

## TABLE DES MATIÈRES

	Page
INTRODUCTION .....	1
<b>CHAPITRE 1 REVUE DE LA LITTÉRATURE .....</b>	<b>3</b>
1.1 Équifinalité et singularité des jeux de paramètres .....	3
1.2 Régionalisation par la méthode de la régression linéaire multiple .....	4
1.3 Régionalisation par la méthode de la proximité spatiale .....	5
1.4 Régionalisation par la méthode de similitude physique .....	6
1.5 Méthodes d'interpolation utilisées lors du transfert des jeux de paramètres .....	7
1.5.1 Méthode 1 : Interpolation des paramètres par la moyenne arithmétique ....	7
1.5.2 Méthode 2 : Interpolation des débits par la moyenne arithmétique .....	8
1.5.3 Méthode 3 : Interpolation des débits par la pondération inverse à la distance .....	8
1.6 Problématiques et objectifs de l'étude .....	9
<b>CHAPITRE 2 MÉTHODOLOGIE.....</b>	<b>13</b>
2.1 Modèle hydrologique GR4J et module de neige CemaNeige.....	13
2.2 Variation du nombre de paramètres libres .....	14
2.3 Région étudiée et base de données géospatiale.....	18
2.4 Bases de données hydrologiques et météorologiques.....	19
2.5 Calibration des modèles hydrologiques .....	21
2.5.1 Algorithmes d'optimisation .....	21
2.5.2 Échantillonnage aléatoire.....	21
2.6 Configurations testées.....	22
2.7 Performance et robustesse des modèles hydrologiques .....	22
2.7.1 Critère d'efficacité de Nash-Sutcliffe .....	23
2.7.2 Taux de succès .....	23
2.8 Équifinalité et méthode de ré-échantillonnage par <i>bootstrap</i> .....	24
2.9 Bassins versants pseudos non jaugés et validation du modèle hydrologique .....	24
2.10 Échantillons évalués et sélection des bassins versants donneurs.....	25
2.11 Régionalisation par la méthode de la régression linéaire multiple .....	26
2.11.1 Modèle régional linéaire .....	26
2.11.2 Évaluation statistique du modèle régional linéaire .....	27
2.12 Régionalisation par la méthode de la proximité spatiale .....	28
2.13 Régionalisation par la méthode de similitude physique .....	29
<b>CHAPITRE 3 RÉSULTATS.....</b>	<b>31</b>
3.1 Identité des paramètres libres du modèle hydrologique à 9 degrés de liberté .....	31
3.2 Performance des cinq configurations MH/OC en calibration.....	33
3.3 Caractéristiques physiques retenues pour construire l'index de similitude .....	33
3.4 Distribution des jeux de paramètres optimisés au sein de l'espace paramétrique .....	35

3.5	Taux de succès obtenu par les méthodes de régionalisation par proximité spatiale et par similitude physique .....	36
3.5.1	Interpolation des paramètres par la moyenne arithmétique .....	37
3.5.2	Interpolation des débits par la moyenne arithmétique .....	38
3.5.3	Interpolation des débits par la pondération inverse à la distance.....	39
3.6	Nombre de caractéristiques physiques sélectionnées pour construire les modèles régionaux linéaires.....	40
3.7	Taux de succès obtenu par la méthode de régionalisation par régression linéaire multiple .....	44
3.8	Analyse statistique des modèles régionaux linéaires construit par régression linéaire multiple .....	44
3.8.1	Coefficient de détermination $R^2$ .....	44
3.8.2	Statistique $F$ .....	46
3.8.3	Seuil descriptif $P$ -value.....	47
3.9	Performance des méthodes de régionalisation.....	49
CHAPITRE 4 DISCUSSION .....		55
4.1	Comparaison des méthodes de régionalisation des paramètres et importance de transférer le jeu de paramètres dans son intégralité .....	55
4.2	Normalisation de l'identité des bassins versants donneurs pour les cinq configurations MH/OC .....	56
4.3	Impact du nombre de bassins versants donneurs .....	57
4.4	Impact de l'échantillon de bassins versants évalué.....	58
4.5	Complexité du modèle hydrologique.....	59
4.5.1	Impact de la complexité du modèle hydrologique sur la performance des configurations MH/OC régionalisées .....	59
4.5.2	Impact de la complexité du modèle hydrologique sur sa robustesse et sur la performance des méthodes de régionalisation des paramètres .....	60
4.5.3	Modèle hydrologique à 9 paramètres libres.....	62
4.6	Impact de la distribution du jeu de paramètres optimisés sur les méthodes de régionalisation des paramètres.....	63
4.7	Information contenue dans les caractéristiques physiques retenues pour construire l'index de similitude .....	63
4.8	Incertitude liée à l'équifinalité .....	66
4.9	Validité statistique du modèle régional linéaire construit pour appliquer la méthode de régionalisation par régression linéaire multiple .....	68
CONCLUSION.....		71
ANNEXE I	DISTRIBUTION DES JEUX DE PARAMÈTRES OPTIMISÉS DES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES.....	75
ANNEXE II	COEFFICIENT DE DÉTERMINATION $R^2$ DES MODÈLES RÉGIONAUX LINÉAIRES CONSTRUITS POUR LES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES .....	77

ANNEXE III	STATISTIQUE $F$ DES MODÈLES RÉGIONAUX LINÉAIRES CONSTRUITS POUR LES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES.....	79
ANNEXE IV	SEUIL DESCRIPTIF $P$ -VALUE DES MODÈLES RÉGIONAUX LINÉAIRES CONSTRUITS POUR LES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES .....	81
ANNEXE V	COMPARAISON DES ÉCHANTILLONS <i>GLOBAL</i> ET <i>DONNEUR</i> SUR LA BASE DU CRITÈRE D'EFFICIENCE $NSE_{MOYEN}$ DES MÉTHODES DE RÉGIONALISATION PAR PROXIMITÉ SPATIALE ET SIMILITUDE PHYSIQUE COUPLÉES AUX MÉTHODES D'INTERPOLATION DES DÉBITS .....	83
ANNEXE VI	PERFORMANCE EN CALIBRATION DES MODÈLES HYDROLOGIQUES À 6, 9 ET 15 PARAMÈTRES LIBRES SELON LES INTERVALLES DE VALEURS DES CARACTÉRISTIQUES PHYSIQUES.....	85
BIBLIOGRAPHIE .....		87



## LISTE DES TABLEAUX

	Page
Tableau 4.1	Comparaison des échantillons <i>global</i> et <i>donneur</i> sur la base du critère d'efficacité $NSE_{\text{moyen}}$ de la méthode de régionalisation par similitude physique couplée à l'interpolation par la PID des débits, présenté pour les trois configurations MH/OC optimisées par l'algorithme d'optimisation SCE-UA .....59
Tableau 4.2	Estimation du degré d'amélioration résultant du gain en complexité des modèles hydrologiques à 9 et 15 paramètres libres .....62



## LISTE DES FIGURES

		Page
Figure 2.1	Schéma conceptuel du modèle hydrologique GR4J .....	15
Figure 2.2	Schéma conceptuel du module de neige CemaNeige .....	16
Figure 2.3	Liste et description des paramètres libres et fixés du modèle hydrologique GR4J et du module de neige CemaNeige.....	17
Figure 2.4	Localisation des 266 bassins versants de la région étudiée .....	19
Figure 2.5	Caractéristiques physiques des 266 bassins versants de la région étudiée.....	20
Figure 3.1	Distribution du critère d'efficacité $NSE_{moyen}$ des 10 jeux de paramètres optimisés pour chacun des 266 bassins versants des cinq configurations MH/OC testées.....	34
Figure 3.2	Rang de la performance obtenue par les caractéristiques physiques lors de la sélection des caractéristiques utilisées pour calculer l'index de similitude de la méthode de régionalisation par similitude physique .	34
Figure 3.3	Distribution des 10 jeux de paramètres optimisés par bassin versant pour les trois configurations MH/OC du modèle hydrologique à 6 paramètres libres .....	37
Figure 3.4	Taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique – Interpolation des paramètres par la moyenne arithmétique .....	41
Figure 3.5	Taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique – Interpolation des débits par la moyenne arithmétique .....	42
Figure 3.6	Taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique – Interpolation des débits par la PID .....	43
Figure 3.7	Taux de succès de la méthode de régionalisation par régression linéaire multiple – Modèle régional linéaire construit avec quatre caractéristiques physiques.....	45
Figure 3.8	Coefficient de détermination $R^2$ des modèles régionaux linéaires construits avec quatre caractéristiques physiques.....	46

Figure 3.9	Statistique $F$ des modèles régionaux linéaires construits avec quatre caractéristiques physiques.....	48
Figure 3.10	Seuil descriptif $P$ -value des modèles régionaux linéaires construits avec quatre caractéristiques physiques .....	49
Figure 3.11	Distribution du critère d'efficacité $NSE_{moyen}$ des 266 bassins versants de l'échantillon global résultant du transfert de jeux de paramètres issus de 12 bassins versants donneurs pour les méthodes de régionalisation par proximité spatiale et similitude physique.....	51
Figure 3.12	Distribution du critère d'efficacité $NSE_{moyen}$ des 167 bassins versants de l'échantillon donneur résultant du transfert des jeux de paramètres issus de 12 bassins versants donneurs pour les méthodes de régionalisation par proximité spatiale et similitude physique.....	52
Figure 3.13	Distribution du critère d'efficacité $NSE_{moyen}$ des échantillons <i>global</i> et <i>donneur</i> résultant du modèle régional linéaire construit avec quatre caractéristiques physiques et 130 bassins versants donneurs pour la méthode de régionalisation par régression linéaire multiple .....	53
Figure 4.1	Localisation des 167 bassins versants de l'échantillon <i>donneur</i> .....	57
Figure 4.2	Performance des modèles hydrologiques à 6, 9 et 15 paramètres libres exprimée selon l'intervalle de valeurs des caractéristiques physiques.....	61
Figure 4.3	Détail des caractéristiques physiques des 266 bassins versants de l'échantillon <i>global</i> .....	65
Figure 4.4	Enveloppe de l'incertitude relative en fonction du nombre de bassins versants donneurs obtenue par les méthodes de régionalisation par proximité spatiale et similitude physique pour les deux échantillons étudiés ( <i>global</i> et <i>donneur</i> ) .....	68



## LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AISH	Association Internationale des Sciences Hydrologiques
CNLWIS	Canadian National Land and Water Information Service
COCG	Conseil canadien de la géomatique
CP	Caractéristique physique
éch.	Échantillon
MH/OC	Modèles hydrologiques couplés aux options de calibration
PBNJ	Prévisions en bassins versants non jaugés
s.d.	Sans dimension



## LISTE DES SYMBOLES ET UNITÉS DE MESURE

$\alpha$	risque posé de rejeter à tort l'hypothèse nulle posée par le test de Fisher
°	degré
°C	degré Celsius
$\Delta$	delta
%	pourcentage
j	jour
km	kilomètre
km <sup>2</sup>	kilomètre carré
m	mètre
m/km	mètre/kilomètre
mm	millimètre
mm/°C	millimètre/degré Celsius
mm/j	millimètre/jour
NSE	critère d'efficience de Nash-Sutcliffe
NSE <sub>moyen</sub>	critère d'efficience moyen de Nash-Sutcliffe
PID	pondération inverse à la distance
<i>P-value</i>	seuil descriptif du test de Fisher
$Q_1$	1 <sup>er</sup> quartile de la boîte à moustache
$Q_3$	3 <sup>ièm</sup> quartile de la boîte à moustache
$R^2$	coefficient de détermination des variables du modèle régional linéaire
Statistique $F$	variable de décision du test de Fisher

XXII

TS            taux de succès

## INTRODUCTION

Le réchauffement du système climatique est sans équivoque. À l'exception des régions subtropicales arides où les précipitations devraient diminuer, la plupart des régions du monde devraient connaître une augmentation des précipitations moyennes annuelles, de l'intensité et de la fréquence des précipitations extrêmes (IPCC, 2013). Les conséquences de ce cycle hydrologique modifié par des eaux de surface plus chaudes, des précipitations plus intenses et des étiages plus sévères, se traduisent par une dégradation de la qualité des cours d'eau ainsi qu'un accroissement des risques d'inondations et de sécheresses (IPCC, 2014). L'exposition et la vulnérabilité face aux risques d'inondation, la sécurité alimentaire, la sécurité énergétique, l'accès à l'eau potable ainsi que la détérioration des écosystèmes sont au nombre des risques hydrologiques exprimés dans le rapport.

Malgré la nécessité d'implanter des politiques de gestion durable de l'eau, dans plusieurs régions du monde, les réseaux de collecte d'observations hydrologiques – qui procurent une information nécessaire aux outils de modélisation hydrologique – sont en déclin ou inexistantes (Sivapalan et al., 2003). La pratique usuelle qui consiste à modéliser le comportement hydrologique d'un bassin versant, en optimisant les paramètres du modèle hydrologique à l'aide de séries d'observations, n'est pas une option pour ces bassins versants dits, *non jaugés* (He, Bárdossy et Zehe, 2011). Un bassin versant est qualifié *non jaugé* lorsque le débit de son exutoire n'est pas observé ou que les observations disponibles sont soit de piètre qualité, soit inadaptées à l'échelle temporelle modélisée (Sivapalan et al., 2003). À cette définition, il faut aussi inclure les bassins versants qui subissent ou subiront des modifications climatiques ou physiques significatives, et pour lesquels les séries d'observations hydrologiques passées ne seront plus représentatives de leur réponse hydrologique (He, Bárdossy et Zehe, 2011).

La régionalisation des paramètres propose et explore différentes méthodes pour transférer les paramètres de bassins versants jaugés et similaires, vers un bassin versant non jaugé qui leur est similaire, dans le but de prédire efficacement son comportement hydrologique. La

décennie (2003-2012) sur les prévisions en bassins versants non jaugés (PBNJ), initiée par l'Association Internationale des Sciences Hydrologiques (AISH), a fortement contribué à créer un engouement pour la thématique. Environ 70 études ont été publiées en une vingtaine d'années sur le sujet, la plupart depuis 2003 (Razavi et Coulibaly, 2013). Les résultats ne sont cependant pas constants. Le degré d'affinité de chaque méthode de régionalisation pour les spécificités climatiques et géographiques des régions étudiées ainsi que l'inconstance des caractéristiques physiques retenues dans les différentes études sont en cause (Oudin et al., 2008).

Les méthodes de régionalisation par régression se distinguent par leur popularité; deux tiers des études présentées par Razavi et Coulibaly (2013) appliquent ces méthodes. Les hypothèses sur lesquelles ces méthodes reposent sont cependant critiquées (Oudin et al., 2008). Les méthodes de proximité spatiale et de similitude physique sont celles qui performant le mieux dans les régions où prévalent des conditions hivernales de froid et de neige, dans les régions chaudes et tempérées et dans les régions arides (Razavi et Coulibaly, 2013). Au Canada, ces trois approches de régionalisation sont évaluées par Samuel, Coulibaly et Metcalfe (2011) sur 94 bassins versants situés en Ontario et par Arsenault et Brissette (2014) sur 268 bassins versants situés au Québec.

La présente étude applique ces trois méthodes de régionalisation. Elle vise à évaluer l'impact de la complexité du modèle hydrologique ainsi que celui de la distribution des jeux de paramètres équifinaux du modèle hydrologique, sur la prédiction du débit à l'exutoire d'un bassin versant non jaugé. L'étude se penche ensuite sur la méthode de régionalisation par régression linéaire multiple et analyse le modèle régional linéaire construit pour transférer les jeux de paramètres.

## CHAPITRE 1

### REVUE DE LA LITTÉRATURE

Deux approches sont généralement explorées par les méthodes de régionalisation. La première approche repose sur l'explication des paramètres du modèle hydrologique par les caractéristiques physiques des bassins versants alors que la seconde consiste à identifier des régions d'homogénéité, au sens géographique ou en similitude, pour transférer les jeux de paramètres entre bassins versants similaires (He, Bárdossy et Zehe, 2011). Le concept d'homogénéité, à la base de la deuxième approche, repose sur l'hypothèse que les bassins versants qui possèdent des caractéristiques physiques similaires – climat, géologie, topographie, végétation, types de sols – ont une réponse hydrologique similaire et il implique que les bassins versants non jaugés et donneurs soient issus de la même région hydrologiquement homogène (Smakhtin, 2001).

#### 1.1 Équifinalité et singularité des jeux de paramètres

Deux limitations sont communes aux méthodes de régionalisation des paramètres. La première est liée au principe d'équifinalité (Beven, 2006) qui établit la non-unicité des jeux de paramètres du modèle hydrologique optimisé. La conséquence de cette équifinalité est qu'il existe une multitude de jeux de paramètres qui offrent une prédiction similaire en calibration mais qui ne sont pas tous équivalents en termes de transférabilité et de sensibilité [Gibbs, Maier et Dandy (2012); Oudin et al. (2008); Bárdossy (2007)]. Cette équifinalité implique qu'il existe une interdépendance entre les paramètres d'un modèle hydrologique dont il faut tenir compte en privilégiant le transfert de jeux de paramètres entiers plutôt que le transfert individuel des paramètres [Bárdossy (2007); Parajka, Merz et Blöschl (2005); Oudin et al. (2008)]. Perrin, Michel et Andréassian (2001) ont démontré que la parcimonie d'un modèle hydrologique tend à limiter cette équifinalité et à augmenter la robustesse des jeux de paramètres transférés. Cette non-unicité affecte aussi la répartition des paramètres au sein d'une structure régionale identifiable puisque les jeux de paramètres de bassins versants

possédant un comportement hydrologique similaire peuvent n'avoir aucune corrélation entre eux [Arsenault et Brissette (2014); Peel et Bloschl (2011)].

La deuxième limitation sera ici appelée la *singularité* du jeu de paramètres optimisés. Un jeu de paramètre est singulier pour trois raisons. Premièrement, il intègre une compensation pour tenir compte de l'erreur contenue dans les données observées et optimiser la relation entre le débit simulé du modèle hydrologique et celui observé à l'exutoire du bassin versant modélisé [Oudin et al. (2006b); Andréassian, Perrin et Michel (2004); Lee et al. (2006); McIntyre et al. (2005); Beven (2006)]. Deuxièmement, il dépend de la période de calibration utilisée [Merz et Blöschl (2004); Brigode, Oudin et Perrin (2013); Beven (2006)] et des conditions climatiques spécifiques liées à cette période (Yapo, Gupta et Sorooshian, 1996). Troisièmement, il est tributaire de l'erreur conceptuelle des modèles hydrologiques globaux – dont les composantes s'adaptent de façon spécifique aux processus hydrologiques dominants selon le type de bassin versant (Beven, 1999) – ce qui implique une variation régionale de la signification des paramètres optimisés [Lee et al. (2006); Duan et al. (2006)].

## 1.2 Régionalisation par la méthode de la régression linéaire multiple

La régionalisation par la méthode de la régression linéaire multiple consiste à construire, à partir d'un groupe de bassins versants donneurs, un modèle régional linéaire qui lie les paramètres optimisés du modèle hydrologique – les variables expliquées – aux caractéristiques physiques des bassins versants modélisés – les variables explicatives – dans le but d'estimer les paramètres du modèle hydrologique du bassin versant non jaugé. Un modèle régional linéaire est construit pour chacun des paramètres libres du modèle hydrologique tel que formalisé par Wagener et Wheeler (2006);

$$\hat{\theta}_L = H_R(\theta_R|\Phi) + V_R \quad (1.1)$$

où  $\hat{\theta}_L$  est la valeur estimée du paramètre du modèle hydrologique du bassin versant non jaugé,  $H_R$  est le modèle régional linéaire construit avec les bassins versants jaugés, qui lie les



valeurs du paramètre estimé  $\theta_R$ , aux caractéristiques physiques des bassins versants  $\Phi$ , et  $V_R$  constitue le terme d'erreur du modèle régional linéaire.

La méthode suppose qu'il existe une relation bien définie entre les caractéristiques physiques du bassin versant et les paramètres optimisés du modèle hydrologique malgré le fait que le concept d'équifinalité évoque la non-unicité du jeu de paramètres optimisés [Oudin et al. (2008); Beven (1999); Bárdossy (2007)]. La deuxième hypothèse de la méthode présume que les caractéristiques physiques sélectionnées pour construire le modèle régional linéaire sont représentatives des processus hydrologiques du bassin versant mais cette représentativité est difficile à vérifier [Oudin et al. (2008); Peel et Blöschl (2011)]. Bien que certains paramètres du modèle hydrologique puissent être fortement corrélés aux caractéristiques physiques, cette corrélation est généralement absente pour la majorité des paramètres [Yadav, Wagener et Gupta (2007); Lee et al. (2006); McIntyre et al. (2005)]. La dernière limitation de la méthode réside dans l'hypothèse de linéarité des relations entre les paramètres du modèle hydrologique et les caractéristiques physiques, qui est une simplification utile (Parajka, Blöschl et Merz, 2007) mais peu probable d'un système naturel (Samuel, Coulibaly et Metcalfe, 2011). Pour Ouarda, St-Hilaire et Bobée (2008), cette structure linéaire limite la sensibilité de la méthode à l'hétérogénéité régionale. Lee et al. (2006) démontrent cependant que ce sont l'erreur sur les données et les limitations de l'information disponible qui affectent la performance du modèle hydrologique; le choix du modèle de régression n'a que peu d'influence sur cette dernière.

### **1.3 Régionalisation par la méthode de la proximité spatiale**

La régionalisation par la méthode de la proximité spatiale consiste à transférer au bassin versant non jaugé, les jeux de paramètres optimisés de bassins versants donneurs qui sont sélectionnés en fonction de la proximité entre leur centroïde respectif et celui du bassin versant non jaugé. Cette proximité est calculée par la distance euclidienne (He, Bárdossy et Zehe, 2011) ;

$$d_{t,d} = \sqrt{(X_t - X_d)^2 + (Y_t - Y_d)^2} \quad (1.2)$$

où  $X_t, Y_t$  sont les coordonnées du bassin versant non jaugé et  $X_d, Y_d$  sont les coordonnées du bassin versant donneur.

L'hypothèse sous-jacente à la méthode est, qu'étant donnée leur proximité spatiale, les bassins versants non jaugés et les bassins versants donneurs sont *de facto* situés dans une région homogène et géographiquement similaire leurs conférant un comportement similaire (Samuel, Coulibaly et Metcalfe, 2011). Or, le fait que deux bassins versants soient contigus n'implique pas nécessairement qu'ils partagent un même comportement et une même réponse hydrologiques [He, Bárdossy et Zehe (2011); Ouarda, St-Hilaire et Bobée (2008)]. Pour Oudin et al. (2008), la performance de la méthode dépend de la densité du réseau de stations de jaugeage; plus les stations de jaugeage sont rapprochées, meilleure est la performance de la méthode.

#### 1.4 Régionalisation par la méthode de similitude physique

La régionalisation par la méthode de similitude physique consiste à transférer au bassin versant non jaugé, les jeux de paramètres optimisés de bassins versants donneurs qui sont sélectionnés sur la base de leur proximité en similitude au bassin versant non jaugé. Cette similitude est estimée à l'aide de caractéristiques physiques et climatiques. La distance en similitude est généralement déterminée en fonction de l'index de similitude ( $\Phi$ ) défini par Burn et Boorman (1993) ;

$$\Phi = \sum_{i=1}^k \frac{|X_i^G - X_i^U|}{\Delta X_i} \quad (1.3)$$

où  $k$  est le nombre de caractéristique physiques contenues dans l'index de similitude,  $X_i^G$  est la valeur de la  $i^{\text{ième}}$  caractéristique physique du bassin versant potentiellement donneur,  $X_i^U$  est la valeur de de la  $i^{\text{ième}}$  caractéristique physique du bassin versant non jaugé et  $\Delta X_i$  est le

domaine des valeurs possibles de la  $i^{\text{èm}}$  caractéristique physique pour l'ensemble des bassins versant potentiellement donneurs.

La difficulté à vérifier si les caractéristiques physiques, sélectionnées pour construire l'index de similitude, sont représentatives des processus hydrologiques constitue, ici aussi, une limitation de la méthode. Le fait de regrouper des bassins versants au sein de régions homogènes, sur la base de leurs caractéristiques physiques et climatiques, n'assure pas qu'ils soient similaires d'un point de vue hydrologique selon Oudin et al. (2010) qui ont investigué la question en étudiant un échantillon composé de 893 bassins versants français et 10 bassins versants anglais.

## 1.5 Méthodes d'interpolation utilisées lors du transfert des jeux de paramètres

Trois méthodes sont utilisées pour transférer l'information contenue dans les jeux de paramètres du modèle hydrologique des bassins versants donneurs, vers celui du bassin versant non jaugé. La première méthode est basée sur le transfert individuel de chacun des paramètres libres du modèle hydrologique alors que les deux autres méthodes transfèrent les jeux de paramètres dans leur intégralité. Ces deux dernières méthodes présentent l'avantage de tenir compte de l'interdépendance des paramètres libres au sein du jeu de paramètres (*Voir* section 1.1) et offrent conséquemment une meilleure prédiction du débit à l'exutoire du bassin versant non jaugé [(Arsenault et Brissette, 2014); (Oudin et al., 2008)].

### 1.5.1 Méthode 1 : Interpolation des paramètres par la moyenne arithmétique

La première méthode consiste à transférer au bassin versant non jaugé, un jeu de paramètres où chacun des paramètres est égal à la moyenne arithmétique des valeurs du paramètre optimisé contenu dans les jeux transférés. La méthode est décrite par Oudin et al. (2008);

$$\hat{Q}_{(j)} = \hat{Q} \left( j, \frac{\sum_{i=1}^m X_i}{m} \right) \quad (1.4)$$

où  $\hat{Q}$  est le débit estimé à l'exutoire du bassin versant non jaugé,  $j$  est le pas de temps journalier,  $m$  est le nombre de bassins versants donneurs et  $X_i$  est le vecteur du jeu de paramètres pour chaque bassin versant donneur  $i$ .

### 1.5.2 Méthode 2 : Interpolation des débits par la moyenne arithmétique

La deuxième méthode consiste à transférer au bassin versant non jaugé, le jeu de paramètres de chacun des bassins versants donneurs et d'évaluer le débit simulé à l'exutoire du bassin versant non jaugé pour chacun des jeux de paramètres transférés. Les débits simulés sont ensuite interpolés à l'aide de la moyenne arithmétique tel que le présentent Oudin et al. (2008);

$$\hat{Q}_{(j)} = \frac{1}{m} \sum_{i=1}^m \hat{Q}(j, X_i) \quad (1.5)$$

où  $\hat{Q}$  est le débit estimé à l'exutoire du bassin versant non jaugé,  $j$  est le pas de temps journalier,  $m$  est le nombre de bassins versants donneurs et  $X_i$  est le vecteur du jeu de paramètres pour chaque bassin versant donneur  $i$ .

### 1.5.3 Méthode 3 : Interpolation des débits par la pondération inverse à la distance

La troisième méthode suit l'approche de la deuxième méthode quant au transfert des jeux de paramètres et à la simulation des débits à l'exutoire du bassin versant non jaugé. Pour cette troisième méthode, les débits sont interpolés à l'aide d'équations adaptées de Samuel, Coulibaly et Metcalfe (2011) qui proposent une formulation des équations par la pondération inverse à la distance (PID) (Shepard, 1968) pour interpoler les paramètres par la PID. L'interpolation des débits par la PID prend la forme;

$$\hat{Q}_{(j)} = \sum_{i=1}^m w_i q_i \quad (1.6)$$

où  $\hat{Q}$  est le débit estimé à l'exutoire du bassin versant non jaugé,  $j$  est le pas de temps journalier,  $m$  est le nombre de bassins versants donneurs,  $q_i$  est le débit simulé pour chaque jeu de paramètres transféré au bassin versant non jaugé et  $w_i$  est la fonction de poids définie telle que;

$$w_{(i)} = \frac{(h_i^{-2})}{\sum_{i=1}^m (h_i^{-2})} \quad (1.7)$$

où  $h_i$  est la distance entre le  $i^{\text{ém}}$  bassins versants donneurs et celui qui est non jaugé. La distance considérée est celle entre les centroïdes dans le cas de l'approche par proximité spatiale et celle déterminée par l'index de similitude pour l'approche par similitude physique.

