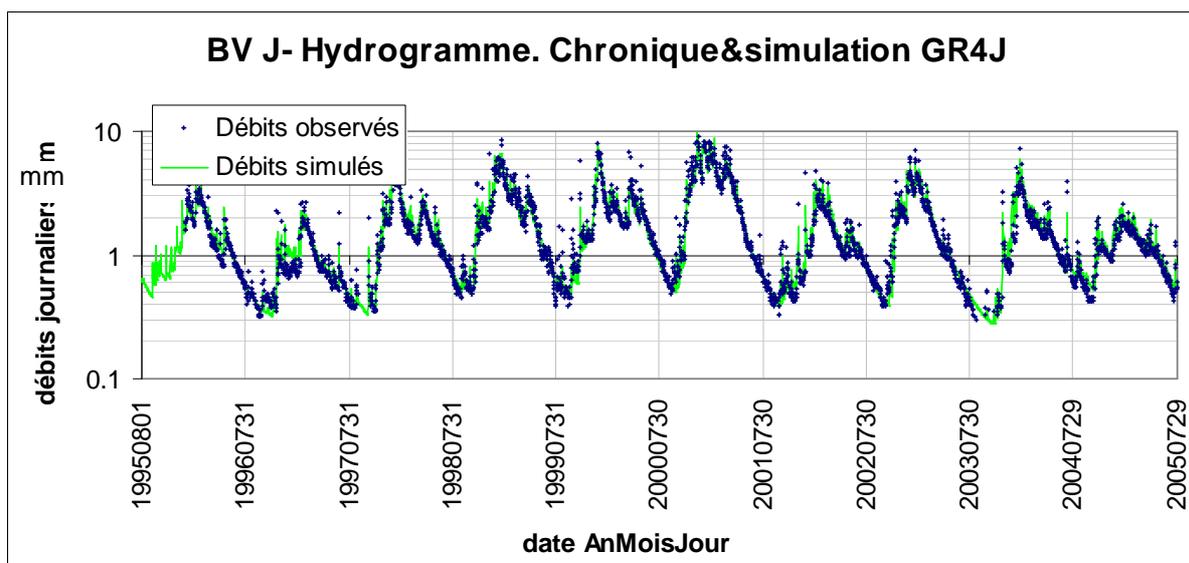


Figure 34: superposition des débits observés et simulés avec GR4J pour toute la chronique 1995-2005 du bassin J

Les *Tableau 17* et *Tableau 18* rapportent respectivement pour les bassins A et J les valeurs des quatre paramètres de GR4J issues des phases de calage sur les deux sous-périodes P1 et P2.

On observe tout d'abord que les valeurs issues des deux sous-périodes sont assez cohérentes pour les deux bassins, ce qui crédibilise le calage en vue d'une extrapolation sur de nouvelles données.

On note par ailleurs que les valeurs des paramètres diffèrent fortement entre les deux bassins. La différence la plus nette est enregistrée pour X3, qui correspond à la capacité du réservoir de routage, c'est-à-dire à la plus grande partie de l'eau facilement disponible pour le réseau hydrographique. La capacité de stockage pour le bassin J est de l'ordre de 15 fois supérieure à celle du bassin A.

Tableau 17 : valeurs après calage des 4 paramètres de GR4J pour le bassin A.

BV A	P1	P2
X1(mm)	355	444
X2(mm)	-0.93	-0.83
X3(mm)	68	64
X4(j)	2.47	2.46

Tableau 18 : valeurs après calage des 4 paramètres de GR4J pour le bassin J.

BV J	P1	P2
X1(mm)	174	177
X2(mm)	-1.64	-3.01
X3(mm)	947	1047
X4(j)	1.28	1.20

La Figure 35 illustre pour l'ensemble la chronique les valeurs du taux de remplissage de ce réservoir de routage, exprimé en %:

$$\tau R\% = R/X3$$

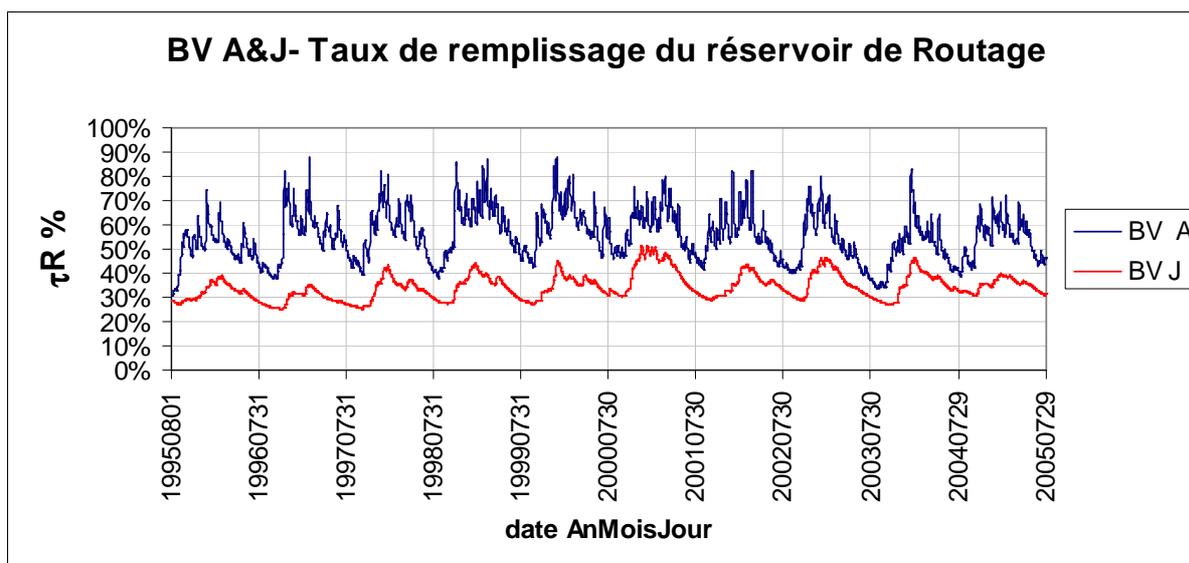


Figure 35: taux de remplissage du réservoir de routage au cours de la chronique 1995-2005 pour les deux bassins tests A et J.

On peut constater d'après cette Figure 35 que le taux de remplissage τ_R pour le bassin A fluctue fortement, ce qui indique sa sensibilité à la pluviométrie, et la labilité de cette réserve. A contrario, du fait de sa capacité nettement plus élevée, le réservoir de routage pour le bassin J voit son taux de remplissage évoluer de manière beaucoup plus lente et limitée. Cela rend les hydrogrammes moins sujets aux variations pluviométriques et davantage déterminés par l'état des réservoirs des jours qui précèdent.

D'un point de vue de prévisionniste, cela devrait se traduire pour J par des potentialités de prévision des débits sur des échelles de temps plus importantes. Dans la suite de ce travail nous qualifierons "d'effet mémoire" cette propriété du bassin à stocker des réserves d'eau suffisantes pour alimenter de manière prédictible le débit à l'exutoire du bassin.

Notons que dans certains cas le réservoir de Production peut aussi jouer un rôle pour la prévision, par exemple après une longue période de sécheresse il est susceptible de se vider fortement de sorte que lors des premières périodes de pluie succédant à la sécheresse l'eau infiltrée contribue majoritairement au remplissage de ce réservoir et n'est plus mise à disposition pour le routage. Dans ces cas, la sensibilité du débit à la pluie devient faible et les débits deviennent prédictibles de manière fiable à plus longue échéance.

XII. ANNEXE B: DETERMINATION DU CRITERE D'ALERTE OPTIMAL

En tant que prévision « probabiliste », le modèle H propose pour chaque jour calendaire j d'une année de la chronique une série de débits possibles (chacun d'entre eux correspond à l'un des débits observés les autres années ce même jour j).

Partant de là il existe une marge dans la façon d'utiliser le modèle pour la détection des événements d'étiage : on peut choisir le pourcentage de ces débits qui serviront effectivement de critère d'alerte. Dans notre cas le modèle H prototype propose 10 valeurs par jour (débits des 11 années de la chronique moins celui de l'année courante). On peut alors chercher à savoir quel est le critère d'alerte le plus efficace : 1 seul débit sous le seuil d'étiage ? 2 ? 3 ? Sachant qu'un compromis sera à trouver entre fréquence des fausses alertes et probabilité de réussite en détection.

Afin d'optimiser le modèle H, nous avons fait le choix de nous baser sur la maximisation des indices ETS et IR décrits plus haut, sachant que ces deux indices sont réputés ne pas être affectés par d'éventuels effets saisonniers (les périodes calendaires sélectionnées sont différentes selon le critère d'alerte choisi) ou de biais à grande échelle (qui varie avec la FAR et la POD). Les *Tableau 19* et *Tableau 20* restituent respectivement les scores ETS et IR obtenus en utilisant 3 critères d'alerte différents, selon que le nombre critique de prévisions situées sous le seuil d'étiage est fixé à 1, 2 ou 3. La dernière colonne des tableaux rapporte un indice global, moyenne des scores mensuels pondérée par les nombres de jours d'étiage propre à chaque mois.

Tableau 19 : Scores de Menace Equitable mensuels et globaux annuels obtenus pour les deux bassins pour trois critères d'alertes différents.

ETS										
nombre critique de prévisions	BV	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Global
1/10	A	6.4%	-1.5%	17.1%	4.5%	10.2%	10.4%	7.9%	-2.1%	8.4%
	J				11.3%	3.2%	16.2%	3.8%	0.9%	7.4%
2/10	A	0.0%	0.0%	8.1%	2.1%	5.7%	8.6%	-5.8%	0.0%	4.2%
	J				-6.4%	-9.3%	13.3%	-1.1%	-8.9%	-0.8%
3/10	A	0.0%	0.0%	-2.5%	-3.5%	2.1%	-9.4%	-1.5%	0.0%	-3.3%
	J				0.0%	-12.5%	5.5%	-7.7%	-5.4%	-3.8%

Tableau 20 : Indices de Rousseau mensuels et globaux annuels obtenus pour les deux bassins pour trois critères d'alertes différents.

IR										
nombre critique de prévisions	BV	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Global
1/10	A	0.04		0.24	-0.24	-0.01	0.07	-0.01		-0.03
	J					-0.34	0.15	-0.22	-0.39	-0.15
2/10	A			0.119	-0.106	-0.03	0.041	-0.17		-0.03
	J					-0.46	0.11	-0.13	-0.26	-0.16
3/10	A			-0.094	0.02	-0.27				-0.07
	J						0.04	-0.17		-0.05

Il apparaît d'après ces résultats que le critère le plus efficace est de déclencher l'alerte dès qu'une seule des 10 prévisions passe sous le seuil d'étiage. Dans le cadre de ce travail nous conserverons ce critère d'alerte, que nous appliquerons aussi au modèle TAGE-GR4J, afin de comparer celui-ci au modèle H dans des conditions équivalentes. Lorsque les chroniques considérées seront plus conséquentes, ou bien lorsque des modèles générant des scénarios plus nombreux seront mis en œuvre, il sera utile de reconsidérer ce critère d'alerte.

XIII. ANNEXE C: SCORES DE PERFORMANCE OBTENUS AVEC TAGE-GR4J POUR LE BASSIN "A" POUR LE CAS "SEUIL FIXE"

A. SCORES DETERMINISTES CONTINUS : ECARTS ENTRE DEBITS OBSERVES ET PREVUS

La Figure 36 présente les écarts relatifs moyens mensuels (cf. définition § V.A.1) pour le bassin A en fonction de l'horizon de prévision. On peut y voir qu'à l'exception de mai, les écarts moyens sont quasi toujours négatifs, ce qui indique que, pour les jours d'étiage, GR4J tend à sous-estimer significativement les débits observés. Par ailleurs, à l'exception de mai, ces écarts moyens n'évoluent que peu en fonction de h.

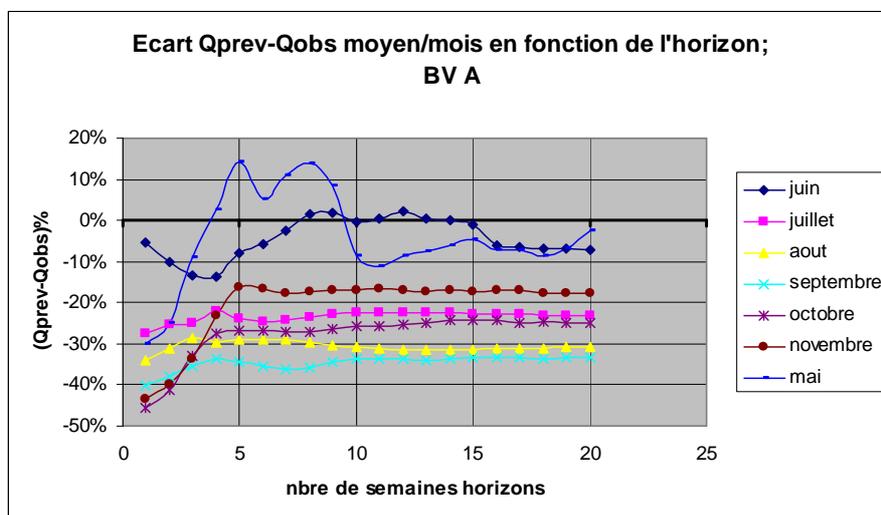


Figure 36: Ecarts mensuels relatifs entre débits prévus et observés pour les mois à étiages (seuil fixe) du bassin A selon l'horizon de prévision

La Figure 37 présente la RMSE relative (cf. définition § V.A.5) ; on y constate que cette dernière part d'une valeur assez élevée (>45%), et (en dehors du mois de novembre) qu'elle croît rapidement pour les premiers horizons, puis qu'il y a une stabilisation nette. Le fait que l'erreur maximale soit atteinte rapidement semble indiquer que l'avantage procuré par la connaissance de l'état initial des réservoirs se perd très vite dans le cas du bassin A.

