



**HAL**  
open science

## Prévision de seuils de vigilance sécheresse: tests de régression logistique avec des données hydro-climatiques

Lila Collet, Guillaume Thirel, Charles Perrin, Joël Gailhard, Jean-Philippe Vidal, Fabienne Rousset, Jean-Pierre Wagner

### ► To cite this version:

Lila Collet, Guillaume Thirel, Charles Perrin, Joël Gailhard, Jean-Philippe Vidal, et al.. Prévision de seuils de vigilance sécheresse: tests de régression logistique avec des données hydro-climatiques. [Rapport de recherche] INRAE. 2020, pp.25. hal-03121126

**HAL Id: hal-03121126**

**<https://hal.inrae.fr/hal-03121126v1>**

Submitted on 22 Jun 2021

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

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



# CHlers - Meuse : Evolution du RégimE hydrologique au 21e siècle

---

Rapport spécifique :  
**Prévision de seuils de vigilance sécheresse : tests de régression logistique  
avec des données hydro-climatiques**

**Autrice : Lila Collet**

**Contributeurs :** Guillaume Thirel, Charles Perrin, Joël Gailhard, Jean-Philippe Vidal,  
Fabienne Rousset, Jean-Pierre Wagner

## Résumé :

Cette étude spécifique du projet CHIMERE21 avait pour objectif de proposer et d'évaluer une méthodologie de prévision de seuils de vigilance sécheresse en utilisant des données hydro-climatiques et un modèle de régression logistique. Nous développons dans ce rapport la méthodologie et différents tests effectués afin d'évaluer la configuration optimale et les performances par rapport à un modèle de persistance naïf. Différents tests sont effectués afin de déterminer la transférabilité spatiale et temporelle de la méthode.

Nous montrons que les performances d'un tel modèle restent faibles pour les différents seuils de vigilance : pour les toutes premières échéances de prévision (jusqu'à un mois), le modèle n'apporte aucune information par rapport au modèle de persistance. A plus longue échéance, le modèle devient meilleur que le modèle naïf, mais avec une forte dégradation des performances absolues. Ces résultats tendent à indiquer que le pouvoir prédictif de ce type de modèle est vraisemblablement insuffisant pour des objectifs opérationnels et qu'il vaudrait mieux s'orienter vers d'autres types d'approches, par exemple basées sur de la modélisation hydrologique.

## Table des matières

Table des matières.....	2
1. Contexte .....	3
2. Données disponibles .....	3
2.1. Données spatiales.....	3
2.2. Données hydrologiques .....	5
2.3. Prise en compte spatiale des cartes d'indicateurs hydro-climatiques .....	6
3. Méthodologie .....	7
3.1. Modèle de régression logistique .....	7
3.2. Tests appliqués.....	7
3.3. Critères de performance.....	10
4. Résultats pour le seuil de vigilance jaune.....	11
4.1. Choix du modèle optimum .....	11
4.2. Performance du modèle optimum aux échéances futures .....	15
5. Résultats pour les seuils de vigilance orange et rouge .....	17
5.1. Prévision du seuil de vigilance orange.....	18
5.2. Prévision du seuil de vigilance rouge.....	20
6. Robustesse spatiale des modèles optimum de vigilance jaune, orange et rouge .....	22
7. Conclusions et perspectives .....	23
Références bibliographiques.....	25

## 1. Contexte

La partie française du bassin de la Meuse a été identifiée comme vulnérable au risque sécheresse par la DREAL Grand-Est (Direction Régionale de l'Environnement, de l'Aménagement et du Logement) et ce risque devrait s'intensifier en termes de durée et de magnitude en contexte de changement climatique. Les gestionnaires ont besoin de nouveaux outils statistiques pour mieux anticiper les sécheresses et souhaitent utiliser les données climatiques spatiales existantes pour compenser des données hydrologiques parfois manquantes sur certains bassins. Ces enjeux sont abordés dans le projet CHIMERE21, soutenu par l'Agence de l'eau Rhin-Meuse, dans lequel les impacts du changement climatique sur les ressources en eau du bassin de la Meuse sont analysés. **Ce rapport spécifique s'attache particulièrement à utiliser et développer des indicateurs simples pour aider à la prise de décision pour la gestion des étiages et sécheresses sur ce bassin.**

En France, Météo-France fournit aux Agences de l'eau des cartes de différents indicateurs aux pas de temps décennaire, particulièrement des cartes de Standardized Soil Wetness Index (SSWI), un indicateur couramment utilisé dans la littérature pour quantifier l'humidité du sol. Les hydrologues de la DREAL Grand-Est, en charge de la gestion des bassins du Rhin et de la Meuse française, aimeraient utiliser ces cartes fournies par Météo-France pour caractériser les sécheresses hydrologiques et mieux appréhender et s'adapter au risque sécheresse. Pour définir ce risque hydrologique, la DREAL Grand-Est utilise généralement une approche de seuil hydrologique, le seuil correspondant à un débit prédéfini. Quand le débit passe sous un seuil défini, un état de vigilance sécheresse, plus ou moins sévère, est atteint et déclaré, et des mesures de gestion préalablement définies sont alors mises en place.

Le travail présenté dans ce rapport implique les gestionnaires et chercheurs travaillant sur le bassin de la Meuse française et cherche à développer une nouvelle approche pour la prévision de différents seuils de vigilance sécheresse. Pour cela, une régression logistique est utilisée pour construire des relations statistiques simples entre des indices hydro-climatiques spatialisés observés au pas de temps décennaire et la probabilité d'atteindre un seuil de vigilance. **On cherche ici à définir un modèle de prévision efficace pour différents niveaux de vigilance sécheresse et à en évaluer la robustesse temporelle et spatiale à différentes échéances de prévision.**

## 2. Données disponibles

### 2.1. Données spatiales

Météo-France produit des cartes d'indices hydro-climatiques (Vidal et al., 2010) sur une grille de 8 km sur toute la France. L'indice d'humidité des sols standardisé (SSWI) est disponible au pas de temps décennaire de 1958 à nos jours (2020). Similairement à l'indice de précipitations standardisé (SPI), un indicateur plus largement utilisé (voir Mc Kee et al., 1993), le SSWI s'appuie sur la moyenne décennaire de l'indice d'humidité des sols (Soil Wetness Index, SWI), et est issu des données journalières de la chaîne Safran-Isba-Modcou (SIM) disponibles dans la base de données BDClim (1990).

Le SSWI est calculé sur chaque maille de la façon suivante. Tout d'abord, les chroniques journalières de SWI sont moyennées au pas de temps décennaire. Ensuite, la distribution statistique décennaire est ajustée pour la période 1981-2010 sur la distribution théorique caractérisée par sa fonction de densité. Une fonction de densité est ajustée pour chaque pas de temps (e.g. à chaque décennie) sur la période 1981-2010. Les distributions sont ensuite projetées sur une loi normale centrée réduite de moyenne 0 et d'écart type 1 pour obtenir les valeurs standardisées de l'indicateur.

Le SSWI caractérise des sécheresses pédologiques, particulièrement sensibles dans le domaine agricole. Cet indicateur met en évidence des périodes plus humides (pour des valeurs positives) ou

plus sèches (pour des valeurs négatives) que la normale sur 1981-2010. Le SSWI intègre l'ensemble des paramètres du bilan hydrique, notamment la température et l'évapotranspiration réelle (ETR) des sols qui peut s'avérer limitée, surtout en période de forte sécheresse des sols. Ainsi, le SSWI est généralement en décalage avec le signal de précipitations : les sols s'assèchent avec une inertie variable par rapport au déficit de précipitations, marquée par le cycle saisonnier (faible évaporation hivernale et forte évaporation estivale).

Les utilisateurs des DREAL ont accès à ces données sous forme cartographique uniquement, sur toute la France (voir Fig. 1). Les codes couleurs utilisés sur les cartes tracées pour les DREAL s'appuient sur différentes périodes de retour (voir Tab. 1) :

**Tableau 1** : Catégorisation des indices hydro-climatiques produits par Météo-France

Période de retour (années)	Valeur de l'indice SSWI	Catégorie
≥25	≥1,75	Extrêmement humide
[10 - 25[	[1,28 – 1,75[	Très humide
[5 - 10[	[0,84 – 1,28[	Modérément humide
[0 - 5[	] -0,84 – 0,84[	Autour de la normale
[5 - 10[	] -1,28 – -0,84[	Modérément sec
[10 - 25[	] -1,75 – -1,28[	Très sec
≥25	≤ -1,75	Extrêmement sec

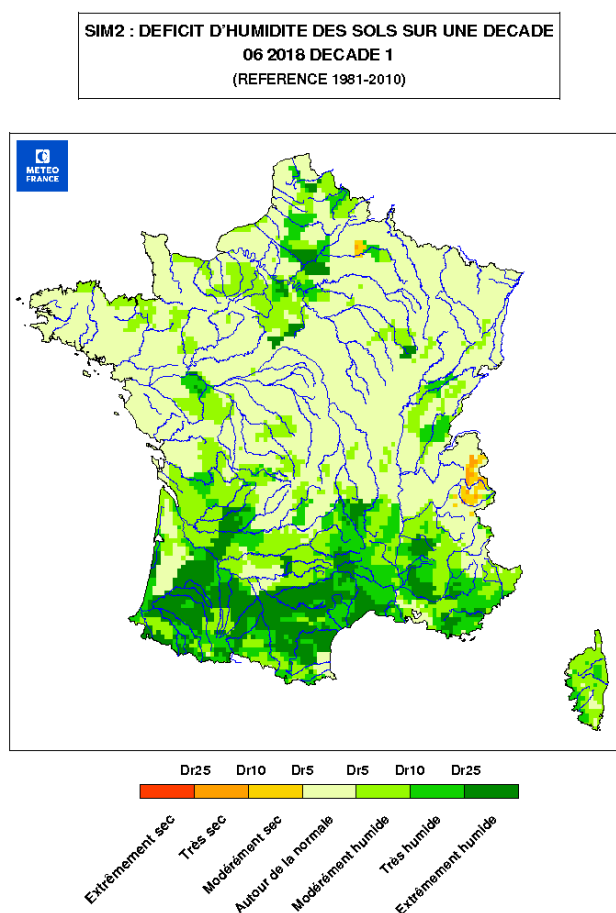


Figure 1 : Carte d'indice d'humidité des sols en France pour la première décade de juin 2018 (Source : Météo-France)

## 2.2. Données hydrologiques

Les débits observés sont disponibles au pas de temps journalier sur la Banque Hydro pour plusieurs stations du bassin de la Meuse (Leleu et al., 2014 ; Delaigue et al., 2020). Dans cette étude, une sélection de huit stations a été faite sur l'ensemble du bassin (Fig. 2) en se basant sur les stations choisies dans le projet CHIMERE21 et en fonction des données disponibles de vigilance sécheresse fournies par la DREAL Grand-Est (voir Tab. 2). Pour ces huit stations, les données de débit sont disponibles à partir des années 50, 60 ou 70, sauf pour la Meuse à Sedan (à partir de 1993).

En se basant sur les chroniques de débit observé, un indicateur hydrologique a été calculé à chaque pas de temps : le VCN3. Cet indicateur est couramment utilisé par les acteurs de l'eau en France pour caractériser les étiages. Il est calculé pour chaque station de la façon suivante : tout d'abord, une moyenne glissante sur trois jours des débits journaliers est appliquée sur l'ensemble de la chronique observée. Ensuite, la valeur minimale sur 10 jours est sélectionnée et attribuée à chaque pas de temps décadaire.