## 1.6 Problématiques et objectifs de l'étude

Dans un contexte d'équifinalité, la complexité du modèle hydrologique – le nombre de paramètres libres du modèle – doit être *suffisante* pour décrire efficacement les processus hydrologiques du bassin versant (Beven, 2006). Typiquement, une complexité supérieure à ce qui est suffisant résulte en une surparamétrisation du modèle hydrologique, une interaction accrue entre les paramètres libres ainsi qu'une perte de robustesse du modèle alors qu'un modèle hydrologique trop parcimonieux peut manquer de complexité pour s'adapter à des bassins versants trop différents (Perrin, Michel et Andréassian, 2003). Un modèle hydrologique parcimonieux est généralement admis comme étant plus robuste (Perrin, Michel et Andréassian, 2001) et plus approprié pour appliquer les méthodes de régionalisation des paramètres (Razavi et Coulibaly, 2013), qu'un modèle hydrologique complexe. Conséquemment, l'objectif principal du présent mémoire est de :

- quantifier l'impact de la complexité du modèle hydrologique sur la prédiction du débit à l'exutoire d'un bassin versant non jaugé soumis à un climat nordique. Pour observer cet impact, trois modèles hydrologiques issus du même modèle conceptuel et possédant respectivement 6, 9 et 15 paramètres libres sont comparés.

D'autre part, un bassin versant constitue un ensemble unique de caractéristiques physiques et de modifications anthropiques et cette unicité se reflète dans le jeu de paramètres optimisés du modèle hydrologique de ce bassin versant (Beven, 1999). Dans un contexte de singularité des jeux de paramètres, un jeu de paramètres optimisés contient donc une part d'information qui est particulière au bassin versant modélisé et non pertinente en termes de transfert vers un autre bassin versant, par exemple l'incertitude liée aux observations de débit. Duan, Sorooshian et Gupta (1992) observent que la méthode d'échantillonnage aléatoire de l'espace paramétrique a pour effet de générer des jeux de paramètres dont la distribution est fortement dispersée au sein de l'espace paramétrique du modèle hydrologique. L'échantillonnage aléatoire permet donc d'identifier plusieurs régions de l'espace paramétrique aptes à décrire les processus hydrologiques du bassin versant modélisé. L'étude pose l'hypothèse que la diversité de l'information contenue dans un jeu de paramètres généré par échantillonnage aléatoire permet de s'affranchir, en partie, de la singularité du jeu de paramètres optimisés. Donc, le premier objectif spécifique de ce travail consiste à :

- évaluer l'impact de la distribution des jeux de paramètres optimisés pour les bassins versants donneurs, sur l'aptitude de ces jeux de paramètres à décrire le comportement hydrologique d'un bassin versant non jaugé. Pour observer cet impact, les jeux de paramètres générés par deux algorithmes d'optimisation et ceux générés par l'échantillonnage aléatoire de l'espace paramétrique sont comparés.

Finalement, malgré sa forte popularité, la méthode de régionalisation par régression linéaire se montre moins performante que les méthodes par proximité spatiale ou similitude physique dans plusieurs études comparatives [Arsenault et Brissette (2014); Bao et al. (2012); Oudin et al. (2008); Parajka, Merz et Blöschl (2005) ; Merz et Blöschl (2004)]. La principale critique de la méthode réside dans la faible corrélation observée entre les caractéristiques physiques des bassins versants et les paramètres optimisés du modèle hydrologique (*Voir* section 1.2). En conséquence, le second objectif spécifique de ce travail consiste à :

- vérifier la pertinence d'appliquer la régression linéaire multiple pour régionaliser les paramètres d'un modèle hydrologique. Pour observer la pertinence de cette méthode de régionalisation, la validité du modèle régional linéaire construit pour appliquer la méthode est statistiquement analysée.





## CHAPITRE 2

### MÉTHODOLOGIE

Les 266 bassins versants de la région étudiée sont soumis à des conditions climatiques nordiques. Un ensemble de caractéristiques physiques est répertorié pour chacun de ces bassins versants. La qualité des séries d'observations météorologiques et hydrologiques des bassins versants est variable et amène la nécessité de les regrouper au sein d'échantillons distincts. Choisi pour sa parcimonie de paramètres, le modèle hydrologique utilisé pour appliquer les méthodes de régionalisation ne simule pas les phénomènes de neige. Il est donc couplé à un module de neige. Trois versions, distinctes en complexité, du modèle hydrologique résultant sont proposées et la version la plus parcimonieuse est optimisée selon trois options. En validation des modèles hydrologiques, l'étude distingue la robustesse de la performance et elle estime l'incertitude liée à l'équifinalité des jeux de paramètres. Le présent chapitre décrit les outils, les bases de données et les méthodes déployées pour atteindre les objectifs de l'étude ainsi que les critères d'évaluation utilisés pour quantifier les résultats.

#### **2.1 Modèle hydrologique GR4J et module de neige CemaNeige**

Le modèle hydrologique GR4J (Perrin, Michel et Andréassian, 2003) est un modèle global au pas de temps journalier à quatre paramètres libres et six paramètres fixés. Son modèle conceptuel est schématisé à la figure 2.1 et la liste ainsi que la description de ses paramètres sont présentées à la figure 2.3. Oudin et al. (2008) observent que l'interdépendance entre les paramètres libres du modèle hydrologique GR4J est plus faible que celle qui existe entre les paramètres libres du modèle hydrologique TOPMO (Michel, Perrin et Andréassian, 2003), inspiré du modèle hydrologique TOPMODEL (Beven et Kirkby, 1979). Cette interdépendance entre les paramètres libres du modèle hydrologique GR4J est cependant démontrée par Andréassian, Perrin et Michel (2004). L'introduction de différents scénarios d'évapotranspiration dans le modèle permet de constater que le modèle s'adapte aux variations en compensant les gains ou les pertes d'évapotranspiration à l'aide des paramètres

$X_2$  et  $X_3$ , tout en gardant une même performance de prédiction du débit. Oudin et al. (2006a) ont étudié, sur 221 bassins versants français, la dépendance en régression linéaire entre les paramètres libres du modèle hydrologique GR4J et la couverture végétale du bassin versant pour conclure qu'il n'existait pas de relation entre les deux. Les performances du modèle hydrologique GR4J se révèlent similaires à celles du modèle hydrologique TOPMO dans une étude qui applique les méthodes de régionalisation des paramètres à 913 bassins versants français, (Oudin et al., 2008). Le modèle hydrologique GR4J ne simulant pas les processus relatifs à l'accumulation et à la fonte de la neige, il est nécessaire de lui coupler un module de neige pour modéliser des bassins versants nordiques.

Le module de neige CemaNeige (Valéry, 2010) est un modèle à deux paramètres libres et trois paramètres fixés. Le schéma conceptuel du modèle est présenté à la figure 2.2 et la liste ainsi que la description de ses paramètres sont présentées à la figure 2.3. Le développement empirique du modèle est basé sur les résultats provenant d'un échantillon contenant 219 bassins versants français; 31 bassins versants suisse; 94 bassins versants suédois; et 36 bassins versants québécois. Le module CemaNeige est utilisé pour modéliser le manteau de neige sur deux bassins versants, l'un québécois et l'autre allemand, dans une étude comparant 20 modèles globaux où le modèle hydrologique GR4J couplé au module CemaNeige (GR4JCN) a été trouvé le plus performant (Seiller, Anctil et Perrin, 2011). Dans la présente étude, le module est implanté sans boucle d'altitude puisque cette fonctionnalité n'améliore pas sa performance pour les bassins versants québécois (Valéry, 2010).

Le modèle GR4JCN à 6 paramètres libres est, à notre connaissance, le modèle hydrologique le plus parcimonieux pour la prédiction des débits en régions nordiques.

## **2.2 Variation du nombre de paramètres libres**

Des modèles hydrologiques de complexité variable sont comparés dans la littérature pour évaluer l'impact de la parcimonie du modèle hydrologique en régionalisation des paramètres [Perrin, Michel et Andréassian (2001); Oudin et al. (2008); Zhang et Chiew (2009)].

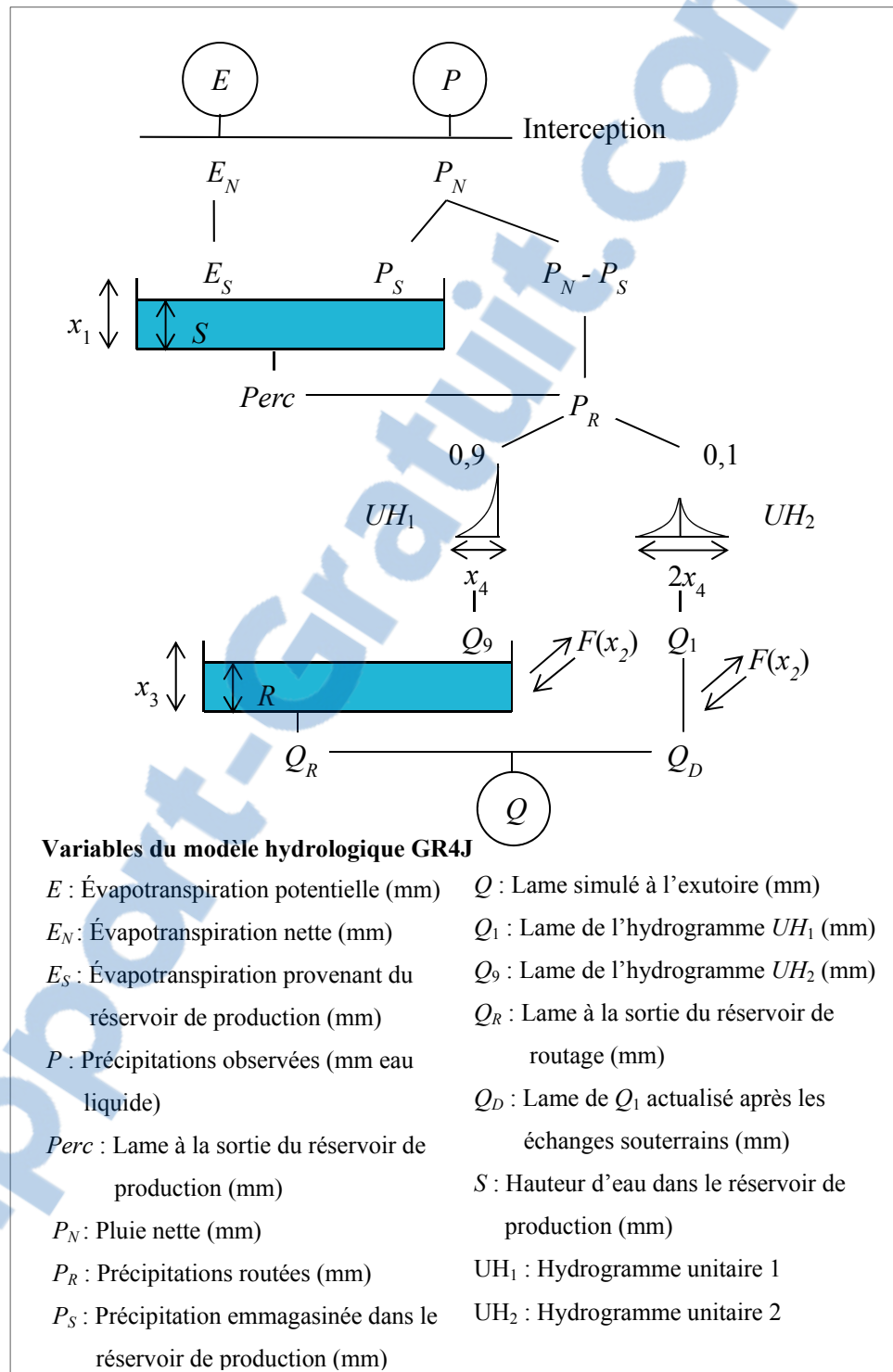


Figure 2.1 Schéma conceptuel du modèle hydrologique GR4J  
Reproduite et adaptée avec l'autorisation de Perrin, Michel et  
Andréassian (2003, p. 277)

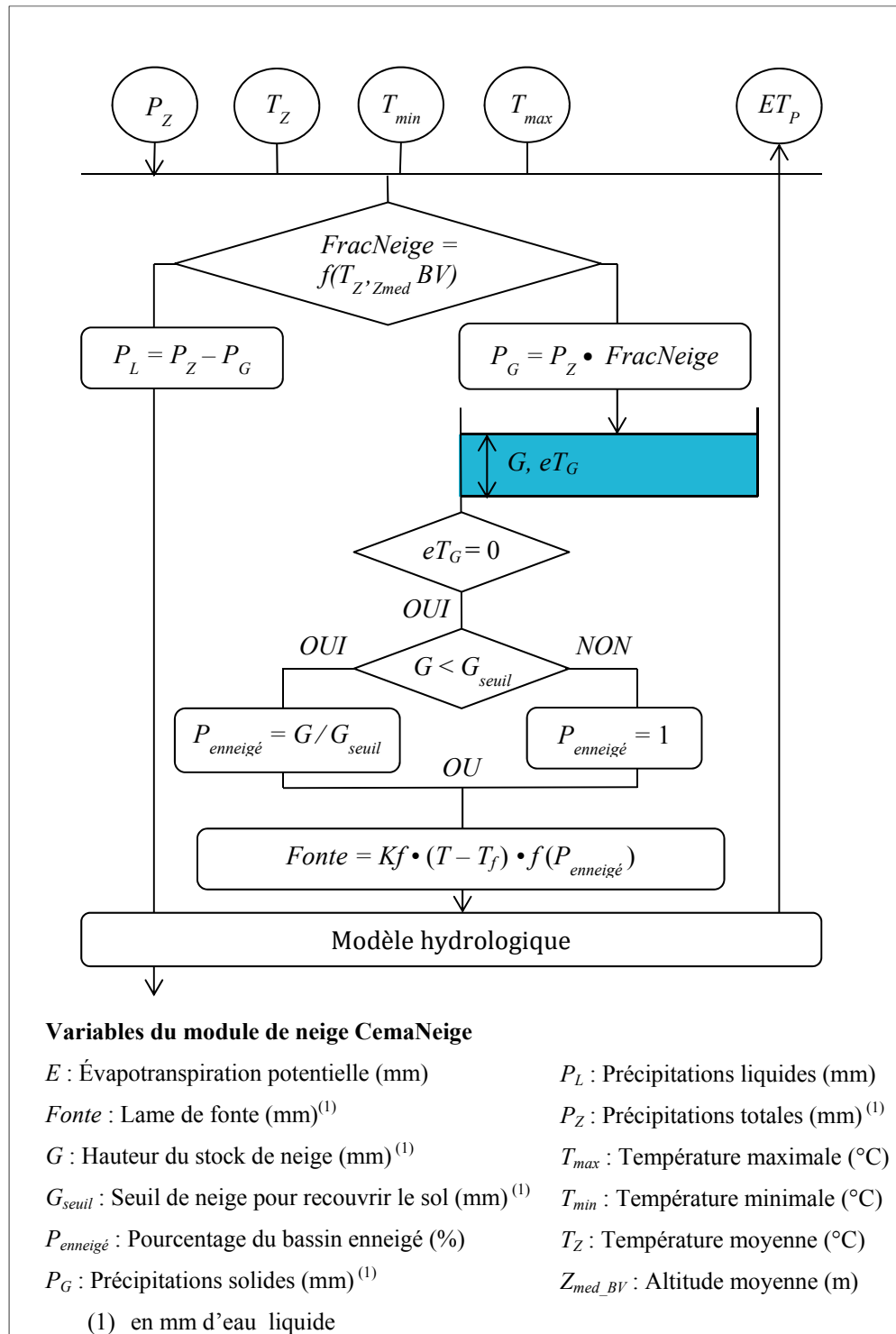


Figure 2.2 Schéma conceptuel du module de neige CemaNeige  
Reproduite et adaptée avec l'autorisation de Valery (2010, p. 191)

<b>Paramètres du modèle hydrologique GR4J</b>	
X1	Niveau maximum de l'eau dans le réservoir de production (mm)
X2	Coefficient d'échange d'eaux souterraines (mm)
X3	Niveau maximum dans le réservoir de routage (mm)
X4	Temps de base de l'hydrogramme unitaire 1 (jour)
Qhyd1	Fraction [0,1] de la quantité de précipitations transférée à l'hydrogramme unitaire 1 (s.d.)
ExpPerc	Exposant de l'équation de percolation (s.d.)
CoeffPerc	Coefficient de l'équation de percolation (s.d.)
ExpQr	Exposant de l'équation du débit à la sortie du réservoir de routage (mm)
ExpFlux	Exposant de l'équation du flux d'eaux souterraines entre les bassins versants (mm)
ExpHU	Exposant des équations SH1 et SH2 (s.d.)
<b>Paramètres du module de neige CemaNeige</b>	
Kf	Facteur de fonte (mm/j /°C)
CTg	Coefficient de pondération de l'état thermique du manteau de neige [0,1] (s.d.)
Tfonte	Température de fonte du manteau de neige (°C)
Vmin	Vitesse minimale de fonte (mm/j)
FractionSeuil	Fraction [0,1] de neige au-dessus de laquelle toute la surface d'un bassin donné est recouverte de neige(s.d.)
Paramètre: <input type="radio"/> Libre <input type="radio"/> Fixé	

Figure 2.3 Liste et description des paramètres libres et fixés du modèle hydrologique GR4J et du module de neige CemaNeige

Ces études comparent cependant des modèles conceptuels différents et leurs résultats ne distinguent pas la variation de performance liée à la complexité du modèle hydrologique, de celle qui découle de l'erreur liée au modèle conceptuel. La présente étude propose d'adapter trois versions du modèle conceptuel GR4JCN, possédant respectivement 6, 9 et 15 paramètres libres. Ces trois versions contiennent chacune les mêmes 6 paramètres libres de base. La version à 9 paramètres libère les 3 paramètres fixes qui améliorent le plus la performance du modèle. Une étude préliminaire – non montrée ici – indique que le gain en performance devient négligeable lorsque l'on augmente la complexité du modèle hydrologique au-delà de 9 paramètres libres. Pour obtenir la combinaison optimale à 9 paramètres libres, les 84 combinaisons possibles, de 3 paramètres parmi les 9 fixées disponibles, sont ajoutées à tour de rôle aux 6 paramètres libres de base puis le jeu résultant est utilisé pour exécuter le modèle du bassin versant. La performance des 84 jeux de paramètres est évaluée et le meilleur jeu de 9 paramètres est retenu. La version à 15 paramètres libres propose une version du GR4JCN où tous les paramètres fixes sont libérés pour tester l'extrême de l'interdépendance possible entre les paramètres du GR4JCN et quantifier son impact.

### **2.3 Région étudiée et base de données géospatiale**

La région étudiée comprend 266 bassins versants qui sont répartis sur l'ensemble du territoire de la province de Québec. La partie amont de quelques-uns de ces bassins versants se trouve dans l'est de la province de l'Ontario. D'une superficie totale de 1 788 950 km<sup>2</sup>, la région est traversée par trois zones de climat : continental humide, subarctique et arctique. L'échantillon des bassins versants étudiés est cartographié à la figure 2.4. Certains bassins versants de l'échantillon sont composés de deux ou plusieurs bassins versants et sont identifiés en conséquence.

Les caractéristiques physiques de chaque bassin versant sont établies à partir de la superficie, de l'occupation du sol, des coordonnées géographiques et de la hauteur moyenne annuelle des précipitations. Les informations géospatiales sont issues de la base de données GéoBase

du Conseil canadien de la géomatique (COCG)<sup>1</sup>. Les 12 caractéristiques physiques retenues sont répertoriées à la figure 2.5 et permettent d’apprécier la forte hétérogénéité des bassins versants étudiés.

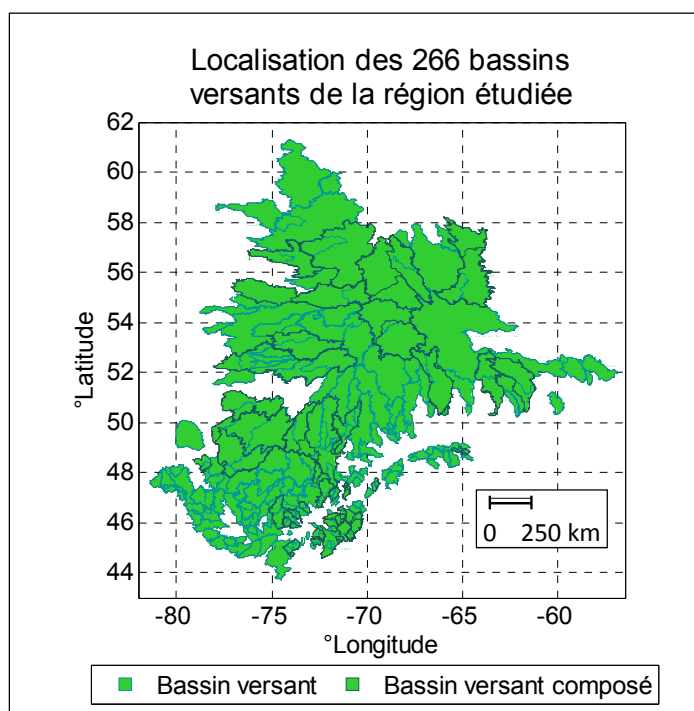


Figure 2.4 Localisation des 266 bassins versants de la région étudiée

## 2.4 Bases de données hydrologiques et météorologiques

Les séries chronologiques de débit journalier proviennent de la base de données  $(cQ)^2$  qui contient les observations de 306 stations de jaugeage réparties sur le territoire québécois. Cette base de données provient de la mise en commun des données de jaugeages de sociétés publiques et privées. Les données météorologiques journalières – températures minimales et maximales, précipitations totales – sont issues de la base de données Canadian National Land and Water Information Service (CNLWIS).

<sup>1</sup> Contient de l'information visée par la Licence du gouvernement ouvert – Canada.

Les séries d'observation des bases de données  $(cQ)^2$  et CNLWIS coïncident temporellement pour 268 bassins versants. Pour deux bassins versants, le modèle hydrologique optimisé ne démontre aucune relation entre les observations météorologiques et hydrologiques en affichant un critère d'efficacité  $NSE_{moyen}$  négatif. Ces bassins versants sont supprimés de l'échantillon. L'échantillon final retenu regroupe donc 266 séries d'observations possédant de 2 à 30 années d'observations réparties sur la période 1961 à 1990.

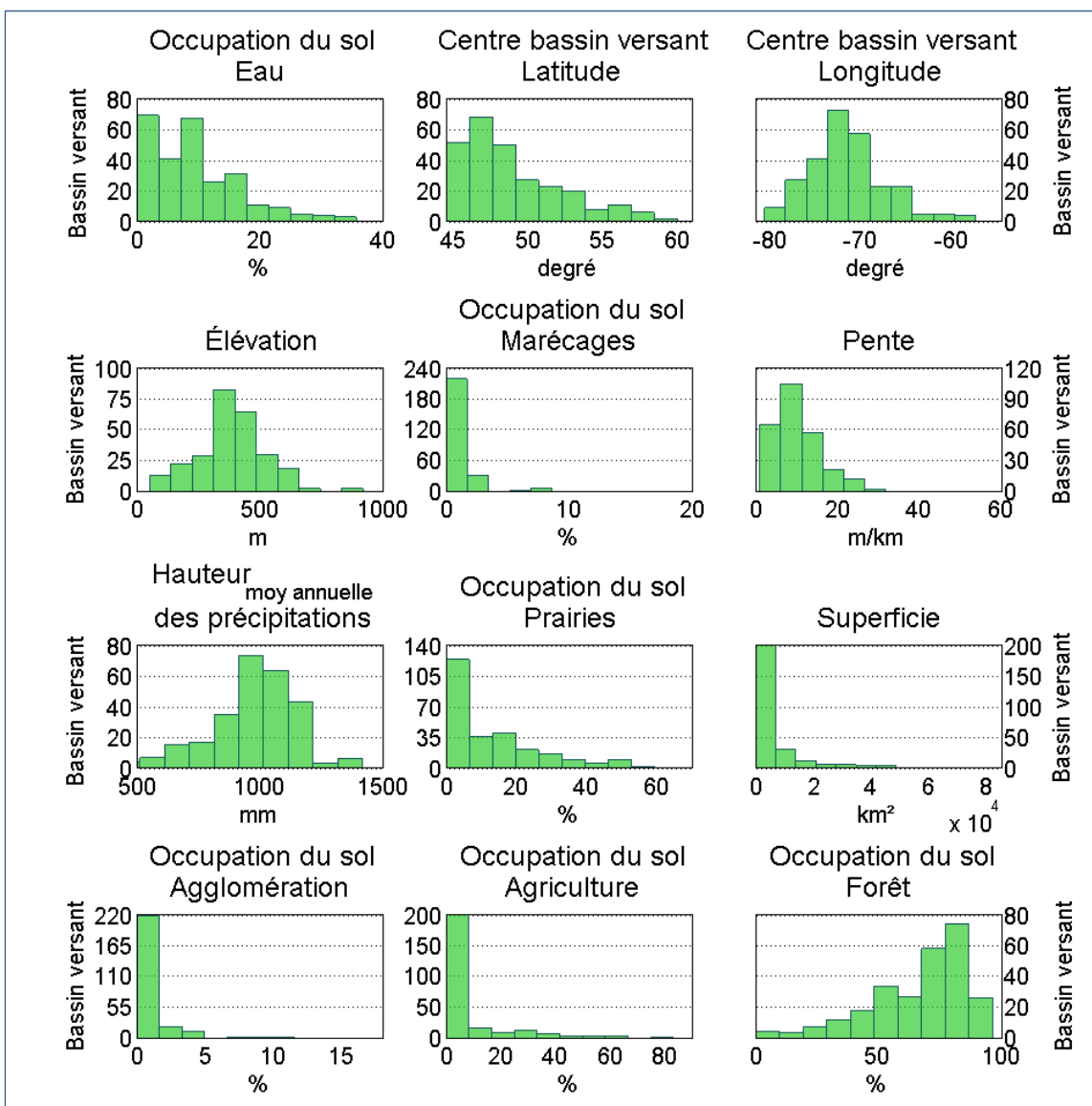


Figure 2.5 Caractéristiques physiques des 266 bassins versants de la région étudiée



## **2.5 Calibration des modèles hydrologiques**

Trois options de calibration des modèles hydrologiques sont testées pour évaluer l'impact de la distribution – au sein de l'espace paramétrique du modèle hydrologique – des jeux de paramètres optimisés sur l'efficacité des méthodes de régionalisation. L'équifinalité de chaque option de calibration est explorée en générant, pour chacune des options, 10 jeux de paramètres pour chacun des 266 bassins versants. La durée totale des séries chronologiques est utilisée pour calibrer les modèles hydrologiques. La première année des débits simulés est omise du calcul du critère de performance (*Voir* section 2.7.1).

### **2.5.1 Algorithmes d'optimisation**

Les deux premières options de calibration utilisent respectivement les algorithmes d'optimisation CMAES (Hansen et Ostermeier, 1996) et SCE-UA (Duan, Sorooshian et Gupta, 1992). Une étude préliminaire – non montrée ici – portant sur la répétabilité des jeux de paramètres générés par les deux algorithmes a permis de constater que, pour un même bassin versant et les mêmes conditions de calibration, le premier génère plus fréquemment un jeu de paramètres similaire, alors que le second – de performance équivalente – génère des jeux de paramètres plus variés.

### **2.5.2 Échantillonnage aléatoire**

La troisième option de calibration vise à tester l'extrême de l'étendue de la distribution paramétrique pouvant être observée tout en maintenant une performance équivalente à celle obtenue par un algorithme d'optimisation. Un échantillonnage aléatoire de 500 000 jeux de paramètres est effectué puis le modèle hydrologique est exécuté avec chacun des jeux de paramètres générés. Le classement subséquent des jeux de paramètres en fonction de leur performance permet de sélectionner les 10 meilleurs d'entre eux. Ceux qui sont retenus offrent une performance en calibration similaire à celle obtenue par les algorithmes d'optimisation. L'échantillonnage aléatoire de l'espace paramétrique se montre plus coûteux en termes d'évaluations du modèle hydrologique qu'un algorithme d'optimisation. La

parcimonie du modèle hydrologique GR4JCN ainsi que sa rapidité d'exécution permettent ici l'utilisation de cette approche, qui serait inapplicable pour un modèle hydrologique plus complexe en termes de paramètres libres.