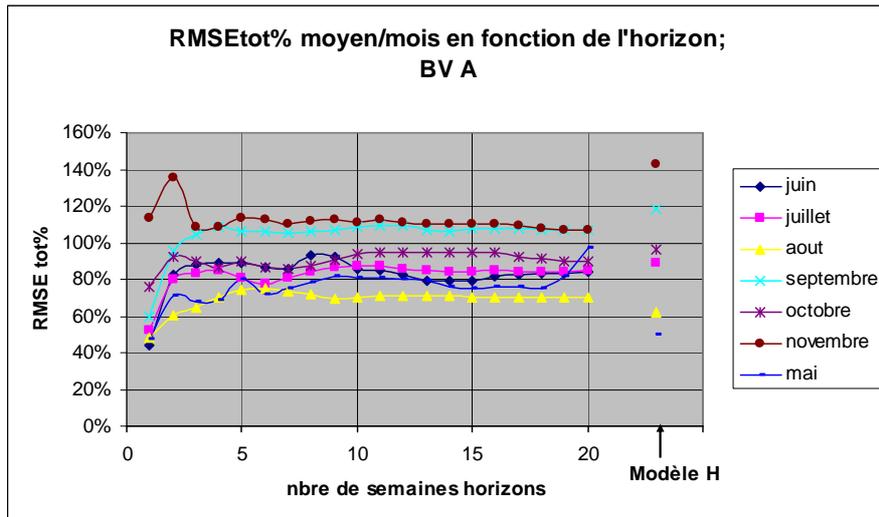


Figure 37: RMSE totales mensuelles relatives pour les mois à étiages du bassin A (seuil fixe) selon l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

Notons que les RMSE ainsi observées pour le bassin A (entre 50% et 120%) sont de valeurs similaires à celles obtenues par le modèle H (cf.§ VII.C.2), ce qui supporte l'idée que le modèle hydrologique n'apporte pas de plus value en matière de prévision d'étiage.

Les Figure 38 et Figure 39 illustrent, pour le bassin A, en fonction des mois de l'année et sous forme de « boîtes à moustaches » de couleur rouge, les distributions des écarts entre prévisions TAGE-GR4J et observations, respectivement pour des horizons de une et deux semaines. Pour comparaison, les boîtes représentées en couleur verte rapportent les distributions d'écarts obtenues avec le modèle H. Pour un horizon d'une semaine (Figure 38), les distributions TAGE-GR4J sont nettement plus resserrées que les boîtes H, et malgré le biais négatif enregistré, les écarts sont nettement plus faibles qu'avec le modèle H. Cet avantage se perd néanmoins très rapidement puisque dans la Figure 39 (horizon de 2 semaines), les amplitudes des boîtes rouges deviennent déjà similaires à celles du modèle H, ce qui signe le fait que pour ce bassin le modèle hydrologique n'apporte plus grand-chose à la prévision au-delà de ces deux semaines.

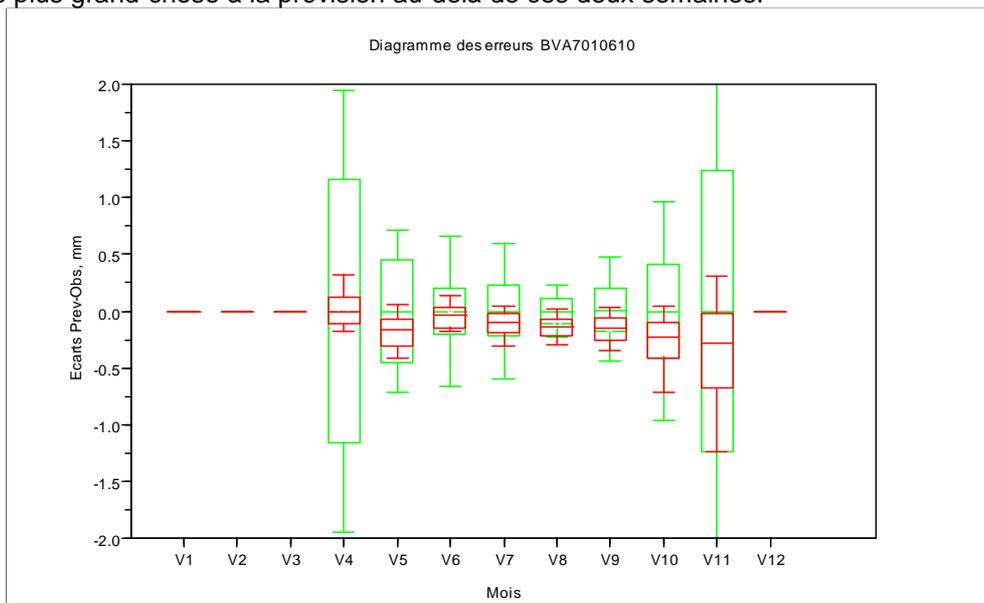


Figure 38: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la **première** semaine d'horizon, bassin A. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.

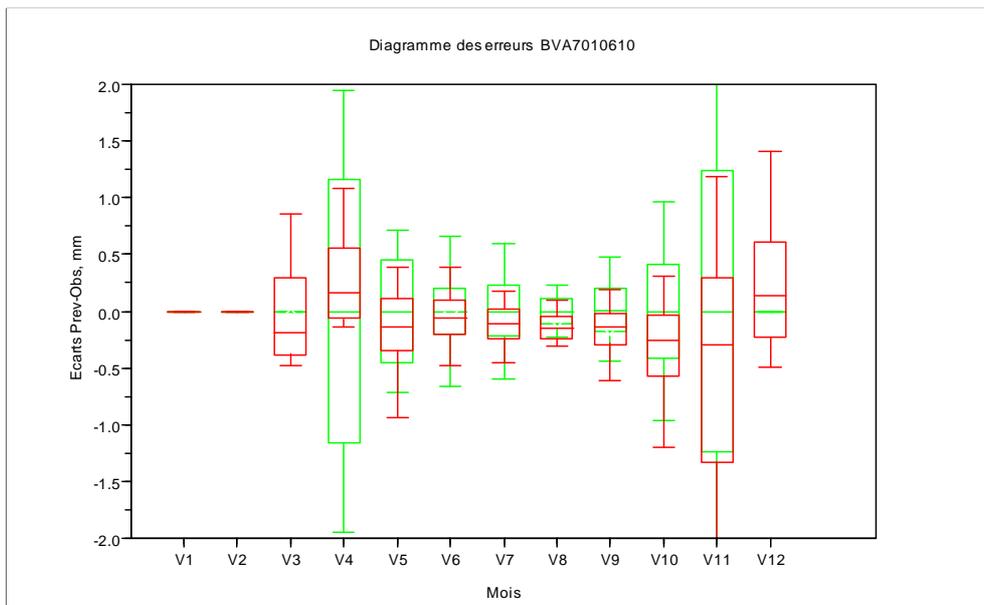


Figure 39: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la **seconde** semaine d'horizon, bassin A. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.

B. SCORE PROBABILISTE : RPS.

Les scores RPS mensuels obtenus pour le bassin A sont illustrés par la Figure 40. On y voit que pour les mois d'avril à juillet, le score se dégrade en croissant fortement les premières semaines d'horizon, puis se stabilise à une valeur entre 0.20 et 0.25.

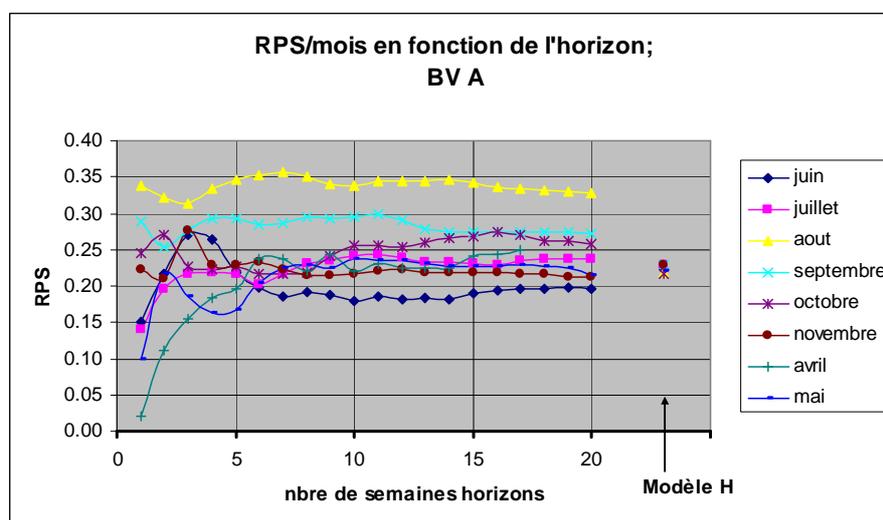


Figure 40: scores RPS mensuels pour le bassin A (seuil fixe) en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

Pour les mois ultérieurs, le RPS varie peu et se situe à des valeurs plus élevées, jusqu'à 0.35 pour le mois d'août. Pour interpréter ces valeurs élevées du RPS en août & septembre il faut se souvenir que le RPS quantifie l'adéquation entre des catégories de débits prévus et observés, et qu'en conséquence ce score sera d'autant plus sensible que les catégories seront fines par rapport aux

écarts enregistrés entre débits observés et prévus. La Figure 41 illustre pour les deux bassins A et J l'amplitude moyenne de la catégorie RPS centrale selon les mois de l'année, cette amplitude étant égale à la différence entre les percentiles 66% et 33% des débits d'étiages. On voit que pour le bassin A, les mois d'août et septembre présentent des amplitudes de classe très fines, essentiellement du fait que les débits d'étiages sont faibles et reproductibles à ces périodes de l'année. Comme par ailleurs pour ces mois TAGE-GR4J sous-estime de manière importante les débits, ainsi que l'exprime par exemple la Figure 38 pour le premier horizon de prévision, la combinaison de ces deux facteurs produit un score RPS fort, plus fort même que celui du modèle H (qui vaut 0.22). En effet le modèle H présente par définition un écart moyen nul, de sorte que seule la dispersion des débits influe sur le RPS.

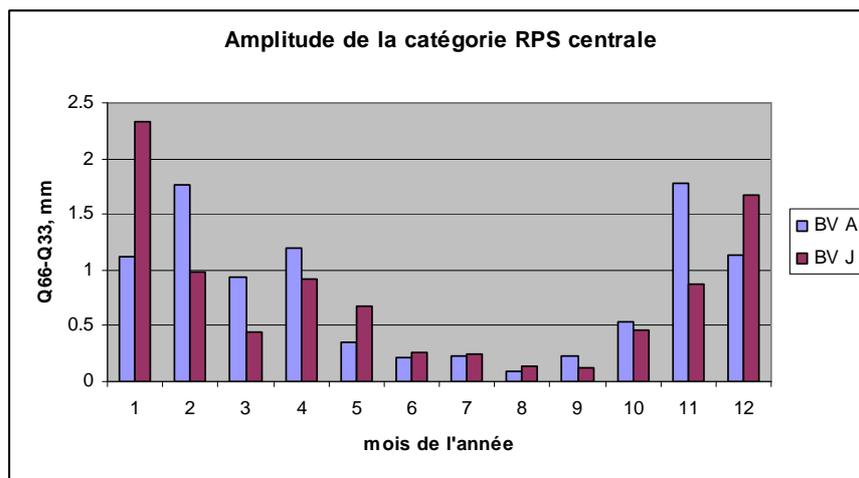


Figure 41: Amplitudes mensuelles moyennes de la catégorie centrale définissant le score RPS pour les deux bassins

Le score RPS constitue donc un indicateur particulièrement sévère et sensible aux biais des modèles pour les périodes sèches, ce qui présente un intérêt pour la prévision d'étiage. Il conviendra néanmoins de conserver cette notion à l'esprit pour interpréter les résultats RPS.

C. SCORES DETERMINISTES PAR CATEGORIES : PERFORMANCES DE LA PREVISION EN TERMES D'ALERTE D'ETIAGE

Pour le bassin A, les Figure 42a - f tracent respectivement les évolutions des indicateurs mensuels suivants, définis au § V.B :

- taux de fausses alertes FAR
- probabilité de détection POD
- indice de succès critique CSI
- score de menace équitable ETS
- biais à grande échelle B.
- indice de Rousseau IR

Concernant le FAR, le taux est toujours élevé (>50%), et est typiquement de 60-70% en été, davantage encore aux autres saisons. Des chiffres tout à fait similaires avaient été enregistrés pour le modèle H (cf. Tableau 9). Le FAR marque une montée pour les deux premiers horizons, puis se

stabilise. La **POD** est quasi toujours à 100%, notamment pour les mois d'été ; là encore le modèle H donnait des résultats semblables.