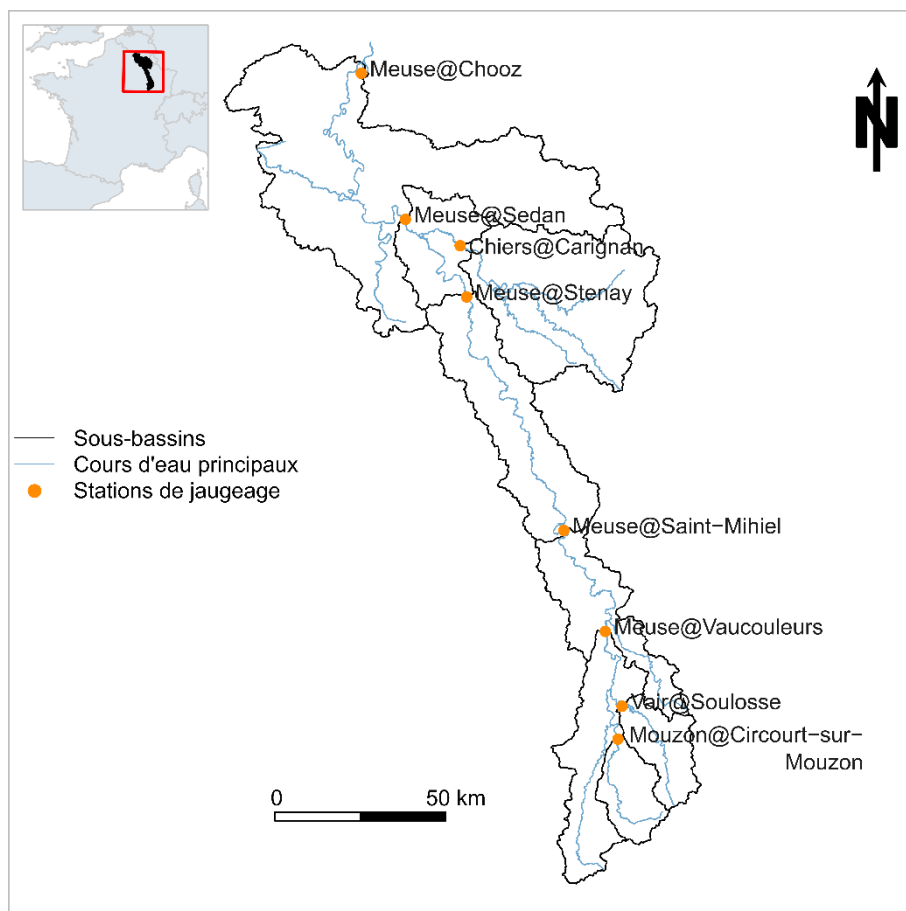


Figure 2 : Stations de la Meuse française utilisées dans cette étude

La DREAL Grand-Est fournit des valeurs de seuil de débit permettant de qualifier l'étiage selon trois niveaux de vigilance correspondant à des seuils réglementaires de débit (voir Tab. 2) associés au code couleur suivant :

- Jaune : les difficultés apparaissent sur certains secteurs, l'étiage étant marqué
- Orange : les difficultés se généralisent, l'étiage étant sévère
- Rouge : des difficultés majeures, l'étiage étant extrêmement sévère.

**Tableau 2 :** Valeur des seuils de vigilance définis par la DREAL Grand-Est pour les stations de la Meuse

Code	Nom	Seuil jaune (m <sup>3</sup> /s)	Seuil orange (m <sup>3</sup> /s)	Seuil rouge (m <sup>3</sup> /s)
B1092010	Le Mouzon à Villars	0.15	0.09	0.02
B1282010	Le Vair à Soulosse	0.50	0.36	0.21
B1340010	La Meuse à Chalaïnes	1.95	1.38	0.80
B2220010	La Meuse à Saint-Mihiel	3.20	2.20	1.20
B3150020	La Meuse à Stenay	8.66	6.40	4.13
B4631010	La Chiers à Carignan	8.60	7.10	5.60
B5020010	La Meuse à Sedan	22.60	18.25	13.90
B7200000	La Meuse à Chooz	30.50	22.25	14.00

### 2.3. Prise en compte spatiale des cartes d'indicateurs hydro-climatiques

Cette étude cherche à établir une relation statistique entre les données hydro-climatiques disponibles à la DREAL Grand-Est et les dépassements de seuils de vigilance : en d'autres termes, entre les cartes de SSWI et les observations de VCN3. Pour ce faire, cette analyse se place à l'échelle des sous-bassins versants de la Meuse. Comme les indices de SSWI spatialisés sont disponibles sur une grille et ne peuvent pas être moyennés à l'échelle des sous-bassins (cela n'a aucun sens physique), ces données sont agrégées à chaque pas de temps en termes de catégories de valeur. En se basant sur les catégories définies par Météo-France (Tab. 1), le pourcentage de chaque sous-bassin qui présente un indice appartenant à chaque catégorie est calculé à chaque pas de temps. On obtient ainsi des chroniques de pourcentage de sous-bassin modérément sec, très sec, et extrêmement sec (voir e.g. Fig. 3 pour l'année 1976). De plus, le dépassement de seuil est défini lorsque le VCN3 passe sous la valeur du seuil d'alerte correspondant. Ainsi pour le seuil de vigilance jaune, une chronique binaire (= 1 si le seuil est dépassé, = 0 sinon) est analysée pour chaque sous-bassin. Ce sont ces chroniques binaires que cette étude vise à modéliser.

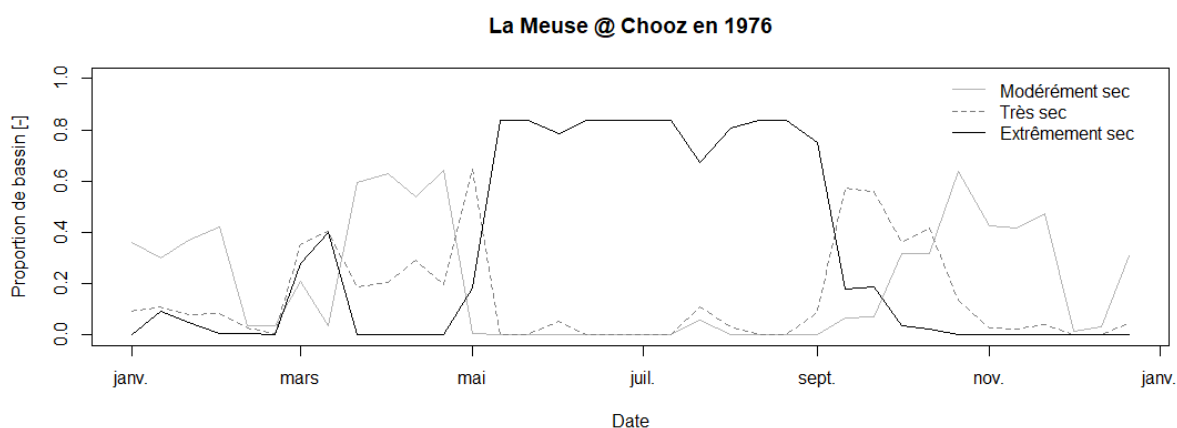


Figure 3 : Evolution du SSWI décadaire de la Meuse à Chooz en 1976 : proportion de bassin modérément, très et extrêmement sec

### 3. Méthodologie

#### 3.1. Modèle de régression logistique

Le modèle de régression logistique (glm) permet d'établir une relation binomiale entre une chronique binaire  $Y$  et un ensemble d'observations continues ou binaires  $X = \{x_1, x_2, \dots, x_j\}$ . On cherche ici à modéliser  $p(Y=1|X)$ , c'est-à-dire la probabilité que la variable  $Y$  ait la valeur 1 en connaissant l'ensemble des observations  $X$ . Le modèle de régression logistique s'écrit de la façon suivante :

$$\ln \frac{p(Y=1|X)}{1-p(Y=1|X)} = b_0 + b_1x_1 + \dots + b_jx_j \quad (1)$$

Et :

$$p(Y = 1|X) = \frac{e^{b_0+b_1x_1+\dots+b_jx_j}}{1+e^{b_0+b_1x_1+\dots+b_jx_j}} \quad (2)$$

Il s'agit donc d'estimer les coefficients  $\{b_0, b_1, \dots, b_j\}$  lors du calage de ce modèle de régression.

Dans cette étude, pour chaque sous-bassin de la Meuse, la variable à modéliser  $Y$  est la probabilité de dépassement du seuil de vigilance jaune (dans un premier temps, puis les seuils plus sévères dans un deuxième temps) à la station et les observations  $X$  sont les indices de SSWI spatialisés regroupés par catégorie (modérément sec, très sec et extrêmement sec). Comme nous nous plaçons en contexte de **prévision** des étiages et du risque sécheresse, la valeur binaire d'évènement de vigilance jaune du pas de temps de prévision ( $t+L$ , avec  $L$  l'échéance de prévision) est confrontée aux observations à un pas de temps donné  $t$  : on cherche à prévoir la probabilité de passage en vigilance jaune dans le futur en se basant sur des observations disponibles jusqu'à l'instant où la prévision est faite. La donnée de dépassement du seuil de vigilance au pas de temps observé est alors aussi intégrée dans les données d'entrée. En d'autres termes, pour prévoir la probabilité de dépassement d'un seuil de vigilance au pas de temps  $t+L$ , on peut aussi utiliser notre connaissance concernant ce seuil de vigilance au pas de temps  $t$ . Ainsi, deux jeux de données d'entrée  $X$  ont été testés (voir section 3.2) :

- un premier n'utilisant que les données de catégories de SSWI,
- un deuxième utilisant en plus les chroniques observées de dépassement de seuil de vigilance.

#### 3.2. Tests appliqués

Nous détaillons dans cette partie les différents tests mis en œuvre afin d'établir la validité d'un modèle de régression linéaire selon les configurations possibles. Ainsi, l'impact de la période de calage, le choix des données d'entrée (observations), le seuil de probabilité et la restriction du calage à différentes périodes estivales seront investigués. Cette section a donc pour objet de définir dans quelles conditions seront effectués les tests permettant d'investiguer ces aspects. Par ailleurs, une extension aux vigilances orange et rouge de la configuration optimale a été menée, avant d'évaluer la transférabilité spatiale des relations obtenues.

Les données sont analysées d'août 1970 à juillet 2016, qui est la période de calage des modèles hydrologiques du projet CHIMERE21 (voir Thirel et al., 2020). Cette période est divisée en deux parts égales (en août 1993) afin d'appliquer un **protocole de calage-validation** (correspondant au Split Sample Test, SST, décrit par Klemeš, 1986) : les modèles sont d'abord calés sur la période 1970-1993 et validés sur 1993-2016, puis le calage est fait sur 1993-2016 et la validation sur 1970-1993. Cette approche permet d'évaluer la transférabilité des modèles dans le temps (i.e. leur robustesse), particulièrement en contexte non-stationnaire (e.g. changement climatique). De plus, afin d'obtenir un modèle générique sur l'ensemble du bassin de la Meuse et de l'appliquer à des sous-bassins non-



jaugés, les données de toutes les stations sélectionnées ont été utilisées et un modèle pour tous les sous-bassins de la Meuse est obtenu par test.

Afin de choisir le modèle le plus robuste et performant, plusieurs tests ont été menés : tout d’abord, **différentes données d’entrée** (observations X) ont été utilisées pour générer plusieurs modèles de régression logistique. Deux configurations ont été testées (Tab. 3) : la première utilisant uniquement les données de SSWI et la seconde utilisant aussi les observations de passage du seuil de vigilance. L’utilisation de cette dernière information observée au pas de temps précédant l’échéance de simulation permet d’utiliser une chronique binaire (=1 si VCN3 < seuil jaune, =0 sinon) qui fournit une information hydrologique, et ainsi potentiellement de mieux contraindre les modèles statistiques. Notez que l’utilisation des données de SSWI implique trois chroniques de données d’entrée : les pourcentages du sous-bassin modérément, très ou extrêmement sec. Ainsi pour le modèle 1, trois chroniques sont utilisées en données d’entrée et pour le modèle 2, quatre chroniques sont utilisées (Tab. 3).

**Tableau 3** : Modèles de régression logistique testés en fonction des données d’entrée utilisées au pas de temps décadaire

Modèle	SSWI	Seuil jaune
1	✓	
2	✓	✓

Ensuite, comme le modèle glm produit une chronique continue de probabilité d’évènement de vigilance jaune (valeurs entre 0 et 1), il faut transformer ces valeurs pour obtenir une chronique binaire. Pour cela, un **seuil de probabilité** doit être choisi pour définir le passage en vigilance jaune. Les seuils de probabilité de 0,2 (i.e. 20 %) à 0,9 (i.e. 90 %) ont été testés ici. Par exemple, pour un seuil de probabilité de 0,2, toute probabilité d’évènement de vigilance jaune simulée par le modèle calé qui est supérieure ou égale à 0,2 se voit attribuer la valeur 1, et les autres évènements (de probabilité < 0,2) se voient attribuer la valeur 0. On obtient ainsi des chroniques binaires d’évènements de vigilance simulés.

De plus, dans les chroniques binaires observées de dépassement de seuil jaune, la proportion de valeurs à 0 (quand le débit est au-dessus du seuil) est très grande : 87 % des chroniques observées d’alerte vigilance jaune présentent des valeurs à 0. Ainsi, le modèle glm peut avoir des difficultés à prédire des évènements (valeurs à 1) pour des chroniques contenant beaucoup de 0. Afin de réduire ce biais, **ces modèles ont aussi été calés sur la période estivale**, ce qui réduit la proportion de valeurs à 0 dans les chroniques binaires observées. En effet, la majorité des étiages sévères ont lieu en été (voir Fig. 4). Par exemple, en 1976, la Meuse à Chooz est passée en alerte jaune, puis orange au mois de mai et en alerte rouge en juillet, pour repasser au-dessus du seuil de vigilance jaune en novembre (voir Fig. 5). De même en 2003 (Fig. 6), la Meuse à Chooz était en vigilance jaune et orange entre les mois de juillet et novembre. Ainsi, deux périodes estivales de calage ont été testées, comprenant le plus d’évènements de vigilance jaune, orange et rouge sur la période observée : de mai à octobre pour une période de 6 mois et de juillet à septembre pour une période de 3 mois. La performance de ces modèles a été calculée sur les périodes de 12, 6, ou 3 mois en calage et sur toute l’année (12 mois) en validation. On force ainsi les modèles sur des périodes où il y a le plus d’observations. On notera par ailleurs sur la Fig. 4 la présence de plus nombreux dépassements de seuils jaune et orange sur la période récente, au contraire de la période la plus ancienne pour le seuil rouge.