## **2.6 Configurations testées**

Pour répondre aux objectifs de l'étude, cinq configurations couplant les modèles hydrologiques (*Voir* section 2.2) aux méthodes de calibration (*Voir* section 2.5) sont testées. Trois configurations modèle hydrologique/option de calibration (MH/OC) testées utilisent le modèle hydrologique à 6 paramètres libres optimisé selon les trois options, soit; SCEUA-6, ALÉA-6 et CMAES-6. Les deux dernières configurations MH/OC testées couplent l'algorithme d'optimisation SCE-UA aux modèles hydrologiques à 9 et 15 paramètres libres, soit ; SCEUA-9 et SCEUA-15.

## **2.7 Performance et robustesse des modèles hydrologiques**

Le fait de comparer des modèles hydrologiques de diverses complexités, deux algorithmes d'optimisation et une méthode d'échantillonnage aléatoire de l'espace paramétrique résulte en plusieurs hydrogrammes distincts pour un même bassin versant optimisé avec les mêmes séries chronologiques. Une aptitude distincte à simuler les débits, selon les modèles hydrologiques optimisés, est conséquemment attendue. Le degré de calibration du modèle hydrologique d'un bassin versant dépend de l'erreur liée aux observations – météorologiques ou hydrologiques – mais aussi de l'aptitude du modèle hydrologique conceptuel à décrire les processus hydrologiques spécifiques au bassin versant (Kuczera et al., 2010). Puisque la robustesse consiste à mettre en relation la performance de la prédiction en validation d'un modèle hydrologique par rapport à sa performance obtenue en calibration, ce critère d'évaluation est relatif. Ainsi, la robustesse accrue des modèles hydrologiques parcimonieux (*Voir* section 3.5) n'implique pas qu'ils affichent une meilleure performance en validation par rapport aux modèles hydrologiques plus complexes (*Voir* section 3.9). Cette robustesse accrue des modèles hydrologiques parcimonieux est plutôt attribuable à leur degré de calibration qui est en-deçà de celui des modèles hydrologiques plus complexes (*Voir* section

3.2). Il s'avère donc important de distinguer la performance en validation du modèle hydrologique – quantifiée par le critère d'efficacité de Nash-Sutcliffe –, de sa robustesse – quantifiée par le taux de succès – pour comparer les résultats de l'étude.

### 2.7.1 Critère d'efficacité de Nash-Sutcliffe

La performance des modèles hydrologiques est définie par le critère d'efficacité de Nash-Sutcliffe (NSE) (Nash et Sutcliffe, 1970) exprimé dans sa forme usuelle;

$$NSE = 1 - \frac{\sum(Q_{obs} - Q_{sim})^2}{\sum(Q_{obs} - \bar{Q}_{obs})^2} \quad (2.1)$$

où,  $Q_{obs}$  est le débit observé,  $\bar{Q}_{obs}$  est le débit observé moyen et  $Q_{sim}$  est le débit simulé.

Une limitation lui est connue; les modèles hydrologiques optimisés à l'aide du critère d'efficacité NSE tendent à sous-estimer les crues (Gupta et al., 2009). Ces auteurs décomposent le critère en trois composantes; soit la corrélation, le biais et une mesure de la variabilité, et observent qu'il est nécessaire de sous-estimer la composante de la variabilité pour maximiser la fonction du critère d'efficacité NSE. Les auteurs mentionnent que comparer différents bassins versants sur la base de leur critère d'efficacité NSE peut se révéler ardu puisque la composante du biais est proportionnelle à l'écart-type estimé pour les observations de chaque bassin versant. Malgré ces limitations, le critère d'efficacité NSE est probablement la métrique la plus utilisée en modélisation hydrologique. Cette large utilisation lui confère un avantage certain; les résultats sont aisément comparables avec la littérature.

### 2.7.2 Taux de succès

La robustesse en validation des modèles hydrologiques est mesurée par le taux de succès (TS) (Arsenault et Brissette, 2014). Le taux de succès offre l'avantage de permettre une comparaison visuelle relativement aisée de la distribution d'un nombre élevé de

configurations MH/OC, testées sur un échantillon de grande dimension. Le taux de succès s'exprime par rapport au critère d'optimisation, offrant ainsi une échelle de mesure *relative* du critère de validation. Arsenault et Brissette (2014) proposent un taux de succès défini selon le rapport entre le nombre de prédictions acceptables et le nombre de prédictions totales. Le seuil pour déterminer le succès d'une prédiction est fixé à 85% de la performance moyenne des 10 jeux de paramètres optimisés du modèle hydrologique du bassin versant évalué. Ce seuil de 85% signifie, que pour un bassin versant dont le modèle hydrologique est optimisé avec un critère d'efficacité NSE égal à 0,70, le succès est attribué si la validation du modèle hydrologique présente un critère d'efficacité NSE égal ou supérieur à 0,60.

## **2.8 Équifinalité et méthode de ré-échantillonnage par *bootstrap***

La méthode de ré-échantillonnage dite par *bootstrap* (Efron, 1979), consiste à effectuer une succession de tirages aléatoires avec remise dans un échantillon, de façon à déterminer la variation de prédiction du modèle qui résulte de la dispersion de cet échantillon. Arsenault et Brissette (2014) adoptent cette méthode statistique pour explorer l'incertitude liée au choix du jeu de paramètres transféré à un bassin versant non jaugé dans un contexte d'équifinalité et fixent le nombre de tirages à 1000. Ainsi, l'échantillon de bassins versants est évalué 1 000 fois pour chacune des méthodes de régionalisation appliquée à chaque configuration MH/OC testée. Pour chacune des évaluations, les jeux de paramètres de chacun des bassins versants donneurs sont tirés aléatoirement parmi leurs 10 jeux de paramètres optimisés respectifs.

## **2.9 Bassins versants pseudos non jaugés et validation du modèle hydrologique**

Les méthodes de régionalisation ont été évaluées par une validation croisée (*leave-one-out cross validation*). Ainsi, à tour de rôle, chaque bassin versant est considéré comme étant non jaugé, les méthodes de régionalisation sont appliquées puis les paramètres sont transférés. La validation du transfert est effectuée en comparant le débit simulé par le modèle hydrologique au débit observé à l'exutoire du bassin versant pseudo non jaugé. La durée totale des séries chronologiques est utilisée pour valider les modèles hydrologiques régionalisés. La première année des débits simulés est omise du calcul du critère de performance (*Voir* section 2.7.1).

## 2.10 Échantillons évalués et sélection des bassins versants donateurs

Deux échantillons de bassins versants sont tirés de la base hydrologique ( $cQ$ )<sup>2</sup>. Le premier, l'échantillon *global*, contient l'ensemble des 266 bassins versants étudiés. Le second, l'échantillon *donneur*, contient l'ensemble des bassins versants sélectionnés pour transférer les jeux de paramètres vers les bassins versant pseudo non jaugés. Pour être inclus dans l'échantillon *donneur*, les bassins versants doivent présenter un critère d'efficacité  $NSE_{\text{moyen}} \geq 0,70$  pour leurs 10 jeux de paramètres optimisés respectifs. Oudin et al. (2008) proposent cette approche, reprise par Arsenault et Brissette (2014), pour répondre au dilemme à savoir si les bassins versants ayant obtenu une mauvaise performance en calibration doivent être conservés puisqu'ils apportent une diversité potentiellement utile pour caractériser le bassin versant pseudo non jaugé ou s'ils doivent, au contraire, être exclus sur la base de leur faible performance en calibration. Il est aussi important de souligner qu'étant donné la tendance observée sur les réseaux de collecte d'observations hydrologiques à l'échelle mondiale, la probabilité d'avoir à sa disposition des séries chronologiques de bonne qualité lorsque que viendra le temps d'appliquer ces méthodes de régionalisation à des études concrètes – et non académiques – est relativement faible. Tester la robustesse des méthodes face à la variation de qualité des données apparaît nécessaire.

Malgré que la performance observée pour les trois options de calibration de la version du modèle hydrologique GR4JCN à 6 paramètres libres soit d'un même ordre de grandeur (*Voir* figure 3.1), une différence subsiste et fait fluctuer le nombre de bassins versants qui se qualifient par rapport au seuil du critère d'efficacité  $NSE_{\text{moyen}} \geq 0,70$ . De même, ce seuil appliqué aux performances plus élevées en calibration des modèles hydrologiques GR4JCN à 9 et 15 paramètres libres (*Voir* figure 3.1) résulte en un nombre supérieur de bassins versants qui sont acceptés dans l'échantillon. De façon à standardiser le nombre et l'identité des bassins versants donateurs utilisés dans l'étude, 167 bassins versants sont sélectionnés pour former l'échantillon *donneur*. Ils répondent au critère de posséder un critère d'efficacité  $NSE_{\text{moyen}} \geq 0,70$  pour les trois options de calibration du modèle hydrologique GR4JCN à 6 paramètres libres ainsi que pour les versions à 9 et 15 paramètres libres du même modèle

hydrologique. Cette standardisation de l'échantillon *donneur* a pour but d'affranchir les résultats de l'étude, de la variation potentielle en qualité des bassins versants donneurs. La limitation de cette approche est discutée à la section 4.2.

## **2.11 Régionalisation par la méthode de la régression linéaire multiple**

La méthode de régionalisation par régression linéaire multiple implique que chaque paramètre libre soit transféré individuellement au bassin versant pseudo non jaugé. Un modèle régional linéaire est donc construit pour chaque transfert dans le but d'expliquer le paramètre libre par les caractéristiques physiques présentées à la figure 2.5, et ce pour l'ensemble des bassins versants donneurs. Suivant l'approche par *bootstrap*, 1000 modèles régionaux linéaires sont construits pour chaque paramètre libre du modèle hydrologique de chacun des bassins versants pseudos non jaugés.

### **2.11.1 Modèle régional linéaire**

Le nombre et l'identité des caractéristiques physiques ainsi que le nombre de bassins versants donneurs retenus pour établir les modèles de régression est déterminé sur la base des résultats obtenus lors d'une étude préliminaire – non montrée ici. Cette étude préliminaire compare les résultats des modèles régionaux linéaires construits en combinant de 1 à 17 groupes de bassins versants donneurs au nombre de caractéristiques physiques transférées. Les 17 groupes de bassins versants donneurs contiennent de 20 à 180 bassins versants – chacun des groupes possédant 10 bassins versants de plus que le groupe précédent. Le nombre de caractéristiques physiques transférées varie de 4 à 12. Puisque la méthode par similitude physique offre une option de classement crédible du degré d'homogénéité (*Voir* sections 3.5 et 3.9), les bassins versants potentiellement donneurs sont classés selon l'index de similitude (*Voir* section 1.4). À titre d'exemple, le premier groupe de bassins versants donneurs est ainsi constitué des 20 premiers bassins de l'index de similitude et le second groupe est constitué des 30 premiers bassins du même index. Les caractéristiques physiques sont ajoutées, une à la fois, selon l'ordre trouvé lors de la première étape de la méthode par similitude physique (*Voir* section 2.13).

Certains modèles régionaux linéaires prédisent des valeurs de paramètres du modèle hydrologique inférieures aux bornes de calibration, conduisant à des aberrations dans le calcul du modèle hydrologique telle la division par zéro ou la manipulation de nombres complexes. Pour remédier à cette situation, toute valeur de paramètre du modèle hydrologique trouvée inférieure à sa borne de calibration est posée égale à cette dernière. Les valeurs de paramètres trouvées supérieures à la borne de calibration ne sont pas modifiées.

### 2.11.2 Évaluation statistique du modèle régional linéaire

Trois critères statistiques sont utilisés pour établir la validité des modèles régionaux linéaires construits dans le cadre de l'étude.

Le premier critère statistique, le coefficient de détermination  $R^2$ , vise à quantifier la corrélation entre les variables explicatives – les caractéristiques physiques des bassins versants – et les variables expliquées – les jeux de paramètres transférés aux modèles hydrologiques – du modèle régional linéaire. Le coefficient de détermination  $R^2$  représente donc la proportion de variance expliquée par le modèle régional linéaire. La qualité de la corrélation entre les variables explicatives et les variables expliquées du modèle régional linéaire est d'autant meilleure que la valeur du coefficient  $R^2$  tend vers l'unité et le coefficient  $R^2 \in [0, 1]$ .

Il existe une limitation au nombre de variables explicatives – utilisées pour construire le modèle régional linéaire – qui affecte le coefficient de corrélation  $R^2$ . L'ajout de variables explicatives a pour conséquence de diminuer la somme des moindres carrés des résidus et donc, d'améliorer le coefficient  $R^2$ . Cette augmentation mécanique du coefficient  $R^2$  est toutefois indépendante de la qualité de la corrélation.

Le second critère statistique, la statistique  $F$ , estime la significativité de la relation entre les paramètres du modèle régional linéaire construit et ses variables explicatives. La statistique  $F$  repose sur le test de Fisher-Snédecour, l'outil décisionnel utilisé pour répondre au test

d'hypothèse. Le test d'hypothèse postule la nullité de tous les paramètres du modèle régional linéaire, et donc l'absence d'interaction entre les paramètres du modèle régional linéaire et ses variables explicatives. Le test de Fisher-Snédecour décide du rejet ou du non-rejet de cette hypothèse nulle. Le fait de ne pas rejeter l'hypothèse suppose qu'on ne peut exclure que tous les paramètres du modèle régional linéaire soient nuls. Le fait de rejeter l'hypothèse nulle implique d'accepter l'hypothèse alternative du test d'hypothèse : il existe au moins un paramètre du modèle régional linéaire qui n'est pas nul. La statistique  $F$ , lorsque calculée pour tester la nullité des paramètres du modèle régional linéaire, suit une loi de Fisher  $F(p, n-p)$ , où  $p$  est le nombre de variables explicatives et  $n$  la taille de l'échantillon. Le seuil de la région critique conduisant au rejet de l'hypothèse nulle est défini par le quantile d'ordre 1 du test de Fisher;  $F_{1-\alpha}(p, n-p)$ , le risque  $\alpha$  étant posé égal à 0,05. L'hypothèse nulle est rejetée lorsque la statistique  $F \geq F_{1-\alpha}(p, n-p)$  et la force de son rejet est d'autant plus grande que la statistique  $F$  est élevée.

Le dernier critère statistique est le seuil descriptif  $P$ -value qui représente la probabilité, en supposant que l'hypothèse nulle est vraie, d'obtenir une valeur au moins aussi grande que la valeur obtenue pour la statistique du test. Pour rejeter l'hypothèse de nulle, le risque exprimé par le seuil descriptif  $P$ -value doit être inférieur au risque posé  $\alpha$  – égal à 0,05. Le niveau de crédibilité de l'hypothèse nulle augmente avec une augmentation du seuil descriptif  $P$ -value et le seuil descriptif  $P$ -value  $\in [0, 1]$ .

## 2.12 Régionalisation par la méthode de la proximité spatiale

Pour la méthode de régionalisation par proximité spatiale, les bassins versants potentiellement donneurs sont classés en fonction de la proximité entre leurs centroïdes respectifs et celui du bassin versant pseudo non jaugé. Dans le but d'identifier le nombre optimal de bassins versants donneurs, la méthode est appliquée en utilisant de 1 à 15 bassins versants donneurs. L'approche par *bootstrap* consiste ici à réaliser 1000 transferts de paramètres pour chaque configuration du nombre de bassins versants donneurs et pour



chaque bassin versant pseudo non jaugé. L'interpolation sur la moyenne des paramètres, la moyenne des débits et la pondération inverse à la distance des débits sont comparées.

### 2.13 Régionalisation par la méthode de similitude physique

La méthode de régionalisation par similitude physique s'implante en deux étapes. La première étape consiste à sélectionner la combinaison de caractéristiques physiques la plus apte à caractériser le comportement hydrologique des bassins versants de l'échantillon. Pour y parvenir, l'étude prend pour point de départ les conclusions de Arsenault et Brissette (2014) qui utilisent la base de données  $(cQ)^2$  et le modèle hydrologique global à 23 paramètres libres HSAMI (Fortin, 2000). Les auteurs concluent qu'un nombre supérieur à 5 bassins versants n'améliore pas statistiquement la performance de la méthode de régionalisation et que l'interpolation des débits par la PID se montre plus performante que l'interpolation par la moyenne des paramètres ou par la moyenne des débits. Dans le but de déterminer la meilleure combinaison de caractéristiques physiques pour la présente étude, la méthode de régionalisation par similitude physique est donc implémentée en fixant le nombre de bassins versants donneurs à 5 et en interpolant les débits par la PID. Chacune des caractéristiques physiques présentées à la figure 2.5 est utilisée à tour de rôle pour déterminer la distance en similitude entre le bassin versant pseudo non jaugé et le reste des bassins versants de l'échantillon *global*, qui sont tous considérés donneurs. Ensuite les jeux de paramètres des cinq bassins versants donneurs les plus près en similitude sont transférés au bassin versant non jaugé. La PID des débits est calculée puis comparée au débit observé. À tour de rôle, chacun des 266 bassins versants est considéré non jaugé et la même procédure est répétée. La caractéristique ayant obtenu la meilleure performance en termes de taux de succès pour l'ensemble des 266 bassins versants est retenue. Ensuite, à tour de rôle, chacune des caractéristiques non-retenues est combinée à la caractéristique retenue de façon à déterminer, selon la même procédure, la deuxième meilleure caractéristique. La méthode est répétée pour la totalité des caractéristiques physiques. La combinaison de caractéristiques physiques qui présente la meilleure performance est retenue pour effectuer le transfert des paramètres lors de la deuxième étape.

La deuxième étape consiste à classer les bassins versants donneurs en fonction de leur proximité en similitude avec le bassin versant pseudo non jaugé selon la combinaison de caractéristiques physiques retenue. Le nombre optimal de bassins versants donneurs est étudié en appliquant la méthode avec un nombre de bassins versants donneurs allant de 1 à 15. L'approche par *bootstrap* a ici aussi consisté à réaliser 1000 transferts de paramètres pour chaque configuration du nombre de bassins versants donneurs et pour chaque bassin versant pseudo non jaugé. L'interpolation sur la moyenne des paramètres, la moyenne des débits et la pondération inverse à la distance des débits sont ici aussi comparées.

## CHAPITRE 3

### RÉSULTATS

Le chapitre des résultats expose dans un premier temps les résultats qui définissent ou évaluent les modèles hydrologiques et les index de similitude utilisés dans le cadre de l'étude. Les résultats qui comparent la complexité des modèles hydrologiques, les méthodes de calibration et les méthodes de régionalisation sont ensuite présentés. Les résultats obtenus par l'analyse statistique du modèle régional linéaire complètent le chapitre.

#### 3.1 Identité des paramètres libres du modèle hydrologique à 9 degrés de liberté

La dualité entre la parcimonie et la complexité *suffisante* du modèle hydrologique (*Voir* section 1.6) est étudiée à l'aide des modèles hydrologiques à 6, 9 et 15 paramètres libres. Les modèles hydrologiques à 6 et 15 paramètres libres proposent respectivement le minimum et le maximum de paramètres libres de GR4JCN. L'identité des paramètres libres du modèle hydrologique à 9 paramètres libres est pour sa part déterminée empiriquement (*Voir* section 2.2).

Rappelons que le modèle hydrologique GR4JCN à 6 paramètres libres comporte 9 paramètres fixés (*Voir* figure 2.3). Parmi ces 9 paramètres fixés, les 3 paramètres sélectionnés pour augmenter la complexité de GR4JCN à 6 paramètres libres et obtenir le modèle hydrologique à 9 paramètres libres sont *CoeffPerc*, *FacteurSeuil* et  $V_{min}$ .

Le paramètre *CoeffPerc* intervient dans l'équation de percolation (*Perc*) qui détermine la lame d'eau à la sortie du réservoir de production du modèle hydrologique GR4J (*Voir* figure 2.1);

$$Perc = S \left\{ 1 - \left[ 1 + \left( CoeffPerc \cdot \frac{S}{X1} \right)^{ExpPerc} \right]^{-1/ExpPerc} \right\} \quad (3.1)$$

où, dans le modèle hydrologique de Perrin, Michel et Andréassian (2003),  $CoeffPerc$  est fixé à 4/9 et l'exposant  $ExpPerc$  est fixé à 4.

Les deux autres paramètres appartiennent au module de neige CemaNeige (Voir figure 2.2). Le paramètre  $FacteurSeuil$  est la fraction du stock de neige ( $G_{50}$ ) à partir de laquelle le bassin versant est entièrement recouvert d'un manteau de neige. Le stock de neige ( $G_{50}$ ) correspond à la moyenne annuelle des précipitations solides. La hauteur seuil de neige pour recouvrir le sol du bassin versant ( $G_{seuil}$ ) est calculée telle que;

$$G_{seuil} = FacteurSeuil \cdot G_{50} \quad (3.2)$$

où le  $FacteurSeuil$  est fixé à 9/10 dans le module de neige de Valéry (2010).

Le paramètre  $V_{min}$  est la vitesse minimale de fonte de la neige, vitesse qui est atteinte lorsque le stock de neige est faible. Dans le module de neige de Valéry (2010), le paramètre  $V_{min}$  est fixé à 1/10 et il intervient dans l'équation de la fonte;

$$Fonte = F_{potentielle} \cdot \left[ (1 - V_{min}) \cdot \text{minimum} \left( \frac{G}{FracSeuil \cdot G_{50}}, 1 \right) + V_{min} \right] \quad (3.3)$$

où, si l'état thermique du manteau de neige ( $eTg$ ) est supérieur à 0 et si la température moyenne observée de l'air ( $T_{moy}$ ) est supérieure à la température de fonte ( $T_{fonte}$ );

$$F_{potentielle} = \text{minimum} [Kf \cdot (T_{moy} - T_{fonte}), G] \quad (3.4)$$

mais, dans le cas où au moins une des deux conditions précédentes n'est pas respectée;

$$F_{potentielle} = 0 \quad (3.5)$$

### 3.2 Performance des cinq configurations MH/OC en calibration

La distribution de la performance moyenne des versions à 6, 9 et 15 paramètres, exprimée par le critère d'efficacité de Nash-Sutcliffe, est présentée à la figure 3.1. Les valeurs exprimées du critère d'efficacité représentent la moyenne des 10 calibrations effectuées pour chacun des bassins versants. La figure 3.1 permet d'observer que la performance en calibration des modèles hydrologiques augmente avec leur complexité. Pour les trois configurations MH/OC du modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6), une densité maximale de bassins versants est recensée dans l'intervalle de  $NSE_{\text{moyen}} \in [0,65; 0,85[$ . Pour les modèles hydrologique à 9 et 15 paramètres libres (SCEUA-9 et SCEUA-15), l'intervalle correspondant est  $NSE_{\text{moyen}} \in [0,75; 0,95[$ . Le modèle hydrologique à 15 paramètres libres (SCEUA-15) compte 27 bassins versants de plus que le modèle hydrologique à 9 paramètres libres dans cet intervalle. Le modèle hydrologique à 9 paramètres libres compte 102 bassins versants de plus que le modèle hydrologique à 6 paramètres libres dans le même intervalle. Parmi les trois configurations MH/OC testées avec le modèle hydrologique à 6 paramètres libres, celle configurée par échantillonnage aléatoire (ALÉA-6) est légèrement moins performante que les deux configurations MH/OC optimisées par un algorithme d'optimisation (SCEUA-6 et CMAES-6), ces dernières affichant une performance similaire.

### 3.3 Caractéristiques physiques retenues pour construire l'index de similitude

La sélection des caractéristiques physiques pour appliquer la méthode de régionalisation par similitude physique est distincte selon les configurations MH/OC (SCEUA-15, SCEUA-9, SCEUA-6, ALÉA-6 et CMAES-6). La figure 3.2 montre la variation des rangs obtenus pour chaque caractéristique physique selon la configuration MH/OC testée. Pour chaque configuration MH/OC, les caractéristiques physiques retenues pour calculer la distance en similitude, entre les bassins versants donneurs et le bassin versant pseudo non jaugé, sont celles qui occupent respectivement les quatre premiers rangs du classement des caractéristiques physiques (*Voir* figure 3.2).

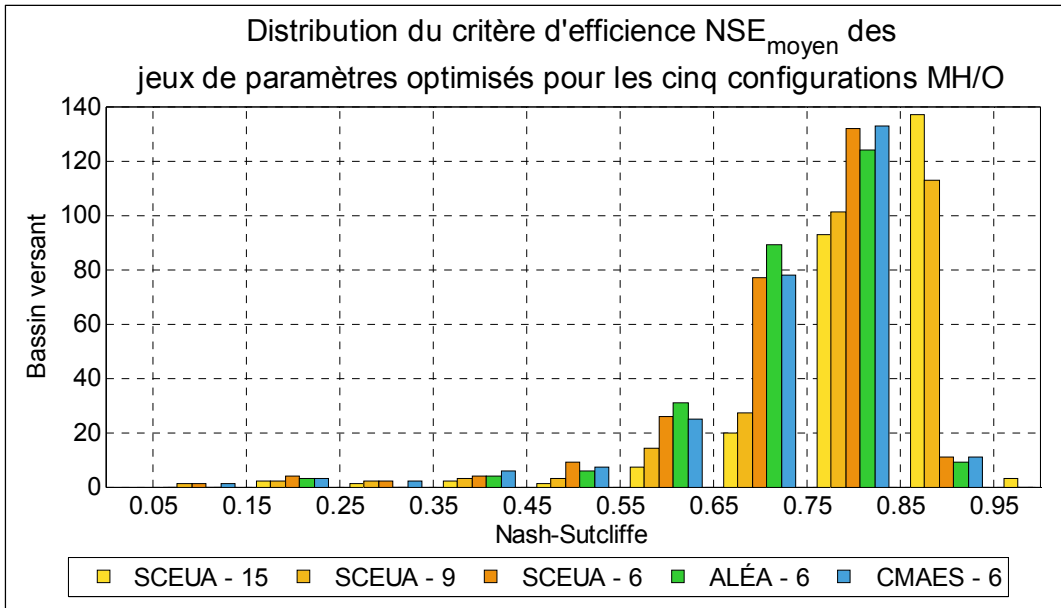


Figure 3.1 Distribution du critère d'efficience  $NSE_{moyen}$  des 10 jeux de paramètres optimisés pour chacun des 266 bassins versants des cinq configurations MH/OC testées

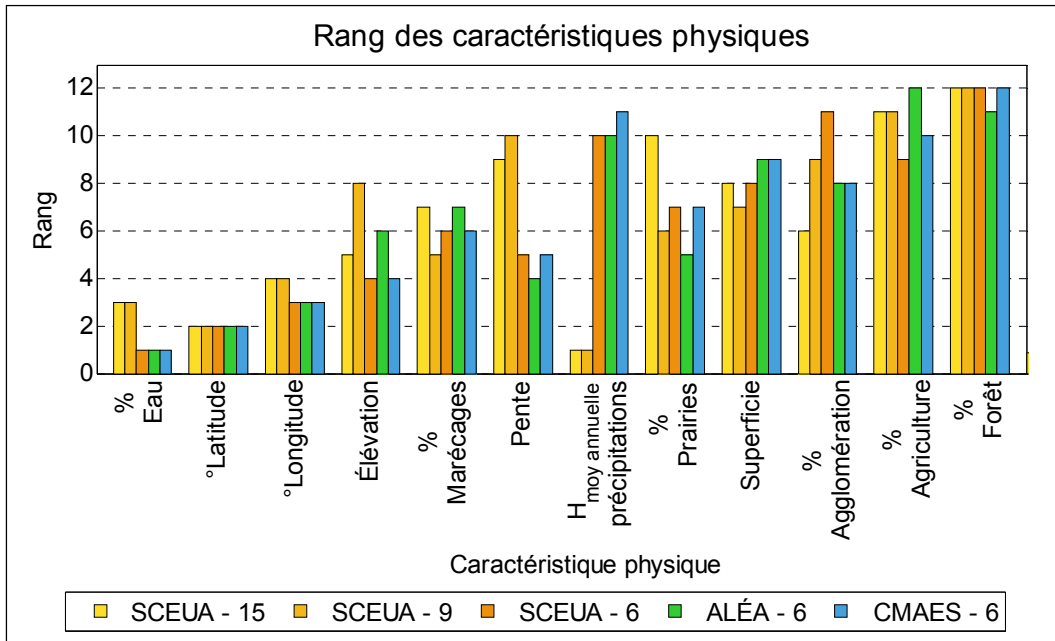


Figure 3.2 Rang de la performance obtenue par les caractéristiques physiques lors de la sélection des caractéristiques utilisées pour calculer l'index de similitude de la méthode de régionalisation par similitude physique

### 3.4 Distribution des jeux de paramètres optimisés au sein de l'espace paramétrique

Les jeux de paramètres optimisés occupent une région distincte de l'espace paramétrique pour chacune des cinq configurations MH/OC testées (SCEUA-15, SCEUA-9, SCEUA-6, ALÉA-6 et CMAES-6). La figure 3.3 montre, pour les configurations MH/OC utilisant le modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6), la distribution des jeux de paramètres au sein de l'espace paramétrique autorisé par les bornes de calibration. La distribution des jeux de paramètres pour les configurations MH/OC utilisant les modèles hydrologiques à 9 et 15 paramètres libres (SCEUA-9 et SCEUA-15) est présentée aux figures –A I–1a et –A I–1b. Les figures 3.3, –A I–1a et –A I–1b présentent l'information de manière identique. Chacun des 10 jeux de paramètres optimisés, pour chacun des 266 bassins versants, est représenté par un point de coordonnées (*bassin versant*, *valeur du paramètre*) et de la couleur attribuée à sa configuration MH/OC.