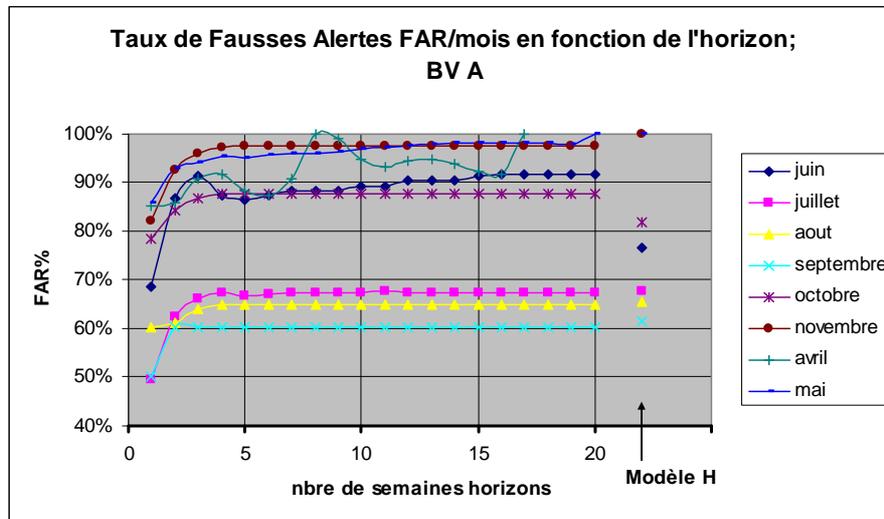


Figure 42a : taux mensuels de fausses alertes pour le bassin A en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

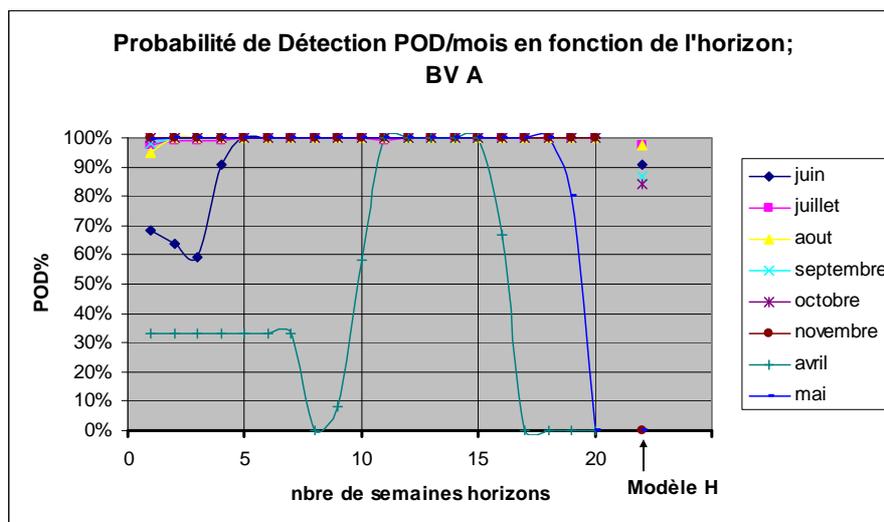


Figure 42b: probabilité de détection des étiages pour le bassin A en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.

Le **CSI** a un comportement miroir de celui du FAR, avec des valeurs moyennes en été, et très faibles sinon. Les valeurs relativement élevées en été sont à considérer avec précaution ; en effet le CSI n'est proprement utilisable que lorsque les événements d'étiage vrais sont rares au regard des autres événements. Or dans notre cas on dénombre aux environs de 80-90 jours d'étiage observés pour les mois d'été (cf. *Tableau 1*) à comparer avec les 300 jours par mois que la chronique de 10 années ; soit un taux d'étiages de près de 30%, de sorte que l'on ne peut plus qualifier ces événements de rares. A l'exception de la première semaine d'horizon, ces valeurs sont similaires à celles obtenues avec le modèle H.

L'ETS décroît rapidement pour tous les mois et se stabilise à des valeurs proches de celles obtenues avec le modèle H, à l'exception du mois de juin. On note que l'ETS est très inférieur au CSI, en

particulier pour les mois d'été ; cela tient au fait que la probabilité d'émettre une alerte vraie par chance y est élevée, du simple fait que les étiages observés y sont fréquents.

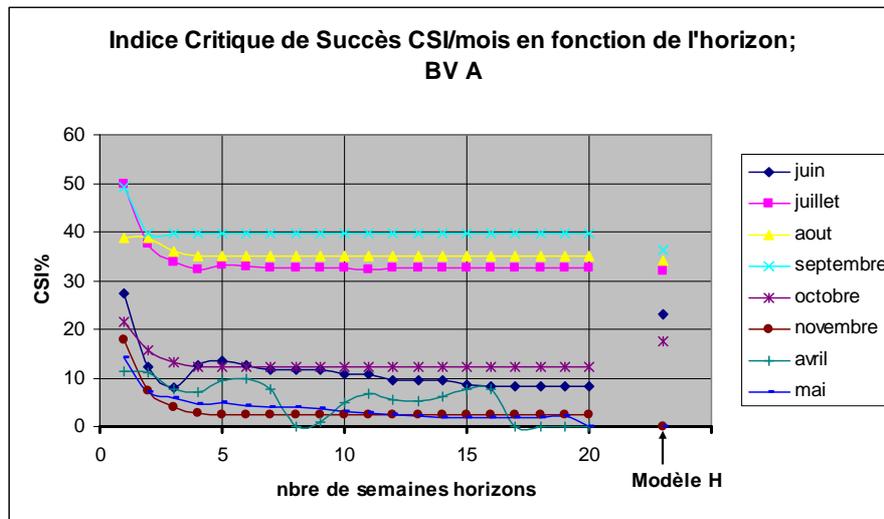


Figure 42c : Indices mensuels de Succès Critique pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin A. Le dernier point de chaque mois correspond à la performance du modèle H.

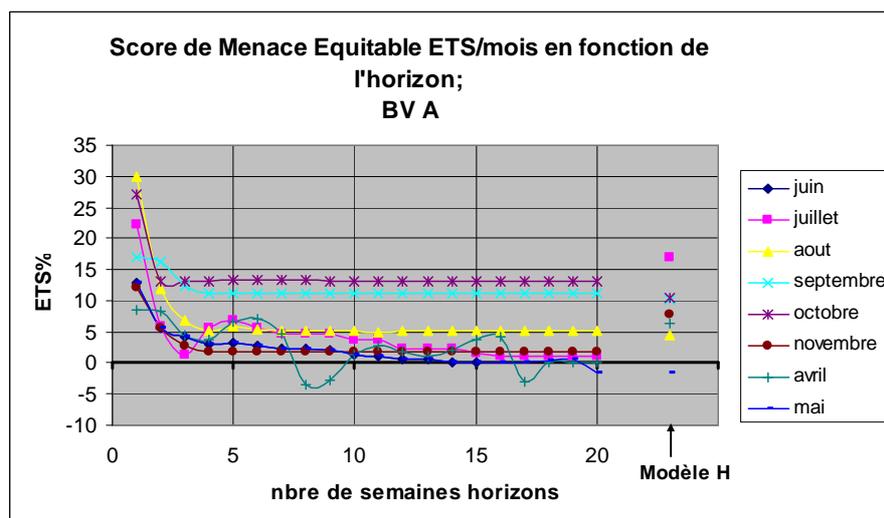


Figure 42d : Scores mensuels de Menaces Equitables pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin A. Le dernier point de chaque mois correspond à la performance du modèle H.

Le biais à grande échelle est similaire à celui obtenu avec le modèle H pour le premier horizon, ainsi que pour tous les horizons des mois d'été. Il atteint par contre des valeurs beaucoup plus élevées que H pour les autres horizons des mois de printemps et d'automne. Le fait que le biais en été est très inférieur aux autres mois s'impute au fait que la densité d'étiage y est très supérieure, favorisant ainsi la détection d'alertes vraies.

Les **indices de Rousseau** sont très faibles et sont quasi tous négatifs au-delà du premier horizon de prévision. Les mois à grand biais sont les plus pénalisés par cet indice.

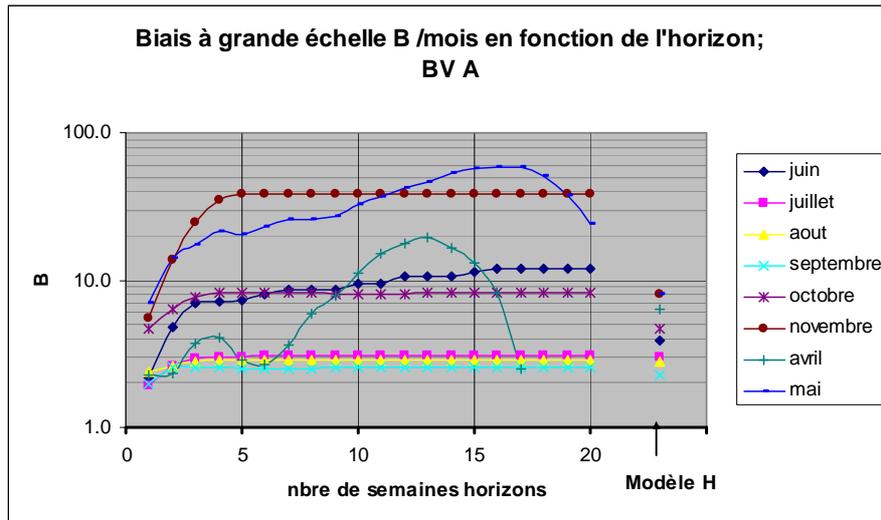


Figure 42e: : Biais à grande échelle mensuels pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin A. Le dernier point de chaque mois correspond à la performance du modèle H.

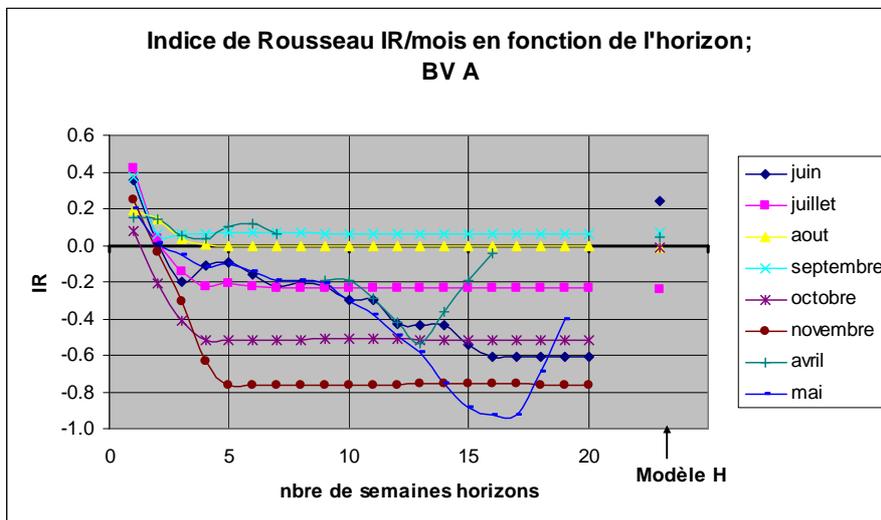


Figure 42f: : Indices mensuels de Rousseau pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin A. Le dernier point de chaque mois correspond à la performance du modèle H.

XIV.ANNEXE D: SCORES DE COMPETENCES AVEC TAGE-GR4J POUR LE CAS "SEUIL FIXE"

La Figure 43: et la Figure 44: rapportent les valeurs des scores de compétence $Comp_{RMSE}$ du modèle TAGE-GR4J en référence au modèle H, respectivement pour les deux bassins A et J. Pour le bassin A, il apparaît nettement qu'au-delà des premiers horizons (4 semaines dans le meilleurs des cas, 1 à 2 semaines sinon), les scores mensuels atteignent leur minimum et y demeurent pour les horizons plus lointains. Qualitativement, cela traduit le peu de mémoire hydrologique de ce bassin, qui pénalise fortement les prévisions à moyen et long termes. La valeur du score au-delà des premiers horizons n'est pas nulle comme on aurait pu s'y attendre dans le cas idéal, mais prend des valeurs comprises entre +20% et -20% selon les mois. Une explication du fait que le score puisse rester indéfiniment positif même une fois l'effet mémoire disparu pourrait être que la dispersion des débits observés, qui fondent le modèle H, est artificiellement exagérée par des erreurs de mesure les mois en question.

Des effets inverses liés au modèle GR4J et au fait de moyenner les pluies réelles sur un pas de temps de 24 heures pourraient jouer en amenuisant artificiellement la RMSE des débits modélisés. Ce point resterait à éclaircir.

Pour le bassin J (Figure 44) le score de compétence a un comportement très différent, le minimum de compétence étant atteint lentement au bout de 2 mois d'horizon, ou plus. Le mois de septembre est une exception, la compétence ne montrant pas de tendance, conformément à ce qui avait déjà été enregistré pour la RMSE dans la Figure 12. Un comportement asymptotique vers des valeurs positives est aussi constaté pour le score aux horizons les plus élevés.