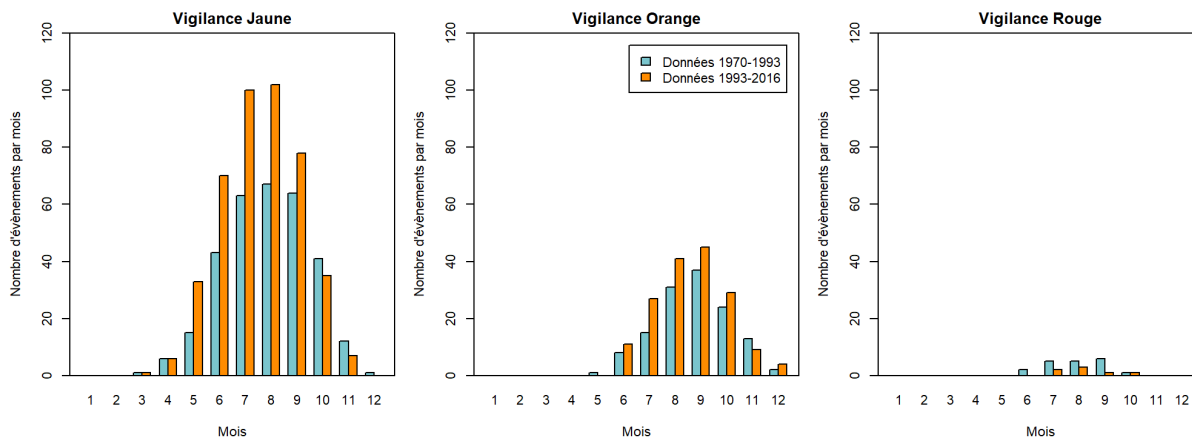


Figure 4 : Nombre d'évènements de passage du VCN3 sous le seuil de vigilance jaune (à gauche), orange (au milieu) et rouge (à droite) par mois sur les périodes 1970-1993 et 1993-2016 pour toutes stations confondues

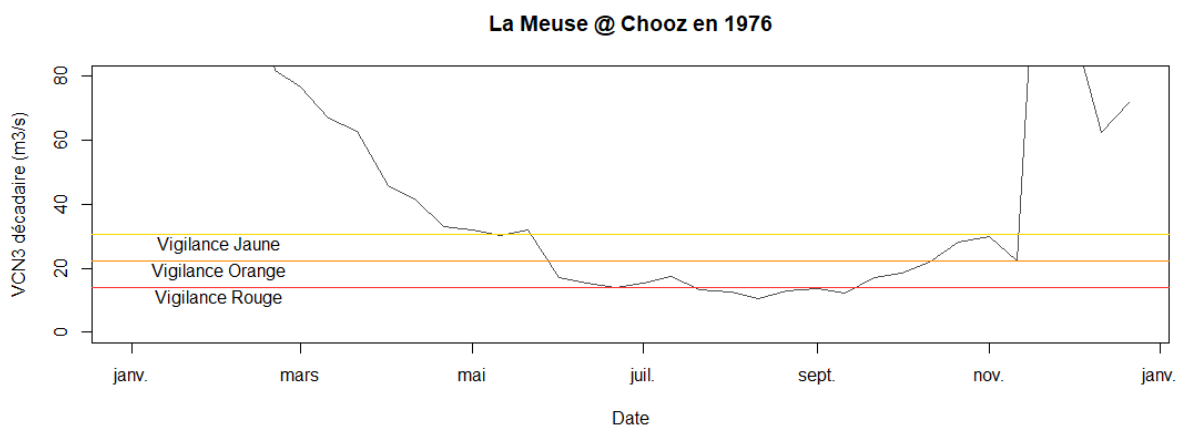


Figure 5 : Evolution du VCN3 décadaire de la Meuse à Chooz en 1976 avec les seuils de vigilance jaune, orange et rouge

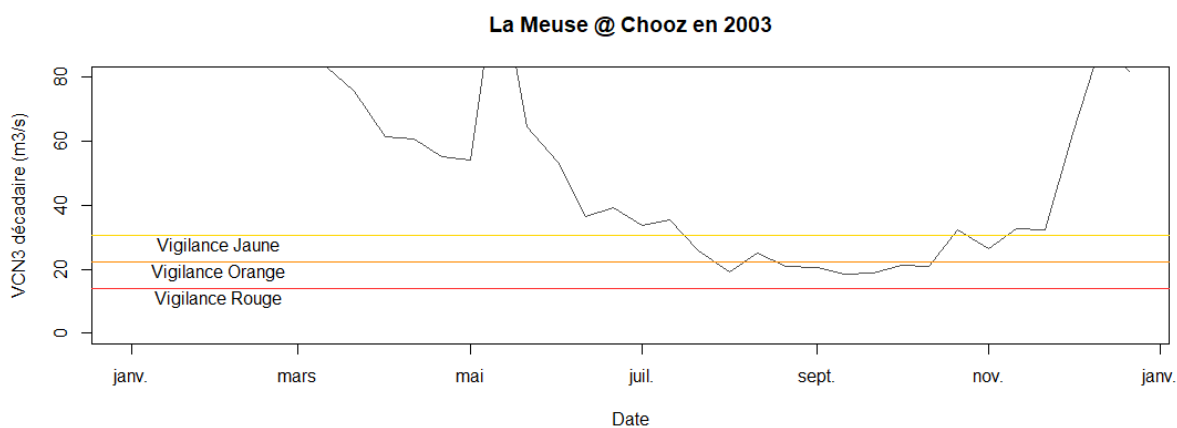


Figure 6 : Evolution du VCN3 décadaire de la Meuse à Chooz en 2003 avec les seuils de vigilance jaune, orange et rouge

Ces trois premières étapes de tests permettent de **définir le jeu de données d'entrée, le seuil de probabilité et la période de calage optimum** pour prédire au pas de temps suivant un évènement de vigilance jaune. La dernière étape utilise ce résultat pour étendre la prévision à des pas de temps plus

lointains (jusqu'à 9 décades), et ainsi définir jusqu'à quelle échéance le modèle obtenu pour le seuil de vigilance jaune est performant.

Puis, afin de **prévoir le passage des seuils de vigilance orange et rouge**, la configuration optimale obtenue pour le seuil de vigilance jaune a été utilisée : la période de calage (12 mois, 6 mois ou 3 mois d'été) et le jeu de données d'entrée (combinaison de SSWI avec ou sans observation du seuil de vigilance avant prévision) qui fournissent les meilleures performances pour la vigilance jaune ont été choisis (nous verrons plus tard quelle configuration a été jugée optimale), et des tests concernant le seuil de probabilité ont été menés afin de choisir le modèle optimum pour chacun des deux autres niveaux de vigilance. La performance de chaque modèle optimum a aussi été évaluée sur des échéances futures (jusqu'à 9 décades).

Enfin, afin d'**évaluer la robustesse spatiale des modèles optimum de prévision de vigilance jaune, orange et rouge**, le jeu de sous-bassins versants a été divisé en deux et un test de validation croisée sur ces deux jeux de sous-bassins indépendants (Proxy Basin Test, PBT, voir Klemeš, 1986) a été mené. Le premier échantillon de sous-bassins regroupe les quatre stations amont (le Mouzon à Villars, le Vair à Soulosse, la Meuse à Chalaines et la Meuse à Saint-Mihiel, voir Fig. 2), et le deuxième échantillon rassemble les quatre stations aval (la Meuse à Stenay, la Chiers à Carignan, la Meuse à Sedan et la Meuse à Chooz, voir Fig. 2). Ce choix regroupe les stations amont d'un côté et les stations aval de l'autre, ce qui représente la répartition la plus contrastée en termes de caractéristiques physiques du bassin. Ceci permet de tester la transférabilité spatiale du modèle dans le cas le plus contrasté. Ainsi pour chaque seuil de vigilance sécheresse, le modèle optimum a été calé sur le premier échantillon de sous-bassins puis validé sur le deuxième échantillon, et vice-versa, et les performances de ces modèles ont été évaluées en calage et validation pour une échéance d'une décade.

### 3.3. Critères de performance

La performance des modèles obtenus est quantifiée à l'aide d'une matrice de contingence (Tab. 4), et plusieurs critères sont calculés :

- Le taux de bonnes prévisions totales (BPT)
- Le taux de bonnes prévisions d'évènements (BPE)
- Le taux de fausses alertes (FA)

Avec (voir terminologie dans le Tab. 4) :

$$BPT = \frac{A+D}{A+B+C+D} \quad (3)$$

$$BPE = \frac{D}{B+D} \quad (4)$$

$$FA = \frac{C}{C+D} \quad (5)$$

**Tableau 4** : Matrice de contingence du modèle de régression logistique. A : nombre d'évènements non simulés et non observés ; B : nombre d'évènements non simulés et observés ; C : nombre d'évènements simulés et non observés ; D : nombre d'évènements simulés et observés.

		Observé	
		Non	Oui
Simulé	Non	A	B
	Oui	C	D

La performance des modèles testés est évaluée en se basant sur les critères suivants :

$$KSS = \frac{(AD)-(BC)}{(A+C)(B+D)} \quad (6)$$

$$PS = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2 \quad (7)$$

$$EFF_L = 1 - \frac{\sum_{i=1}^N (o_{i+L} - p_{i+L})^2}{\sum_{i=1}^N (o_{i+L} - o_i)^2} \quad (8)$$

Avec N le nombre de pas de temps,  $p_i$  la variable (continue pour PS et binaire pour  $EFF_L$ ) modélisée au pas de temps  $i$ ,  $o_i$  la variable binaire observée au pas de temps  $i$ , et L l'échéance de prévision.

Le **KSS** (Hanssen-Kuiper skill score, voir Hanssen and Kuipers, 1965) est optimal pour une valeur à 1, soit quand  $B = C = 0$ . Il tient compte à la fois de la bonne représentation des évènements observés, mais aussi des évènements manqués et des fausses alertes. En revanche, si A est grand, le KSS tend à minimiser le rôle des fausses alertes. Le choix de ce score s'appuie néanmoins sur l'expérience des prévisionnistes de la difficulté à obtenir un taux de fausses alertes proche de zéro. C'est pourquoi une emphase limitée sur les fausses alertes semble acceptable dans cette étude.

**PS**, le score de probabilité, ou score de Brier (Brier, 1950 ; Eq. 7), est utilisé couramment pour évaluer la performance de modèles probabilistes. Cet indicateur compare la prévision probabiliste modélisée d'un évènement (variable continue, dont la valeur est comprise entre 0 et 1) avec l'observation binaire de cet évènement (= 1 si l'évènement est observé, = 0 sinon). La prévision est parfaite quand  $PS = 0$ .

Le score de persistance, **EFF<sub>L</sub>**, compare la performance d'un modèle de prévision avec celle du modèle de persistance à l'échéance L (Kitanidis et Bras, 1980 ; Eq. 8). Quand  $EFF_L = 0$ , le modèle de prévision est aussi performant que le modèle de persistance, et quand  $EFF_L > 0$  ( $< 0$ ), le modèle de prévision est plus (moins) performant que le modèle de persistance à l'échéance L. Si le score PS est utilisé sur les chroniques brutes de sorties du modèle glm (i.e. sur des chroniques continues de probabilité), les scores KSS et  $EFF_L$  sont eux appliqués aux chroniques binaires de prévision, i.e. transformées après application du seuil de probabilité.

## 4. Résultats pour le seuil de vigilance jaune

### 4.1. Choix du modèle optimum

La Fig. 7 montre les observations binaires et les simulations continues brutes de prévision de vigilance jaune à l'échéance d'une décade en sortie du modèle glm pour les évènements de 1976 et 2003, obtenus avec les modèles 1 et 2 (voir Tab. 3 pour la description des modèles) en calage sur l'année complète à l'exutoire (la Meuse à Chooz). Les scores de Brier (PS) obtenus pour chacun des modèles en calage et en validation en SST sont présentés dans le Tab. 5. Les deux modèles sont robustes puisque leurs performances ne diffèrent que peu entre les périodes de calage et de validation. Les sorties brutes (c'est-à-dire avant utilisation d'un seuil) montrent de meilleures performances pour le modèle 2, qui utilise les observations de vigilance jaune avant la prévision, et ce en calage et en validation, avec un PS proche de 0. Ceci est illustré sur la Fig. 7 où les sorties brutes représentent mieux les observations pour les deux évènements de sécheresse. On note cependant un décalage temporel d'un pas de temps entre les observations et les simulations, le modèle étant en retard. Ces différents résultats semblent indiquer que l'essentiel de l'information vient du seuil de vigilance observé. Nous évaluerons dans quelle mesure cela est vrai en comparant par la suite la performance du modèle de régression linéaire à un modèle de persistance.