La méthode d'échantillonnage aléatoire permet d'estimer les régions *possibles* de l'espace paramétrique du modèle hydrologique; celles qui modélisent efficacement les processus hydrologiques des bassins versants modélisés. Conséquemment, elle apporte une estimation de l'amplitude des valeurs possibles des paramètres équifinaux. L'observation de la configuration MH/OC configurée par échantillonnage aléatoire (ALÉA-6) à la figure 3.3 permet de constater que cette amplitude est forte pour les paramètres  $X_2$ ,  $X_3$  et, dans une moindre mesure, pour le paramètre  $CTg$ . En revanche, pour le paramètre  $Kf$  et, dans une moindre mesure, pour les paramètres  $X_1$  et  $X_4$ , les jeux de paramètres occupent une région restreinte de l'espace paramétrique. Les valeurs des paramètres optimisés à l'aide de l'algorithme d'optimisation CMAES (CMAES-6) apparaissent fortement groupées : l'algorithme possède une aptitude pour identifier la même région de l'espace paramétrique à chaque calibration. À l'inverse, l'algorithme d'optimisation SCE-UA (SCEUA-6) présente des valeurs des paramètres optimisés plus dispersées : l'algorithme possède une répétabilité moins élevée et identifie des régions souvent distinctes de l'espace paramétrique. La surface paramétrique des bassins versants no. 110 à 220 de l'échantillon diffère sensiblement de celle des autres bassins versants de l'échantillon, laissant présager un comportement hydrologique

distinct pour ces bassins versants. Les bassins versants no. 110 à 220 correspondent aux bassins versants possédant les latitudes les plus élevées de l'étude (*Voir* figure 4.3). Les conséquences de leur situation géographique sont discutées à la section 4.7. Rappelons que les jeux de paramètres présentés à la figure 3.3 sont équi-finaux et donc, de performance similaire.

### **3.5 Taux de succès obtenu par les méthodes de régionalisation par proximité spatiale et par similitude physique**

Les taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique sont présentés de manière à comparer les méthodes de régionalisation sur la base des trois méthodes d'interpolation utilisées dans l'étude pour transférer les jeux de paramètres; soit, la moyenne des paramètres, la moyenne des débits et la PID des débits. Pour ces deux méthodes de régionalisation, la figure 3.4 compare la méthode d'interpolation par la moyenne des paramètres, la figure 3.5 compare la méthode d'interpolation par la moyenne des débits et la figure 3.6 compare la méthode d'interpolation par la PID des débits. Les figures 3.4 à 3.6 présentent le taux de succès des méthodes de régionalisation couplées aux méthodes d'interpolation en fonction du nombre de bassins versants utilisés pour transférer les jeux de paramètres et ce, pour chacune des cinq configurations MH/OC testées. Chaque taux de succès est exprimé sous la forme d'une boîte à moustache qui schématise la distribution des résultats; le trait central de la boîte identifie la médiane, les arrêtes inférieure et supérieure identifient respectivement les 1<sup>er</sup> et 3<sup>iem</sup> quartiles ( $Q1$  et  $Q3$ ), les moustaches délimitent la région des résultats non-extrêmes et les résultats extrêmes sont identifiés par des points à l'extérieur des moustaches. Les valeurs sont considérées comme étant extrêmes lorsqu'elles sont inférieures à  $Q1-1,5(Q3-Q1)$ , ou supérieures  $Q3+1,5(Q3-Q1)$  soit, respectivement,  $\alpha \approx 0,027$  et  $\alpha \approx 0,993$  en assumant une distribution normale des résultats.



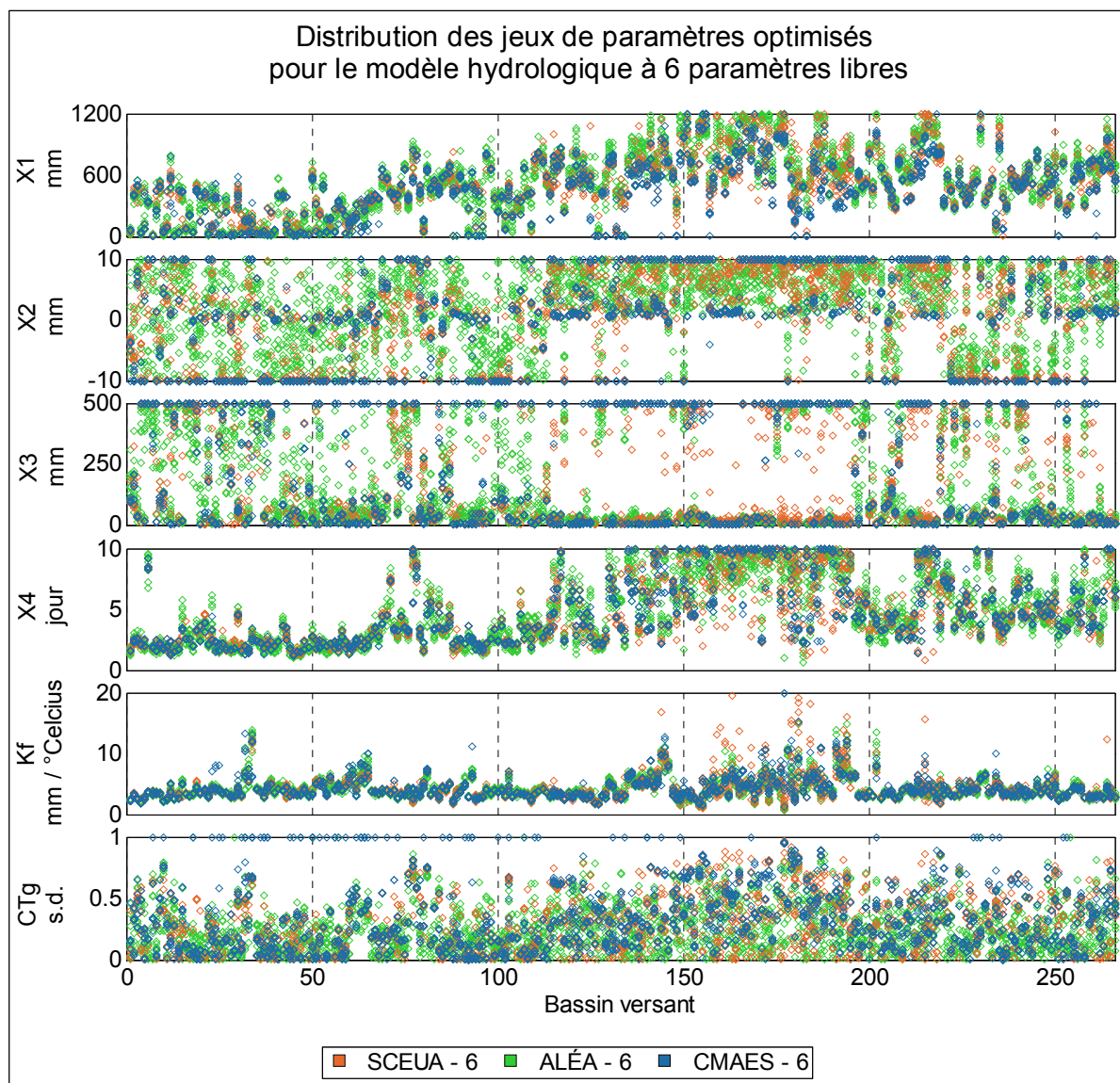


Figure 3.3 Distribution des 10 jeux de paramètres optimisés par bassin versant pour les trois configurations MH/OC du modèle hydrologique à 6 paramètres libres

### 3.5.1 Interpolation des paramètres par la moyenne arithmétique

La méthode d'interpolation des paramètres par la moyenne arithmétique (*Voir* figure 3.4) obtient un taux de succès nettement moins élevé que les deux autres méthodes d'interpolation (*Voir* figures 3.5 et 3.6). Pour cette méthode d'interpolation, la configuration MH/OC utilisant le modèle hydrologique à 9 paramètres libres (SCEUA-9) se distingue des autres

configurations MH/OC par son taux de succès nettement plus élevé. Parmi les trois configurations MH/OC qui utilisent le modèle hydrologique à 6 paramètres libres, celle configurée par échantillonnage aléatoire (ALÉA-6) détient un taux de succès nettement supérieur aux deux autres configurations MH/OC. Parmi ces deux dernières configurations MH/OC, la configuration MH/OC optimisée à l'aide de l'algorithme d'optimisation SCE-UA (SCEUA-6) affiche, globalement, un taux de succès légèrement supérieur à celle optimisée à l'aide de l'algorithme d'optimisation CMAES (CMAES-6). La configuration MH/OC qui utilise le modèle hydrologique à 15 paramètres libres (SCEUA-15) se révèle la moins robuste des configurations MH/OC en proximité spatiale. Cette dernière configuration MH/OC gagne cependant en robustesse lorsque la méthode de régionalisation par similitude physique est appliquée, notamment sur l'échantillon *global*; son taux de succès atteint alors celui de la configuration MH/OC ALÉA-6.

Tant pour la méthode de régionalisation par proximité spatiale que pour celle par similitude physique, le taux de succès obtenu sur l'échantillon *donneur* est plus élevé que celui obtenu sur l'échantillon *global*. Pour ces deux échantillons de bassins versants, la méthode de régionalisation par similitude physique s'avère celle qui obtient le meilleur taux de succès. Cette méthode améliore particulièrement les configurations MH/OC SCEUA-15, SCEUA-6 et CMAES-6.

### **3.5.2 Interpolation des débits par la moyenne arithmétique**

La méthode d'interpolation du débit par la moyenne arithmétique (*Voir figure 3.5*) présente deux tendances distinctes selon la complexité du modèle hydrologique. Les configurations MH/OC qui utilisent le modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6) affichent un taux de succès nettement supérieur à celui des configurations MH/OC qui utilisent les modèles hydrologiques plus complexes (SCEUA-15 et SCEUA-9). Parmi les trois premières configurations MH/OC, la configuration MH/OC ALÉA-6 est celle dont le taux de succès est le plus élevé. Les configurations MH/OC SCEUA-6 et CMAES-6 obtiennent un taux de succès globalement similaire, le taux de succès de la configuration

MH/OC SCEUA-6 étant cependant toujours supérieur ou égal à celui de la configuration MH/OC CMAES-6. La robustesse de la méthode d'interpolation décroît avec l'augmentation de la complexité du modèle hydrologique. La configuration MH/OC SCEUA-9 présente un taux de succès légèrement supérieur à celui de la configuration MH/OC SCEUA-15.

Pour les deux méthodes de régionalisation, l'échantillon *donneur* obtient un meilleur taux de succès que l'échantillon *global*. Pour ces deux échantillons de bassins versants, la méthode de régionalisation par similitude physique présente un taux de succès supérieur à celui de la méthode de régionalisation par proximité spatiale. La première méthode améliore particulièrement la configuration MH/OC ALÉA-6 lorsque les jeux de paramètres sont issus de 11 bassins versants donneurs ou plus.

### **3.5.3 Interpolation des débits par la pondération inverse à la distance**

La méthode d'interpolation du débit par la PID (*Voir* figure 3.6) affiche un taux de succès légèrement supérieur, et plus groupé pour les cinq configurations MH/OC, que celui de la méthode d'interpolation du débit par la moyenne arithmétique (*Voir* figure 3.5). Bien que son taux de succès soit parfois surclassé par celui des configurations MH/OC SCEUA-6 et CMAES-6, lorsque la méthode de régionalisation par similitude physique est appliquée, la configuration MH/OC ALÉA-6 affiche généralement le plus haut taux de succès. La configuration MH/OC SCEUA-9 affiche une tendance distincte selon la méthode de régionalisation appliquée. Lorsque la méthode de régionalisation par similitude physique est appliquée, son taux de succès est légèrement supérieur à celui de la configuration MH/OC SCEUA-15, qui est la configuration MH/OC la moins robuste globalement. Par contre, lorsque la méthode de régionalisation par proximité spatiale est appliquée, son taux de succès, bien qu'inférieur, tend à rejoindre le taux de succès similaire des configurations MH/OC SCEUA-6 et CMAES-6. Pour ces deux derniers, lorsque la méthode de régionalisation par proximité spatiale est appliquée, le taux de succès de la configuration MH/OC SCEUA-6 se montre généralement supérieur ou égal à celui de la configuration MH/OC CMAES-6. À l'inverse, le taux de succès de la configuration MH/OC CMAES-6 se

montre généralement supérieur ou égal à celui de la configuration MH/OC SCEUA-6 lorsque la méthode de régionalisation par similitude physique est appliquée.

Pour les deux méthodes de régionalisation, l'échantillon *donneur* est affecté d'un taux de succès plus élevé que l'échantillon *global*. Pour ces deux échantillons de bassins versants, la méthode de régionalisation par similitude physique se montre plus robuste que celle par proximité spatiale et ce, peu importe le nombre de bassins versants donneurs utilisés pour transférer les jeux de paramètres vers le bassin versant pseudo non jaugé.

### **3.6 Nombre de caractéristiques physiques sélectionnées pour construire les modèles régionaux linéaires**

L'analyse – non montrée ici – du taux de succès des modèles régionaux linéaires en fonction du nombre de caractéristiques physiques, avec lesquels ils sont construits, ne permet pas de discerner une tendance globale nette. Le taux de succès augmente légèrement lorsque le nombre de caractéristiques physiques du modèle régional linéaire augmente de 4 à 10 caractéristiques. Il redescend cependant lorsque 11 ou 12 caractéristiques physiques sont utilisées pour construire le modèle régional linéaire. Le taux de succès est similaire pour les modèles régionaux linéaires construits avec respectivement 4 et 12 caractéristiques physiques. Pour cette raison et puisque la parcimonie de variables explicatives du modèle régional linéaire est importante sur la signification du coefficient de détermination  $R^2$ , le nombre de caractéristiques physiques utilisées pour construire les modèles régionaux linéaires dans la présente étude est de quatre caractéristiques. La sélection des quatre caractéristiques physiques pour chacune des configurations MH/OC s'effectue similairement à la sélection faite dans le cadre de la méthode de régionalisation par similitude physique (*Voir section 3.3*).

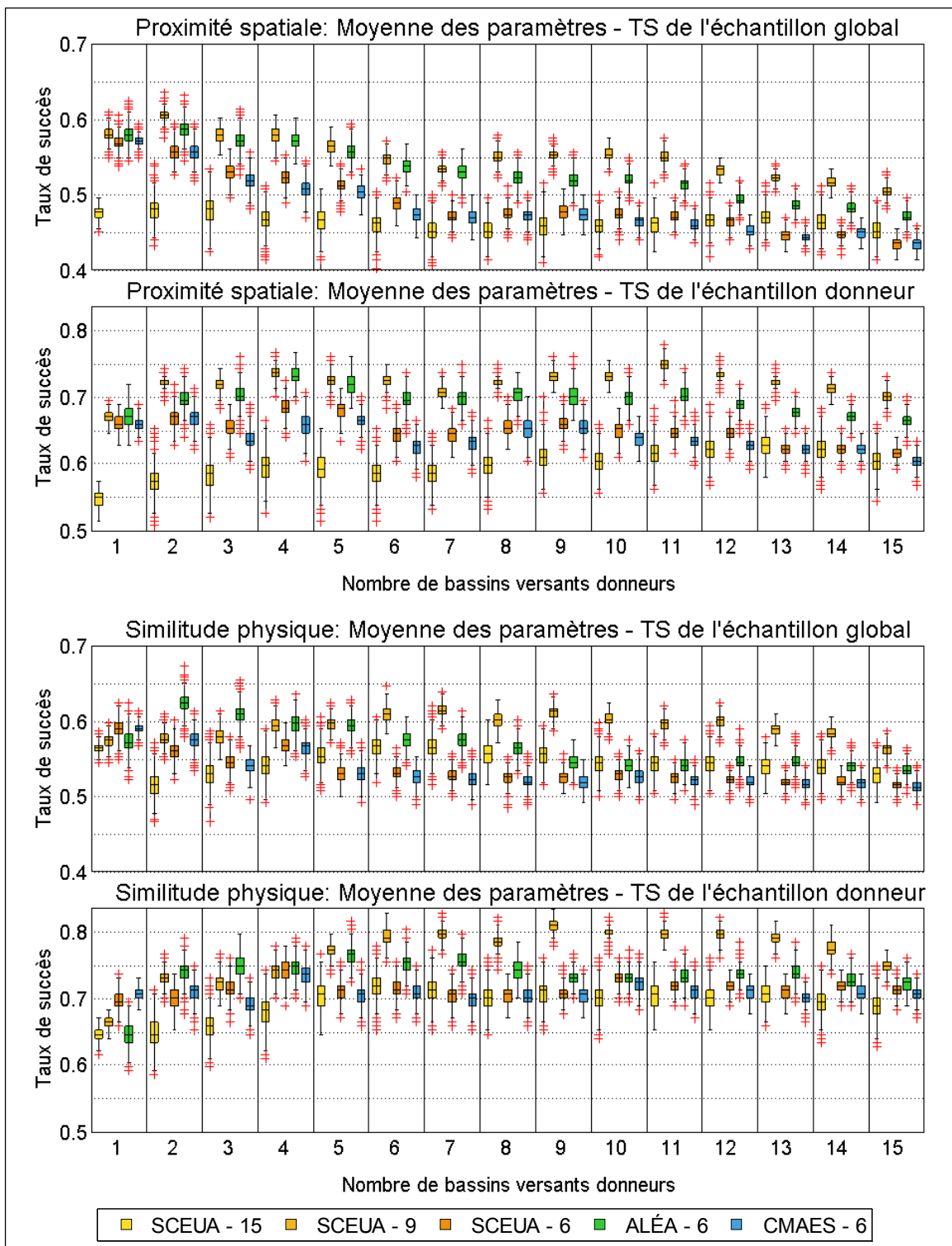


Figure 3.4 Taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique – Interpolation des paramètres par la moyenne arithmétique

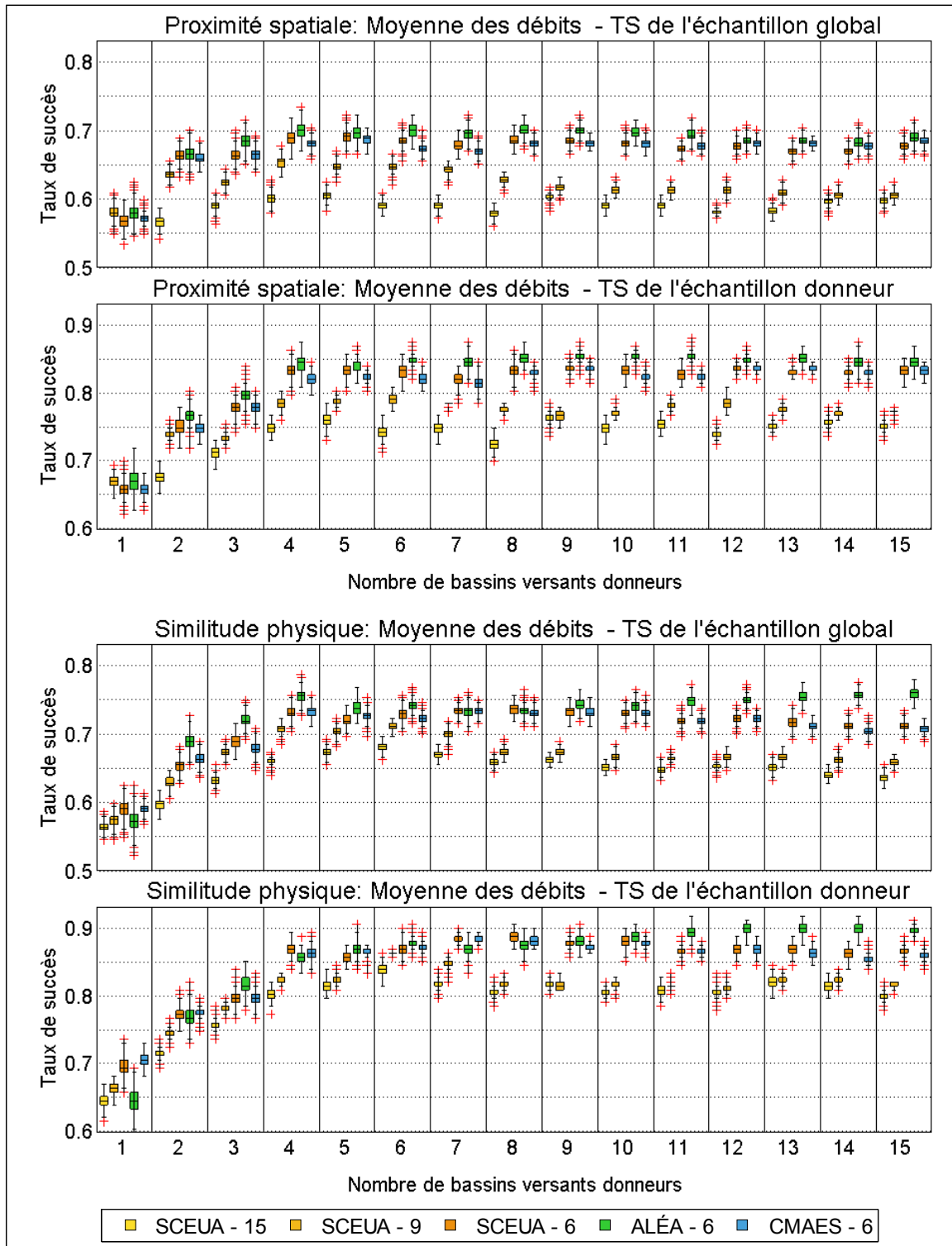


Figure 3.5 Taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique – Interpolation des débits par la moyenne arithmétique

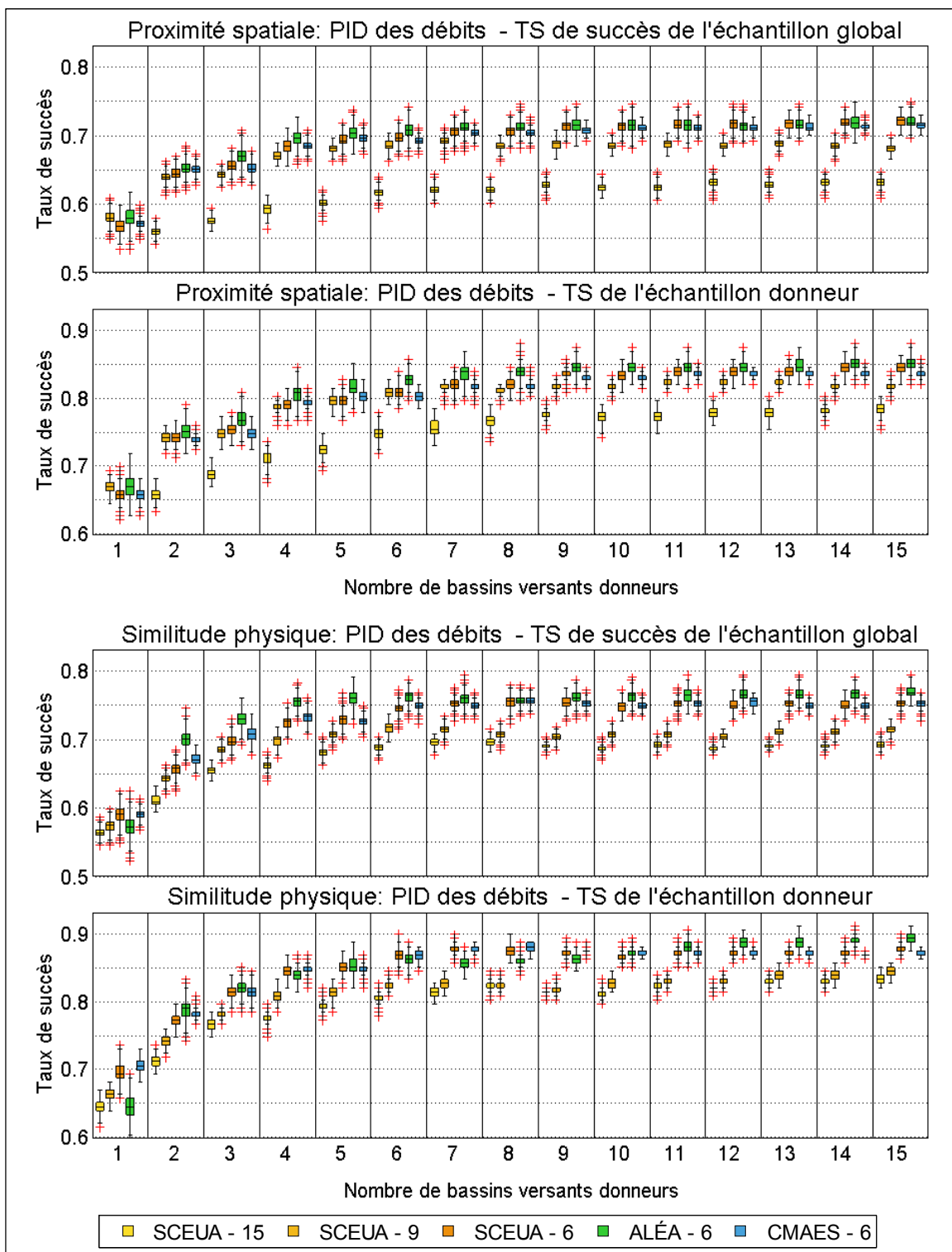


Figure 3.6 Taux de succès des méthodes de régionalisation par proximité spatiale et similitude physique – Interpolation des débits par la PID

### **3.7 Taux de succès obtenu par la méthode de régionalisation par régression linéaire multiple**

Le taux de succès de la méthode de régionalisation par régression linéaire multiple est présenté à la figure 3.7. Ce taux de succès est plus faible que celui affiché par les deux autres méthodes de régionalisation couplées aux trois méthodes d'interpolation testées (*Voir* figures 3.4 à 3.6). À l'image des méthodes de régionalisations couplées à la méthode d'interpolation des paramètres par la moyenne arithmétique (*Voir* figure 3.4), la méthode de régionalisation par régression linéaire multiple affiche une tendance à présenter un taux de succès relativement groupé pour les cinq configurations MH/OC testées. La configuration MH/OC qui utilise le modèle hydrologique à 9 paramètres libres (SCEUA-9) se distingue par son taux de succès globalement plus élevé que celui des autres configurations MH/OC. Parmi les configurations MH/OC utilisant le modèle hydrologique à 6 paramètres libres, celle configurée par échantillonnage aléatoire (ALÉA-6) présente généralement un meilleur taux de succès que celles optimisées à l'aide d'un algorithme d'optimisation (SCEUA-6 et CMAES-6).

### **3.8 Analyse statistique des modèles régionaux linéaires construit par régression linéaire multiple**

L'analyse statistique effectuée ici, présente la médiane du critère statistique évalué, qui correspond à la médiane des valeurs du critère obtenues pour les 1000 transferts de jeux de paramètres répétés en *bootstrap* pour un bassin versant.

#### **3.8.1 Coefficient de détermination $R^2$**

Le coefficient de détermination  $R^2$  du modèle régional linéaire, pour les 6 paramètres libres du modèle hydrologique qui sont communs aux cinq configurations MH/OC, est présenté à la figure 3.8. Celui obtenu par les paramètres libres supplémentaires des configurations MH/OC utilisant un modèle hydrologique plus complexe (SCEUA-15 et SCEUA-9) est présenté à la figure–A II–1.



