

**Université Pierre et Marie Curie, École des Mines de Paris  
& École Nationale du Génie Rural des Eaux et des Forêts**

---

**Master 2 Sciences de l'Univers, Environnement, Ecologie  
Parcours Hydrologie-Hydrogéologie**

**Tentative d'amélioration des techniques de régionalisation des  
modèles hydrologiques pluie-débit**

**Abderrahmen KHALIFA**

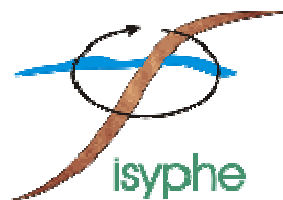
**Encadré par : Vazken ANDREASSIAN (Cemagref Antony)**

**Ludovic OUDIN (UMR-Sisyphe)**



**Cemagref, UR-HBAN**

**1 rue Pierre Gilles de Gennes  
92761 Antony Cedex, France**



**UMR-Sisyphe**

**4 place Jussieu  
75252 Paris Cedex 05, France**

**Septembre 2011**



## **REMERCIEMENTS**

*Pour avoir eu la chance d'effectuer un stage extrêmement enrichissant, et sur un sujet particulièrement intéressant, je tiens à remercier...*

*Mes maîtres de stage, Vazken Andréassian (chef d'équipe HBAN au Cemagref Antony) et Ludovic Oudin (maître de conférence à l'UPMC paris 6), qui m'ont donné de précieux conseils tout au long de ce stage. Leur disponibilité, leur patience et le partage de leurs connaissances m'ont permis d'étudier des voies de recherche particulièrement intéressantes pour la modélisation hydrologique. Ils ont su se rendre disponibles quand j'avais besoin d'eux et ils ont toujours su répondre avec pédagogie aux questions que je me posais. Leur regard et leur expérience m'ont été d'une grande aide.*

*Charles Perrin et Maria-Helena Ramos pour leurs conseils et leur aide, notamment en matière de programmation.*

*Je tiens par ailleurs à exprimer ma reconnaissance à l'ensemble des équipes Hydro du Cemagref d'Antony et de Sisyphe, pour leur enthousiasme et leur accueil au sein du service.*

*L'ASCCR pour les parties de frisbee enflammées et disputées.*

*Merci également pour l'ambiance plus qu'agréable qui règne au Cemagref au sein de l'équipe.*

*Enfin, merci aux enseignants de la formation à l'Université Pierre et Marie Curie.*

*Ce stage m'a apporté bien plus que des compétences scientifiques.*

## Dédicaces

*A mes chers parents Qui ont tant attendu ma  
réussite,  
Et qui ont peiné et sacrifié pour moi durant ces  
dernières années.  
Que Dieu les garde pour moi.*