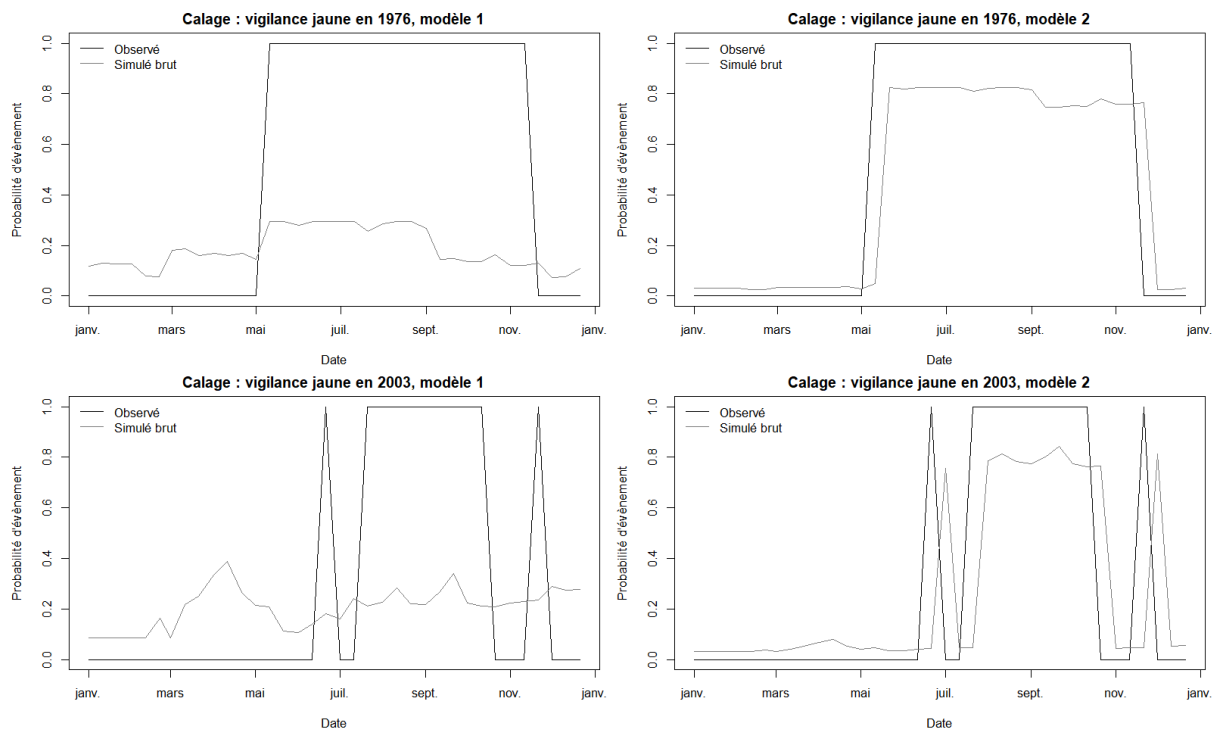


Figure 7 : Observations binaires (en noir) et simulations brutes continues (en gris) des évènements de vigilance jaune de 1976 (en haut) et 2003 (en bas) pour les modèles 1 (à gauche) et 2 (à droite) en calage sur l'année complète pour la Meuse à Chooz

**Tableau 5** : Scores de Brier (PS) en calage et validation pour les modèles 1 et 2 pour le seuil de vigilance jaune pour toutes les stations confondues

Période	Modèle 1		Modèle 2	
	1970-1993	1993-2016	1970-1993	1993-2016
<b>Calage</b>	0.09	0.10	0.04	0.05
<b>Validation</b>	0.09	0.11	0.04	0.05

La Fig. 8 montre les performances des modèles 1 et 2 en termes de KSS pour des seuils de probabilité testés de 0,2 à 0,9 en calage et validation sur l'année complète. Ces résultats illustrent la performance des modèles testés une fois que les sorties continues brutes du modèle de glm (voir Fig. 6) sont transformées en chroniques binaires via l'utilisation d'un seuil. Le modèle 2 est plus performant que le modèle 1, avec les meilleures performances obtenues pour les seuils de probabilité entre 0,2 et 0,7 (KSS compris entre 0,7 et 0,8), et ce en calage et en validation. Notez que les zones blanches correspondent à des seuils de probabilité pour lesquels toutes les valeurs brutes ont été converties en 0, et pour lesquels les performances ne sont donc pas calculées.

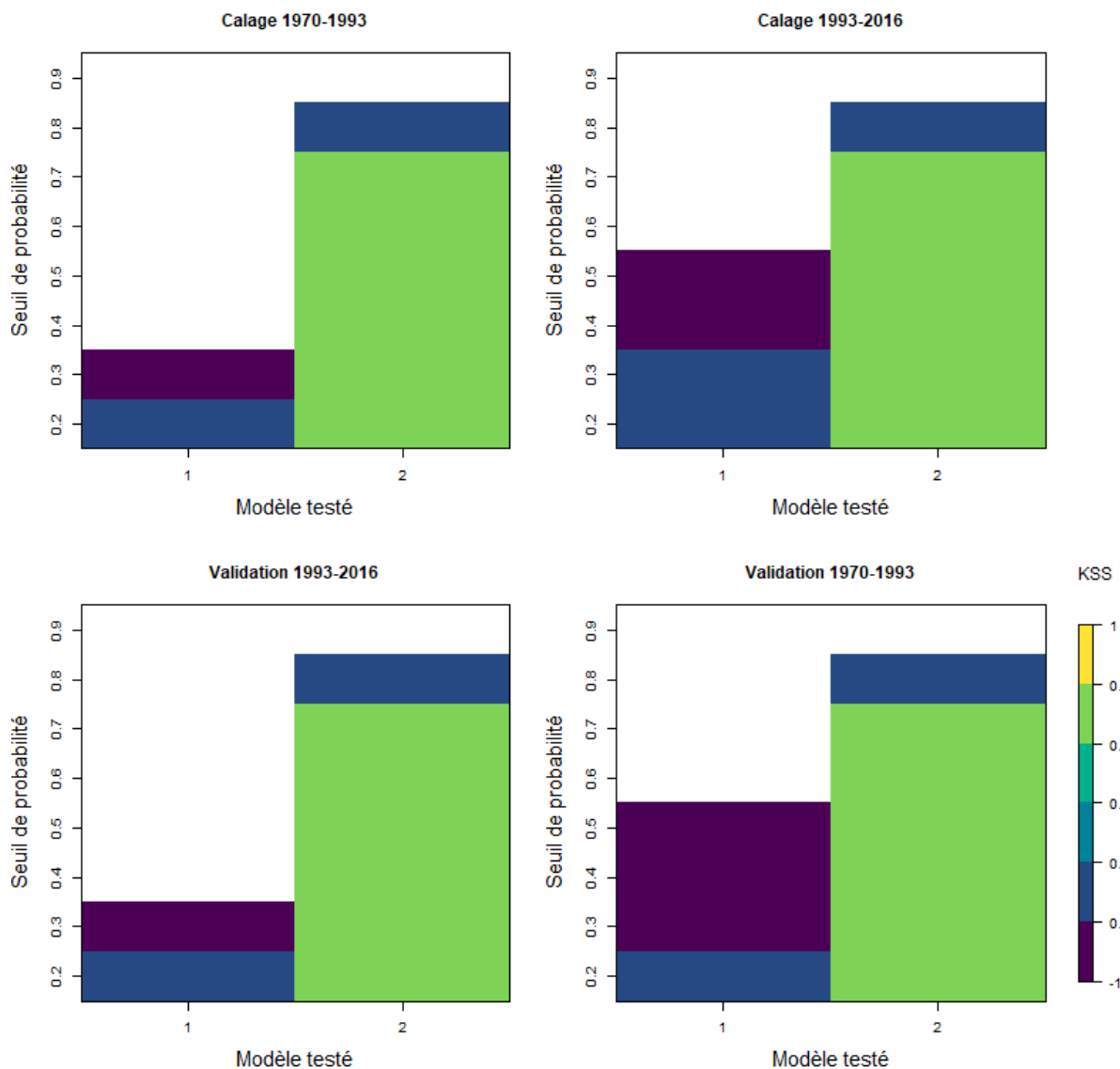


Figure 8 : Performance des modèles 1 et 2 en calage et validation croisée (SST) sur l'année en termes de score de KSS pour des seuils de probabilité de 0,2 à 0,9 pour toutes les stations confondues

Comme le modèle 2 est plus performant que le modèle 1, le calage a aussi été testé sur des périodes de trois et six mois en été. Ainsi la Fig. 9 montre les résultats obtenus en calage et en validation SST en termes de KSS pour le modèle 2, avec un calage sur l'année complète, 6 mois d'été, et 3 mois d'été. Comme précédemment, les résultats sont illustrés pour les seuils de probabilité de 0,2 à 0,9. En termes de KSS, les calages sur l'année complète et sur six mois d'été montrent des scores équivalents en calage et en validation pour les seuils de probabilité de 0,2 à 0,7, avec des KSS compris entre 0,70 et 0,71. Le calage sur trois mois d'été montre des scores de KSS moins bons ( $<0,70$ ). Afin de départager les deux périodes de calage, les scores de BPT, BPE et FA ont aussi été analysés (non montré ici). Il en ressort des performances plus satisfaisantes pour le calage sur la période de six mois d'été, notamment avec un BPE entre 0,75 et 0,76 (contre 0,74 pour le calage sur l'année complète), et un FA entre 0,21 et 0,22 (contre 0,25 sur l'année complète). **On sélectionne donc pour la suite le modèle utilisant comme données d'entrée les chroniques continues de SSWI et les chroniques binaires de passage de seuil de vigilance jaune, calé sur une période de six mois avec un seuil de probabilité de 0,5.** Les performances de ce modèle en calage et en validation en termes de BPT, BPE et FA sont présentées dans le Tab. 6. Ce modèle est robuste puisque les performances sont équivalentes en

calage et en validation, particulièrement en termes de BPT et BPE. On note toutefois une dégradation du FA en validation, de 0,02 pour les deux périodes.

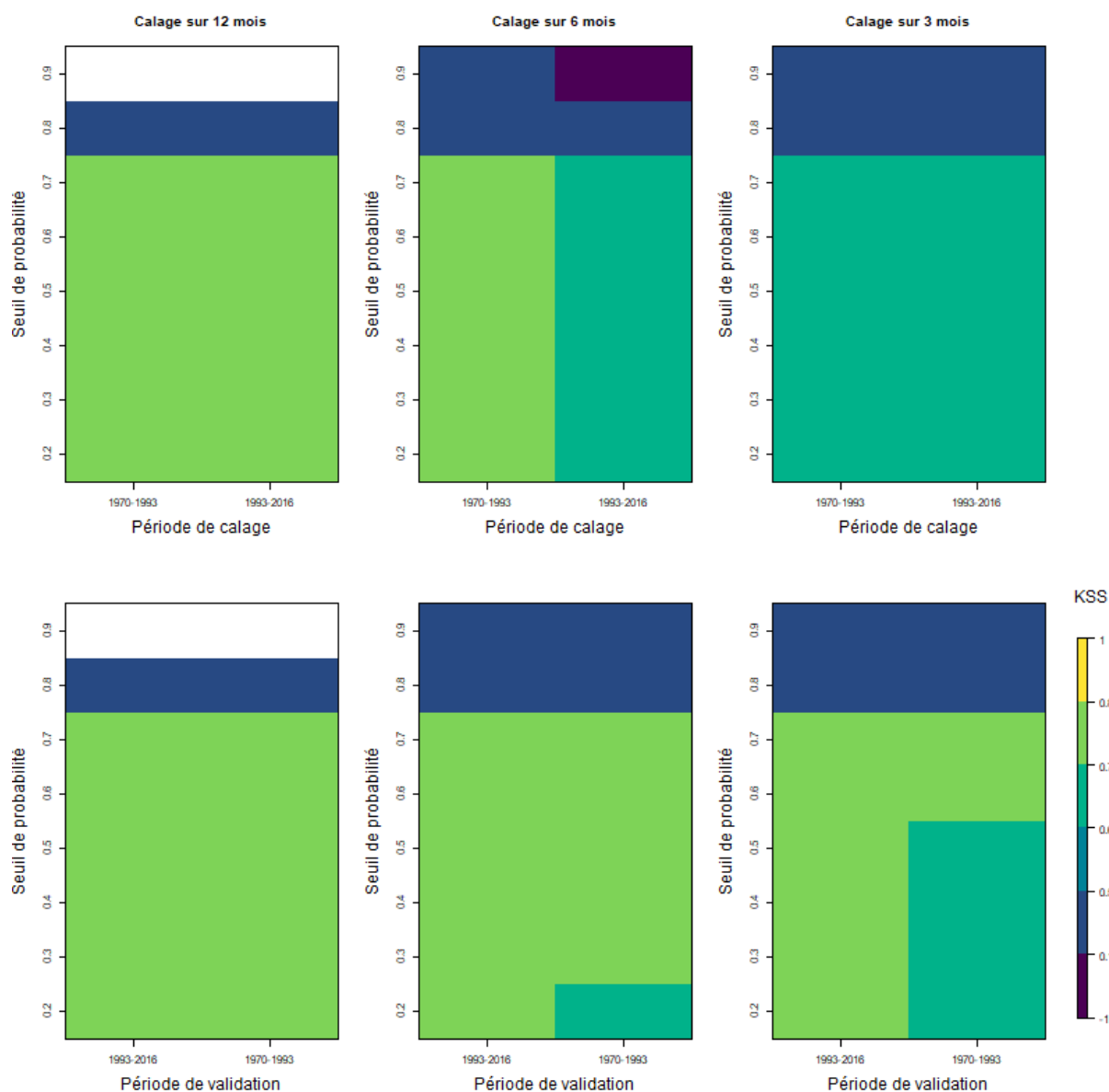


Figure 9 : Performance du modèle 2 en termes de KSS en calage (en haut) et validation (en bas) avec un calage sur l'année complète (à gauche), six mois d'été (au milieu) et trois mois d'été (à droite). Les scores sont calculés sur la période de calage et sur toute l'année en validation, pour toutes les stations confondues