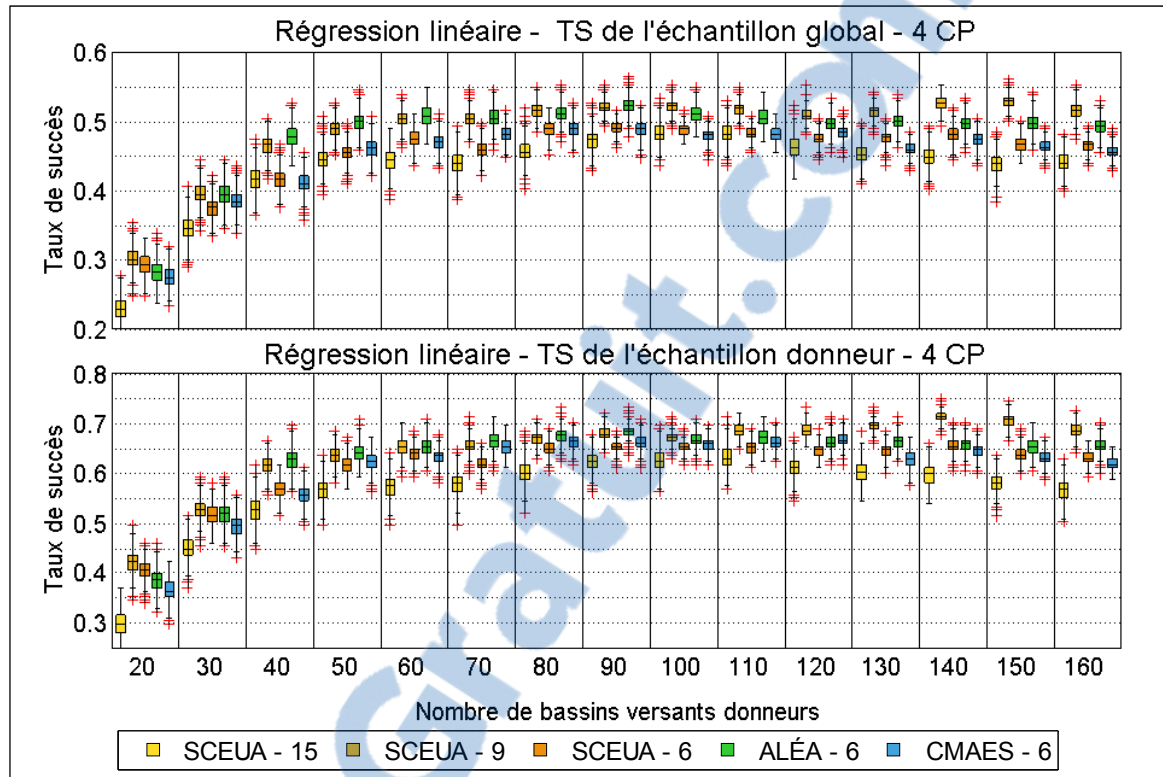


Figure 3.7 Taux de succès de la méthode de régionalisation par régression linéaire multiple – Modèle régional linéaire construit avec quatre caractéristiques physiques

La première observation qui s'impose à la figure 3.8 est la valeur élevée du coefficient  $R^2$  lorsque le modèle régional linéaire est construit avec peu de bassins versants donneurs; particulièrement pour un nombre de 20 bassins versants donneurs où il est généralement supérieur à 0,7. À l'exception des paramètres libres  $X1$  et  $X4$ , la médiane des boîtes à moustaches affiche un coefficient  $R^2$  inférieur à 0,5 lorsque plus de 30 bassins versants donneurs sont utilisés pour construire le modèle régional linéaire. Ces deux derniers paramètres conservent cependant un coefficient  $R^2$  globalement supérieur à 0,5 pour les différentes configurations étudiées de bassins versants donneurs. Le coefficient  $R^2$  affiche une distribution distinctement plus large pour le paramètre libre  $Kf$  que pour les autres paramètres libres. Globalement, l'amplitude de la distribution des boîtes à moustache diminue avec une augmentation du nombre de bassins versant donneurs. Pour les paramètres libres  $X1$  et, dans une moindre mesure,  $X3$ , le coefficient  $R^2$  obtenu avec le modèle hydrologique à 9 paramètres libres (SCEAU-9) se montre légèrement supérieur à celui

qu'obtiennent les autres configurations MH/OC. De manière générale, la complexité du modèle hydrologique et le choix de la méthode de calibration ne semblent pas influencer la valeur du coefficient  $R^2$ .

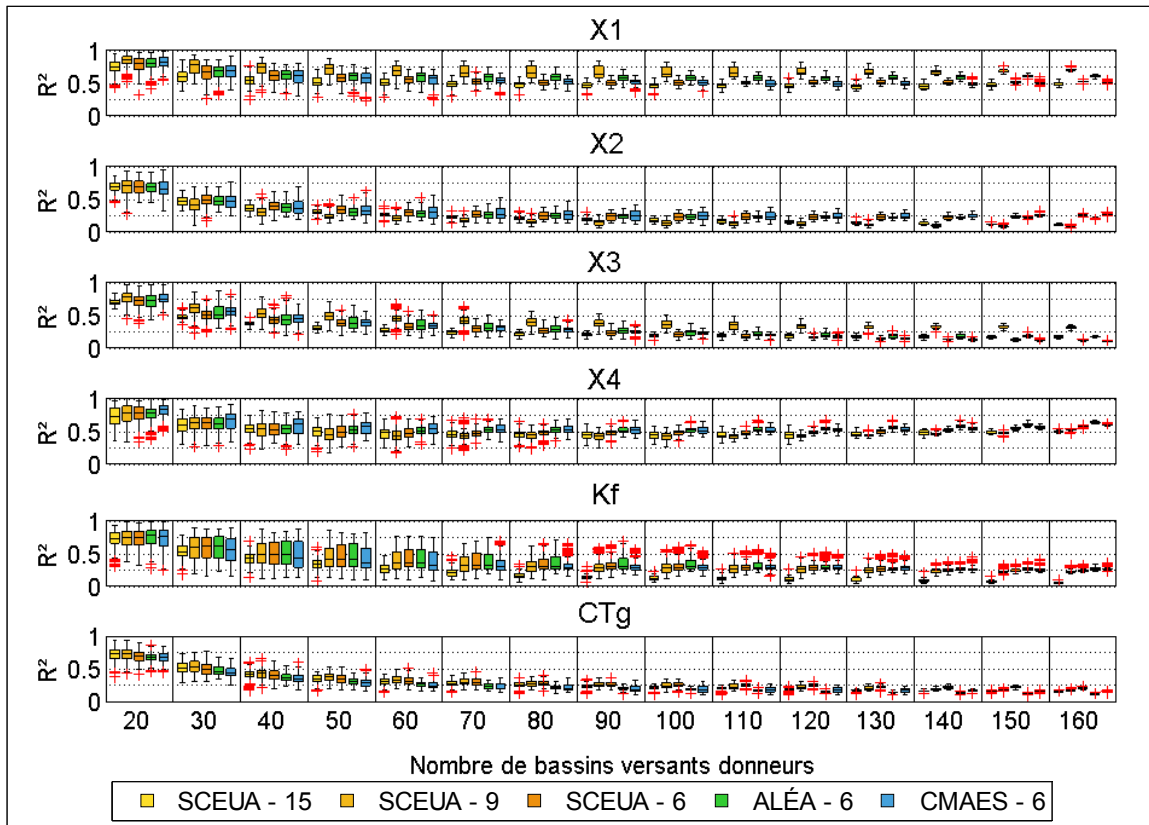


Figure 3.8 Coefficient de détermination  $R^2$  des modèles régionaux linéaires construits avec quatre caractéristiques physiques

### 3.8.2 Statistique $F$

La statistique  $F$  du modèle régional linéaire, pour les 6 paramètres libres du modèle hydrologique qui sont communs aux cinq configurations MH/OC, est présentée à la figure 3.9. Celle obtenue par les paramètres libres supplémentaires des configurations MH/OC utilisant un modèle hydrologique plus complexe (SCEUA-15 et SCEUA-9) est présentée à la figure–A III–1.

Le seuil de la région critique conduisant au rejet de l'hypothèse nulle, calculé pour un modèle régional linéaire construit sur la base de quatre variables explicatives, est à son minimum pour un nombre de 160 bassins versants donneurs,  $F_{1-0,05}(4, 160) \approx 2,4282$  et atteint son maximum pour un nombre de 20 bassins versants donneurs  $F_{1-0,05}(4, 20) \approx 2,8661$ . La figure 3.9 montre que la significativité de la relation entre les paramètres du modèle régional linéaire et ses variables explicatives augmente avec le nombre de bassins versants donneurs. L'hypothèse nulle des paramètres du modèle régional linéaire est fortement contredite par la statistique  $F$  des paramètres libres du modèle hydrologique  $X1$  et  $X4$ . Généralement, pour les cinq configurations MH/OC, les autres paramètres libres du modèle hydrologique ne présentent pas une valeur de la statistique  $F$  suffisamment élevée pour rejeter cette hypothèse nulle. Cette dernière remarque ne s'applique cependant pas à la configuration MH/OC utilisant le modèle hydrologique à 9 paramètres libres (SCEUA-9) qui se distingue pour le paramètre libre  $X3$ . Le paramètre libre  $X1$  de ce dernier modèle hydrologique affiche les valeurs les plus élevées obtenues pour la statistique  $F$ . Les modèles régionaux linéaires issus de la configuration MH/OC qui utilise le modèle hydrologique à 15 paramètres libres (SCEAU-15) présentent des valeurs de la statistique  $F$  relativement faible. Les valeurs de la statistique  $F$  des modèles régionaux linéaires issus des configurations MH/OC utilisant le modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6) se montrent relativement équivalentes.

### 3.8.3 Seuil descriptif $P$ -value

Le seuil descriptif  $P$ -value du modèle régional linéaire, pour les 6 paramètres libres du modèle hydrologique qui sont communs aux cinq configurations MH/OC, est présenté à la figure 3.10. Celui obtenu par les paramètres libres supplémentaires des configurations MH/OC utilisant un modèle hydrologique plus complexe (SCEUA-15 et SCEUA-9) est présentée à la figure–A IV–1.

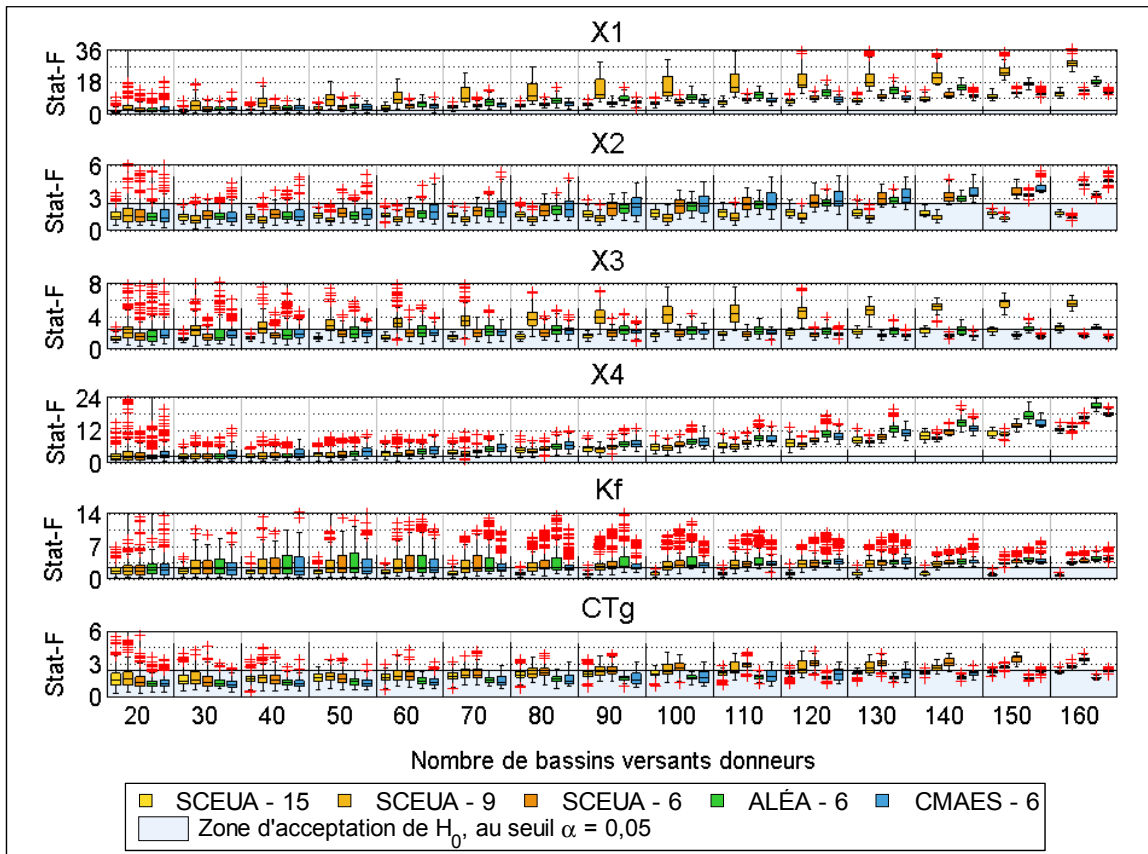


Figure 3.9 Statistique  $F$  des modèles régionaux linéaires construits avec quatre caractéristiques physiques

La figure 3.10 montre que le risque exprimé par le seuil descriptif  $P$ -value de rejeter l'hypothèse nulle de tous les paramètres du modèle régional linéaire alors qu'elle est vraie décroît avec une augmentation du nombre de bassins versants donneurs. Pour les paramètres libres du modèle hydrologique  $X1$  et  $X4$  un nombre de 50 bassins versants donneurs présente généralement un risque associé au seuil descriptif  $P$ -value qui justifie le rejet de l'hypothèse nulle. Les configurations MH/OC utilisant les modèles hydrologiques à 9 et 15 paramètres libres (SCEUA-9 et SCEUA-15), présentent un seuil descriptif  $P$ -value distinctement élevé pour le paramètre libre du modèle hydrologique  $X2$ . Cette dernière configuration MH/OC (SCEUA-15) présente le même caractère distinct pour le paramètre libre  $Kf$ . La distribution des boîtes à moustache est particulièrement étendue pour les paramètres libres du modèle

hydrologique  $Kf$  et, dans une moindre mesure,  $X2$ . Globalement, le choix de la méthode de calibration ne semble pas influencer le résultat du seuil descriptif  $P$ -value.

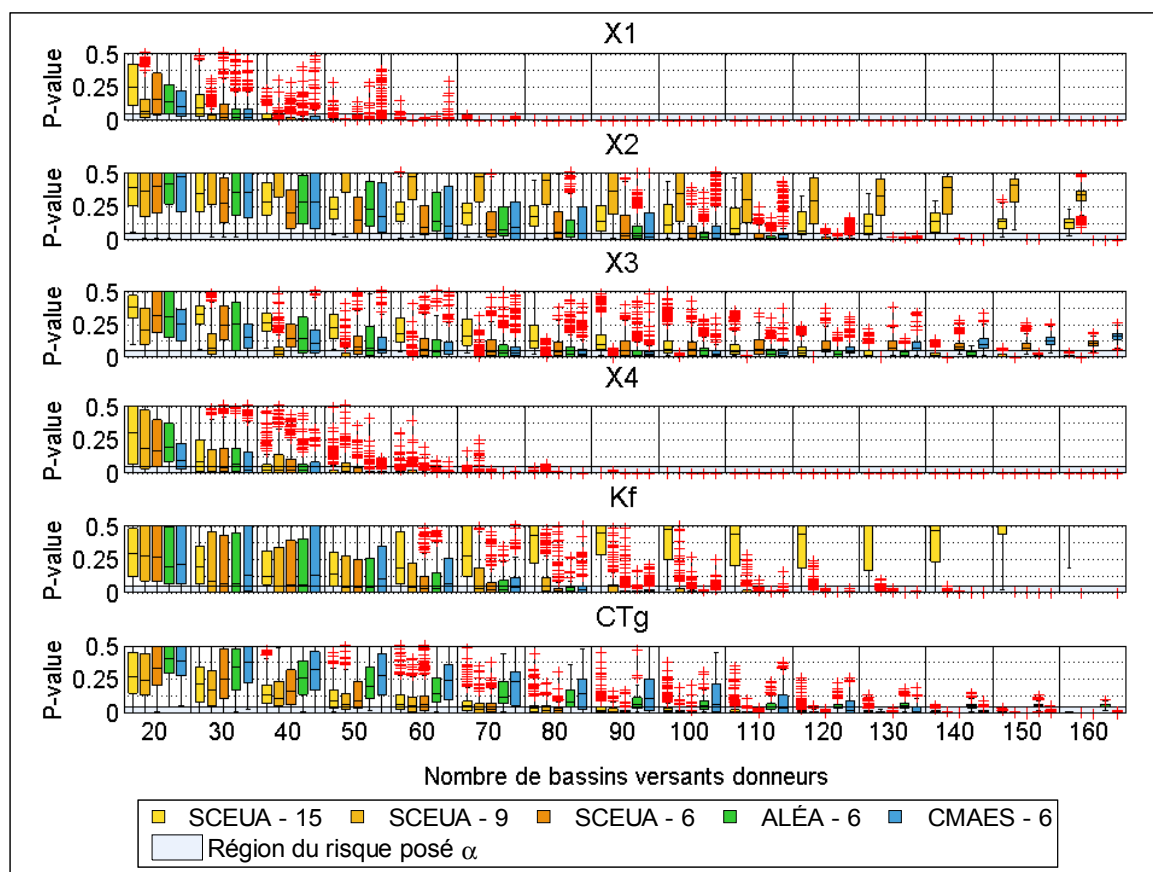


Figure 3.10 Seuil descriptif  $P$ -value des modèles régionaux linéaires construits avec quatre caractéristiques physiques

### 3.9 Performance des méthodes de régionalisation

La performance, exprimée selon de critère d'efficacité de Nash-Sutcliffe, des méthodes de régionalisation par proximité spatiale et par similitude physique – appliquées sur les 266 bassins versants de l'échantillon *global* pour les trois méthodes d'interpolation étudiées – est comparée à la figure 3.11. La performance de ces deux méthodes de régionalisation appliquées sur les 167 bassins versants de l'échantillon *donneur* pour les trois méthodes d'interpolation étudiées est comparée à la figure 3.12. La performance des méthodes de

régionalisation par proximité spatiale et par similitude physique exprime la performance obtenue pour le transfert de jeux de paramètres issus de 12 bassins versants donneurs. La performance de la méthode de régionalisation par régression linéaire multiple appliquée sur les échantillons *global* et *donneur* est présentée à la figure 3.13. Pour cette dernière méthode, la performance est exprimée pour le modèle régional linéaire construit sur la base de quatre variables explicatives et 130 bassins versants donneurs. La justification du nombre de bassins versants donneurs se trouve aux figures 3.4 à 3.7 qui montrent que, globalement, le taux de succès des cinq configurations MH/OC pour ces trois méthodes de régionalisation avoisine son optimum pour ces nombres sélectionnés de bassins versants donneurs.

Pour les trois méthodes de régionalisation présentées, les configurations MH/OC utilisant le modèle hydrologique à 9 et 15 paramètres libres (SCEUA-9 et SCEUA-15) se distinguent des configurations MH/OC utilisant le modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6) par un nombre nettement supérieur de bassins versants dont la performance est incluse dans la classe  $NSE_{\text{moyen}} [0,85; 0,95[$ . La configuration MH/OC SCEUA-9 se montre légèrement plus performante que la configuration MH/OC SCEUA-15 pour les méthodes de régionalisation par proximité spatiale et par régression linéaire alors que la configuration MH/OC SCEUA-15 se révèle celle des deux configurations MH/OC qui est légèrement plus performante pour la méthode de régionalisation par similitude physique.

La configuration MH/OC ALÉA-6 affiche une performance légèrement supérieure à celle des configurations MH/OC SCEUA-6 et CMAES-6 lorsque la méthode d'interpolation par la moyenne des paramètres est appliquée aux méthodes de régionalisation par proximité spatiale et similitude physique mais l'inverse est constaté lorsque la méthode de régionalisation par régression linéaire multiple est appliquée. Ces trois dernières configurations MH/OC offrent globalement une performance similaire pour les méthodes de régionalisation testées.

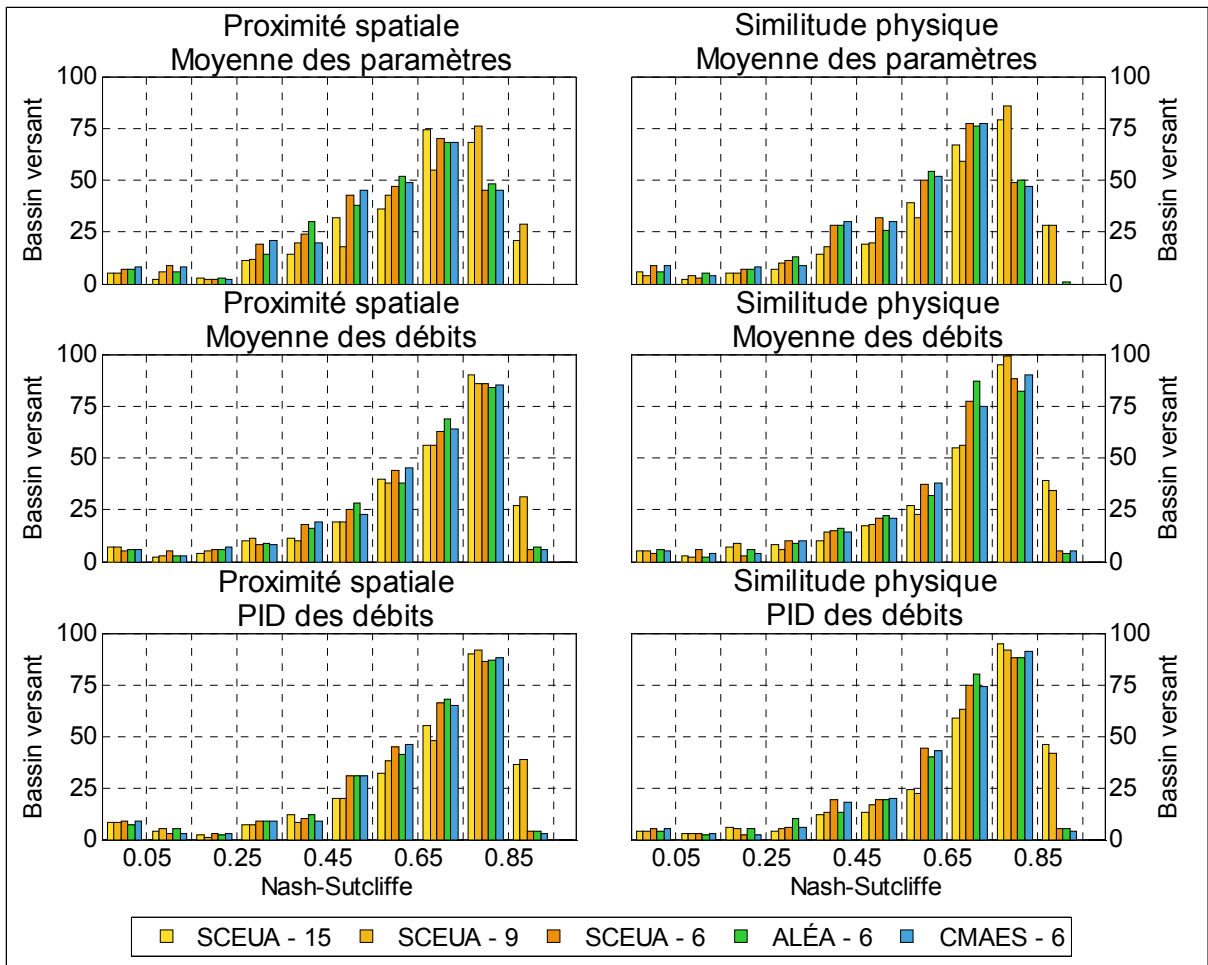


Figure 3.11 Distribution du critère d'efficacité  $NSE_{moyen}$  des 266 bassins versants de l'échantillon global résultant du transfert de jeux de paramètres issus de 12 bassins versants donneurs pour les méthodes de régionalisation par proximité spatiale et similitude physique

La comparaison de la distribution des performances obtenues pour les méthodes de régionalisation par proximité spatiale et par similitude physique montre que cette dernière méthode de régionalisation possède une densité légèrement plus importante de bassins versants dans les classes supérieures de l'histogramme. La méthode de régionalisation par régression linéaire multiple affiche une performance plus faible que les deux méthodes de régionalisation précédentes. La méthode d'interpolation par la moyenne des paramètres est nettement la moins performante des méthodes d'interpolation. Celles qui interpolent les débits par la moyenne et la PID se révèlent globalement aussi performantes. Pour toutes les

méthodes de régionalisation testées, l'amélioration de la performance lorsque les méthodes sont appliquées sur l'échantillon *donneur* plutôt que sur l'échantillon *global* est notable.