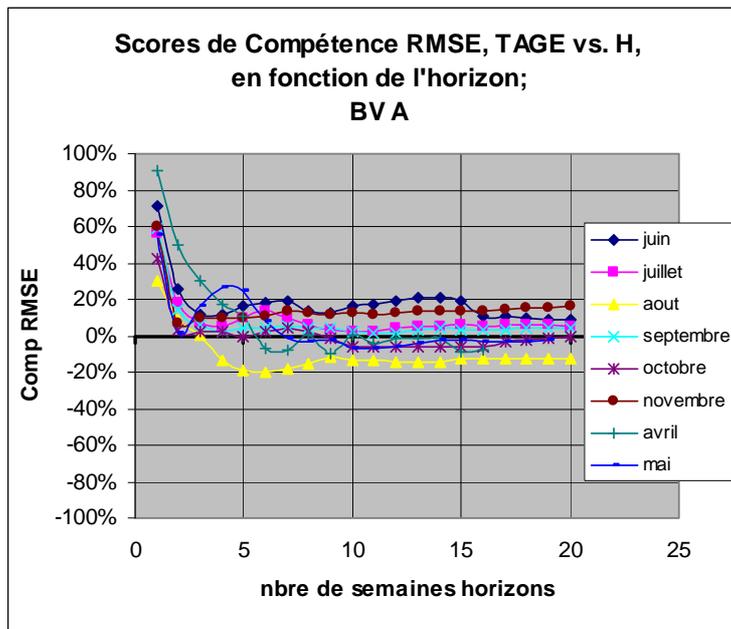


Figure 43: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score RMSE pour le bassin A.

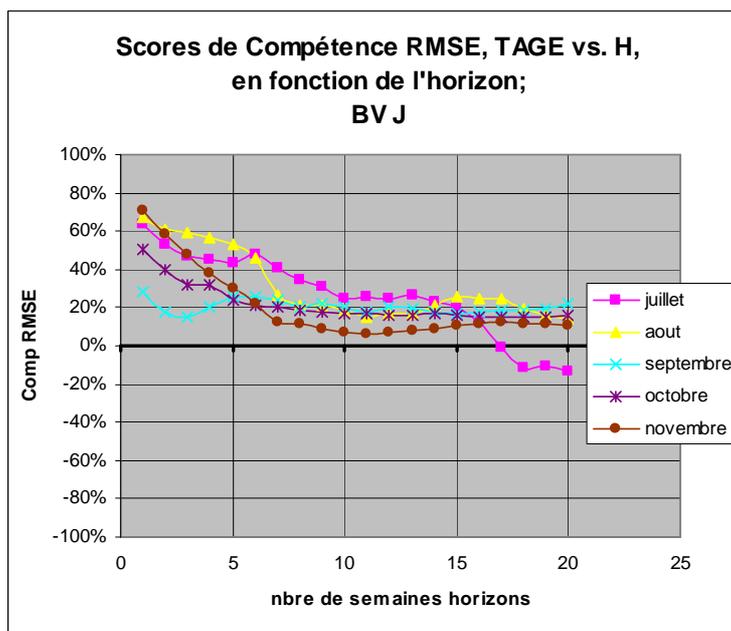


Figure 44: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score RMSE pour le bassin J.

Les scores de compétence associés au score ETS sont présentés dans la Figure 45 et la Figure 46. Pour le bassin A, la compétence est toujours très faible, et fluctue tout de suite autour de zéro ou moins au-delà du premier horizon.

Pour le bassin J, $Comp_{ETS}$ reste par contre positif plusieurs mois concernant les mois d'été.

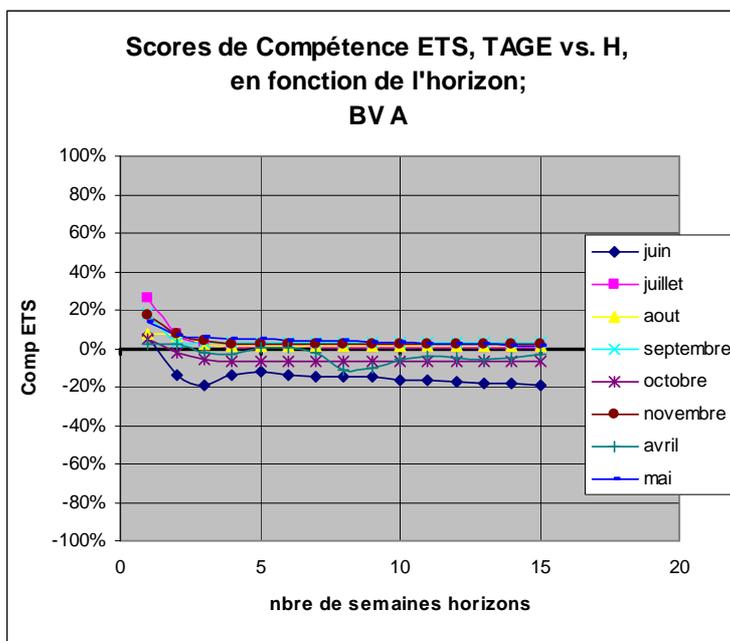


Figure 45: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score ETS pour le bassin A.

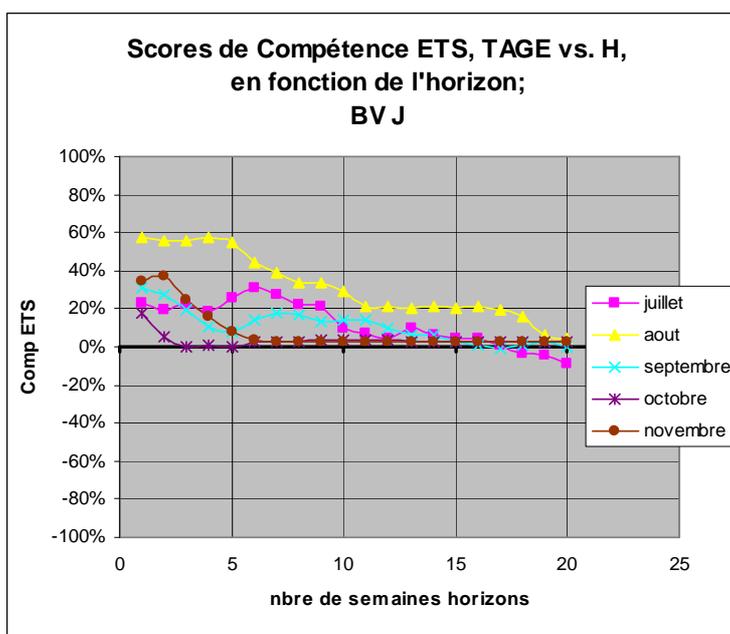


Figure 46: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score ETS pour le bassin J.

En synthèse, on constate une grande diversité des performances selon les mois, visible surtout pour le bassin J. TAGE-GR4J est nettement plus performant que H sur le bassin J, notamment en été. La compétence RPS semble exacerber davantage les différences que les autres scores (sensibilité). La compétence sur RMSE a parfois un comportement asymptotique et reste >0 . Parfois il ne montre pas de tendance (ex :septembre pour le bassin J). La compétence ETS est moins sévère que celle basée sur RPS (cf mois d'août pour le bassin J), probablement car moins sensible aux écarts entre débits prévus et observés.

XV. ANNEXE E: RESULTATS DE LA PREVISION PAR TAGE-GR4J D'ETIAGES DEFINIS COMME EXTREMES SAISONNIERS ("SEUIL VARIABLE")

A. CHOIX DU SEUIL D'ETIAGE

Comme expliqué au §VI, le seuil de débit définissant l'étiage est cette fois fixé au minimum du quantile 20% des débits historiques pour le jour j considéré et du quantile 25% interannuel.

Un tel choix se traduit par la sélection des jours d'étiages donnés dans le *Tableau 21*, et par les débits moyens d'étiage donnés dans le *Tableau 22* pour les deux bassins A et J.

Tableau 21: nombre de jours d'étiages à seuil variable observés chaque mois pour les deux bassins

Etiages observés	Jan	Fev	Mar	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Dec	Total
BV A	0	2	0	37	35	39	38	31	32	30	19	7	270
BV J	7	3	0	0	11	49	60	33	30	30	30	20	273

Tableau 22 : débits moyens des jours correspondants à des étiages à seuil variable observés pour les deux bassins

Débits moyens observés pendant les étiages , mm	Jan	Fev	Mar	Avr	Mai	Juin	Juil	Aout	Sept	Oct	Nov	Dec
BV A		0.52		0.40	0.43	0.40	0.31	0.28	0.26	0.34	0.40	0.44
BV J	0.71	0.72			0.66	0.62	0.53	0.37	0.38	0.41	0.40	0.63

Comparativement au choix d'un seuil fixe (se référer au § VII.D.2.a), le choix du seuil variable que nous avons fait se traduit par la sélection d'évènements en nombre quelque peu inférieur mais mieux répartis dans l'année (davantage d'évènements printaniers et automnaux), avec des débits plus variés, variant d'un facteur 2. Le total mensuel des évènements d'étiages représente au plus 1/5^{ème} des jours du mois en question, et typiquement 1/10^{ème}.

B. JOURS D'ETIAGES PREVUS PAR TAGE-GR4J.

La Figure 47 indique la variation du nombre de jours d'étiages prévus par TAGE-GR4J en fonction de l'horizon de prévision, pour les deux bassins, en comparant les comportements à seuil variable avec ceux obtenus précédemment à seuil fixe. Pour le bassin A, le nombre de jours d'étiage prévus croît encore plus vite qu'à seuil fixe, ce qui préjuge de taux de fausses alertes encore plus élevés dès les premiers horizons. Pour le bassin J, les comportements sont similaires pour les deux types de seuil.

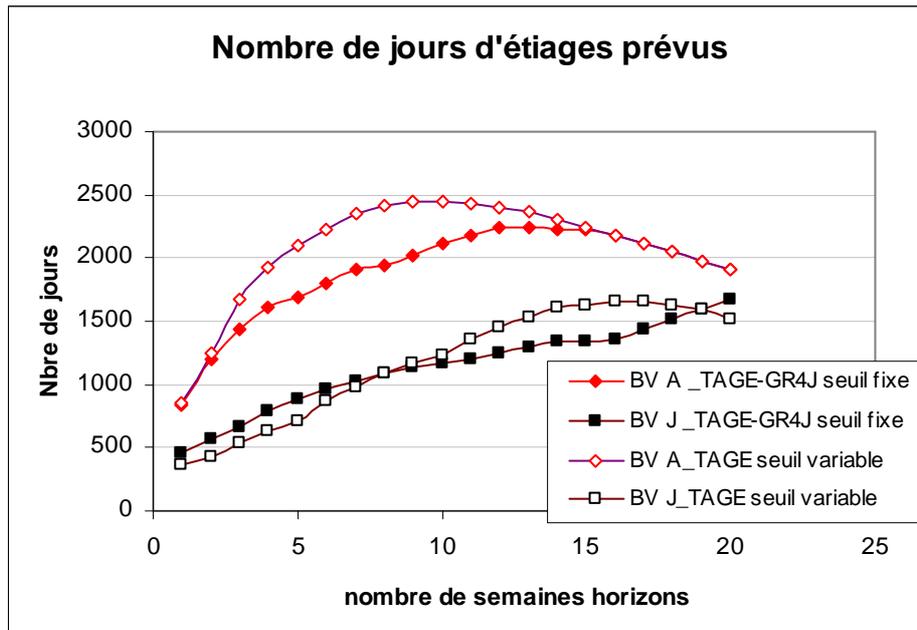


Figure 47: Nombre de jours d'étiages prévus par TAGE-GR4J pour les deux bassins en fonction de l'horizon de prévision. Comparaison entre les cas « seuil variable » et « seuil fixe ».

C. SCORES DETERMINISTES CONTINUS : ECARTS ENTRE DEBITS OBSERVES ET PREVUS.

Pour le bassin A, les écarts algébriques $q^p - q^{obs}$ moyens et les RMSE moyennes mensuels sont rapportées respectivement dans la Figure 48 et la Figure 49. Par comparaison avec les figures équivalentes restituées au § VIID.2.b avec un seuil fixe, les différences sont faibles, voire insignifiantes. Ceci indique que le choix du seuil n'influence pas notablement les performances des prévisions pour ce bassin.