**Table 6** : Performances du modèle optimum pour la vigilance jaune en calage et validation en termes de BPT (taux de bonnes prévisions totales), BPE (taux de bonnes prévisions d'évènements) et FA (taux de fausses alertes) pour toutes les stations confondues

	Calage		Validation	
	1970-1993	1993-2016	1993-2016	1970-1993
Taux de bonnes prévisions totales	0.91	0.88	0.94	0.95
Taux de bonnes prévisions d'évènements	0.75	0.76	0.76	0.74
Taux de fausses alertes	0.21	0.23	0.24	0.25

## Résumé :

Dans cette partie, nous avons montré que pour la prévision de la vigilance jaune à l'échéance d'une décade, la configuration la plus performante était observée pour le modèle logistique 2, qui associe SSWI et état observé de vigilance jaune au pas de temps précédent, et un calage sur les 6 mois estivaux. Nous conserverons cette configuration pour la suite.

### 4.2. Performance du modèle optimum aux échéances futures

Les Figures 10 et 11 présentent la performance du modèle de vigilance jaune optimum en calage et en validation, respectivement, pour des échéances d'une à neuf décades (soit jusqu'à trois mois). Quelle que soit la période de calage (avant ou après 1993), les performances sont bonnes à très bonnes jusqu'à un mois de prévision (trois décades) avec des BPT > 0,7, des BPE > 0,6 et des FA < 0,4. A partir de la quatrième et la cinquième décade, la performance chute sur la période de calage 1993-2016 et les deux périodes, respectivement (BPE < 0,3). Le taux de fausses alertes reste stable (< 0,4 et 0,5 en calage et validation respectivement) jusqu'à l'échéance de 7 décades (après 2 mois), et augmente rapidement dans la fin du troisième mois (> 0,6 en calage et jusqu'à 1 en validation). **En résumé, le modèle de prévision de vigilance jaune est performant et robuste jusqu'à une échéance d'un mois, que ce soit en calage ou sur une période indépendante de validation.**

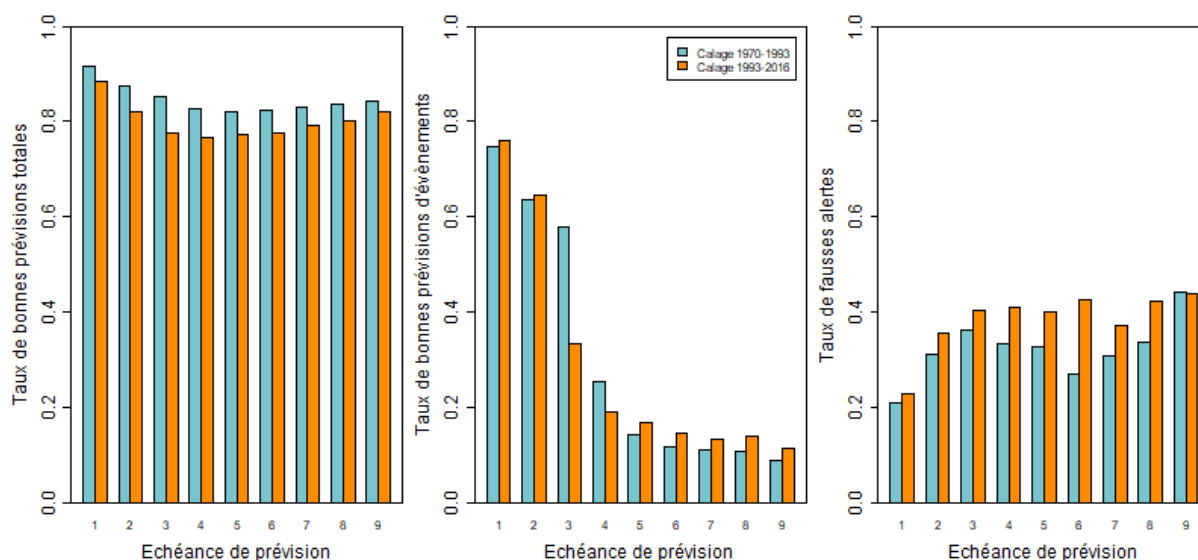


Figure 10 : Performance du modèle optimum sur la période de calage de six mois d'été : taux de bonnes prévisions totales (à gauche, valeur optimale égale à 1), taux de bonnes prévisions d'évènements (au milieu, valeur optimale égale à 1) et taux de fausses alertes (à droite, valeur optimale égale à 0) pour les échéances de prévision de 1 à 9 décades pour toutes les stations confondues



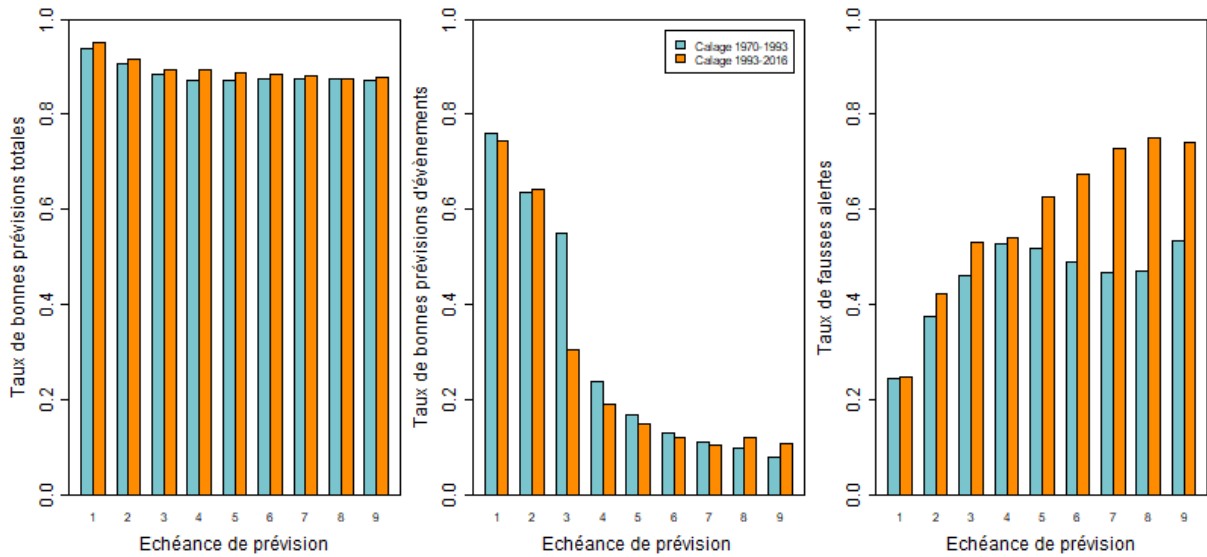


Figure 11 : Performance du modèle optimum sur la période de validation sur l'année complète : taux de bonnes prévisions totales (à gauche, valeur optimale égale à 1), taux de bonnes prévisions d'événements (au milieu, valeur optimale égale à 1) et taux de fausses alertes (à droite, valeur optimale égale à 0) pour les échéances de prévision de 1 à 9 décades pour toutes les stations confondues

Pour approfondir l'analyse de performance du modèle optimum de vigilance jaune, la Fig. 12 montre les scores de persistance ( $EFF_1$ ) en calage pour les échéances allant jusqu'à neuf décades. A l'échéance une décade, le modèle optimum performe exactement comme le modèle de persistance ( $EFF_1 = 0$ ), quelle que soit la période de calage. Pour les échéances de deux et trois décades, le modèle optimum se comporte aussi comme le modèle de persistance ou de manière très similaire, avec un score légèrement négatif en calage sur 1993-2016. A partir d'une échéance de quatre décades, le modèle optimum montre de meilleures performances que le modèle de persistance, alors que c'est l'échéance à partir de laquelle le modèle de persistance se dégrade plus rapidement que celle du modèle testé. A titre d'illustration, les données binaires de vigilance jaune observées et simulées avec le modèle optimum en validation pour les échéances une à trois décades sont montrées sur la Fig. 13 pour les événements de 1976 et 2003. On observe le décalage temporel entre les observations et les simulations pour chaque nouvelle échéance de prévision. Pour ces deux événements, le modèle optimum de vigilance jaune se comporte comme le modèle de persistance.

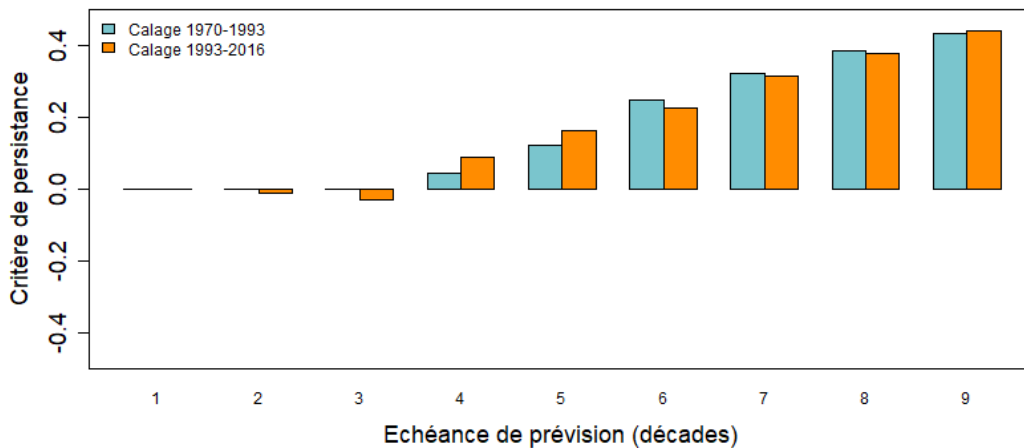


Figure 12 : Performance du modèle optimum en calage : scores de critère de persistance pour les échéances de prévision de 1 à 9 décades pour toutes les stations confondues