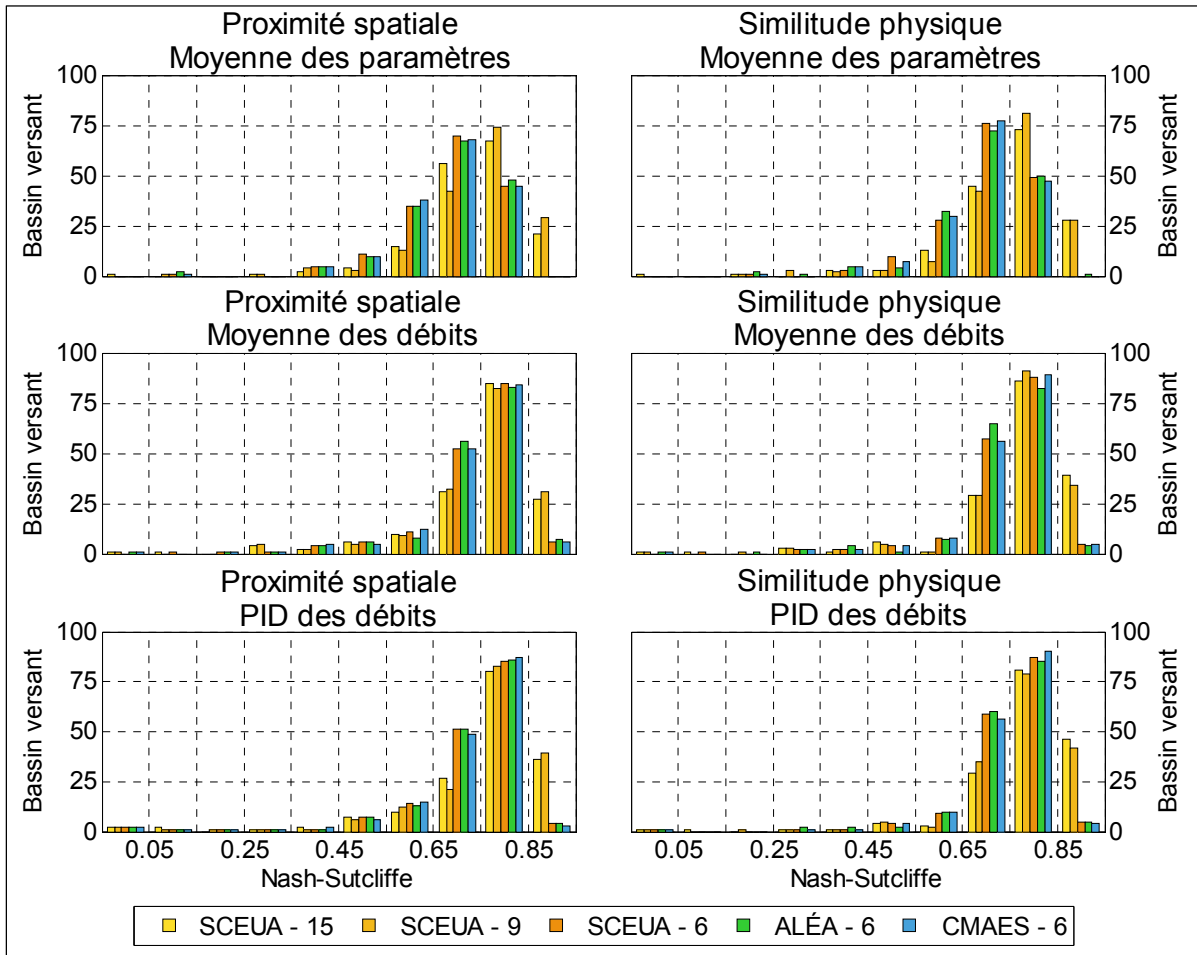


Figure 3.12 Distribution du critère d'efficacité  $NSE_{moyen}$  des 167 bassins versants de l'échantillon donneur résultant du transfert des jeux de paramètres issus de 12 bassins versants donneurs pour les méthodes de régionalisation par proximité spatiale et similitude physique



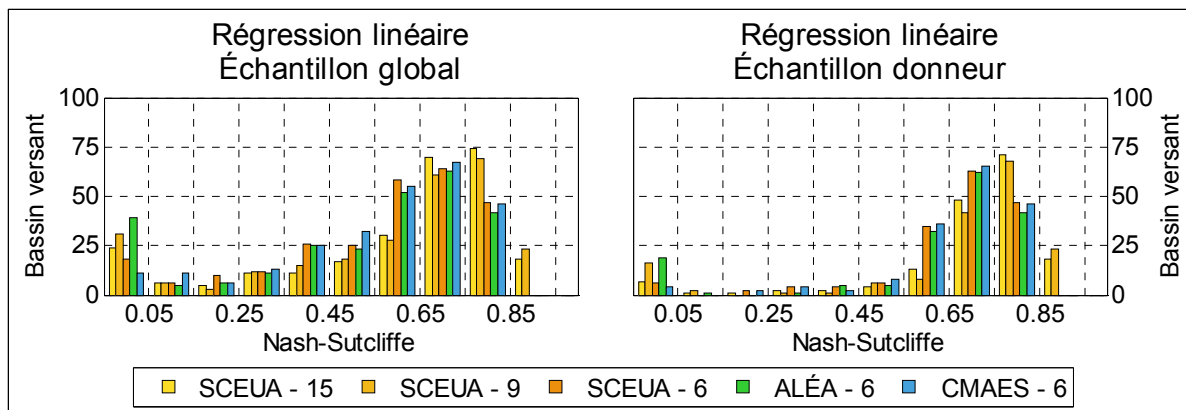


Figure 3.13 Distribution du critère d'efficacité NSEmoyen des échantillons *global* et *donneur* résultant du modèle régional linéaire construit avec quatre caractéristiques physiques et 130 bassins versants donneurs pour la méthode de régionalisation par régression linéaire multiple



## CHAPITRE 4

### DISCUSSION

La discussion porte d'abord sur la comparaison des méthodes de régionalisation des paramètres et souligne l'importance de transférer le jeu de paramètres entier tels que le démontrent les résultats obtenus. Bien que la méthodologie de l'étude soit élaborée dans l'esprit de comparer d'une façon objective les modèles hydrologiques régionalisés, certains choix méthodologiques influencent ces résultats. Le présent chapitre aborde cet aspect limitatif. Pour répondre à l'objectif principal de l'étude, la discussion traite de la complexité du modèle hydrologique et met en évidence le modèle hydrologique à 9 paramètres libres. Pour répondre à l'objectif spécifique no. 1, l'impact de la distribution du jeu de paramètres est mis en évidence. Pour répondre au deuxième objectif spécifique, les critères statistiques du modèle régional linéaire sont mis en relation et discutés. La connaissance de l'incertitude liée à l'équifinalité du jeu de paramètres revêt une importance particulière dans une étude de régionalisation des paramètres. Elle est aussi discutée dans ce chapitre.

#### **4.1 Comparaison des méthodes de régionalisation des paramètres et importance de transférer le jeu de paramètres dans son intégralité**

La méthode de régionalisation par similitude physique se montre plus robuste que celle par proximité spatiale, ce qui est conforme aux résultats obtenus par Arsenault et Brissette (2014), Samuel, Coulibaly et Metcalfe (2011) et Oudin et al. (2008). La figure 4.1 localise les 167 bassins versants de l'échantillon *donneur* et permet de distinguer deux régions relativement homogènes; celle des bassins versants donneurs et celle des bassins versants exclus dudit échantillon. Dans la présente étude, il est possible que la performance de la méthode de régionalisation par proximité spatiale soit limitée par la répartition spatiale des bassins versants donneurs.

La robustesse des méthodes de régionalisation par proximité spatiale et par similitude physique est améliorée lorsque les méthodes d'interpolation des débits (*Voir* figures 3.5 et

3.6) sont préférées à la méthode d'interpolation des paramètres (*Voir* figure 3.4). Ce résultat confirme l'importance de transférer le jeu de paramètres dans son intégralité plutôt que de transférer les paramètres individuellement. Il existe dans un jeu de paramètres, une interdépendance qui doit être respectée pour que les processus hydrologiques d'un bassin versant non jaugé soient représentés aussi efficacement que le permettent le modèle hydrologique et les données. Notons que la méthode de régionalisation par régression linéaire multiple offre une robustesse similaire à la méthode d'interpolation des paramètres couplée aux deux autres méthodes de régionalisation étudiées. Soulignons aussi le gain en robustesse du modèle hydrologique à 9 paramètres libres par rapport à celui à 6 paramètres libres pour la méthode d'interpolation par la moyenne des paramètres et la méthode de régionalisation par régression linéaire multiple (*Voir* figures 3.4 et 3.7). Cette dernière méthode de régionalisation se révèle la moins robuste et la moins performante des trois méthodes de régionalisation étudiées. Cette observation est similaire pour Arsenault et Brissette (2014), Samuel, Coulibaly et Metcalfe (2011) et Oudin et al. (2008).

Les méthodes d'interpolation de la moyenne arithmétique et de la PID des débits, couplées aux méthodes de régionalisation par proximité spatiale et par similitude physique, se révèlent ici, globalement similaires. Arsenault et Brissette (2014) observent pour leur part une légère amélioration de la performance lorsque la PID est préférée à la moyenne arithmétique pour interpoler les débits.

#### **4.2 Normalisation de l'identité des bassins versants donneurs pour les cinq configurations MH/OC**

Dans la présente étude, le fait d'employer le même groupe de bassins versants donneurs pour transférer les paramètres des bassins versants donneurs, pour les cinq configurations MH/OC testées, peut limiter la performance en régionalisation des configurations MH/OC SCEUA-15 et SCEUA-9, qui affichent une performance nettement plus élevée en calibration. Ces deux dernières configurations MH/OC possèdent donc un nombre plus élevé de bassins versants, qualifiés par rapport au seuil posé du critère d'efficience  $NSE_{\text{moyen}} \geq 0,70$ , et qui se trouvent *de facto* inclus dans l'échantillon *donneur*. Il est possible qu'un nombre accru de bassins

versants donneurs réduise la distance, spatiale ou en similitude, entre le bassin versant pseudo non jaugé et les bassins versants donneurs, améliorant ainsi la performance du modèle hydrologique régionalisé, ce qui ne peut être constaté dans la présente étude.

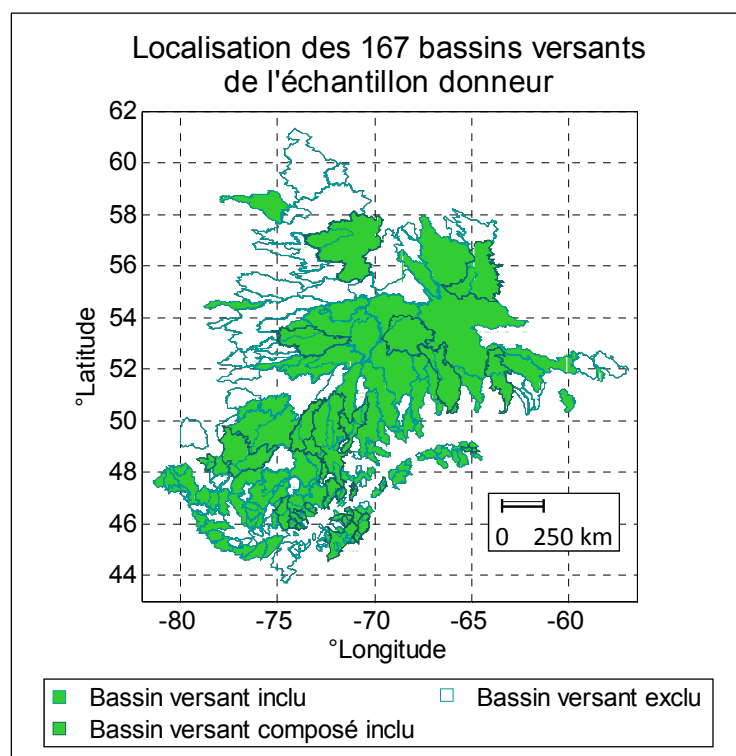


Figure 4.1 Localisation des 167 bassins versants de l'échantillon *donneur*

### 4.3 Impact du nombre de bassins versants donneurs

Pour les méthodes de régionalisation par proximité spatiale et par similitude physique, un nombre de 5 à 8 bassins versants donneurs est généralement suffisant pour atteindre la robustesse optimale de la méthode. À l'exception de la méthode d'interpolation par la moyenne des paramètres, l'augmentation du nombre de bassins versants donneurs n'entraîne généralement pas une diminution tangible de la robustesse lorsque le nombre optimal de bassins versants donneurs est dépassé. Arsenault et Brissette (2014), Bao et al. (2012) et Oudin et al. (2008) obtiennent des courbes de robustesse en fonction du nombre de bassins versants donneurs similaires à celles présentées aux figures 3.4 à 3.6. Notons que, dans

l'étude de Oudin et al. (2008), la courbe obtenue pour le modèle hydrologique GR4J – à 4 paramètres libres – présente cependant une plus forte diminution de sa robustesse une fois dépassé le nombre optimal de bassins versants donneurs. Pour la méthode de régionalisation par régression linéaire et les configurations MH/OC testées, un nombre de 110 bassins versants donneurs est généralement suffisant pour atteindre la robustesse optimale de la méthode.

#### 4.4 Impact de l'échantillon de bassins versants évalué

Le tableau 4.1 atteste du gain de performance de la méthode de régionalisation par similitude physique, couplée à l'interpolation par la PID des débits, lorsque l'échantillon *donneur* est préféré à l'échantillon *global*. Les critères d'efficacité NSE présentés correspondent à la moyenne des critères d'efficacité NSE des bassins versants contenus dans les deux échantillons respectifs pour chacune des configurations MH/OC. Les tableaux –A V–1 à –A V–3 présentent les critères d'efficacité NSE de la méthode de régionalisation par similitude physique, couplée à la méthode d'interpolation par la moyenne arithmétique des débits, et par la méthode de régionalisation par proximité spatiale, couplée aux deux méthodes d'interpolation des débits. Les tableaux 4.1 et –A V–1 à –A V–3 permettent d'observer que le gain en performance, lorsque l'échantillon *donneur* est préféré à l'échantillon *global*, diminue avec l'augmentation de la complexité du modèle hydrologique. Bien que cette tendance soit faible, elle est néanmoins nette. La méthode de régionalisation par similitude physique affiche la meilleure performance moyenne pour les deux échantillons de bassins versant donneurs évalués. Les deux méthodes d'interpolation du débit présentent des performances similaires pour ces deux échantillons. Pour les trois configurations MH/OC régionalisées selon les quatre configurations présentées aux tableaux 4.1 et –A V–1 à –A V–3, un gain moyen du critère d'efficacité NSE de 0,0904 est observé lorsque l'échantillon *donneur* est préféré à l'échantillon *global*.

#### 4.5 Complexité du modèle hydrologique

Rappelons les trois paramètres libres sélectionnés pour obtenir le modèle hydrologique à 9 paramètres libres : *CoeffPerc*, *FacteurSeuil* et *Vmin* (Voir section 3.1). Les processus de fonte du stock de neige et de vidange du réservoir de production du modèle hydrologique sont les principaux facteurs limitatifs à la parcimonie du modèle hydrologique à 6 paramètres libres. Les bassins versants de l'étude sont soumis à des climats allant de continental humide à arctique. Les conditions climatiques et/ou caractéristiques physiques particulières auxquelles sont soumis certains bassins versants de l'étude affectent leurs processus hydrologiques d'une façon que ne peut simuler adéquatement le modèle hydrologique à 6 paramètres libres. Les six degrés de liberté supplémentaires du modèle hydrologique à 15 paramètres libres sur celui à 9 paramètres libres ne lui permettent pas de se distinguer d'une façon notable et globale.

Tableau 4.1 Comparaison des échantillons *global* et *donneur* sur la base du critère d'efficacité  $NSE_{moyen}$  de la méthode de régionalisation par similitude physique couplée à l'interpolation par la PID des débits, présenté pour les trois configurations MH/OC optimisées par l'algorithme d'optimisation SCE-UA

Configuration MH/OC	Échantillon global (NSE)	Échantillon donneur (NSE)	$\Delta$ NSE
SCEUA-15	0,6936	0,7735	0,0799
SCEUA-9	0,6893	0,7741	0,0848
SCEUA-6	0,6500	0,7405	0,0905

##### 4.5.1 Impact de la complexité du modèle hydrologique sur la performance des configurations MH/OC régionalisées

Les résultats de la figure 3.1 confirment que la performance du modèle hydrologique augmente avec une augmentation de sa complexité, ce qui est aussi observé par Perrin, Michel et Andréassian (2001).

La figure 4.2 présente la performance des modèles hydrologiques optimisés à l'aide de l'algorithme d'optimisation SCE-UA (SCEUA-15, SCEUA-9 et SCEUA-6), modulée selon la valeur des caractéristiques physiques des bassins versants. Le découpage des intervalles de valeurs des caractéristiques physiques est effectué par une technique de classification (*clustering*) qui calcule la distance entre ces valeurs en effectuant la moyenne pondérée de la distance euclidienne. La figure montre les sept caractéristiques physiques pour lesquelles un gain de performance du modèle hydrologique à 15 paramètres libres (SCEUA-15) sur le modèle hydrologique à 9 paramètres libres (SCEUA-9) est notable. Les cinq caractéristiques physiques pour lesquelles la performance de ces deux derniers modèles hydrologiques est similaire sont présentées à la figure –A VI–1. Les figures 4.2 et –A VI–1 mettent en évidence la forte proportion de bassins versants qui sont modélisés avec la même efficacité par les modèles hydrologiques à 9 et 15 paramètres. Le modèle hydrologique à 15 paramètres libres (SCEUA-15) apparaît principalement mieux adapté pour décrire les bassins versants de grande superficie, ceux fortement urbanisés et ceux situés au nord de 54,8° de latitude.

#### **4.5.2 Impact de la complexité du modèle hydrologique sur sa robustesse et sur la performance des méthodes de régionalisation des paramètres**

Pour les trois configurations MH/OC optimisées à l'aide de l'algorithme d'optimisation SCE-UA (SCEUA-15, SCEUA-9 et SCEUA-6), les résultats obtenus par les deux méthodes d'interpolation sur les débits, couplées aux méthodes de régionalisation par proximité spatiale et par similitude physique, confortent l'hypothèse selon laquelle la robustesse du modèle hydrologique en validation est proportionnelle à la parcimonie de ses paramètres libres (Perrin, Michel et Andréassian, 2001). Cependant la configuration MH/OC utilisant le modèle hydrologique à 9 paramètres libres (SCEUA-9) se révèle la plus robuste pour les méthodes de régionalisation par proximité spatiale et par similitude physique, couplées à la méthode d'interpolation des paramètres, et pour la méthode de régionalisation par régression linéaire multiple.

Les modèles hydrologiques à 9 et 15 paramètres libres sont globalement aussi performants en validation (*Voir* figures 3.11 à 3.13). Bien que le modèle hydrologique à 6 paramètres libres



soit généralement le plus robuste, sa performance inférieure à celle des deux autres modèles hydrologiques en calibration se traduit par une performance en validation, en-deçà de la leur.

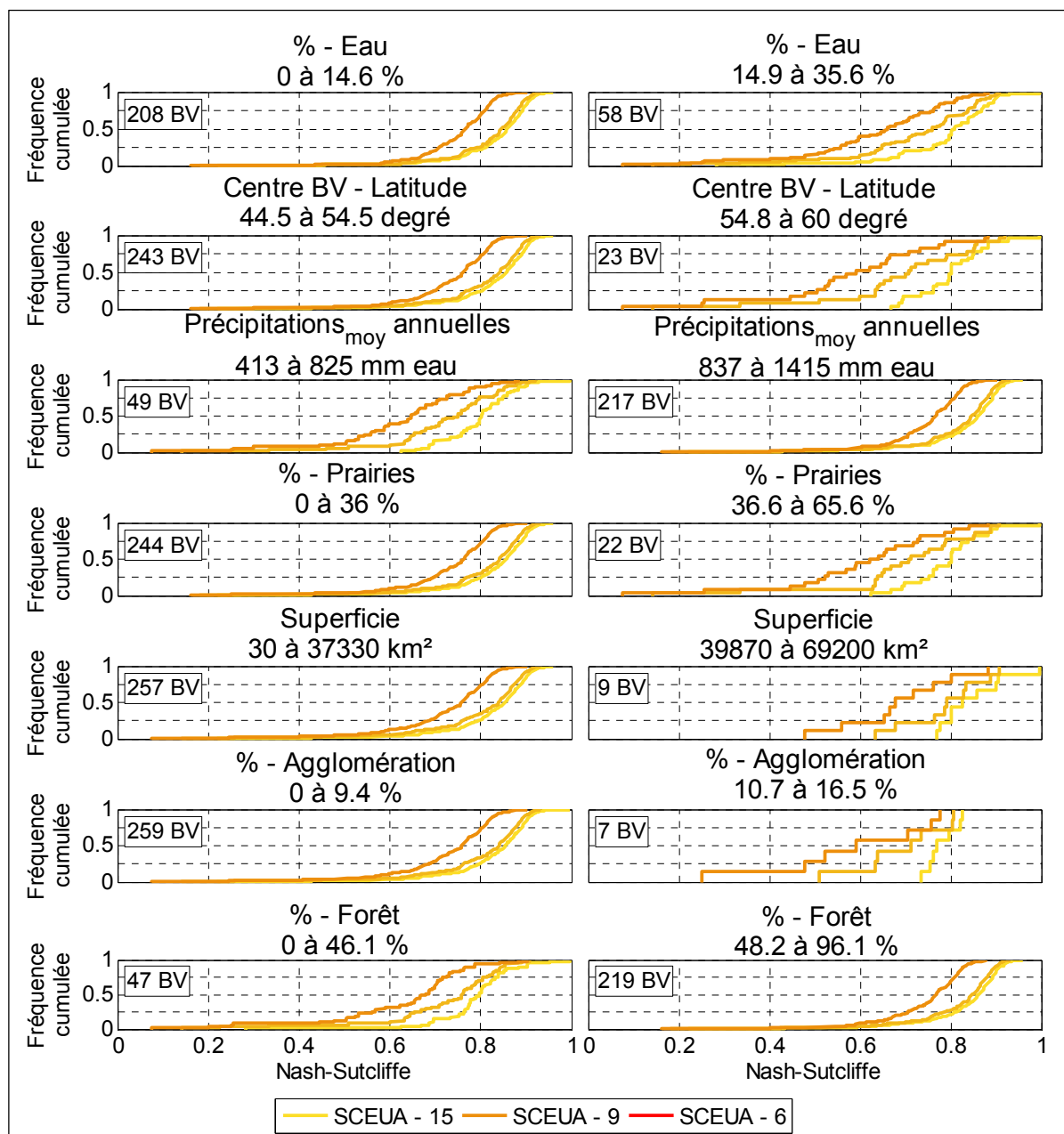


Figure 4.2 Performance des modèles hydrologiques à 6, 9 et 15 paramètres libres exprimée selon l'intervalle de valeurs des caractéristiques physiques

### 4.5.3 Modèle hydrologique à 9 paramètres libres

Perrin, Michel et Andréassian (2003), à l'instar de Senbeta, Shamseldin et O'Connor (1999), proposent le critère  $r^2$  (Nash et Sutcliffe, 1970) pour évaluer le degré d'amélioration d'un modèle hydrologique modifié par rapport au modèle hydrologique de référence;

$$r^2 = \frac{R_2 - R_1}{1 - R_1} \quad (4.1)$$

où,  $R_1$  est le critère d'efficacité NSE du modèle hydrologique de référence et  $R_2$  celui du modèle hydrologique amélioré.

Le tableau 4.2 constate que le modèle hydrologique à 9 paramètres libres est significativement meilleur que le modèle à 6 paramètres libres mais que le modèle hydrologique à 15 paramètres libres n'offre pas un gain de performance aussi important par rapport au modèle hydrologique à 9 paramètres libres. Les variables  $R_1$  et  $R_2$  – utilisés pour déterminer  $r^2$  (Voir équation 4.1) – représentent la moyenne, en calibration, du critère d'efficacité NSE des 266 bassins versants de l'échantillon *global* pour leur configuration MH/OC respective (SCEUA-15, SCEUA-9 ou SCEUA-6).

Tableau 4.2 Estimation du degré d'amélioration résultant du gain en complexité des modèles hydrologiques à 9 et 15 paramètres libres

Modèle hydrologique de référence	Modèle hydrologique amélioré	$r^2$
6 paramètres libres	9 paramètres libres	0,112
9 paramètres libres	15 paramètres libres	0,014

L'hypothèse est ici posée que le modèle hydrologique à 9 paramètres libres est, parmi les modèles étudiés, celui qui possède la complexité s'approchant le plus de la complexité

*suffisante* (Voir section 1.6) pour reproduire adéquatement les processus hydrologiques dominants des bassins versants. La comparaison de la distribution des jeux de paramètres optimisés des configurations MH/OC SCEUA-6 (Voir figure 3.3), SCEUA-9 et SCEUA-15 (Voir figures –A I–1a et –A I–1b) permet de constater que les jeux de paramètres générés pour la configuration MH/OC SCEUA-9 occupent une région similaire à ceux générés pour la configuration MH/OC SCEUA-6; l'interaction entre les 6 paramètres libres de base n'est pas significativement augmentée. À l'inverse, le modèle hydrologique à 15 paramètres libres augmente l'amplitude de la région occupée par les 6 paramètres libres de base, principalement pour trois d'entre eux. La majorité des 15 paramètres libres du modèle hydrologique se trouvent fortement dispersés au sein de l'espace paramétrique, ce qui indique une interaction augmentée entre les paramètres libres du modèle.

#### **4.6 Impact de la distribution du jeu de paramètres optimisés sur les méthodes de régionalisation des paramètres**

La région de l'espace paramétrique occupée par les jeux de paramètres équi-finaux des configurations MH/OC du modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6) est distincte selon la configuration MH/OC observée (Voir figure 3.3). L'impact sur la robustesse est tangible pour la configuration MH/OC configurée par échantillonnage aléatoire (ALÉA-6) qui se montre globalement la plus robuste des trois configurations MH/OC (Voir figures 3.4 à 3.7). La performance légèrement plus élevée obtenue en calibration par les configurations MH/OC SCEUA-6 et CMAES-6 leur permet cependant d'obtenir une performance en validation globalement similaire à celle de la configuration MH/OC ALÉA-6, malgré leur moindre robustesse. Les deux algorithmes d'optimisation affichent une performance similaire en validation.

#### **4.7 Information contenue dans les caractéristiques physiques retenues pour construire l'index de similitude**

Deux des quatre caractéristiques physiques retenues dans la présente étude pour transférer les jeux de paramètres avec la méthode de régionalisation par similitude physique sont des

informations géographiques, soient la latitude et la longitude du centre du bassin versant. La méthode de régionalisation par similitude physique est donc, de fait, une méthode hybride de similitude physique et de proximité spatiale. L'importance de la proximité géographique dans le regroupement des bassins versants en région homogène est mise en évidence par ce résultat, ce qui conforte l'hypothèse sous-jacente à la méthode de régionalisation par proximité spatiale (*Voir* section 1.3).

Trois caractéristiques physiques sont communes à toutes les configurations MH/OC testées : l'occupation du sol par l'eau, la latitude et la longitude du bassin versant (*Voir* figure 3.2). Les processus hydrologiques des bassins versants sont représentés différemment selon le degré de complexité du modèle hydrologique; il paraît donc acceptable que la quatrième caractéristique physique diffère selon la complexité du modèle hydrologique. Pour le modèle hydrologique à 6 paramètres libres, cette quatrième caractéristique physique est similaire pour les configurations MH/OC SCEUA-6 et CMAES-6 – optimisées à l'aide d'un algorithme d'optimisation – mais différente de celle de la configuration MH/OC ALÉA-6 – configurée par échantillonnage aléatoire. La région de l'espace paramétrique, occupée par les jeux de paramètres transférés apparaît donc, en partie, responsable de la sélection des caractéristiques physiques.

La figure 4.3 présente le détail des 12 caractéristiques physiques pour chacun des 266 bassins versants de l'échantillon *global* et permet de constater une certaine interdépendance entre ces caractéristiques physiques. Dans le contexte climatique du Québec, une latitude plus nordique signifie, entre autres, une diminution des précipitations moyennes annuelles, une augmentation de l'occupation du sol par les prairies et une diminution de l'occupation du sol par la forêt. Notons la similitude entre la courbe de la superficie et l'intervalle [10,40] % de la courbe de l'occupation du sol par l'eau, qui peut expliquer le fait que la superficie du bassin versant – qui *a priori* joue un rôle prépondérant dans le comportement du bassin versant – ne soit pas une caractéristique physique retenue pour construire l'index de similitude, alors que l'occupation du sol par l'eau l'est.

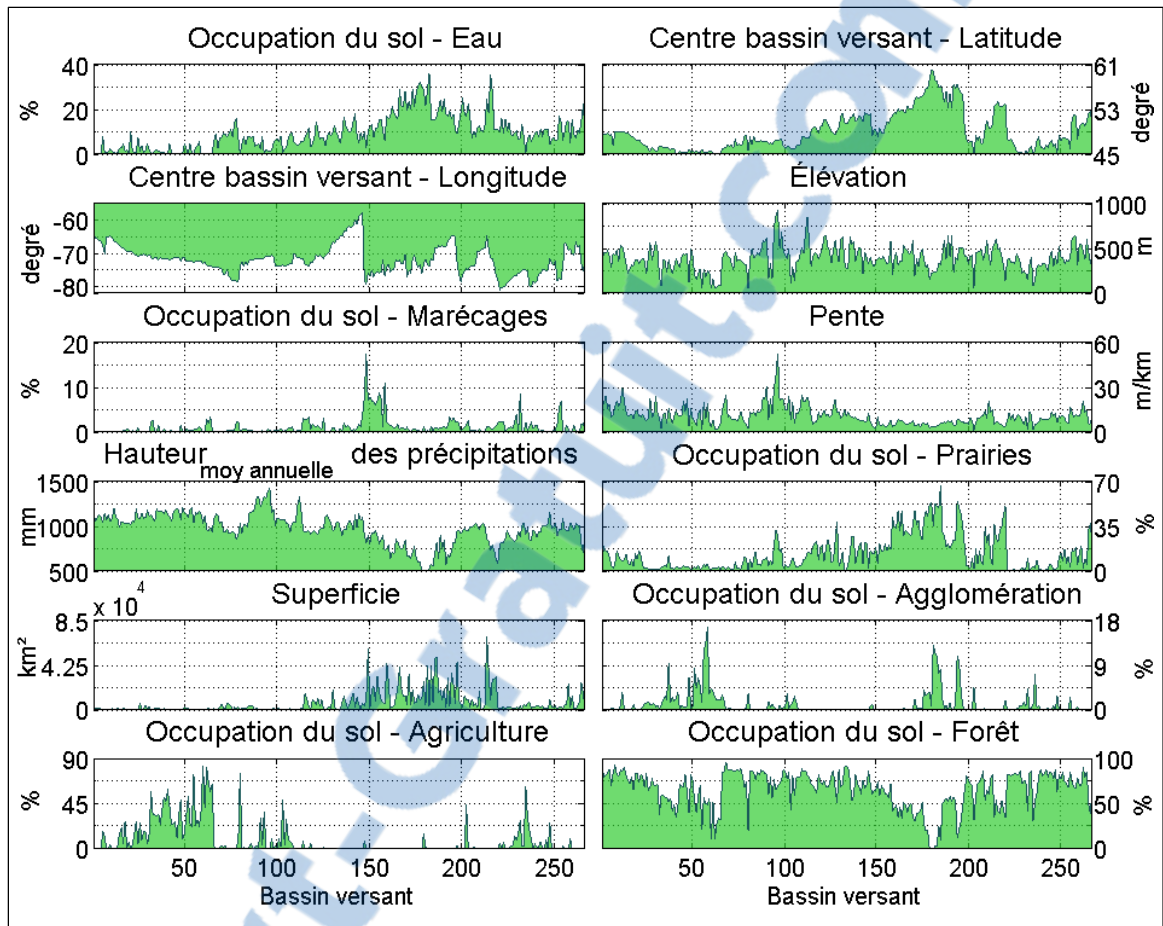


Figure 4.3 Détail des caractéristiques physiques des 266 bassins versants de l'échantillon *global*

Les caractéristiques physiques contiennent donc, *à priori*, une certaine redondance d'information. Le fait que, lors de la sélection des caractéristiques physiques, la performance du transfert des jeux de paramètres n'augmente pas significativement lorsque plus de quatre caractéristiques physiques sont utilisées pour construire l'index de similitude, semble confirmer que l'information additionnelle est en partie redondante. Cependant, le fait de les classer selon la figure 4.2 permet de constater que les caractéristiques physiques fournissent au contraire une information complémentaire. Pour des caractéristiques physiques dont l'information paraît *à priori* redondante – par exemple, l'occupation du sol par l'eau et la superficie –, la distribution des performances obtenues par les configurations MH/OC SCEUA-15 et SCEUA-9 n'est pas toujours identique.