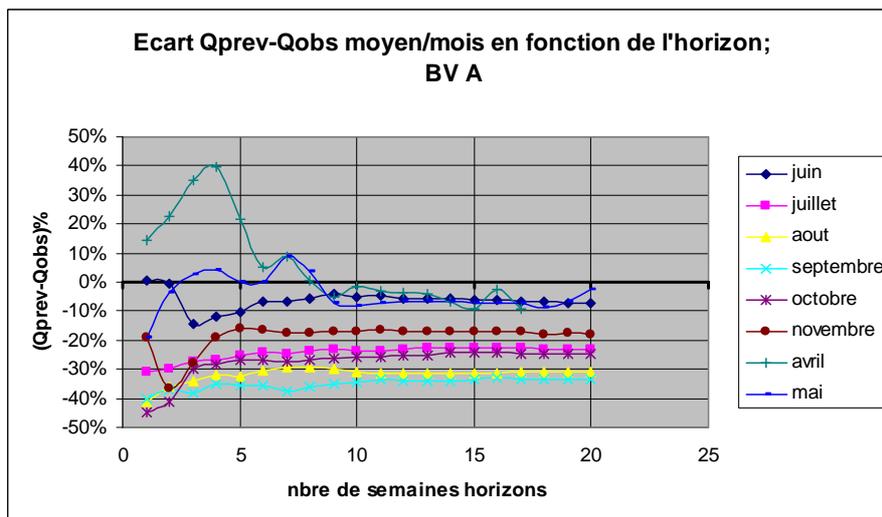


Figure 48: Ecart mensuels relatifs entre débits prévus et observés pour les mois à étiages (seuil variable) du bassin A selon l'horizon de prévision

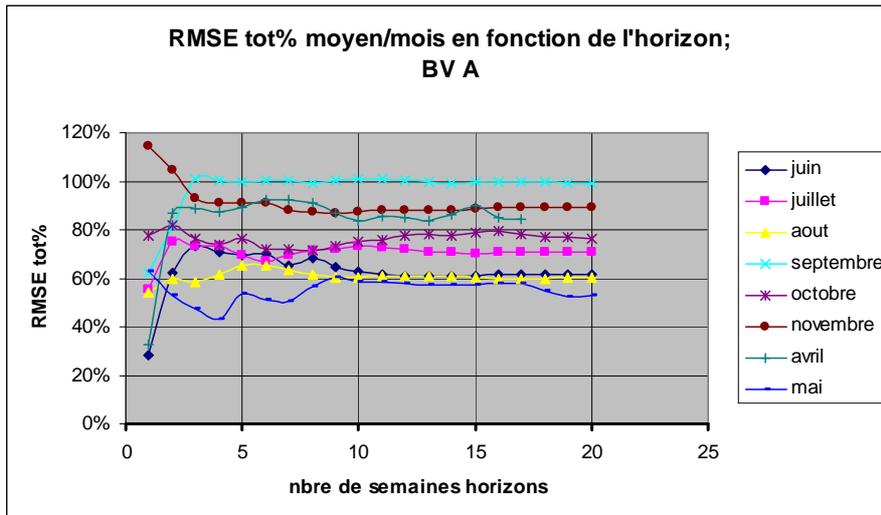


Figure 49: RMSE totales mensuelles relatives pour les mois à étiages du bassin A (seuil variable) selon l'horizon de prévision

Les mêmes quantités pour le bassin J sont rapportées dans les Figure 50 et Figure 51. Cette fois, par comparaison avec les figures équivalentes restituées au § VIID.2.b avec un seuil fixe, quelques différences apparaissent, qui restent néanmoins limitées. Elles concernent le mois d'août pour lequel les écarts restent positifs quelque soit l'horizon, et les RMSE d'août et septembre qui croissent plus régulièrement avec l'horizon.

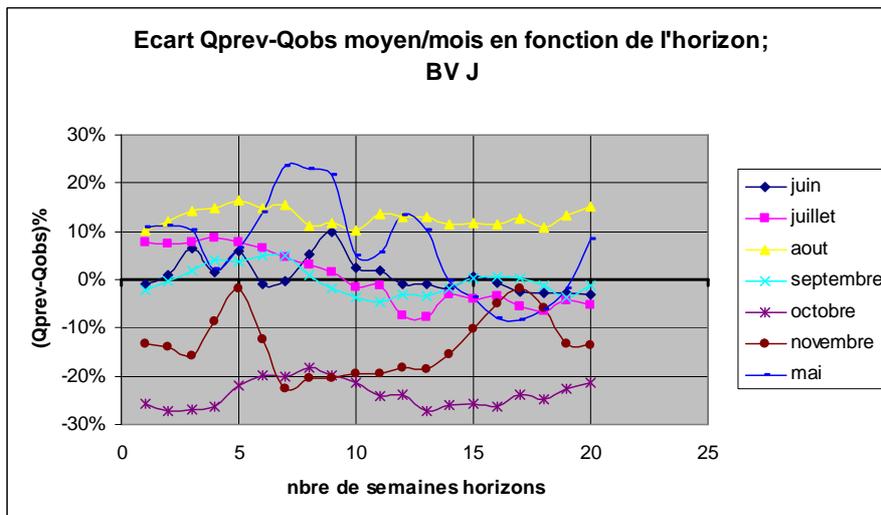


Figure 50: Ecart mensuels relatifs entre débits prévus et observés pour les mois à étiages (seuil variable) du bassin J selon l'horizon de prévision

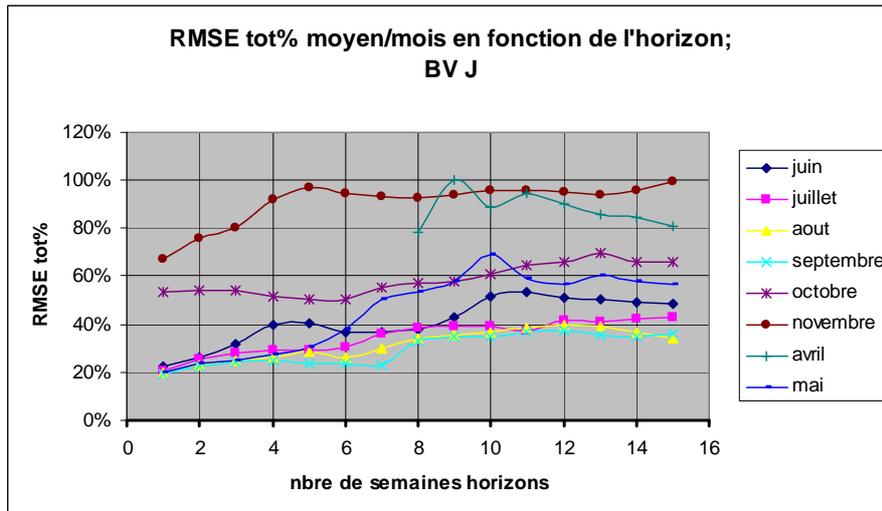


Figure 51: RMSE totales mensuelles relatives pour les mois à étiages du bassin J (seuil variable) selon l'horizon de prévision

D. SCORES RPS.

Les résultats du score RPS pour le bassin A, rapportés dans la Figure 52 sont quasi identiques à ceux observés précédemment à seuil fixe.

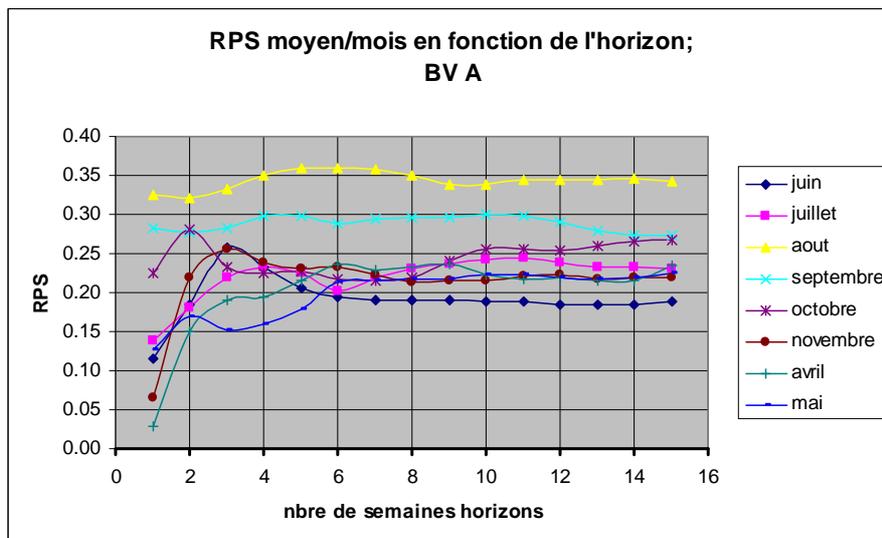


Figure 52: scores RPS mensuels pour le bassin A (seuil variable) en fonction de l'horizon de prévision

Les résultats du score RPS pour le bassin J, rapportés dans la Figure 53 sont différents de ceux observés précédemment à seuil fixe pour les mois d'août et septembre, la montée du score RPS étant plus tardive et mieux marquée à seuil variable. Par ailleurs des résultats sont disponibles pour mai et juin, qui se comportent similairement à ceux de l'été, à savoir une montée nette vers une valeur limite proche de 0.2, là encore comme cela avait été observé à seuil fixe pour les autres mois.

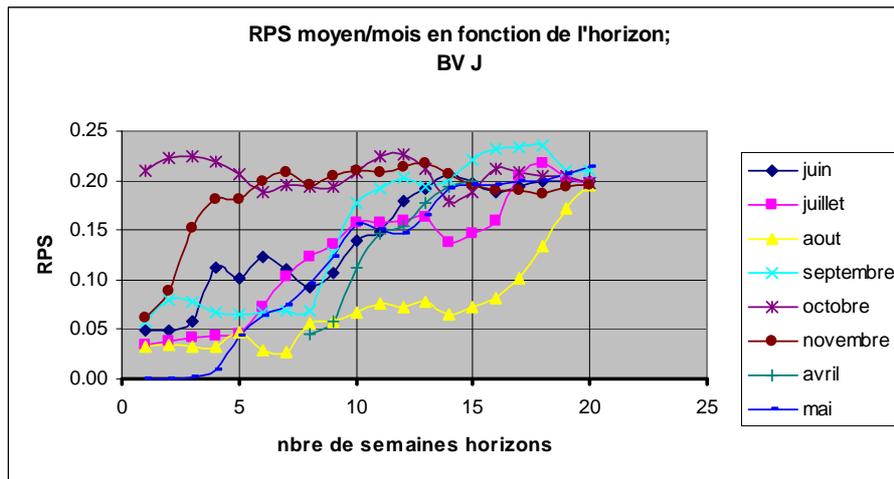


Figure 53: scores RPS mensuels pour le bassin J (seuil variable) en fonction de l'horizon de prévision

E. SCORES DETERMINISTES PAR CATEGORIES : PERFORMANCES DE LA PREVISION EN TERMES D'ALERTE D'ETIAGE.

Les résultats du score CSI pour le bassin A sont rapportés dans la Figure 54. Des différences sont immédiatement perceptibles, le score à seuil variable étant nettement moins contrasté entre mois d'été et autres saisons. Cela est à mettre au compte du fait que les jours d'étiages sont répartis plus équitablement dans l'année, alors qu'à seuil fixe leur nombre élevé en été y favorisait artificiellement le score CSI. Alors qu'à seuil fixe nous comptons jusque 30% de taux d'étiages pour les mois d'été, ce taux est ici typiquement inférieur à 13%, de sorte que le CSI, utilisable uniquement pour des situations à « événements rares », devient pertinent.

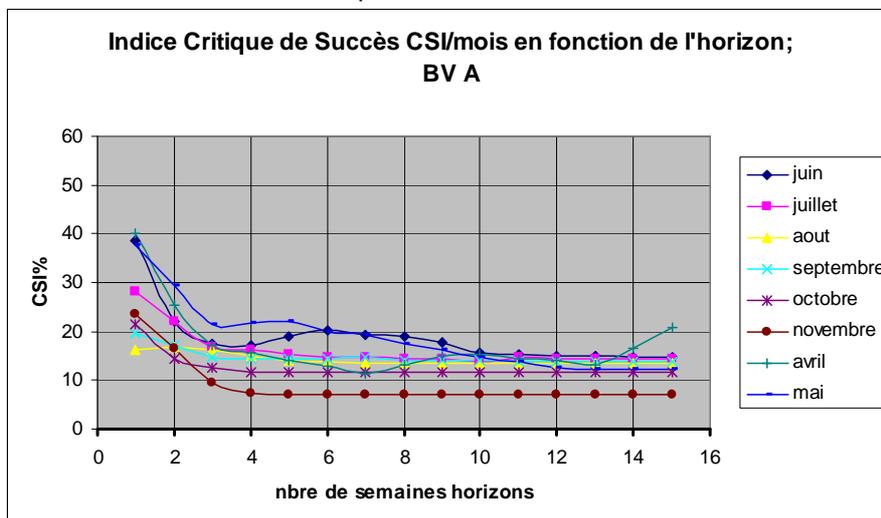


Figure 54: Indices mensuels de Succès Critique pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A

Une observation semblable peut être faite pour les scores ETS (cf. Figure 55).