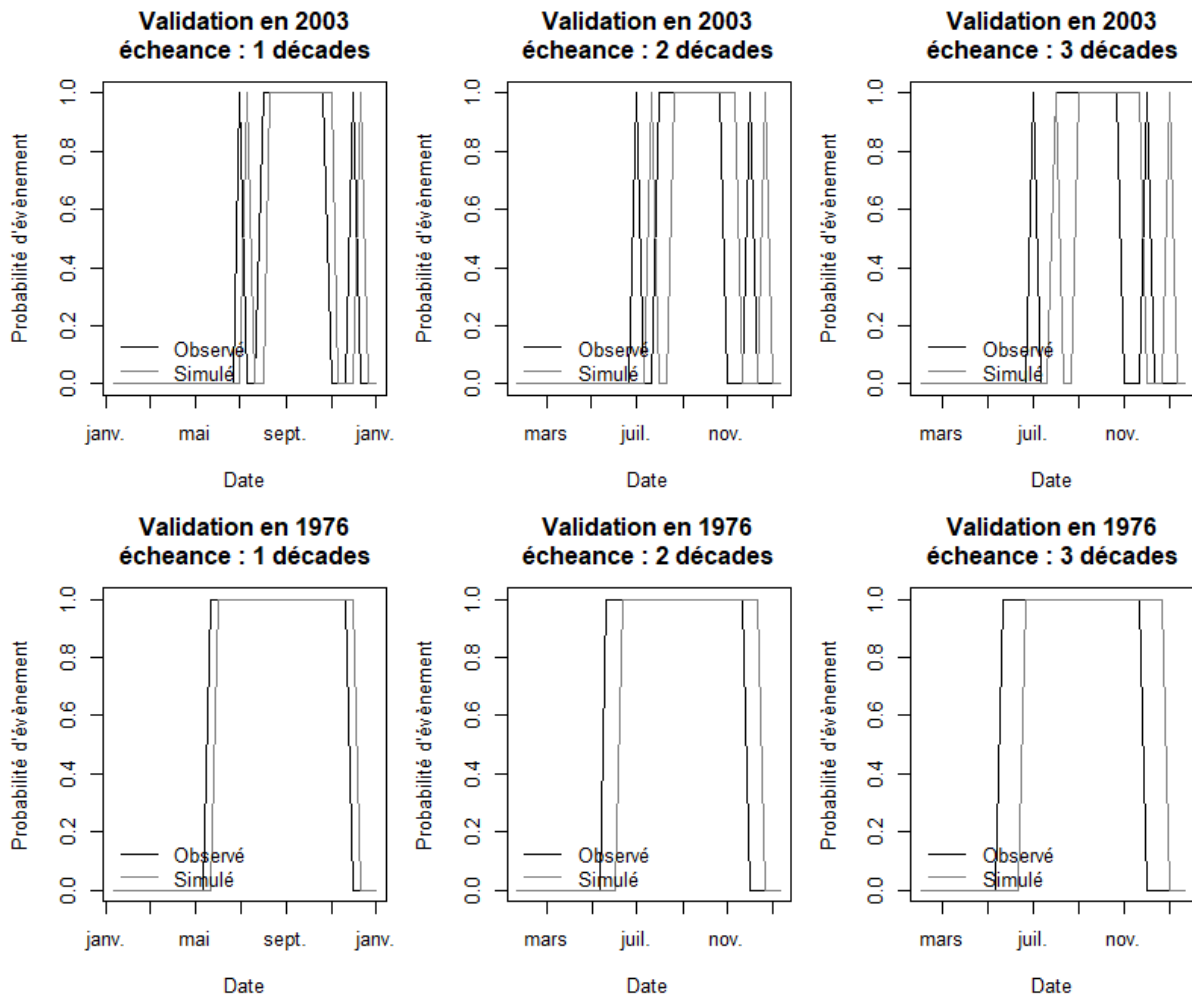


Figure 13 : Observation et simulations avec le modèle optimum des évènements de vigilance jaune en 1976 et 2003 pour les échéances de prévision de 1 à 3 décades pour la Meuse à Chooz

#### Résumé :

Dans cette partie, nous avons pu comparer notre modèle à un modèle de persistance, afin de déterminer dans quelle mesure le modèle de régression logistique apportait une meilleure prévision que la simple prise en compte de la dernière observation de vigilance disponible. Nous avons montré que si la performance était équivalente au modèle de persistance jusqu'à 3 décades, au-delà, la performance des deux modèles chute rapidement, celle du modèle testé se dégradant un peu moins vite que celle du modèle de persistance.

## 5. Résultats pour les seuils de vigilance orange et rouge

En se basant sur le modèle le plus performant pour la prévision du niveau de vigilance jaune, les modèles de prévision des vigilances orange et rouge ont été testés, en calant les données sur trois et six mois d'été et en utilisant une combinaison des chroniques continues de SSWI et des chroniques binaires de passage du seuil de vigilance (orange ou rouge) observé au pas de temps précédant la prévision. Pour ces deux niveaux de vigilance, de meilleures performances ont été obtenues en calant les modèles sur trois mois d'été. Les résultats suivants montrent les performances des modèles pour plusieurs seuils de probabilité testés, afin de choisir le modèle optimum pour chacun de ces niveaux de vigilance, et pour plusieurs échéances de prévision, afin d'évaluer la performance du modèle optimum choisi en prévision.

### 5.1. Prédiction du seuil de vigilance orange

La Figure 14 montre les performances du modèle de prédiction de vigilance orange en calage et en validation sur une période de trois mois d'été pour les seuils de probabilité de 0,2 à 0,9. L'indicateur de performance (KSS) est meilleur en calage et en validation, quelle que soit la période, pour des seuils de probabilité de 0,2 à 0,6. En regardant plus en détail la table de contingence, on peut aussi analyser les performances en termes de BPT, BPE et FA en fonction du seuil de probabilité testé (non montré ici). En termes de BPT, les seuils de probabilité de 0,2 à 0,6 donnent les mêmes performances et sont très hauts par rapport au seuil de vigilance jaune car il y a moins d'évènements de vigilance observés (et donc en proportion plus de zéros dans la chronique). En revanche en termes de BPE, le seuil de 0,6 est moins performant. Bien que le FA du seuil de 0,2 soit le meilleur sur les deux périodes de validation, le seuil de probabilité de 0,5, très similaire, sera privilégié car c'est un niveau de probabilité plus satisfaisant pour la transformation d'une chronique probabiliste continue en chronique binaire. **Ainsi, pour le seuil de vigilance orange, le modèle optimum est calé sur trois mois d'été, utilisant la combinaison des chroniques continues de SSWI et binaires de vigilance orange observée et un seuil de probabilité de 0,5.**

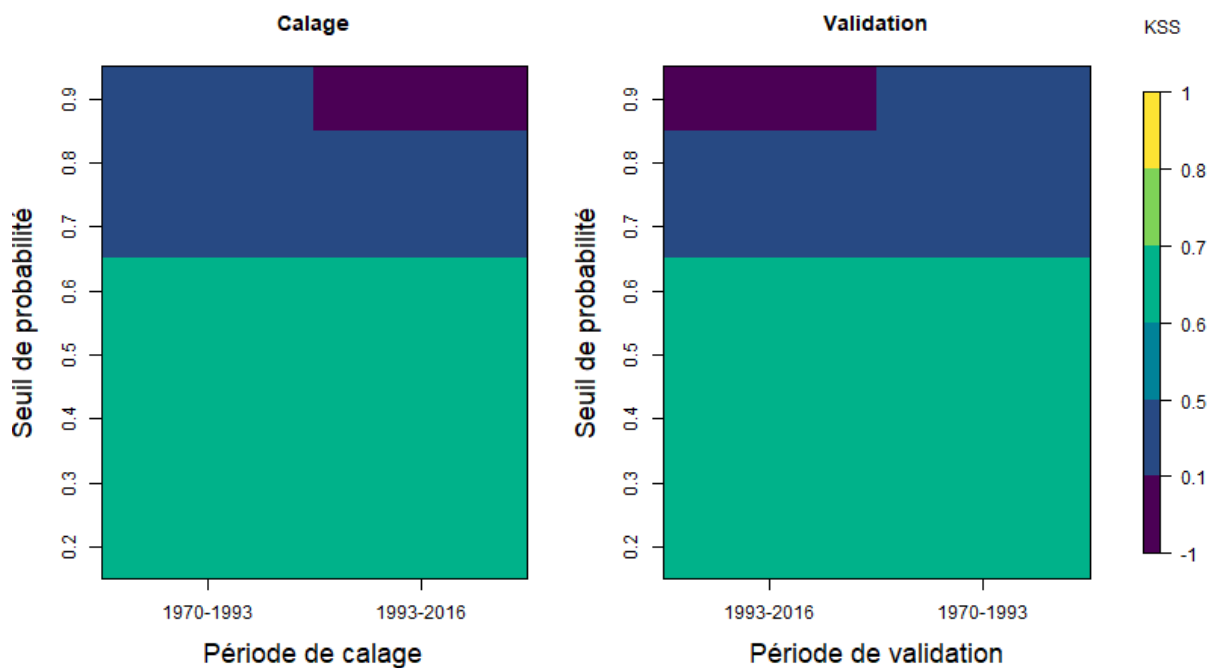


Figure 14 : Performance des modèles de vigilance orange pour les seuils de probabilité de 0,2 à 0,9 sur la période de calage (à gauche) et sur la période de validation (à droite) sur trois mois d'été pour toutes les stations confondues

La Fig. 15 montre les performances de ce modèle optimum sur des échéances de 1 à 9 décades. Bien que le BPT soit toujours très satisfaisant, dès la deuxième décade, le BPE chute sous 0,4 tandis que le FA monte au-dessus de 0,5 dès la 5<sup>e</sup> décade. Dès la 6<sup>e</sup> décade (en calage sur 1993-2016) et la 7<sup>e</sup> décade (sur les deux périodes de calage), le modèle n'est plus capable de prévoir d'évènement de vigilance orange. Enfin, la Fig. 16 illustre la performance du modèle de prédiction en termes de critère de persistance pour les échéances de prévision d'une à neuf décades. Le modèle optimum est identique au modèle de persistance pour la première échéance de prévision ( $Eff_1 = 0$ ), puis montre de meilleures performances jusqu'à 6 décades de prévision, quelle que soit la période de calage ( $Eff_1 > 0$ ).

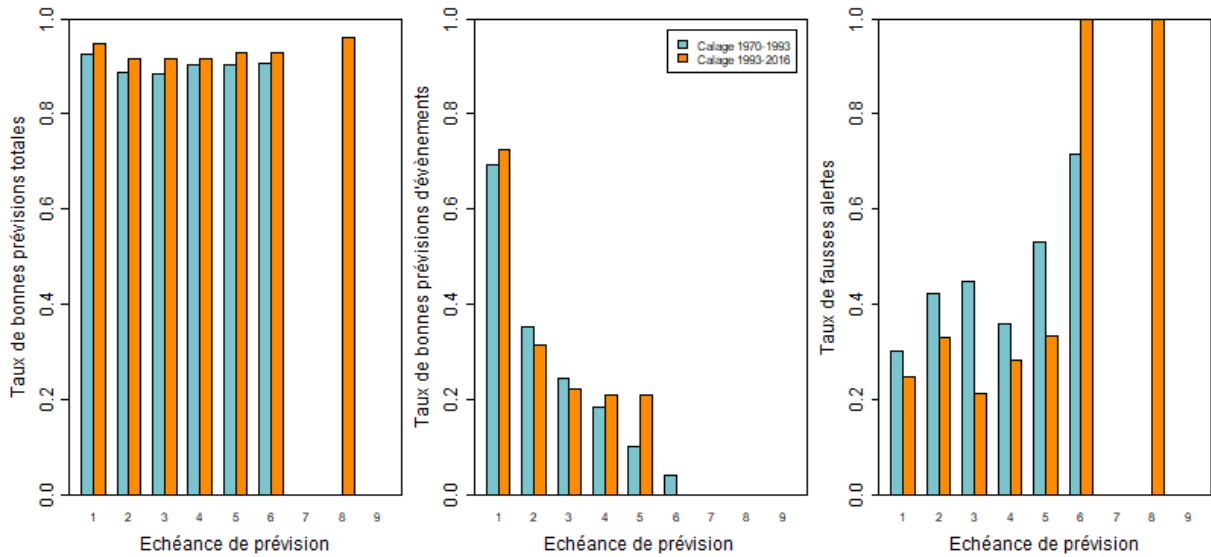


Figure 15 : Performance du modèle optimum de vigilance orange sur la période de validation pour les échéances de prévision de 1 à 9 décades : taux de bonnes prévisions (à gauche, valeur optimale égale à 1), taux de bonnes prévisions d'événements (au milieu, valeur optimale égale à 1) et taux de fausses alertes (à droite, valeur optimale égale à 0) pour toutes les stations confondues

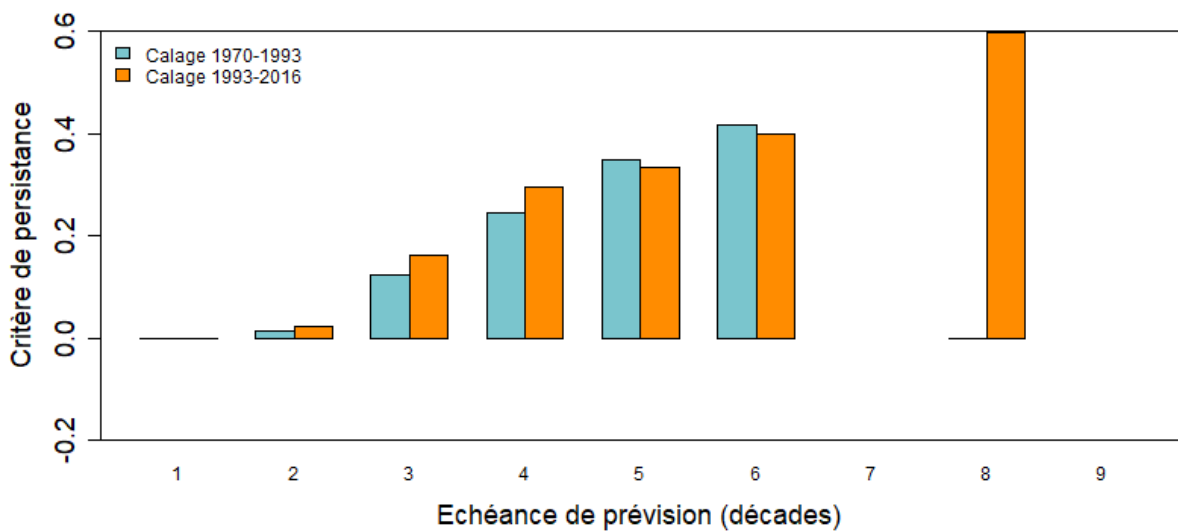


Figure 16 : Performance du modèle optimum de vigilance orange en calage : scores de critère de persistance pour les échéances de prévision de 1 à 9 décades pour toutes les stations confondues

**Résumé :**

Dans cette section, nous avons montré que pour la vigilance orange, le modèle de régression logistique montre des performances qui sont assez vite faibles en fonction de l'échéance de prévision. Elles ne sont pas meilleures que celles du modèle de persistance à court horizon, ce dernier voyant ses performances se dégrader plus vite que celles du modèle testé à mesure que l'horizon augmente. Les meilleures performances sont obtenues pour un seuil de 0,5.