Dans le cas où cette complémentarité est avérée, le choix de construire l'index de similitude avec quatre caractéristiques physiques pour appliquer la méthode de régionalisation par similitude physique est susceptible de limiter la performance de la méthode. Il apparaît, *a posteriori*, préférable de construire cet index de similitude à l'aide de toutes les caractéristiques physiques qui sont disponibles et qui ne diminuent pas la performance du transfert de paramètres. Plutôt que de privilégier la parcimonie, il semble préférable de considérer la complémentarité de l'information contenue dans les caractéristiques physiques.

#### 4.8 Incertitude liée à l'équifinalité

Il est difficile d'observer une tendance sur l'incertitude des résultats affichés aux figures 3.4 à 3.7. L'incertitude observée sur le taux de succès des méthodes de régionalisation par proximité spatiale et par similitude physique, couplées aux méthodes d'interpolation des débits (*Voir* figures 3.5 et 3.6), peut être exprimée sous la forme d'un écart relatif pour chacun des critères statistiques de la boîte à moustache (*Voir* section 3.5) tel que montré pour le critère  $Q1$ ;

$$Q1_{relatif} = \frac{Q1 - \text{médiane}}{\text{médiane}} \cdot 100 \quad (4.2)$$

où  $Q1$  représente le 1<sup>er</sup> quartile de la boîte à moustache et la médiane est celle du taux de succès.

Pour les échantillons *global* et *donneur*, la figure 4.4 présente l'écart relatif des critères statistiques de la boîte à moustache (*Voir* section 3.5) en fonction du nombre de bassins versants donneurs, pour ces deux dernières méthodes de régionalisation couplées à la méthode d'interpolation par la PID des débits. L'enveloppe des moustaches – de couleur verte – est délimitée par les écarts relatifs des moustaches supérieure et inférieure, soit respectivement  $Q3_{relatif} + 1,5(Q3_{relatif} - Q1_{relatif})$  et  $Q1_{relatif} - 1,5(Q3_{relatif} - Q1_{relatif})$ . L'enveloppe de la boîte – de couleur turquoise – est délimitée par  $Q3_{relatif}$  et  $Q1_{relatif}$ ,

respectivement le 3<sup>ième</sup> et 1<sup>er</sup> quartile. Ces enveloppes représentent donc le pourcentage d'écart des variables statistiques respectives par rapport à la médiane.

La figure 4.4 montre que l'écart relatif sur l'enveloppe de la boîte reste généralement stable alors que l'enveloppe de moustaches est celle qui fluctue selon la configuration MH/OC testée. Cela signifie que, pour la moitié des bassins versants de chaque configuration MH/OC testée (définie par la distance entre  $Q3_{relatif}$  et  $Q1_{relatif}$ ), l'écart relatif fluctue peu ou pas en fonction de l'échantillon de bassins versants donneurs ou de la méthode de régionalisation. Bien que l'écart relatif soit maximal pour un nombre de bassins versants donneurs inférieur ou égal à 2 mais ne peut être minimal qu'après 5 bassins versants donneurs, il n'existe pas de relation générale entre le nombre de bassins versants donneurs et l'incertitude du taux de succès. L'écart relatif de la méthode de régionalisation par proximité spatiale appliquée sur l'échantillon *donneur* pour la configuration MH/OC SCEUA-15 présente un comportement distinct des autres écarts relatifs présentés à la figure 4.4. La tendance générale qui se dégage de la figure 4.4 montre que l'écart relatif sur le taux de succès est égal ou diminue lorsque l'échantillon donneur est préféré à l'échantillon global. Le même effet est constaté lorsque la méthode de régionalisation par similitude physique est préférée à celle par proximité spatiale. Le fait que l'écart relatif présenté à la figure 4.4 pour les configurations MH/OC SCEAU-15, SCEUA-9 et SCEAU-6 ne présente pas de tendance en fonction de la complexité indique que la complexité du modèle hydrologique n'a pas d'incidence sur l'incertitude du taux de succès.

Parmi les trois configurations MH/OC utilisant le modèle hydrologique à 6 paramètres libres (SCEUA-6, ALÉA-6 et CMAES-6), la configuration MH/OC ALÉA-6 affiche la plus grande incertitude, ce qui est expliqué par l'amplitude de la région de l'espace paramétrique occupée par les jeux de paramètres de cette dernière configuration MH/OC (*Voir* figure 3.3). L'algorithme d'optimisation CMAES engendre une incertitude moindre que l'algorithme d'optimisation SCE-UA (CMAES-6 et SCEUA-6), ce qui indique que la répétabilité du premier algorithme (*Voir* section 3.4) permet de réduire l'incertitude du modèle hydrologique régionalisé.

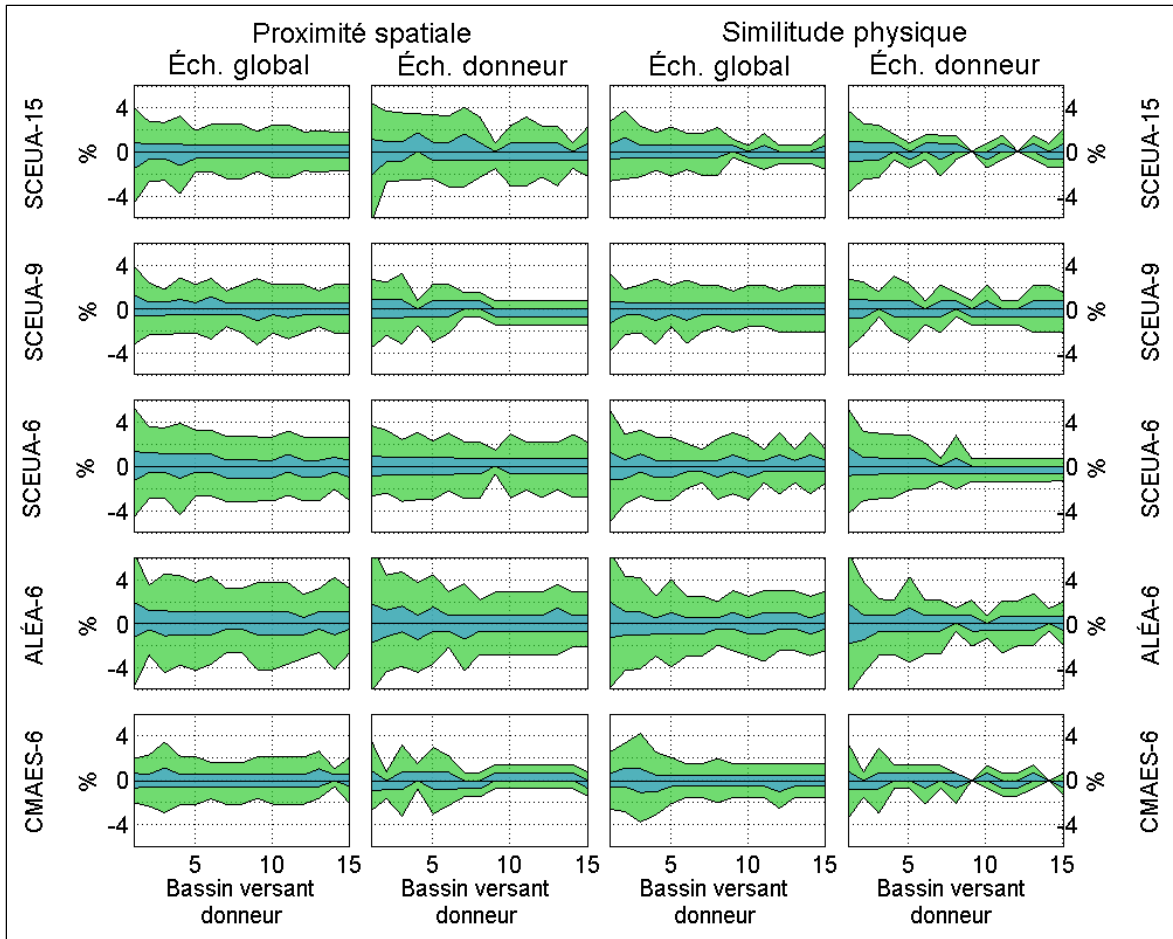


Figure 4.4 Enveloppe de l'incertitude relative en fonction du nombre de bassins versants donateurs obtenue par les méthodes de régionalisation par proximité spatiale et similitude physique pour les deux échantillons étudiés (*global* et *donneur*)

#### 4.9 Validité statistique du modèle régional linéaire construit pour appliquer la méthode de régionalisation par régression linéaire multiple

Le nombre de caractéristiques physiques avec lequel est construit le modèle régional linéaire n'influence pas significativement la performance du modèle (*Voir* section 3.6). La pertinence d'utiliser ces caractéristiques physiques est mise en évidence par la performance obtenue par la méthode de régionalisation par similitude physique, qui regroupe les bassins versants en régions homogènes sur la base d'un index de similitude issu de ces caractéristiques physiques. Ces dernières démontrent donc une aptitude à représenter la variation – au sein de



l'échantillon étudié – des processus hydrologiques qui contrôlent la réponse des bassins versants.

La robustesse affichée par la configuration MH/OC ALÉA-6 pour cette méthode de régionalisation (*Voir* figure 3.7) est surprenante au sens où la forte dispersion de ses jeux de paramètres, au sein de l'espace paramétrique du modèle hydrologique (*Voir* figure 3.3), pourrait laisser présager un modèle régional linéaire moins apte à prédire le jeu de paramètres du bassin versant pseudo non jaugé que les modèles régionaux linéaires issus des configurations MH/OC SCEUA-6 et CMAES-6. Cette robustesse de la configuration MH/OC ALÉA-6 n'est cependant pas suffisante; sa performance en validation (*Voir* figure 3.13) reste inférieure à celle des deux autres configurations MH/OC utilisant le modèle hydrologique à 6 paramètres libres (SCEUA-6 et CMAES-6), en raison notamment du nombre important de bassins versants dont le critère d'efficacité de Nash-Sutcliffe en validation est à toute fin pratique nul ( $NSE \leq 0,05$ ). Il est intéressant de souligner la similitude, en robustesse comme en performance, de la configuration MH/OC SCEUA-9 affichée par la méthode de régionalisation par régression linéaire multiple et par la méthode d'interpolation des paramètres couplée aux deux autres méthodes de régionalisation.

Pour chacune des configurations MH/OC testées et pour chacun des paramètres libres, il est possible d'obtenir un coefficient de détermination  $R^2$  supérieur à 0,5 lorsque 20 bassins versants donneurs sont utilisés pour construire le modèle régional linéaire (*Voir* figures 3.8 et –A II–1). La robustesse exprimée par le taux de succès en fonction du nombre de bassins versants donneurs est cependant en contradiction avec le coefficient de corrélation  $R^2$  des paramètres du modèle régional linéaire. Alors que cette robustesse croît puis se stabilise au-delà de 50 bassins versants donneurs, la corrélation entre les variables explicatives et les variables expliquées du modèle régional linéaire décroît généralement. L'hypothèse nulle posée par le test d'hypothèse ne peut généralement être rejetée (statistique  $F$ ) lorsque les valeurs du coefficient de détermination  $R^2$  sont maximales. La statistique  $F$  croît cependant généralement avec une augmentation du nombre de bassins versants donneurs mais la force du rejet de l'hypothèse nulle reste généralement faible lorsque l'hypothèse est rejetée. Le

seuil descriptif *P-value* est supérieur au risque posé  $\alpha$  lorsque le coefficient  $R^2$  est à son maximum. Le seuil descriptif *P-value* diminue avec une augmentation du nombre de bassins versants donneurs et se trouve généralement inférieure au risque posé ( $\alpha = 0,05$ ) lorsqu'un grand nombre de bassins versants donneurs est utilisé pour construire le modèle régional linéaire. Trois paramètres du modèle hydrologique –  $X1$ ,  $X4$  et  $ExpHU$  – font cependant exception et maintiennent un coefficient de détermination suffisant pour offrir une certaine qualité de corrélation entre les variables explicatives et les variables expliquées du modèle régional linéaire ( $R^2 \geq 0,5$ ) lorsque ce dernier modèle est statistiquement pertinent.

Ces résultats démontrent qu'il n'existe pas de relation proportionnelle entre le coefficient de détermination  $R^2$  et la robustesse de la prédiction du débit simulé à l'exutoire du bassin versant non jaugé. Ils démontrent, de plus, que le modèle régional linéaire ne remplit généralement pas les conditions pour être validé (statistique  $F$  et *P-value*) lorsque le coefficient de détermination  $R^2$  indique qu'une forte corrélation existe entre les variables explicatives et les variables expliquées. Dans un contexte de régionalisation des paramètres, la méthode de régionalisation par régression linéaire multiple ne démontre pas l'aptitude nécessaire à produire un modèle prédictif statistiquement valide.

## CONCLUSION

La présente étude compare les méthodes de régionalisation des paramètres de modèles hydrologiques par proximité spatiale, par similitude physique et par régression linéaire multiple. Pour les deux premières méthodes de régionalisation, les jeux de paramètres sont transférés selon trois options, soit : la moyenne arithmétique des paramètres, la moyenne arithmétique des débits et la moyenne par PID des débits. Les jeux de paramètres régionalisés proviennent de trois modèles hydrologiques – comportant respectivement 6, 9 et 15 paramètres libres – qui sont adaptés du modèle hydrologique GR4J couplé au module de neige CemaNeige. Trois configurations du modèle hydrologique à 6 paramètres libres sont comparées pour quantifier l'impact de la distribution des paramètres, au sein de l'espace paramétrique du modèle hydrologique, sur les méthodes de régionalisation des paramètres. Ces trois configurations sont optimisées respectivement avec l'algorithme d'optimisation SCE-UA, avec l'algorithme d'optimisation CMAES et par une méthode d'échantillonnage aléatoire de l'espace paramétrique. Pour quantifier l'impact de la complexité du modèle hydrologique sur les méthodes de régionalisation des paramètres, chacun des modèles hydrologiques à 9 et 15 paramètres libres est aussi configuré avec l'algorithme d'optimisation SCE-UA. La validité du modèle régional linéaire – construit pour appliquer la méthode de régionalisation des paramètres par régression linéaire multiple – est déterminée à l'aide de l'analyse du coefficient de détermination  $R^2$ , de la statistique  $F$  du test de Fisher et du seuil descriptif  $P$ -value. L'incertitude liée au contexte d'équifinalité des jeux de paramètres est évaluée par une méthode de ré-échantillonnage en *bootstrap* pour chaque configuration testée.

Les résultats montrent que la méthode de régionalisation des paramètres par similitude physique se révèle légèrement meilleure que la méthode par proximité spatiale. L'hypothèse sous-jacente à la méthode de régionalisation des paramètres par proximité spatiale pose que c'est la proximité géographique entre deux bassins versants qui leur confère un comportement hydrologique similaire. Dans la présente étude, le fait que les bassins versants donneurs se trouvent majoritairement dans la partie sud de la province est susceptible de

nuire à la performance des bassins versants pseudo non jaugés situés dans la partie nord de la province. La distance entre un bassin versant pseudo non jaugé et ses bassins versants donneurs peut atteindre des centaines de kilomètres. La méthode par régression linéaire multiple se révèle nettement la moins performante des trois méthodes de régionalisation étudiées. Il apparaît clairement qu'il est préférable de transférer les jeux de paramètres dans leur intégralité; les méthodes d'interpolation des débits sont nettement plus performantes que la méthode d'interpolation des paramètres. L'incertitude de la méthode d'interpolation par la PID des débits est généralement inférieure à celle de la méthode par la moyenne arithmétique des débits.

Malgré la robustesse en régionalisation généralement plus faible des modèles hydrologiques à 9 et 15 paramètres libres par rapport au modèle hydrologique à 6 paramètres libres, la performance nettement plus élevée des deux premiers modèles hydrologiques en calibration leur permet d'afficher une performance supérieure au modèle hydrologique à 6 paramètres libres en validation. Un modèle hydrologique complexe offrant une forte performance en calibration est donc préférable à un modèle hydrologique parcimonieux possédant une forte robustesse. Au-delà de la complexité *suffisante* du modèle hydrologique, les jeux de paramètres générés dans un contexte d'équifinalité tendent à se disperser au sein de l'espace paramétrique. Un modèle hydrologique d'une complexité supérieure à la complexité *suffisante* est susceptible de produire une performance légèrement inférieure à celle du modèle qui s'approche le plus de cette complexité *suffisante*, particulièrement lorsque la méthode de régionalisation par proximité spatiale est appliquée.

Dans un contexte d'équifinalité, la méthode d'échantillonnage aléatoire permet d'identifier les régions *possibles* de l'espace paramétrique d'un modèle hydrologique alors qu'un algorithme d'optimisation ne permet d'identifier qu'une région restreinte de cet espace paramétrique. La performance en calibration des modèles hydrologiques configurés par échantillonnage aléatoire est cependant légèrement inférieure à celle obtenue par les modèles hydrologiques optimisés à l'aide d'un algorithme d'optimisation. La performance en

validation des trois méthodes de calibration est globalement similaire mais la méthode d'échantillonnage aléatoire génère plus d'incertitude.

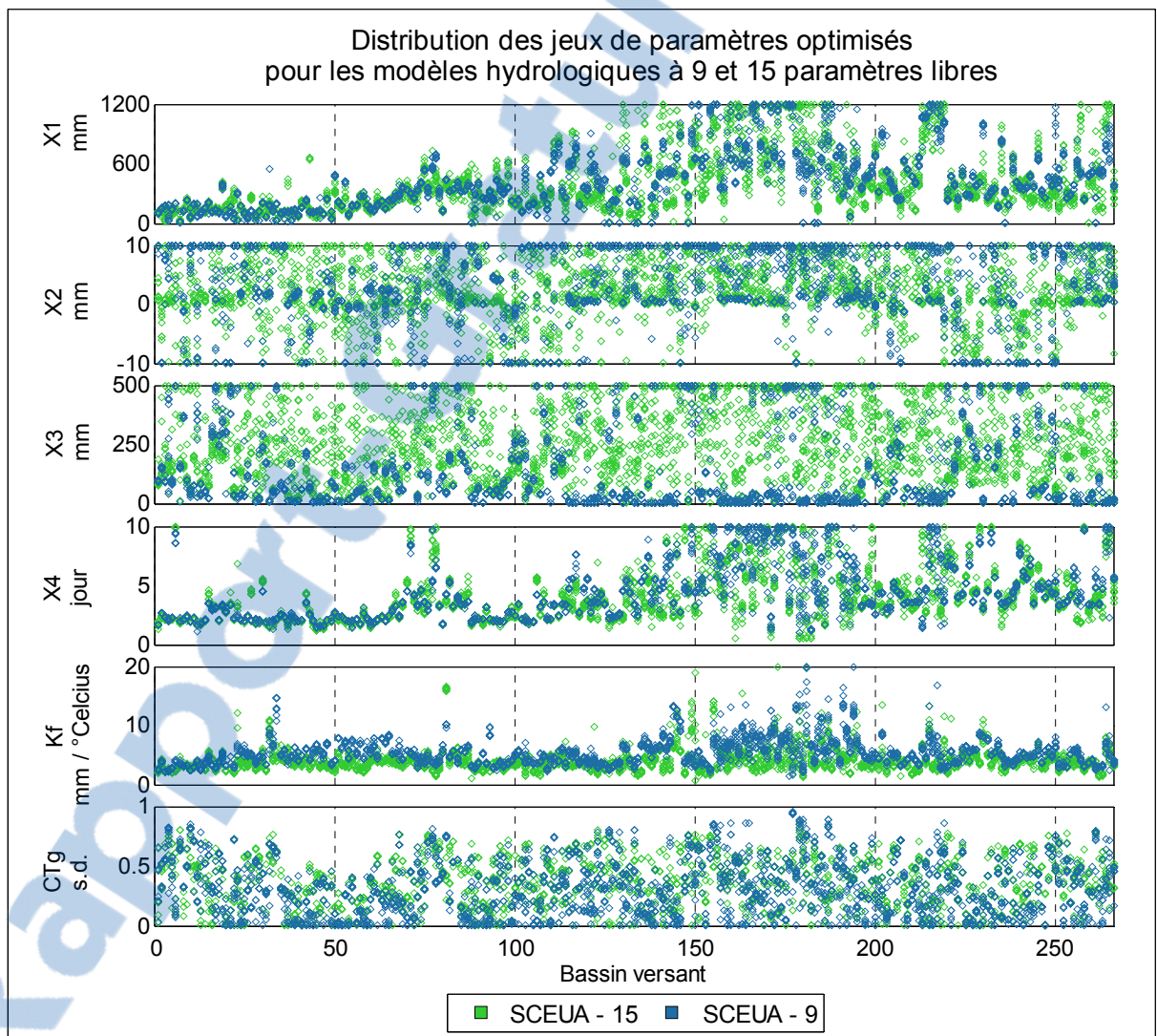
Le modèle régional linéaire construit pour appliquer la méthode de régionalisation par régression linéaire multiple n'est statistiquement pas valide. À l'exception de trois paramètres, le coefficient de détermination  $R^2$  est inversement proportionnel à la statistique  $F$  et au seuil descriptif  $P$ -value. De plus, la robustesse du modèle régional linéaire est généralement, elle aussi, inversement proportionnelle à la grandeur du coefficient de détermination  $R^2$ . En régionalisation des paramètres, la méthode de régression linéaire multiple ne présente pas l'aptitude nécessaire pour établir une relation pertinente entre les variables explicatives et les variables expliquées du modèle régional linéaire. Pour cette méthode de régionalisation, la forte robustesse affichée par le modèle hydrologique à 6 paramètres libres optimisé par la méthode d'échantillonnage aléatoire de l'espace paramétrique est surprenante; bien qu'aucun lien régional n'existe entre les jeux de paramètres transférés, le modèle régional linéaire résultant est aussi efficace qu'un modèle régional linéaire construit avec des jeux de paramètres optimisés à l'aide d'un algorithme d'optimisation – qui possèdent une certaine structure régionale.



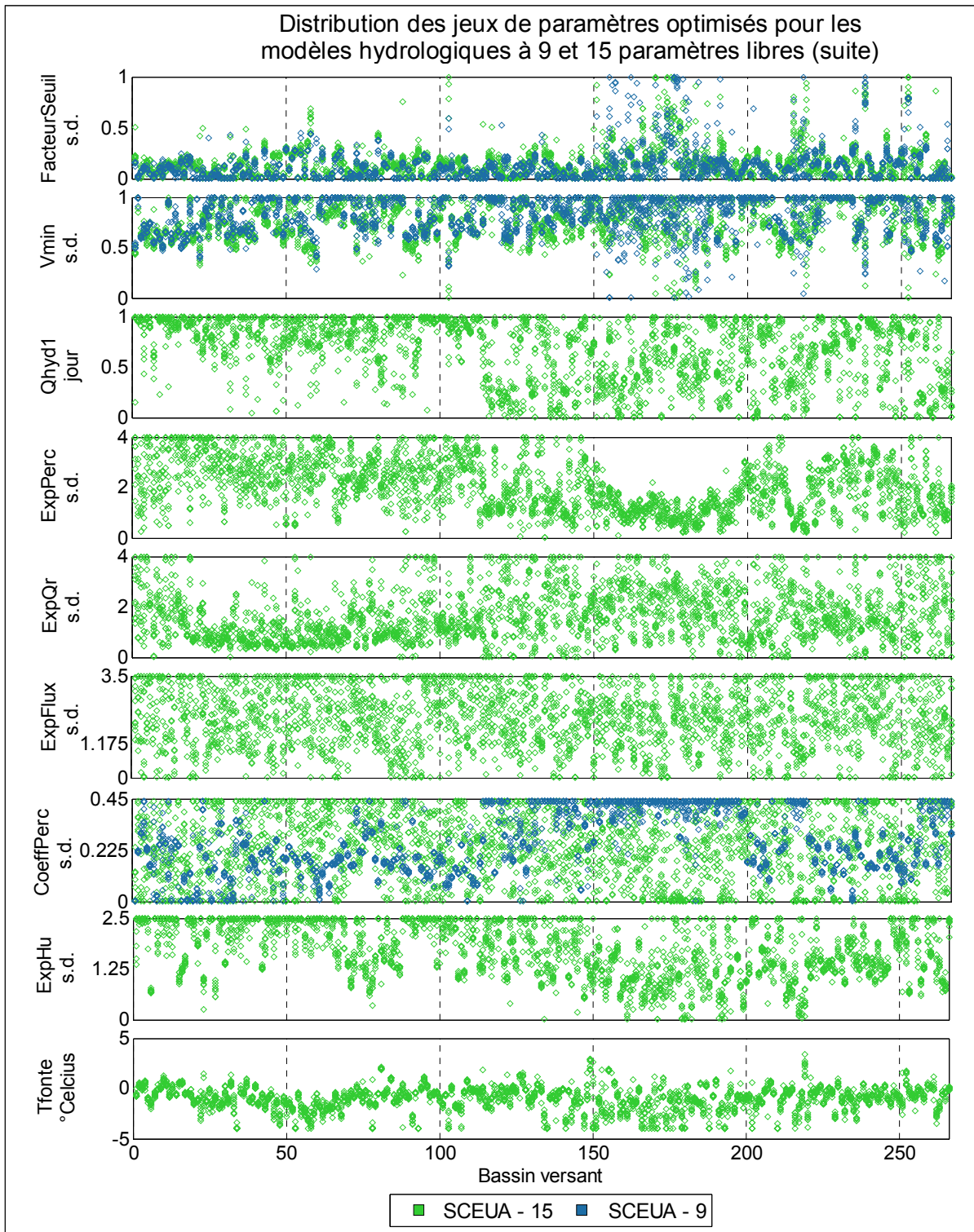
## ANNEXE I

### DISTRIBUTION DES JEUX DE PARAMÈTRES OPTIMISÉS DES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES

L'annexe 1 présente la distribution des jeux de paramètres optimisés des modèles hydrologiques à 9 et 15 paramètres libres. Elle complète l'information de la figure 3.3.



Figure–A I–1a Distribution des 10 jeux de paramètres optimisés par bassin versant pour les configurations MH/OC des modèles hydrologique à 9 et 15 paramètres libres



Figure– A I–1b Distribution des 10 jeux de paramètres optimisés par bassin versant pour les configurations MH/OC des modèles hydrologique à 9 et 15 paramètres libres (suite)



## ANNEXE II

### COEFFICIENT DE DÉTERMINATION $R^2$ DES MODÈLES RÉGIONAUX LINÉAIRES CONSTRUITS POUR LES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES

L'annexe 2 présente le coefficient  $R^2$  des modèles régionaux linéaires construits avec quatre caractéristiques physiques pour les modèles hydrologiques à 9 et 15 paramètres libres. Elle fait suite à la figure 3.8 qui présente les coefficients  $R^2$  des six premiers paramètres libres.