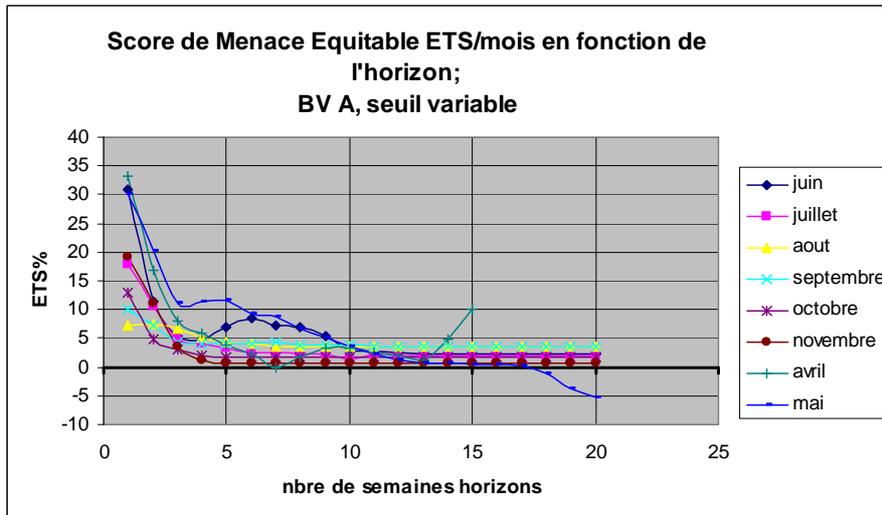


Figure 55: Scores mensuels de Menaces Equitables pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A

Toutefois, dans tous les cas, on retrouve le résultat déjà constaté à seuil fixe pour ce bassin A, à savoir une nette dégradation des scores après le premier horizon de prévision, caractéristique d'un bassin doté d'une mémoire hydrologique très faible.

Le biais à grande échelle (cf. Figure 56) se dégrade lui aussi très vite après le premier horizon, comme cela avait été observé à seuil fixe. Quoique toujours élevées, les valeurs de biais sont toutefois sensiblement différentes du cas seuil fixe, du fait de l'harmonisation des densités d'étiages sur l'année qui ne favorise plus particulièrement les mois d'été comme c'était le cas à seuil fixe.

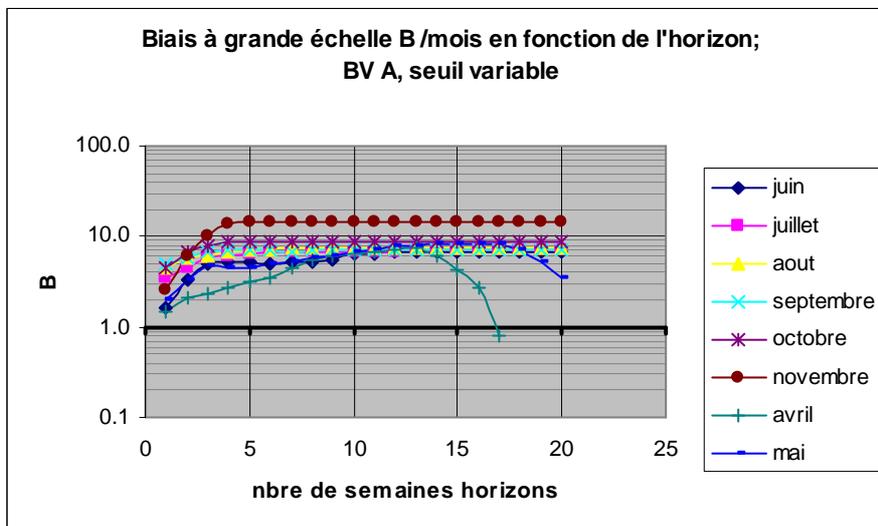


Figure 56: : Biais à grande échelle mensuels pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A

A l'instar de ce qui avait été constaté à seuil fixe, **l'indice de Rousseau** sanctionne par des valeurs négatives les prévisions faites au-delà des tous premiers horizons cf. Figure 57). Cette fois les mois d'été présentent tous des valeurs négatives aux horizons élevés.

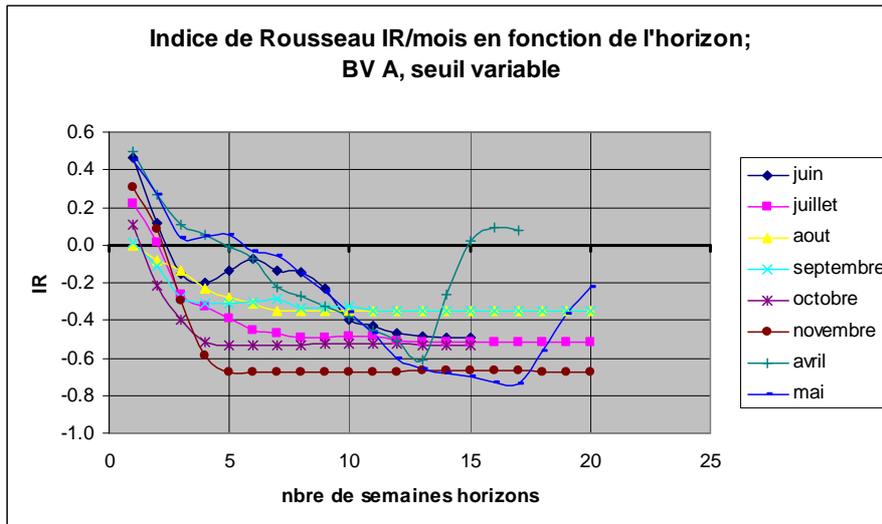


Figure 57: Indices mensuels de Rousseau pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A

Pour le bassin J, les scores CSI (cf. Figure 58) diffèrent de ceux obtenus à seuil fixe surtout pour les mois de juillet et août, le score étant doublé pour le premier, et réduit aux premiers horizons pour le second. On peut avancer que cette différence de comportement serait liée aux gammes de débits prises en compte, les débits d'étiages impliqués en juillet à seuil variable étant sensiblement supérieurs à ceux sélectionnés à seuil fixe, alors que la situation contraire se présente pour août (cf. *Tableau 2* et *Tableau 22*).

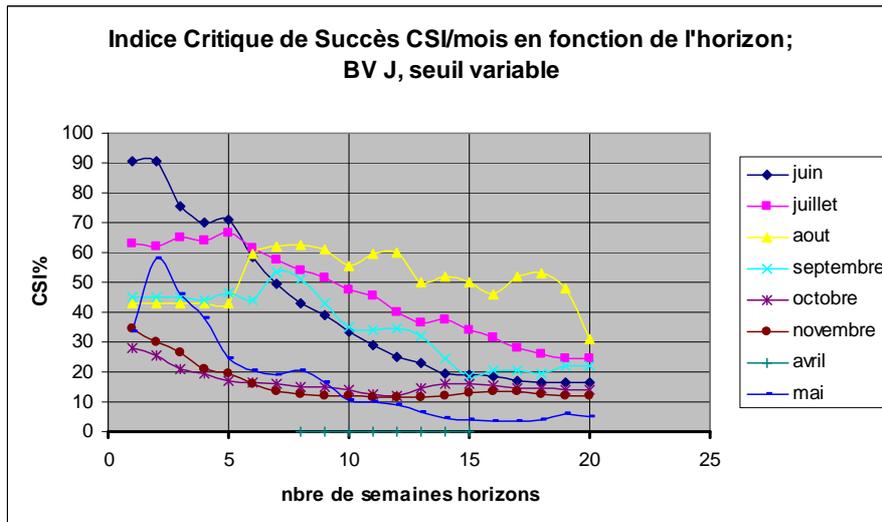


Figure 58: Indices mensuels de Succès Critique pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin J

Les scores ETS (cf. Figure 59) reproduisent la tendance déjà observée à seuil fixe d'une dégradation lente des performances lorsque l'horizon s'accroît, avec une valeur quasi nulle au-delà de 2 à 4 mois d'horizon. Le mois d'août fait toutefois exception, les performances ETS restant élevées sur toute la gamme des horizons.

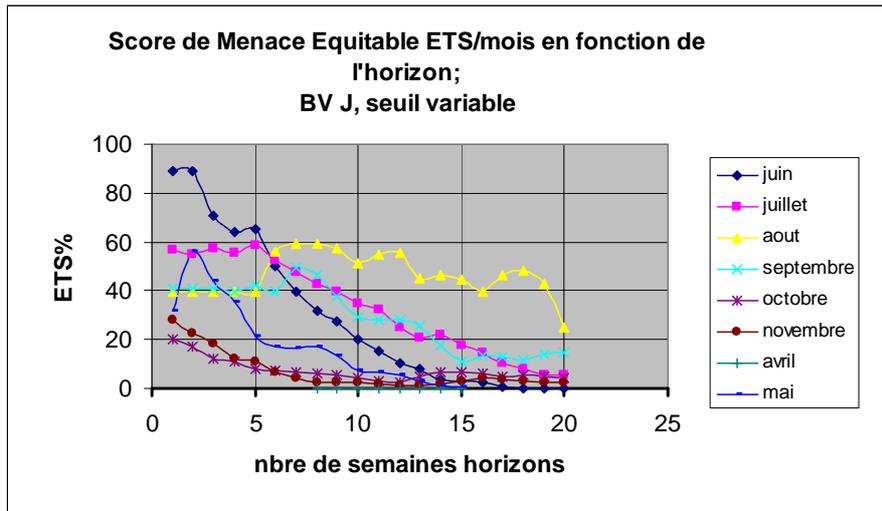


Figure 59: Scores mensuels de Menaces Equitables pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin J

Les biais à grande échelle (cf. Figure 60) s'approchent de ceux obtenus à seuil fixe, avec des valeurs faibles en été et une nette tendance à croître avec l'horizon de prévision.

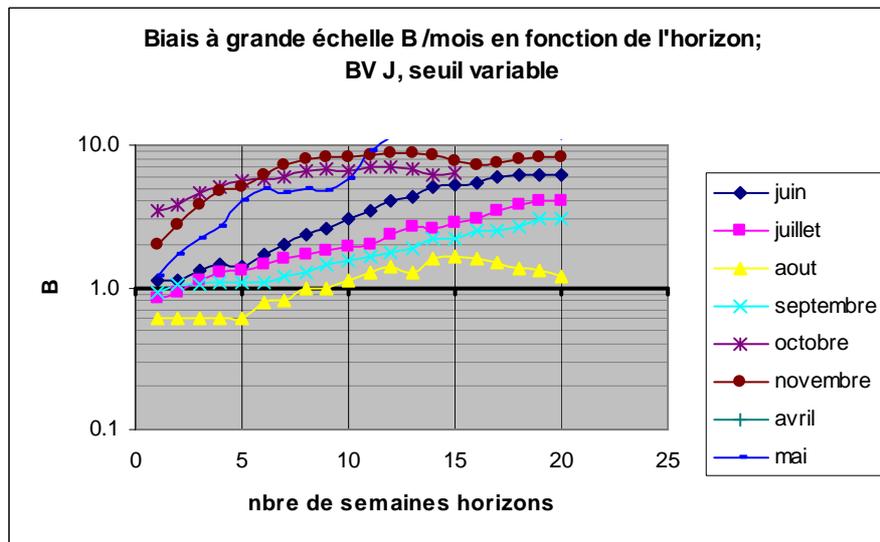


Figure 60: Biais à grande échelle mensuels pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin J

Les Indices de Rousseau (cf. Figure 61) affichent des tendances similaires à celles observées à seuil fixe, avec toutefois l'exception du mois d'août que l'indice ne parvient pas à abattre sous zéro, même aux horizons les plus élevés, du fait que le biais en août se maintient à une valeur faible.

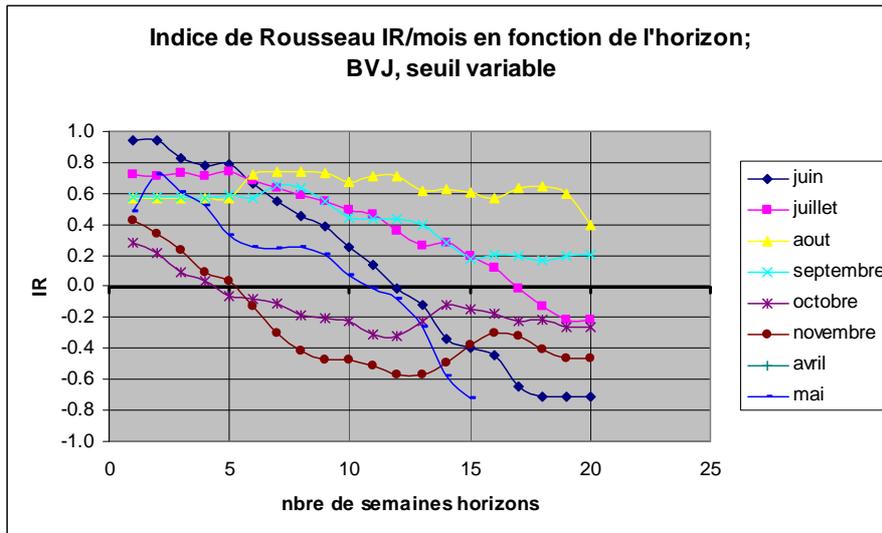


Figure 61: Indices mensuels de Rousseau pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin J

En regard de l'étude que nous avons fait à seuil fixe concernant les garanties offertes par une valeur de l'indice de Rousseau >0.1 (cf. Figure 24) nous présentons dans la Figure 62 le même diagramme à partir des données obtenues à seuil variable. On y vérifie que le critère $IR > 0.1$ permet de garantir des biais à grande échelle < 4 dans la quasi totalité des cas.

La Figure 63 montre que la condition $IR > 0.1$ nous fait également bénéficier de valeurs de $CSI > 20\%$ pour la quasi totalité des cas, valeur minimale que nous avons défini comme critère d'utilité pour les prévisions.