## 5.2. Pr evision du seuil de vigilance rouge

Les r esultats obtenus pour le niveau de vigilance rouge en testant les seuils de probabilit e de 0,2  a 0,9 en calage et en validation sont pr esent es sur la Figure 17. En termes de KSS, les seuils de probabilit e les plus bas donnent les meilleures performances en calage et en validation, et sont meilleurs en calage qu'en validation. Ces mod eles sont donc moins robustes que ceux de pr evision des niveaux de vigilance jaune et orange. Pour le seuil de probabilit e de 0,9, aucun  v enement de vigilance rouge n'est simul e par ces mod eles. En calage, le seuil de 0,2 entra ne les meilleures performances ( $KSS > 0,8$ ), suivi des seuils de 0,3  a 0,5 ( $0,7 < KSS < 0,8$ ) et des seuils de 0,6  a 0,7 ( $0,6 < KSS < 0,7$ ). En validation, les seuils de 0,2 et 0,3 sont les seuls  a fournir des r esultats satisfaisants ( $0,6 < KSS < 0,7$ ) sur les deux p eriodes. Les r esultats de ces diff erents tests en termes de PBT, PBE et FA ont aussi  e t e analys es (non montr e ici). Bien que le PBT soit toujours excellent (gr ace  a une plus grande proportion de valeurs  a 0 dans la chronique binaire de vigilance rouge), le PBE chute sous 0,5 d es le seuil de 0,4 en calage sur 1993-2016, tandis que le FA reste stable ( $< 0,3$ ) pour les seuils de 0,2  a 0,4 puis augmente ( $> 0,4$ ) sur la deuxi eme p eriodes de calage. **Ainsi, pour le niveau de vigilance rouge, le mod ele optimum est cal e sur trois mois d' ete, utilisant la combinaison des chroniques continues de SSWI et binaires de vigilance rouge observ ee et un seuil de probabilit e de 0,3.**

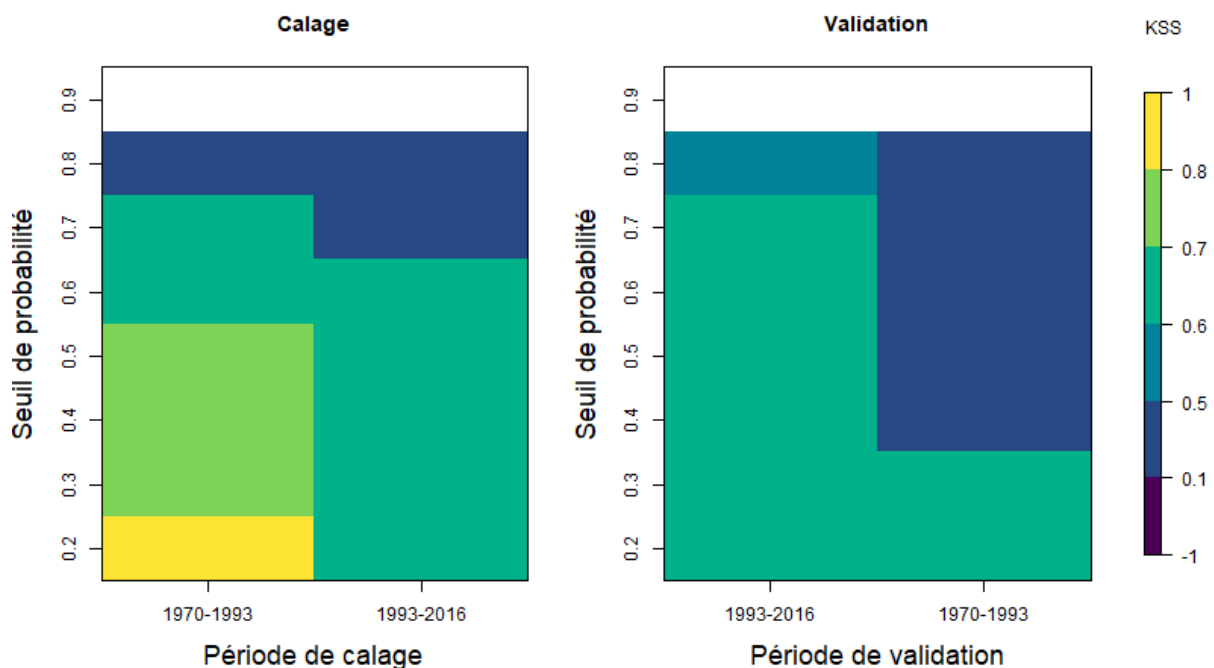


Figure 17 : Performance des mod eles de vigilance rouge en calage ( a gauche) et en validation ( a droite) sur trois mois d' ete pour les seuils de probabilit e de 0,2  a 0,9 pour toutes les stations confondues

Les performances du mod ele optimum  a diff erentes  ech eances sont illustr ees sur la Figure 18. D es la troisi eme d ecade, le BPE passe sous 0,5 et le FA au-dessus de 0,5. Ce mod ele est donc performant sur deux d ecades de pr evision uniquement. En termes de score de persistance (Fig. 19), le mod ele optimum montre de meilleures performances que le mod ele de persistance ( $Eff_1 > 0$ ), sauf  a  ech eance une d ecade en calage sur la p eriodes 1970-1993 ( $Eff_1 < 0$ ).

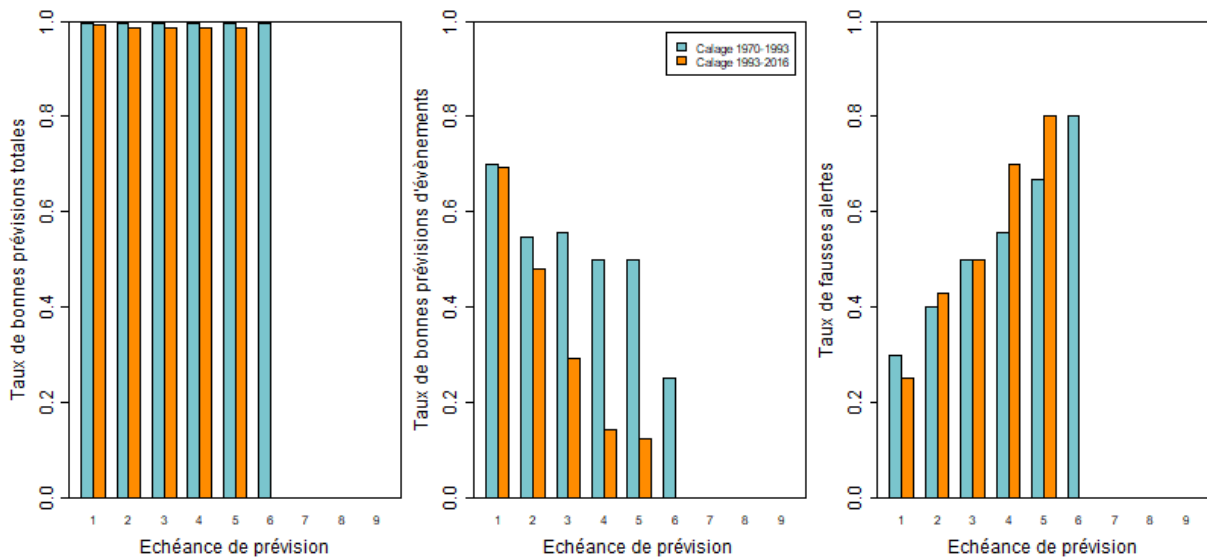


Figure 18 : Performance du modèle optimum de vigilance rouge sur la période de validation pour les échéances de prévision de 1 à 9 décades : taux de bonnes prévisions totales (à gauche, valeur optimale égale à 1), taux de bonnes prévisions d'évènements (au milieu, valeur optimale égale à 1) et taux de fausses alertes (à droite, valeur optimale égale à 0) pour toutes les stations confondues

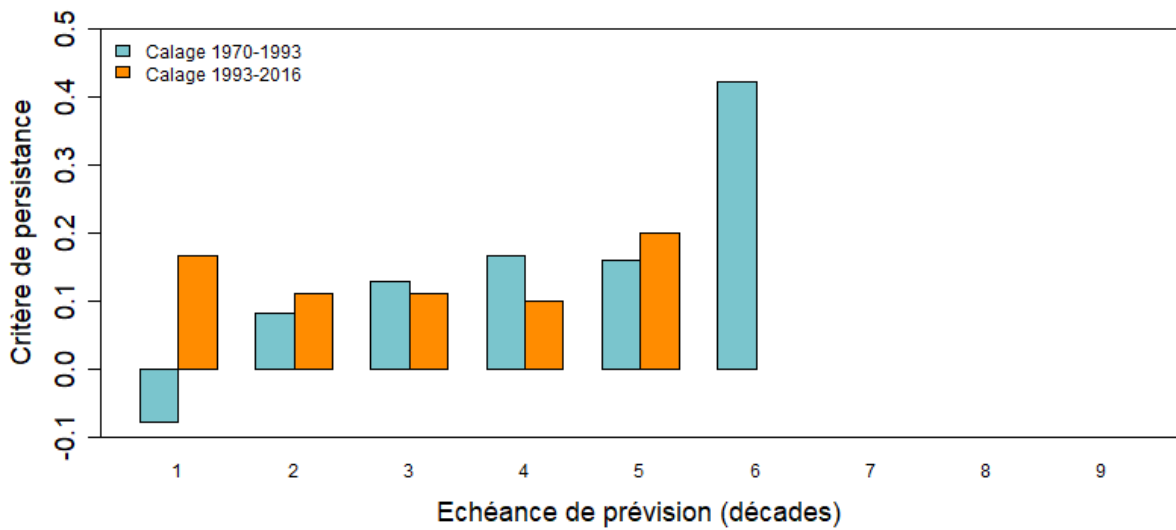


Figure 19 : Performance du modèle optimum de vigilance rouge en calage : scores de critère de persistance pour les échéances de prévision de 1 à 9 décades pour toutes les stations confondues

**Résumé :**

Dans cette section, nous avons montré que pour la vigilance rouge, le modèle de régression logistique montre des performances qui sont assez vite faibles en fonction de l'échéance de prévision, mais qui restent meilleures que le modèle de persistance sauf pour la première échéance en calage sur 1970-1993. Les meilleures performances sont obtenues pour un seuil de 0,3.

## 6. Robustesse spatiale des modèles optimum de vigilance jaune, orange et rouge

La Fig. 20 montre les performances des modèles optimum de vigilances jaune, orange et rouge en PBT calés sur les bassins amont et aval sur la période de calage et sur la période de validation. Ces performances ont été évaluées pour une échéance à une décade. En termes de prévision totale, les trois seuils de vigilance montrent de très bonnes performances ( $> 0,8$ ), autant en calage qu'en validation. Les plus hauts scores sont obtenus pour le seuil rouge, qui contient en fait moins d'évènements d'alerte en proportion (soit plus de valeurs à 0), ce qui est plus simple à prédire. On voit par contraste que c'est le seuil rouge qui montre plus de difficultés à prédire les évènements d'alertes ( $< 0,65$  en validation, voir colonne du milieu), et c'est la vigilance jaune qui montre les meilleures performances sur les périodes de calage et validation ( $> 0,7$ ). On retrouve les mêmes résultats en terme de fausses alertes (colonne de droite) avec les taux les plus bas pour la vigilance jaune (entre 0,13 et 0,25) et les plus hauts pour la vigilance rouge (entre 0,12 et 0,36).