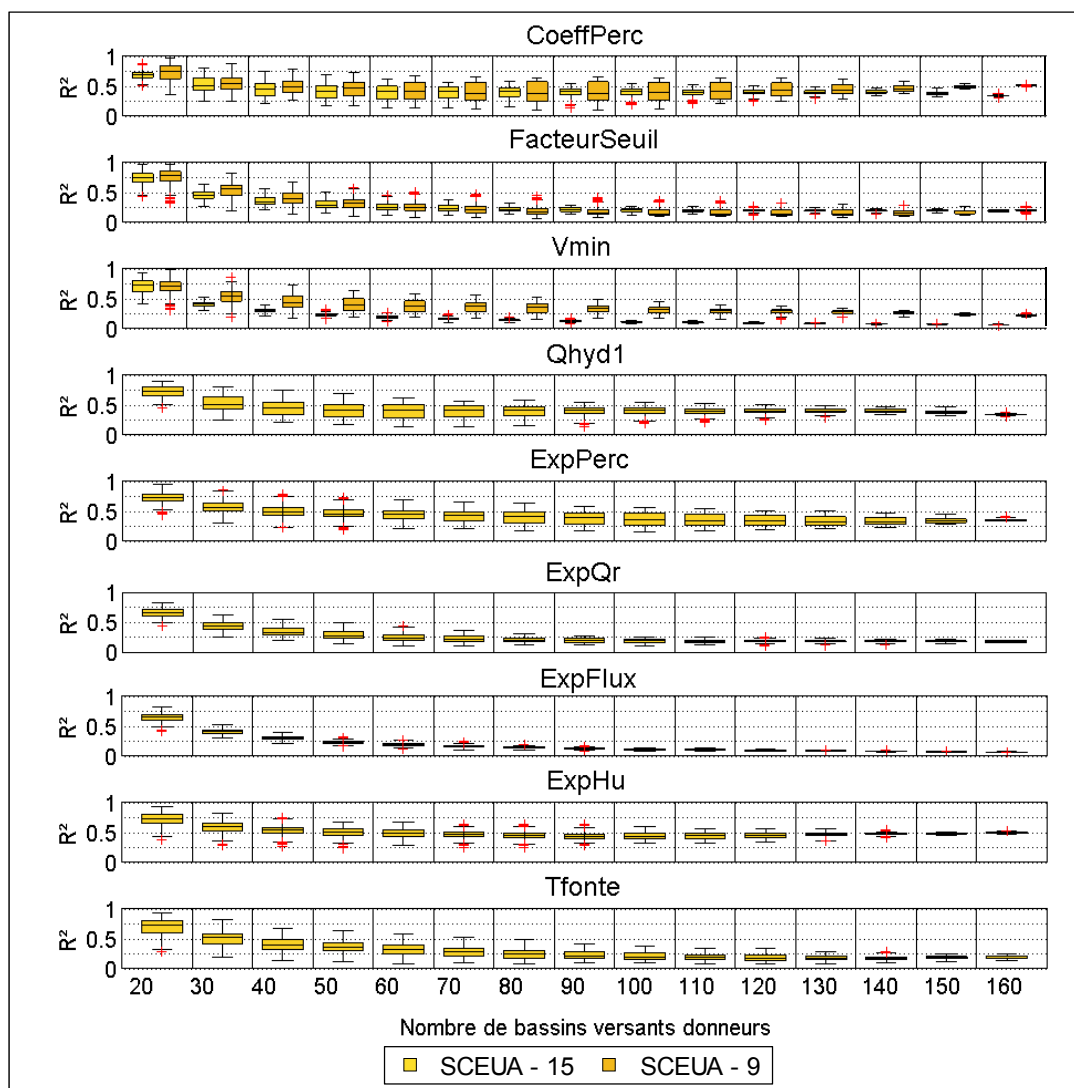


Figure-A II-1 Coefficient de détermination  $R^2$





### ANNEXE III

#### STATISTIQUE $F$ DES MODÈLES RÉGIONAUX LINÉAIRES CONSTRUITS POUR LES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES

L'annexe 3 présente la statistique  $F$  des modèles régionaux linéaires construits avec quatre caractéristiques physiques pour les modèles hydrologiques à 9 et 15 paramètres libres. Elle fait suite à la figure 3.9 qui présente la statistique  $F$  pour les 6 premiers paramètres libres.

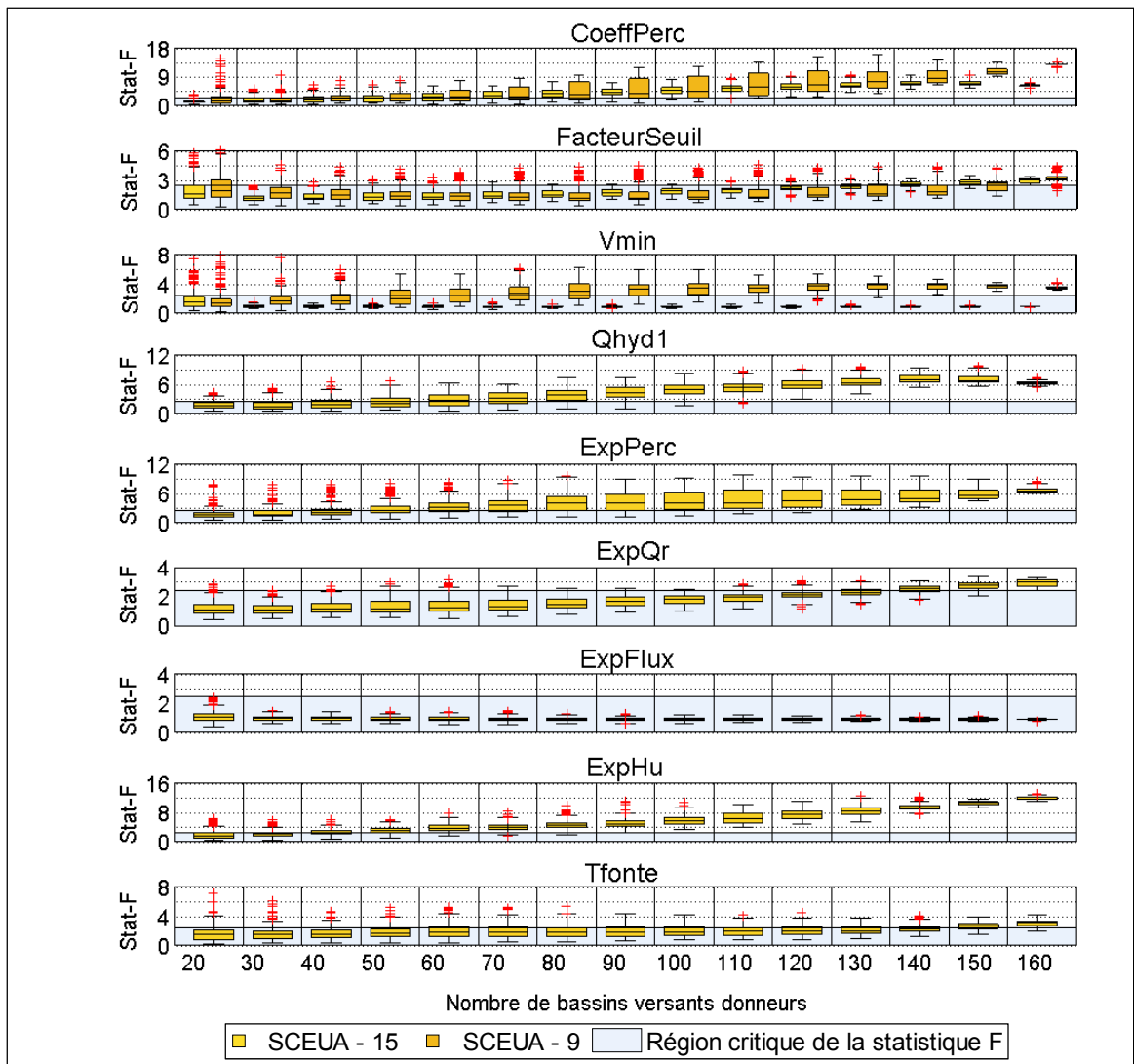


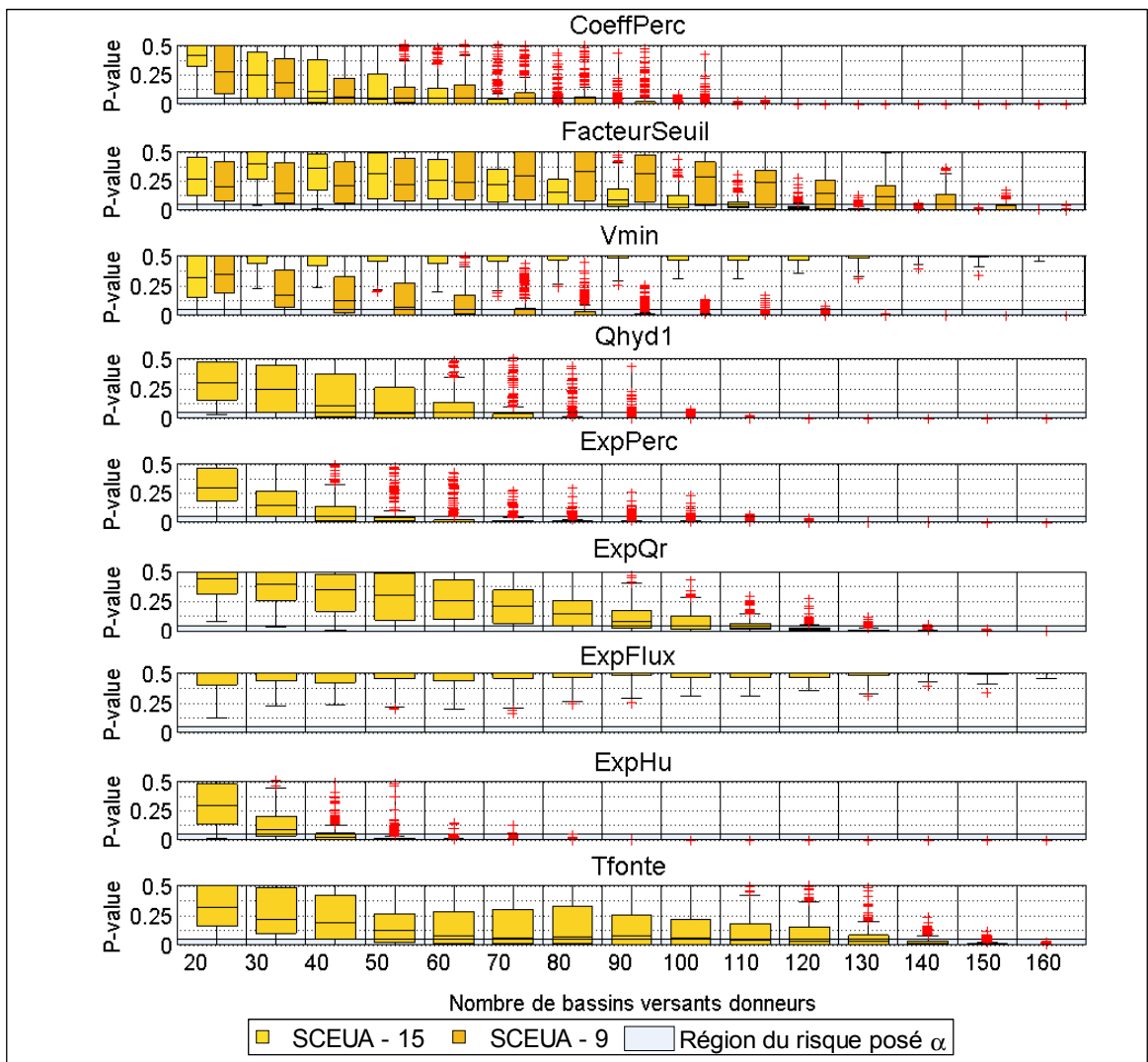
Figure-A III-1 Statistique  $F$  du test de Fisher



## ANNEXE IV

### SEUIL DESCRIPTIF *P-VALUE* DES MODÈLES RÉGIONAUX LINÉAIRES CONSTRUITS POUR LES MODÈLES HYDROLOGIQUES À 9 ET 15 PARAMÈTRES LIBRES

L'annexe 4 présente le seuil *P-value* des modèles régionaux linéaires construits avec quatre caractéristiques physiques pour les modèles hydrologiques à 9 et 15 paramètres libres. Elle fait suite à la figure 3.10 qui présente le seuil *P-value* pour les 6 premiers paramètres libres.



Figure–A IV–1 Seuil descriptif *P-value*



## ANNEXE V

### COMPARAISON DES ÉCHANTILLONS *GLOBAL* ET *DONNEUR* SUR LA BASE DU CRITÈRE D'EFFICIENCE $NSE_{MOYEN}$ DES MÉTHODES DE RÉGIONALISATION PAR PROXIMITÉ SPATIALE ET SIMILITUDE PHYSIQUE COUPLÉES AUX MÉTHODES D'INTERPOLATION DES DÉBITS

L'annexe 5 présente le critère d'efficacité NSE moyen obtenu par les trois configurations MH/OC optimisées à l'aide de l'algorithme d'optimisation SCE-UA en fonction de l'échantillon de bassins versants donneurs. Cette annexe est complémentaire au tableau 4.1.

Tableau–A V–1 Proximité spatiale – Interpolation par la moyenne arithmétique des débits

<b>Configuration MH/OC</b>	<b>Échantillon global (NSE)</b>	<b>Échantillon donneur (NSE)</b>	<b><math>\Delta</math> NSE</b>
SCEUA-15	0,6634	0,7500	0,0866
SCEUA-9	0,6583	0,7538	0,0955
SCEUA-6	0,6280	0,7271	0,0991

Tableau–A V–2 Proximité spatiale – Interpolation par la PID des débits

<b>Configuration MH/OC</b>	<b>Échantillon global (NSE)</b>	<b>Échantillon donneur (NSE)</b>	<b><math>\Delta</math> NSE</b>
SCEUA-15	0,6647	0,7492	0,0845
SCEUA-9	0,6641	0,7573	0,0931
SCEUA-6	0,6195	0,7169	0,0973

Tableau-A V-3 Similitude physique – Interpolation par la moyenne arithmétique des débits

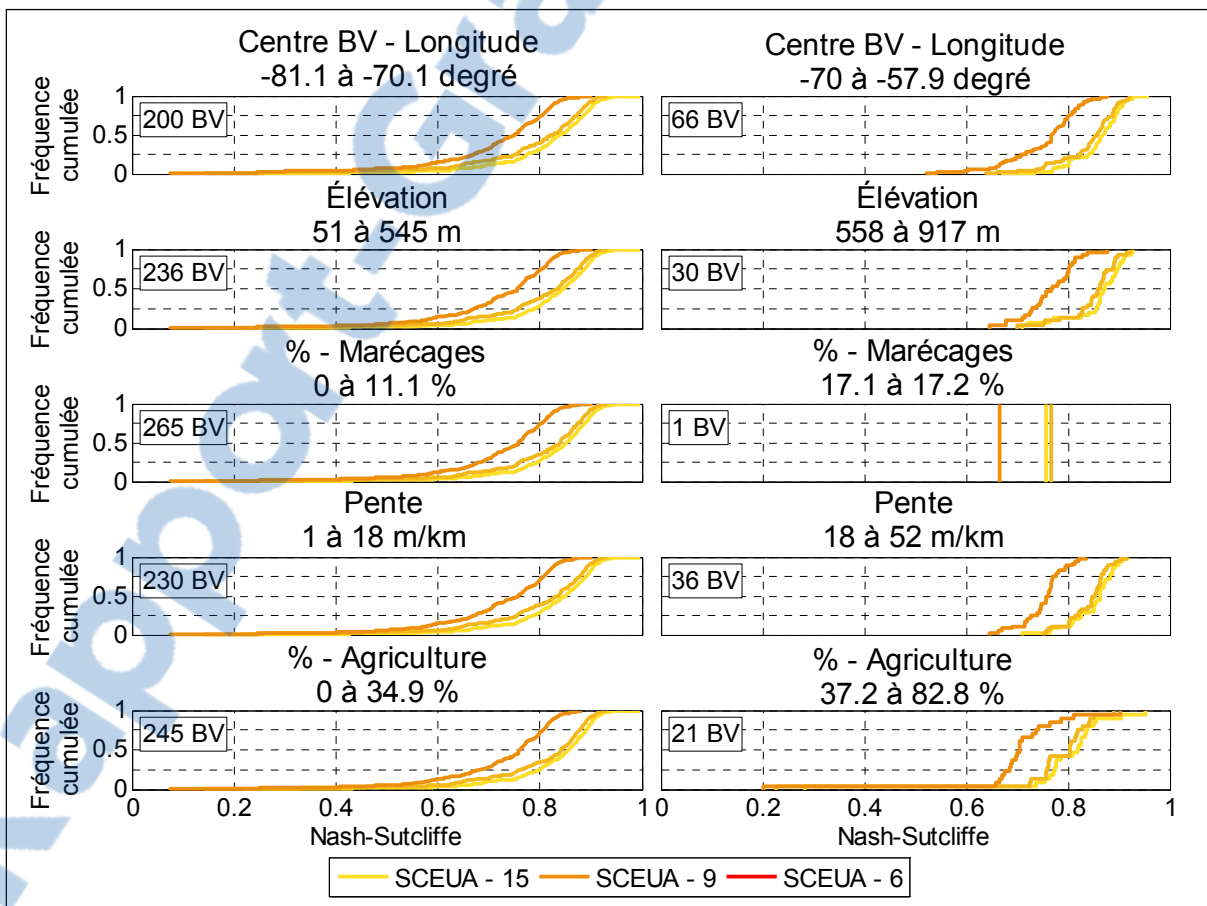
<b>Configuration MH/OC</b>	<b>Échantillon global (NSE)</b>	<b>Échantillon donneur (NSE)</b>	<b><math>\Delta</math> NSE</b>
SCEUA-15	0,6768	0,7643	0,0875
SCEUA-9	0,6728	0,7644	0,0915
SCEUA-6	0,6428	0,7377	0,0949



## ANNEXE VI

### PERFORMANCE EN CALIBRATION DES MODÈLES HYDROLOGIQUES À 6, 9 ET 15 PARAMÈTRES LIBRES SELON LES INTERVALLES DE VALEURS DES CARACTÉRISTIQUES PHYSIQUES

L'annexe 6 présente la performance en calibration des modèles hydrologiques à 6, 9 et 15 paramètres libres, exprimée selon les intervalles de valeurs des caractéristiques physiques. L'annexe présente les caractéristiques physiques pour lesquelles la performance des modèles hydrologiques à 9 et 15 paramètres libres est similaire et fait suite à la figure 4.2, qui présente celles pour lesquelles la performance est distincte pour ces deux modèles hydrologiques.



Figure–A VI–1 Performance des modèles hydrologiques à 6, 9 et 15 paramètres libres exprimée selon l'intervalle de valeurs des caractéristiques physiques



## BIBLIOGRAPHIE

- Andréassian, Vazken, Charles Perrin et Claude Michel. 2004. « Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models ». *Journal of Hydrology*, vol. 286, n° 1-4, p. 19-35.
- Arsenault, Richard, et François Brissette. 2014. « Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches ». *Water Resources Research*, p. n/a-n/a.
- Bao, Zhenxin, Jianyun Zhang, Jiufu Liu, Guobin Fu, Guoqing Wang, Ruimin He, Xiaolin Yan, Junliang Jin et Hongwei Liu. 2012. « Comparison of regionalization approaches based on regression and similarity for predictions in ungauged catchments under multiple hydro-climatic conditions ». *Journal of Hydrology*, vol. 466–467, n° 0, p. 37-46.
- Bárdossy, A. 2007. « Calibration of hydrological model parameters for ungauged catchments ». *Hydrol. Earth Syst. Sci.*, vol. 11, n° 2, p. 703-710.
- Beven, K. J. 1999. « Uniqueness of place and process representations in hydrological modelling ». *Hydrol. Earth Syst. Sci.*, vol. 4, n° 2, p. 203-213.
- Beven, K. J., et M. J. Kirkby. 1979. « A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant ». *Hydrological Sciences Bulletin*, vol. 24, n° 1, p. 43-69.
- Beven, Keith. 2006. « A manifesto for the equifinality thesis ». *Journal of Hydrology*, vol. 320, n° 1–2, p. 18-36.
- Brigode, Pierre, Ludovic Oudin et Charles Perrin. 2013. « Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? ». *Journal of Hydrology*, vol. 476, p. 410-425.
- Burn, Donald H., et David B. Boorman. 1993. « Estimation of hydrological parameters at ungauged catchments ». *Journal of Hydrology*, vol. 143, n° 3-4, p. 429-454.
- Duan, Q., J. Schaake, V. Andréassian, S. Franks, G. Goteti, H. V. Gupta, Y. M. Gusev, F. Habets, A. Hall, L. Hay, T. Hogue, M. Huang, G. Leavesley, X. Liang, O. N. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener et E. F. Wood. 2006. « Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops ». *Journal of Hydrology*, vol. 320, n° 1–2, p. 3-17.

- Duan, Qingyun, Soroosh Sorooshian et Vijai Gupta. 1992. « Effective and efficient global optimization for conceptual rainfall-runoff models ». *Water Resources Research*, vol. 28, n° 4, p. 1015-1031.
- Efron, B. 1979. « Bootstrap Methods: Another Look at the Jackknife ». p. 1-26.
- Fortin, Vincent. 2000. « Le modèle météo-apport HSAMI: historique, théorie et application ». *Institut de recherche d'Hydro-Québec, Varennes*.
- Gibbs, M. S., H. R. Maier et G. C. Dandy. 2012. « A generic framework for regression regionalization in ungauged catchments ». *Environmental Modelling and Software*, vol. 27-28, p. 1-14.
- Gupta, Hoshin V., Harald Kling, Koray K. Yilmaz et Guillermo F. Martinez. 2009. « Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling ». *Journal of Hydrology*, vol. 377, n° 1-2, p. 80-91.
- Hansen, Nikolaus, et Andreas Ostermeier. 1996. « Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation ». In *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation, ICEC'96, May 20, 1996 - May 22, 1996*. (Nagoya, Jpn), p. 312-317. Coll. « Proceedings of the IEEE Conference on Evolutionary Computation »: IEEE.
- He, Y., A. Bárdossy et E. Zehe. 2011. « A review of regionalisation for continuous streamflow simulation ». *Hydrol. Earth Syst. Sci.*, vol. 15, n° 11, p. 3539-3553.
- IPCC, 2013. *Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- IPCC, 2014. *Summary for policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.: Intergovernmental Panel on Climate Change.

- Kuczera, George, Benjamin Renard, Mark Thyer et Dmitri Kavetski. 2010. « There are no hydrological monsters, just models and observations with large uncertainties! ». *Hydrological Sciences Journal*, vol. 55, n° 6, p. 980-991.
- Lee, Hyosang, Neil R. McIntyre, Howard S. Wheater et Andy R. Young. 2006. « Predicting runoff in ungauged UK catchments ». *Proceedings of the Institution of Civil Engineers: Water Management*, vol. 159, n° 2, p. 129-138.
- McIntyre, Neil, Hyosang Lee, Howard Wheater, Andy Young et Thorsten Wagener. 2005. « Ensemble predictions of runoff in ungauged catchments ». *Water Resources Research*, vol. 41, n° 12, p. W12434.
- Merz, Ralf, et Günter Blöschl. 2004. « Regionalisation of catchment model parameters ». *Journal of Hydrology*, vol. 287, n° 1-4, p. 95-123.
- Michel, Claude, Charles Perrin et Vazken Andréassian. 2003. « The exponential store: a correct formulation for rainfall—runoff modelling ». *Hydrological Sciences Journal*, vol. 48, n° 1, p. 109-124.
- Nash, J. E., et J. V. Sutcliffe. 1970. « River flow forecasting through conceptual models part I — A discussion of principles ». *Journal of Hydrology*, vol. 10, n° 3, p. 282-290.
- Ouarda, Taha B.M.J. , André St-Hilaire et Bernard Bobée. 2008. « Synthèse des développements récents en analyse régionale des extrêmes hydrologiques ». *Revue des sciences de l'eau / Journal of Water Science*, vol. 21, n° 2, p. 219-232.
- Oudin, Ludovic, Vazken Andréassian, Cecile Loumagne et Claude Michel. 2006a. « How informative is land-cover for the regionalization of the GR4J rainfall-runoff model? Lessons of a downward approach ». *IAHS-AISH Publication*, n° 307, p. 246-255.
- Oudin, Ludovic, Vazken Andréassian, Charles Perrin, Claude Michel et Nicolas Le Moine. 2008. « Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments ». *Water Resources Research*, vol. 44, n° 3.
- Oudin, Ludovic, Alison Kay, Vazken Andréassian et Charles Perrin. 2010. « Are seemingly physically similar catchments truly hydrologically similar? ». *Water Resources Research*, vol. 46, n° 11, p. W11558.
- Oudin, Ludovic, Charles Perrin, Thibault Mathevet, Vazken Andréassian et Claude Michel. 2006b. « Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models ». *Journal of Hydrology*, vol. 320, n° 1-2, p. 62-83.

- Parajka, J., R. Merz et G. Blöschl. 2005. « A comparison of regionalisation methods for catchment model parameters ». *Hydrol. Earth Syst. Sci. Discuss.*, vol. 2, n° 2, p. 509-542.
- Parajka, Juraj, Günter Blöschl et Ralf Merz. 2007. « Regional calibration of catchment models: Potential for ungauged catchments ». *Water Resources Research*, vol. 43, n° 6, p. W06406.
- Peel, M. C., et G. Bloschl. 2011. « Hydrological modelling in a changing world ». *Progress in Physical Geography*, vol. 35, n° 2, p. 249-261.
- Perrin, C., C. Michel et V. Andréassian. 2001. « Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments ». *Journal of Hydrology*, vol. 242, n° 3-4, p. 275-301.
- Perrin, Charles, Claude Michel et Vazken Andréassian. 2003. « Improvement of a parsimonious model for streamflow simulation ». *Journal of Hydrology*, vol. 279, n° 1-4, p. 275-289.
- Razavi, Tara, et Paulin Coulibaly. 2013. « Streamflow prediction in ungauged basins: Review of regionalization methods ». *Journal of Hydrologic Engineering*, vol. 18, n° 8, p. 958-975.
- Samuel, Jos, Paulin Coulibaly et Robert A. Metcalfe. 2011. « Estimation of continuous streamflow in ontario ungauged basins: Comparison of regionalization methods ». *Journal of Hydrologic Engineering*, vol. 16, n° 5, p. 447-459.
- Seiller, G., F. Anctil et C. Perrin. 2011. « Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions ». *Hydrol. Earth Syst. Sci. Discuss.*, vol. 8, n° 6, p. 10895-10933.
- Senbeta, D. A., A. Y. Shamseldin et K. M. O'Connor. 1999. « Modification of the probability-distributed interacting storage capacity model ». *Journal of Hydrology*, vol. 224, n° 3-4, p. 149-168.
- Shepard, Donald. 1968. « A two-dimensional interpolation function for irregularly-spaced data ». In *Proceedings of the 1968 23rd ACM national conference*. p. 517-524. 810616: ACM.
- Sivapalan, M., K. Takeuchi, S. W. Franks, V. K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J. J. McDonnell, E. M. Mendiondo, P. E. O'Connell, T. Oki, J. W. Pomeroy, D. Schertzer, S. Uhlenbrook et E. Zehe. 2003. « IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences ». *Hydrological Sciences Journal*, vol. 48, n° 6, p. 857-880.

- Smakhtin, V. U. 2001. « Low flow hydrology: a review ». *Journal of Hydrology*, vol. 240, n° 3–4, p. 147-186.
- Valéry, A. 2010. « Modélisation précipitations – débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants. ». Paris, AgroParisTech, 405 p.
- Wagener, Thorsten, et Howard S. Wheater. 2006. « Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty ». *Journal of Hydrology*, vol. 320, n° 1–2, p. 132-154.
- Yadav, Maitreya, Thorsten Wagener et Hoshin Gupta. 2007. « Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins ». *Advances in Water Resources*, vol. 30, n° 8, p. 1756-1774.
- Yapo, Patrice O., Hoshin Vijai Gupta et Soroosh Sorooshian. 1996. « Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data ». *Journal of Hydrology*, vol. 181, n° 1–4, p. 23-48.
- Zhang, Yongqiang, et Francis H. S. Chiew. 2009. « Relative merits of different methods for runoff predictions in ungauged catchments ». *Water Resources Research*, vol. 45, n° 7, p. W07412.