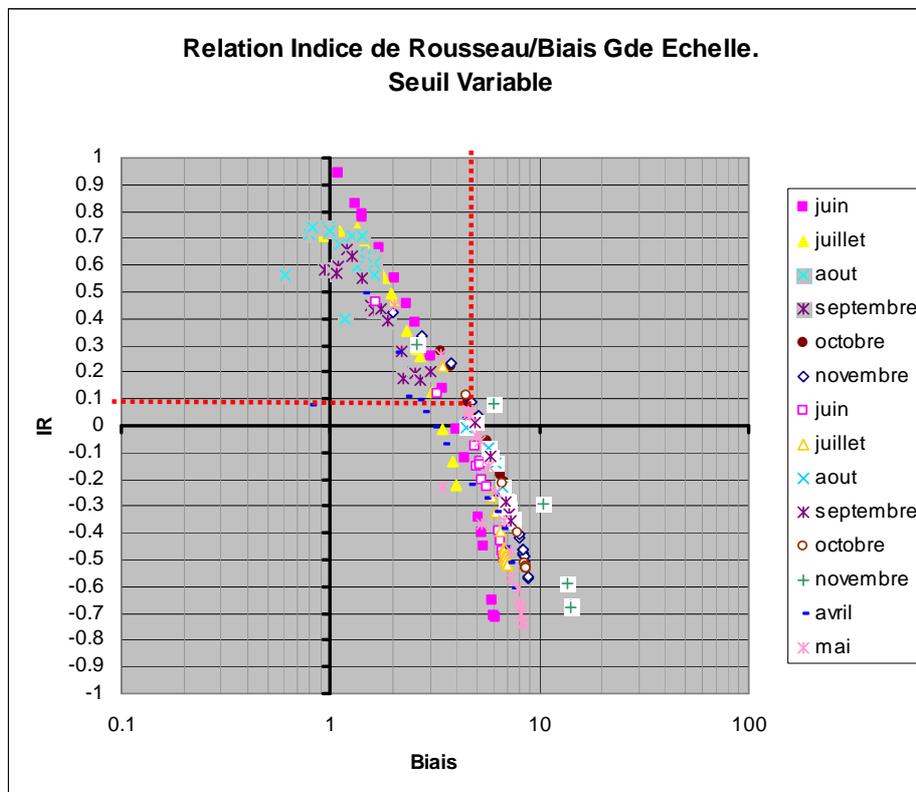


Figure 62: diagramme de corrélation entre l'indice de Rousseau et le biais à grande échelle pour les deux bassins A et J (seuil variable). La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau >0.1

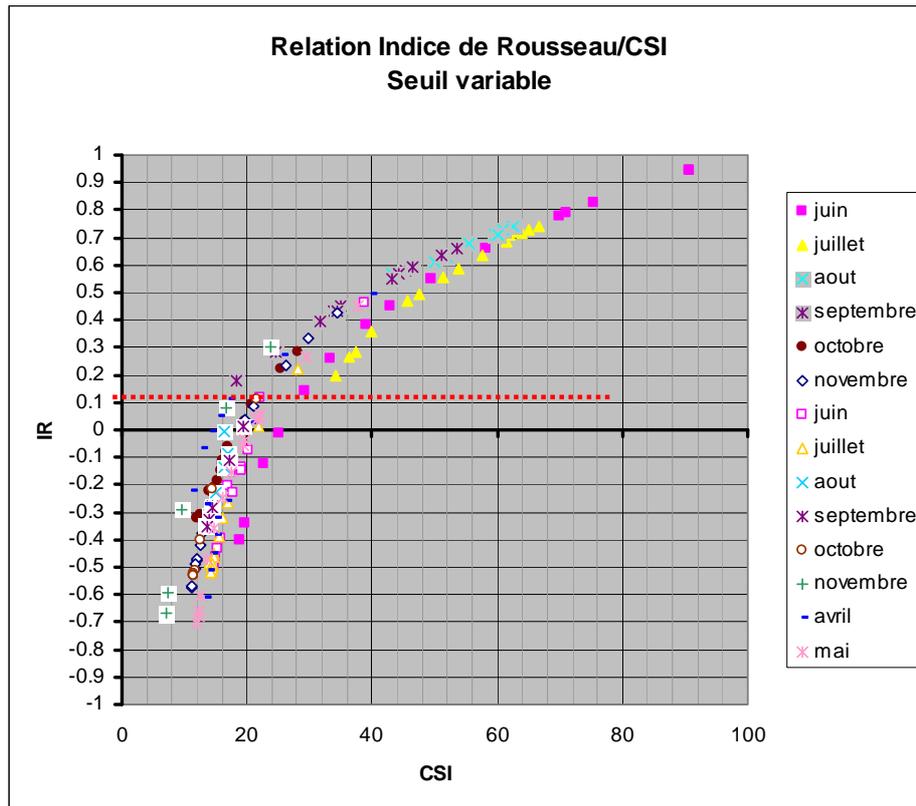


Figure 63: diagramme de corrélation entre l'indice de Rousseau et l'indice de succès critique pour les deux bassins A et J (seuil variable). La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau > 0.1

XVI. ANNEXE H : ETUDE BIBLIOGRAPHIQUE SUR LES GENERATEURS DE PLUIE

Quels sont les divers types de méthodes probabilistes existantes?

i) Prédiction d'ensemble basée sur les archives historiques

Il s'agit dans ce cas de rejouer les scénarios historiques (Day, 1983), quitte à introduire une pondération différentielle entre divers scénarios selon leur degré de vraisemblance, soit en fonction de l'analogie entre l'année courante et les années passées (comme le suggérait déjà Day, 1983), soit en fonction des prévisions climatiques (Wilks, 1992). Druce, 2001, ou Kim, 2001 présentent des exemples d'utilisation des prédictions d'ensemble pour des applications hydrologiques au pas de temps mensuel.

ii) Prévisions Stochastiques

Les méthodes stochastiques génèrent des scénarios de pluies différents de ceux observés historiquement, mais en tâchant de reproduire les propriétés statistiques de ces derniers.

- Méthodes stochastiques paramétriques

Dans ces schémas l'occurrence et la quantité de pluie sont générées indépendamment, typiquement à partir de chaînes de Markov à l'ordre 1. Par exemple dans Perrin, 2001, les pluies sont générées à partir de 3 états qui s'enchainent selon une chaîne de Markov. Les deux états non secs génèrent des pluies selon des distributions différentes: Transitions entre états et distributions sont calibrées sur les chroniques de précipitation.

Utilisée dans le cadre de la génération de chroniques climatiques, la méthode paramétrique produit les précipitations indépendamment des autres variables, ces dernières étant générées à partir des précipitations. Un problème identifié avec ce type de méthode est un manque de fiabilité pour la génération des séquences sèches et humides (Sharif, 2007).

Hayhoe, 2000, propose des améliorations en passant à l'ordre 2 dans les chaînes de Markov. Les paramètres de la chaîne sont estimés à partir des fréquences relatives des jours humides et secs. Les distributions statistiques dans lesquelles tirer (fréquences cumulées pour la quantité de pluie...) sont empiriques, construites sur des intervalles à 2 mois de l'année, et propres à chaque site, permettant ainsi de respecter la saisonnalité.

Notons que pour s'affranchir de la complexité engendrée par le nombre de paramètres induits par l'utilisation des chaînes de Markov, certains auteurs, tel Lawal, 1997, ont recours à une autre théorie, le "alternating renewal process", permettant de reproduire des événements de précipitation de façon stochastique.

- Méthodes stochastiques non paramétriques

Ces méthodes présentent l'avantage initial de ne pas requérir de connaissance a priori des distributions statistiques selon lesquelles les phénomènes climatiques se manifestent. Par ailleurs la corrélation temporelle entre les diverses variables climatiques est respectée. La méthode la plus employée est celle des "K Nearest Neighbors" (Rajagopalan and Lall, 1999) qui consiste à construire des "feature vectors" constitués de plusieurs paramètres climatiques significatifs propres à chaque jour des chroniques de données, puis à générer les vecteurs des jours futurs à partir des K vecteurs ayant les plus de similitude avec les jour précédent. Dans la méthode originale (cf. Buishand and Brandsma, 2001) le vecteur final est tiré au sort parmi les K plus proches voisins. Une publication plus récente propose une alternative consistant à prendre un vecteur moyen parmi un sous-ensemble des K voisins (Bannayan, 2007), ce qui s'avère intéressant pour simuler les faibles précipitations, mais pas les plus élevées. Sharif 2007 propose une version de la méthode KNN avec perturbation des scénarios historiques pour produire des nouvelles valeurs au-delà de celles observées historiquement et ainsi pouvoir prévoir des scénarios extrêmes. La méthode KNN a également été employée pour tenter de sélectionner dans la chronique historique "l'année la plus proche" de celle pour laquelle il s'agit de faire des prédictions, sur la base des 90 à 150 premiers jours précédemment enregistrés dans cette année de prédiction (Bannayan, 2007).

Comment évaluer la pertinence des générateurs stochastiques de pluie pour la prévision d'étéage?

Deux types de propriétés sont favorables pour la sélection d'un générateur de pluie:

- i) les pluies générées doivent respecter les caractéristiques statistiques des pluies historiques qu'il exploite. Hayhoe 2000 décrit les tests statistiques usuels pour détecter les écarts à la statistique historique, en termes de valeurs moyennes de précipitation, d'écart types, et de forme des distributions.

En général un nombre d'années conséquent est requis pour faire fonctionner le générateur (15 années, d'après Soltani, 2003), les générateurs non-paramétriques semblant donner de meilleurs résultats que les méthodes paramétriques si ces 15 années de référence sont disponibles (Lerat, 2007).

Les générateurs vertueux sont ceux qui réussissent :

- (1) à reproduire la persistance de certaines séquences sèches ou humides, ce qui est un point critique pour la prévision d'étéage.
- (2) à reproduire la variabilité interannuelle (Mavromatis, 2001)
- (3) à générer des événements extrêmes au-delà des valeurs observées dans les chroniques passées (Sharif, 2007).

- ii) il doit être possible d'incorporer facilement des informations climatiques externes traduisant des tendances à moyen ou long terme. Un certain nombre de travaux ont

démontré comment prendre en compte des prédictions climatiques (Global Climate Change, prédictions saisonnières du International Research Institute [IRI] for Climate Predictions), dans le cadre de générateur stochastique (Hayhoe, 2000 [méthode paramétrique]; Wilks, 1992, 2001 [paramétrique]; Yates, 2003 [non-paramétrique], Sharif, 2006 [non-paramétrique]) ou bien en termes de probabilité d'occurrence de données basses, moyennes ou hautes, pour mettre à jour la pondération des scénarios historiques dans une stratégie de prédictions d'ensemble (Stedinger, 2002; Croley, 2003).

XVII. BIBLIOGRAPHIE

- Abi-Zeid, I., 1999. La modélisation stochastique des étiages: une revue bibliographique. *Revue des sciences de l'eau* 12(3): 459-484.
- Bannayan, M. , 2008. Weather analogue: A tool for real-time prediction of daily weather data realizations based on a modified k-nearest neighbor approach. *Environmental modelling & software* 23: 703-713.
- 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, Institut National Polytechnique de Grenoble, Ecole Doctorale Terre, Univers, Environnement, Grenoble, 262 pp.
- Croley, T. E., 2003. Weighted-Climate Parametric Hydrologic Forecasting. *Journal of Hydrologic Engineering* July-August: 171-180.
- Day, G.N. , 1983. Extended Streamflow Forecasting using NWSRFS. *Journal of water resources planning and management*. 111(2): 157-170.
- Druce, D. J., 2001. Insights from a history of seasonal inflow forecasting with a conceptual hydrologic model. *Journal of Hydrology* 249: 102-112.
- Epstein, E.S., 1969. A scoring system for probability forecast of ranked categories. *Journal of Applied Meteorology* 8:985-987
- Hayhoe, H. N., 2000. Improvements of stochastic weather data generators for diverse climates. *Climate Research* 14: 75-87.
- Kim, Y-O., Hyoung-Seop Ki, 2001. Improving Water Supply Outlook in Korea with Ensemble Streamflow Prediction. *Water International* 26(4): 563-568.
- Klemes, V., 1986. "Operational testing of hydrological simulation models." *Hydrological sciences* 31(1): 13-24.
- Lawal, S.A., Watts, D.G. , 1997. A stochastic model of low flows. *Stochastic Hydrology and Hydraulics* 11: 303-321.
- Mavromatis, T. , 2001. Interannual variability characteristics and simulated crop response of four stochastic weather generators. *Agricultural and Forest Meteorology* 109: 283-296.
- Michel , C., 1989. Modèle pluie-débit de prévision des étiages en valeurs mensuelles. Rapport d'étude, Secrétariat d'Etat à l'Environnement, Cemagref, 24 p.
- Nash, J.E. and Sutcliffe, J.V. 1970. River flow forecasting through conceptual models. Part I - A discussion of principles. *J. Hydrol.*, 27(3), 282-290.
- Oudin, L., Michel, C. and Anctil, F., 2005. Which potential evapotranspiration input for a lumped rainfall-runoff model? - Part 1 - Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? *Journal of Hydrology*, 303(1-4): 275-289.
- Perrin, C., Michel, C., Andréassian, V., 2001. Long-term low flow forecasting for French rivers by continuous rainfall-runoff modelling, Meeting of the British Hydrological Society on Continuous River Flow Simulation. BHS Occasional Paper n°13, Wallingford, UK, 5th July 2001, pp. 21-29.
- Perrin, C., 2002. Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative *La Houille Blanche* 279(1-4) : 275-289.