Les scores en calage et validation sont stables, sauf pour la prévision de vigilance rouge qui montre des résultats contrastés : en effet, en calage sur les bassins amont, les résultats de BPT et BPE sont stables mais le FA est nul (et donc meilleur) en validation sur les bassins aval ; tandis qu'en calage sur les bassins aval, les BPE et FA sont très satisfaisants en calage mais se dégradent significativement en validation sur les bassins amont (de 0,88 à 0,63 et de 0,12 à 0,36 respectivement). Ces résultats peuvent être expliqués par le seuil de probabilité choisi pour la vigilance rouge (0,3). En calage sur les bassins amont ce seuil est trop bas et génère trop de fausses alertes, mais les performances s'améliorent en validation sur les bassins aval. Au contraire, il entraîne d'excellentes performances en calage sur les bassins aval mais est trop bas pour les bassins amont en validation.

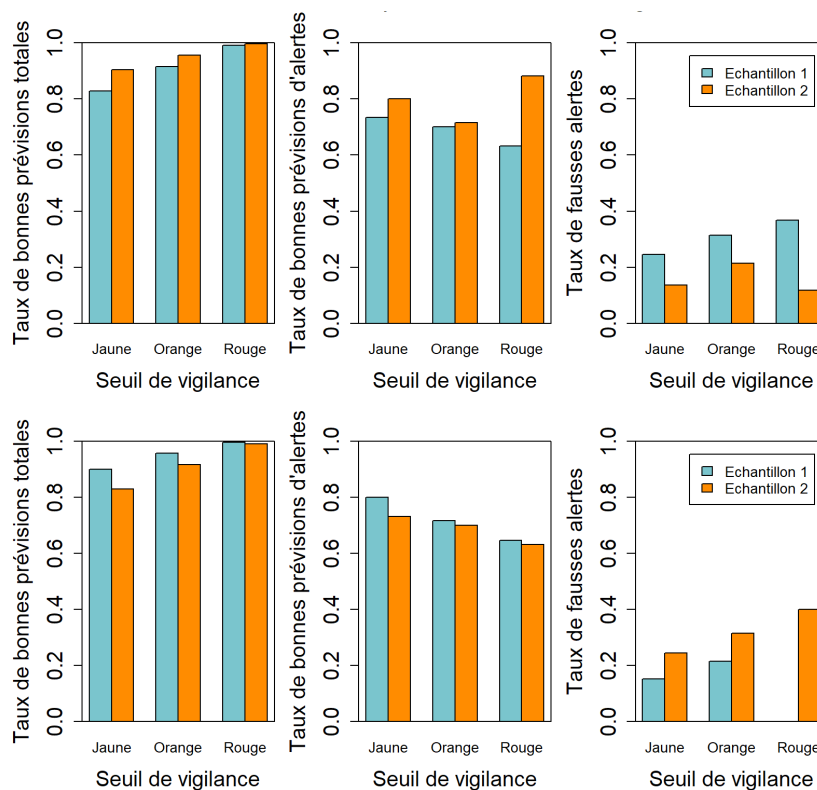


Figure 20 : Performance à une décade d'échéance des modèles optimum de vigilance jaune, orange et rouge sur la période de calage (en haut) et de validation (en bas) : taux de bonnes prévisions totales (à gauche, valeur optimale égale à 1), taux de bonnes prévisions d'évènements (au milieu, valeur optimale égale à 1) et taux de fausses alertes (à droite, valeur optimale égale à 0) pour les modèles calés sur les bassins amont (en bleu, échantillon 1) et aval (en orange, échantillon 2)

## Résumé :

Dans cette section, nous avons montré que pour les vigilances jaune et orange les modèles de régression logistique optimum identifiés dans les sections précédentes sont robustes spatialement. Pour la vigilance rouge, le modèle optimum montre une faible robustesse spatiale, avec notamment une dégradation du score de fausses alertes.

## 7. Conclusions et perspectives

Cette étude avait pour objectif de développer sur le bassin de la Meuse française un modèle de prévision de vigilance sécheresse pour différents niveaux d'alerte (jaune, orange et rouge) en se basant sur des données hydro-météorologiques spatialisées. En effet, la DREAL Grand-Est reçoit de Météo-France des cartes d'indicateurs pluviométriques (SPI) et d'humidité du sol (SSWI) aux pas de temps décadaire et mensuel et souhaite utiliser ces données pour prévoir le passage des différents niveaux d'alerte. Une approche de régression logistique a été utilisée avec les données de débit observé, via la détermination du niveau de vigilance observé, et les cartes d'indicateurs hydro-météorologiques pour modéliser les chroniques binaires de passage de niveau de vigilance en prévision (i.e. au pas de temps suivant). Pour cela, **plusieurs tests ont été menés afin d'évaluer le meilleur jeu de données d'entrée du modèle, la meilleure période de calage au cours de l'année et le seuil de probabilité optimum** pour transformer les sorties du modèle en chroniques binaires. Ces tests ont été réalisés en Split Sample Test, c'est-à-dire en analysant les performances des modèles en calage et en validation sur deux périodes de temps (1970-1993 et 1993-2016), pour évaluer la transférabilité temporelle des modèles. Pour chaque niveau de vigilance, le modèle optimum identifié a été évalué sur plusieurs échéances de prévision.

Ces tests ont tout d'abord été effectués pour le niveau de vigilance jaune afin de définir le modèle optimum pour ce niveau d'alerte. Il a été montré que l'ajout des chroniques de dépassement du niveau de vigilance jaune observé au pas de temps précédant la prévision et le calage des modèles sur la période estivale améliorent significativement les performances. **Le modèle de vigilance jaune le plus performant retenu est celui qui utilise en données d'entrée les chroniques de SSWI et de vigilance observé au pas de temps précédant la prévision, qui est calé sur six mois d'été et avec un seuil de probabilité de 0,5.** Ce modèle est capable de prédire de façon satisfaisante, et jusqu'à une échéance d'un mois (3 décades), les alertes de vigilance jaune en limitant le taux de fausses alertes entre 20 % et 40 % en calage et validation. En se basant sur le modèle optimum de vigilance jaune, des tests ont été effectués avec les données de SSWI et de vigilance observée pour la prévision des niveaux de vigilance orange et rouge. Ces tests ont permis de définir le meilleur seuil de probabilité qui transforme la chronique continues de probabilité de dépassement en chronique binaire de vigilance. Ainsi **les modèles de vigilance orange et rouge les plus performants sont calés sur trois mois d'été et utilisent un seuil de probabilité de 0,5 et 0,3 respectivement, avec pour données d'entrée les chroniques de SSWI et la connaissance de vigilance observée.** Ces modèles sont moins performants pour des échéances de prévision plus lointaines. Enfin, les trois modèles optimums ont été testés en calage et validation sur deux échantillons indépendants de stations pour évaluer leur robustesse spatiale. **Les modèles de vigilance jaune et orange montrent des résultats constants et robustes en calage et validation sur les deux échantillons de bassins, tandis que le modèle de vigilance rouge montre de plus faibles performances sur les bassins amont.**

Bien que les critères de performance aient montré des résultats satisfaisants pour les modèles optimums de vigilance jaune, orange et rouge, les valeurs de critère de persistance tempèrent cette analyse. En effet, **ces modèles sont aussi performants que le modèle de persistance pour des courtes échéances de prévision.** Les modèles proposés se montrent néanmoins plus performants que le



modèle de persistance à des échéances plus lointaines, mais à ces échéances leur performance n'est plus satisfaisante. Comme Météo-France produit aussi des cartes de données de SPI, au pas de temps mensuel uniquement, on peut se demander si l'utilisation de cette nouvelle source de données, en plus des données de SSWI, pourrait améliorer la prévision des niveaux de vigilance sécheresse. Une étude préliminaire, qui n'est pas présentée ici, a produit les mêmes tests que ceux présentés dans ce rapport, en cherchant le modèle optimum de prévision de vigilance sécheresse en utilisant les données spatiales de SPI, SSWI et de vigilance observée au pas de temps mensuel. Les résultats sont moins satisfaisants que ceux obtenus au pas de temps décadaire sans donnée de SPI, ce qui n'est pas encourageant, et les données de SPI ne sont pas disponibles au pas de temps décadaire. Même s'il est légitime de tenter de modéliser les seuils de vigilance sécheresse en se basant sur des cartes d'indicateurs hydro-climatiques, la difficulté de développer ce type d'approche est soulevée dans d'autres études de la littérature. Par exemple récemment, Kumar et al. (2016) ont cherché une corrélation entre SPI et SGI (*Standardized Groundwater Index* – indice d'eau souterraine standardisé) afin de prévoir les sécheresses hydrogéologiques. Cet article montre l'incapacité des auteurs à démontrer un tel lien, ce qui s'explique par la complexité physique des environnements que traversent les chemins de l'eau, i.e. l'atmosphère et les territoires des bassins versants et des réservoirs souterrains. La modélisation hydrologique développée pour la prévision des étiages comme le modèle PREMHYCE (Nicolle et al., 2014), bien que plus lourde à prendre en main en opérationnel, semble donc plus adaptée et performante pour répondre aux besoins de prévision d'évènements de vigilance sécheresse. Elle présente en effet l'avantage de pouvoir tenir compte explicitement des conditions météorologiques futures, qui conditionnent généralement l'évolution des débits de basses eaux.

Enfin, comme il a été montré que la prévision des différents seuils de vigilance sécheresse était significativement plus performante en utilisant la connaissance du seuil de vigilance au pas de temps précédant la prévision, il semble difficile d'utiliser ces résultats directement sur des bassins non jaugés. En effet, la méthodologie appliquée requiert de longues chroniques de débit observé pour l'estimation d'un seuil de vigilance (comme fourni dans le Tableau 2) et du seuil de vigilance au pas de temps précédent la prévision. Une possibilité pour lever ce verrou serait d'utiliser un modèle hydrologique distribué, semi-distribué ou global régionalisé afin de reconstituer les chroniques de débits historiques et de la simulation en temps réel de débit aux stations non jaugées. Ce type de solution demanderait néanmoins aux gestionnaires de prendre en main et de maintenir dans leur service une telle approche de modélisation et avait été écartée par le commanditaire de l'étude lors de la phase de construction du projet.

## Références bibliographiques

- Base de Données Climatologiques - BDCLim (1990) <https://sextant.ifremer.fr/record/1298c4cd-b0a8-4900-a002-e98d8c509243/>
- Delaigue, O., Génot, B., Lebecherel, L., Brigode, P. and Bourgin, P.Y. (2020) Base de données d'observations hydroclimatiques à l'échelle de bassins versants français. Université Paris-Saclay, INRAE, UR HYCAR, Équipe Hydrologie des bassins versants, Antony. URL: <https://webgr.inrae.fr/base-de-donnees>.
- Hanssen, A. W., and Kuipers, W. J. A. (1965) On the relationship between the frequency of rain and various meteorological parameters. Koninklijk Ned. Meteor. Instit., Meded. Verhand., 81, 2–15.
- Kitanidis, P. K., & Bras, R. L. (1980). Real-time forecasting with a conceptual hydrologic model: 2. Applications and results. *Water Resources Research*, 16(6), 1034-1044.
- Klemeš, V. (1986) Operational testing of hydrological simulation models, *Hydrological Sciences Journal*, 31:1, 13-24, DOI: 10.1080/02626668609491024
- Kumar, R., Musuuza, J. L., Van Loon, A. F., Teuling, A. J., Barthel, R., Ten Broek, J., ... & Attinger, S. (2016). Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator. *Hydrology and Earth System Sciences*, 20(3), 1117.
- Leleu, I., Tonnelier, I., Puechberty, R., Gouin, P., Viquendi, I., Cobos, L., Foray, A., Baillon, M., Ndimba, P.-O. (2014) La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La Houille Blanche*, (1) 25-32, DOI: 10.1051/lhb/2014004
- McKee, T. B., Doesken, N. J., and Kleist, J. (1993) The relationship of drought frequency and duration of time scales. Eighth Conference on Applied Climatology, American Meteorological Society, Jan17-23, 1993, Anaheim CA, pp.179-186.
- Nicolle, P., Perrin, C., Andréassian, V., Besson, F., François, D., Le Lay, M., Mathevet, T., Regimbeau, F., Soubeyroux, J.-M., Thiéry, D., Viel, C., Augeard, B., Maugis, P., and Morice, E. (2014): Prévion des Etiages par des Modèles Hydrologiques : Comparaison et Evaluation (Projet PREMHYCE) – Rapport de synthèse (Low-flow forecasting by hydrological modelling : comparison and evaluation (*PREMHYCE project – Final report; In French*). *Rapport final ONEMA*, 44 pp, <http://www.documentation.eaufrance.fr/notice/prevision-des-etiages-par-des-modeles-hydrologiques-comparaison-et-evaluation-projet-premhyce-rappor0>
- Thirel, G., Collet, L. and Rousset, F. (2020) Rapport final du projet CHIMERE21.
- Vidal, J.-P., Martin, E., Franchistéguy, L., Habets, F., Soubeyroux, J.-M., Blanchard, M., Baillon, M. (2010) Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite. *Hydrology and Earth System Sciences*, 14(3), 459-478. DOI : 10.5194/hess-14-459-2010