Rousseau, D. 1980. A new skill score for the evaluation of yes/no forecasts.. WMO (Ed.), *WMO Symposium on probabilistic and statistical method in weather forecasting*, Nice, France, 8-12 Septembre..

Sauquet, E., Lerat, J., & Prudhomme, C. -La prévision hydro-météorologique à 3-6 Mois - Etat des connaissances et applications. *Congrès SHF- Gestion sociale et économique de l'eau - Comment agir sur la demande ?* Paris, 17-18 octobre 2007

Sharif, M., 2007. Improved K-Nearest Neighbor Weather Generating Model *Journal of Hydrologic Engineering* 12(1): 42-51.

Smakthin, V.Y., Hughes, D.A., 1998. Evaluating the performance of a deterministic daily rainfall-runoff model in low-flow context, *Hydrological Processes* 12: 797-811.

Smakthin, V. Y., 2001. Low flow hydrology: a review. *Journal of Hydrology* 240: 147-186.

Soltani, A., 2003. A statistical comparison of the stochastic weather generators WGEN and SIMMETEO. *Climate Research* 24: 215-230.

Stedinger, J. 2002. Updating Ensemble Probabilities Based on Climate Forecasts. *Proc., Water Resources Planning and Management (19-22 May, Roanoke, Virginia)*.

Wilks, D. S., 1992. Adapting stochastic weather generation algorithms for climate change studies. *Climate Change* 22: 67-84.

Wilks, D. S., 2001. Realizations of Daily Weather in Forecast Seasonal Climate. *Journal of Hydrometeorology* 3: 196-207.

Yates, D., 2003. A technique for generating regional climate scenarios using a nearest-neighbor algorithm. *Water Resources Research* 39(7): swc 7 1-14.

INDEX DES FIGURES

Figure 1: schéma du modèle GR4J décrivant les entrées, les variables internes, les réservoirs et paramètres associés.	8
Figure 2 : diagramme générique de fonctionnement de l'outil de prévision TAGE.....	12
Figure 3: représentation schématique de la stratégie de prévision adoptée dans TAGE. La courbe bleue symbolise un hydrogramme. Pour chaque date de prévision j_p une série de n_{max} prévisions sont émises concernant autant de dates cibles j , chacune d'elle étant associée au débit lissé sur la période cible correspondante.	15
Figure 4: diagramme représentant la stratégie du « split-sample test » utilisé lors de la phase de calage-contrôle du modèle hydrologique sur une chronique de $n_{An}+1$ années de mesures (P,E,q). ...	18
Figure 5: illustration pour toute l'année 1997 sur le bassin J des seuils d'étiages variable (trait rouge continu) et fixe (rouge interrompu) superposés sur les hydrogrammes observé (bleu) et simulé (vert). Les triangles noirs matérialisent les jours associés à des alertes d'étiage.	25
Figure 6: zoom de la figure précédente sur la période mai-août 1997.....	25
Figure 7: distribution des écarts entre débits prévus par le modèle H et observés pour les mois à étiage pour le bassin A. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.....	27
Figure 8: distribution des écarts entre débits prévus par le modèle H et observés pour les mois à étiage pour le bassin J. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.....	28
Figure 9: analyse des étiages prévus par le modèle H par comparaison avec les étiages observés pour le bassin A.	30
Figure 10: Nombre de jours d'étiages prévus par TAGE-GR4J pour les deux bassins en fonction de l'horizon de prévision. Comparaison avec le modèle H.	31
Figure 11: Nombres mensuels de jours d'étiage observés et prévus par TAGE-GR4J en fonction de l'horizon de prévision, pour le bassin J.	32
Figure 12: RMSE totales mensuelles relatives pour les mois à étiages du bassin J selon l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.....	33
Figure 13: Ecart mensuels relatifs entre débits prévus et observés pour les mois à étiages du bassin J selon l'horizon de prévision	33
Figure 14: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la première semaine d'horizon, bassin J. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.	34
Figure 15: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la huitième semaine d'horizon, bassin J. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.	34
Figure 16: scores RPS mensuels pour le bassin J en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.....	35
Figure 17 : taux mensuels de fausses alertes pour le bassin J en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.	36
Figure 18 : probabilité de détection des étiages pour le bassin J en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.	36
Figure 19 : <i>Indices mensuels de Succès Critique pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.</i>	37
Figure 20 : <i>Scores mensuels de Menaces Equitables pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.</i>	37
Figure 21 : <i>Biais à grande échelle mensuels pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.</i>	38
Figure 22 : <i>Indices mensuels de Rousseau pour la détection d'étiages en fonction de l'horizon de prévision pour le bassin J. Le dernier point de chaque mois correspond à la performance du modèle H.</i>	38

Figure 23 : Indices mensuels de Rousseau pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin A. Le dernier point de chaque mois correspond à la performance du modèle H.	39
Figure 24: diagramme de corrélation entre l'indice de Rousseau et le biais à grande échelle pour les deux bassins A et J (seuil fixe). Les chiffres en encadré représente le nombre d'étiages observés pour les mois correspondants. La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau>0.1	40
Figure 25: diagramme de corrélation entre l'indice de Rousseau et l'indice de succès critique pour les deux bassins A et J (seuil fixe). La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau>0.1	41
Figure 28: Comparaison des Horizons de Prévision Utiles Maximums mensuels obtenus avec TAGE-GR4J pour les deux bassins et les deux types d'étiages considérés dans ce travail.	46
Figure 29: Diagramme de répartition des horizons de prévision utiles maximaux caractéristiques pour les 261 bassins étudiés.	48
Figure 30: Carte des 261 bassins pour lesquels des prévisions ont été émises. La couleur des bassins est codée en référence à la valeur de HPUM(BV), lequel est exprimé en semaines.	49
Figure 31: Carte des 261 bassins pour lesquels des prévisions ont été émises. La couleur des bassins est codée en référence à la valeur de Nash-LogQ (sans unité, variant entre 0 et 1).	51
Figure 32: Carte des 261 bassins pour lesquels des prévisions ont été émises. La couleur des bassins est codée en référence à la valeur du logarithme du paramètre X3, lequel exprime la capacité du réservoir de routage.	52
Figure 33: superposition des débits observés et simulés avec GR4J pour toute la chronique 1995-2005 du bassin A.	55
Figure 34: superposition des débits observés et simulés avec GR4J pour toute la chronique 1995-2005 du bassin J	56
Figure 35: taux de remplissage du réservoir de routage au cours de la chronique 1995-2005 pour les deux bassins tests A et J.	57
Figure 36: Ecarts mensuels relatifs entre débits prévus et observés pour les mois à étiages (seuil fixe) du bassin A selon l'horizon de prévision	60
Figure 37: RMSE totales mensuelles relatives pour les mois à étiages du bassin A (seuil fixe) selon l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H. .	61
Figure 38: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la première semaine d'horizon, bassin A. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.	61
Figure 39: distributions mensuelles des écarts entre débits d'étiage observés et prévus pour la seconde semaine d'horizon, bassin A. Le trait central correspond à la médiane, les ordonnées de la boîte sont les quantiles 25% et 75%, les extrémités étant les quantiles 10% et 90%.	62
Figure 40: scores RPS mensuels pour le bassin A (seuil fixe) en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.	62
Figure 41: Amplitudes mensuelles moyennes de la catégorie centrale définissant le score RPS pour les deux bassins	63
Figure 42a : taux mensuels de fausses alertes pour le bassin A en fonction de l'horizon de prévision. Le dernier point de chaque mois correspond à la performance du modèle H.	64
Figure 43: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score RMSE pour le bassin A.	67
Figure 44: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score RMSE pour le bassin J.	67
Figure 45: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score ETS pour le bassin A.	68
Figure 46: scores de compétences mensuels de TAGE-GR4J vis-a-vis du modèle H obtenus avec le score ETS pour le bassin J.	68
Figure 47: Nombre de jours d'étiages prévus par TAGE-GR4J pour les deux bassins en fonction de l'horizon de prévision. Comparaison entre les cas « seuil variable » et « seuil fixe ».	71
Figure 48: Ecarts mensuels relatifs entre débits prévus et observés pour les mois à étiages (seuil variable) du bassin A selon l'horizon de prévision	71
Figure 49: RMSE totales mensuelles relatives pour les mois à étiages du bassin A (seuil variable) selon l'horizon de prévision	72
Figure 50: Ecarts mensuels relatifs entre débits prévus et observés pour les mois à étiages (seuil variable) du bassin J selon l'horizon de prévision	72

Figure 51: RMSE totales mensuelles relatives pour les mois à étiages du bassin J (seuil variable) selon l'horizon de prévision	73
Figure 52: scores RPS mensuels pour le bassin A (seuil variable) en fonction de l'horizon de prévision	73
Figure 53: scores RPS mensuels pour le bassin J (seuil variable) en fonction de l'horizon de prévision	74
Figure 54: <i>Indices mensuels de Succès Critique pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A</i>	74
Figure 55: Scores mensuels de Menaces Equitables pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A	75
Figure 56: <i>: Biais à grande échelle mensuels pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A</i>	75
Figure 57: <i>Indices mensuels de Rousseau pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin A</i>	76
Figure 58: <i>Indices mensuels de Succès Critique pour la détection d'étiages (seuil variable) en fonction de l'horizon de prévision pour le bassin J</i>	76
Figure 59: Scores mensuels de Menaces Equitables pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin J	77
Figure 60: Biais à grande échelle mensuels pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin J	77
Figure 61: Indices mensuels de Rousseau pour la détection d'étiages (seuil fixe) en fonction de l'horizon de prévision pour le bassin J	78
Figure 62: diagramme de corrélation entre l'indice de Rousseau et le biais à grande échelle pour les deux bassins A et J (seuil variable). La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau>0.1	78
Figure 63: diagramme de corrélation entre l'indice de Rousseau et l'indice de succès critique pour les deux bassins A et J (seuil variable). La zone délimitée par des traits rouges correspond au domaine des cas présentant un indice de Rousseau>0.1	79

INDEX DES TABLEAUX

<i>Tableau 1 : nombre de jours d'étiages à seuil fixe observés chaque mois pour les deux bassins</i>	26
<i>Tableau 2 : débits moyens des jours correspondants à des étiages à seuil fixe observés pour les deux bassins</i>	26
<i>Tableau 3 : comparaison des nombres mensuels de jours d'étiage observés et prévus par H pour le bassin J</i>	27
<i>Tableau 4 : comparaison des nombres mensuels de jours d'étiage observés et prévus par H pour le bassin A</i>	27
<i>Tableau 5 : scores d'erreur relatives mensuelles obtenus avec le modèle H pour le bassin A</i>	28
<i>Tableau 6 : scores d'erreur relatives mensuelles obtenus avec le modèle H pour le bassin J</i>	28
<i>Tableau 7 : scores RPS mensuels obtenus par le modèle H pour le bassin A</i>	29
<i>Tableau 8 : scores RPS mensuels obtenus par le modèle H pour le bassin J</i>	29
<i>Tableau 9 : scores mensuels d'alertes pour le bassin A obtenus avec le modèle H</i>	29
<i>Tableau 10 : :scores mensuels d'alertes pour le bassin J obtenus avec le modèle H</i>	30
<i>Tableau 11 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin A à seuil fixe, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM et avec le modèle H</i>	44
<i>Tableau 12 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin J à seuil fixe, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM et avec le modèle H</i>	44
<i>Tableau 13 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin A à seuil variable, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM</i>	45
<i>Tableau 14 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin J à seuil variable, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM</i>	45

<i>Tableau 15 : Horizons de prévision mensuels obtenus avec TAGE-GR4J pour le Bassin A à seuil variable, dont le HPUM. Le tableau indique également les scores CSI et Biais à grande échelle obtenus pour l'horizon HPUM.....</i>	<i>50</i>
<i>Tableau 16 : critère de Nash-Sutcliffe obtenus pour les deux bassins pendant la phase de calage-contrôle.....</i>	<i>55</i>
<i>Tableau 17 : valeurs après calage des 4 paramètres de GR4J pour le bassin A.....</i>	<i>56</i>
<i>Tableau 18 : valeurs après calage des 4 paramètres de GR4J pour le bassin J.....</i>	<i>56</i>
<i>Tableau 19 : Scores de Menace Equitable mensuels et globaux annuels obtenus pour les deux bassins pour trois critères d'alertes différents.....</i>	<i>58</i>
<i>Tableau 20 : Indices de Rousseau mensuels et globaux annuels obtenus pour les deux bassins pour trois critères d'alertes différents.....</i>	<i>58</i>
<i>Tableau 21: nombre de jours d'étiages à seuil variable observés chaque mois pour les deux bassins.....</i>	<i>70</i>
<i>Tableau 22 : débits moyens des jours correspondants à des étiages à seuil variable observés pour les deux bassins.....</i>	<i>70</i>