*A toute ma famille Qui m'a entouré  
d'encouragement et d'amour*

*A ma sœur Sondes qui m'a soutenu jusqu'au bout*

## Table des matières

1	Introduction .....	7
2	Synthèse bibliographique .....	8
2.1	Qu'est-ce qu'un bassin versant non jaugé ? .....	8
2.2	Quelle approche pour les bassins non jaugés ?.....	8
2.3	Classification des techniques de régionalisation? .....	8
2.3.1	Classification fondée sur l'information utilisée .....	8
2.3.2	Classification fondée sur le mode d'estimation des paramètres de modèle .....	9
2.4	La méthode de transposition des paramètres .....	10
2.4.1	Principe général.....	10
2.4.2	Choix des bassins donateurs.....	10
2.4.3	Synthèse des descripteurs utilisés .....	10
2.4.4	Combinaison des paramètres.....	11
2.5	Discussions des résultats de différentes approches .....	12
2.6	Positionnement de notre étude.....	13
3	Les modèles hydrologiques utilisés.....	14
3.1	Introduction .....	14
3.2	Le modèle GR4J .....	15
3.3	Le modèle TOPMO .....	15
3.4	Le modèle HBV0.....	16
3.5	Le modèle IHA .....	17
3.6	Critères d'évaluation.....	17
4	Base de données .....	19
4.1	Les données de modélisation.....	19
4.2	Traitement des données hydrométéorologiques .....	20
4.2.1	Origine et différents types d'erreurs.....	20
4.2.2	Traitement des données de débit .....	22
4.3	Influence de la procédure de traitement sur les résultats de calage.....	25
4.4	Collecte et traitement des descripteurs physico-climatique .....	26
4.4.1	Les descripteurs morphologiques (base de données SRTM) .....	26
4.4.2	Les descripteurs Climatiques (Base de données SAFRAN) .....	27
4.4.3	Les descripteurs pédologiques : (Base de données ESDB).....	27
4.4.4	Les descripteurs d'occupation de sol (base de données ECOCLIMAP).....	29
4.4.5	Étude de la corrélation entre les différents descripteurs.....	30
5	Régionalisation des modèles hydrologiques .....	32
5.1	La régionalisation par proximité géographique.....	32
5.1.1	Comment définit-on la similarité géographique?.....	32

5.1.2	Méthode d'évaluation.....	33
5.1.3	Résultats de la régionalisation spatiale.....	33
5.1.4	La régionalisation dépend-elle de type de modèle adopté .....	36
5.1.5	Conclusion relative.....	36
5.2	La régionalisation par similarité physique.....	36
5.2.1	Comment définit-on la similarité physique? .....	36
5.2.2	Différentes approches envisagées pour l'étude la similarité physiques .....	37
5.2.3	Résultats et discussion.....	38
5.2.4	Optimisation des poids des descripteurs pris en compte.....	41
5.2.5	Synthèse .....	43
5.3	Combinaison de la similarité physique et de la proximité spatiale .....	43
6	Conclusion.....	46
7	Référence Bibliographique.....	47

## Liste des figures

<b>Figure 1</b> : Schéma conceptuel du modèle GR4J et les quatre paramètres calés.....	15
<b>Figure 2</b> : Schéma conceptuel et le calcul effectué au sein du modèle TOPMO .....	16
<b>Figure 3</b> : Schéma conceptuel et le calcul effectué au sein du modèle HBV0.....	16
<b>Figure 4</b> : Schéma conceptuel et le calcul effectué au sein du modèle IHA .....	17
<b>Figure 5</b> : Relation entre le critère NSE et le critère C2M.....	18
<b>Figure 6</b> : Carte représentant les 1056 bassins étudiés en France .....	19
<b>Figure 7</b> : Etapes de la chaîne d'acquisition et de traitement des données .....	20
<b>Figure 8</b> : Erreur due au changement d'observateur .....	21
<b>Figure 9</b> : Erreur due à l'interpolation des données de débits .....	22
<b>Figure 10</b> : Les séquences des débits interpolés sont en rouge et en vert .....	23
<b>Figure 11</b> : Exemple de traitement des séries de débit : a) chronique avant traitement, b) chronique après traitement .....	24
<b>Figure 12</b> : Répartition de 973 bassins .....	25
<b>Figure 13</b> : Résultats de calage avec le modèle GR4J en considérant les données brutes et traités .....	26
<b>Figure 14</b> : Répartition des 10 points dans le bassin utilisé par la distance géographique .....	32
<b>Figure 15</b> : Résultats de régionalisation par proximité géographique avec le modèle GR4J ..	33
<b>Figure 16</b> : Résultats d'optimisation de la proximité géographiques.....	35

## Liste des tableaux

<b>Tableau 1</b> : Classification des descripteurs suivant le domaine d'étude.....	11
<b>Tableau 2</b> : Calage des modèles avec les données traitées et les données brutes .....	25
<b>Tableau 3</b> : Descripteurs morphologiques calculés à partir de la base de données SRTM .....	27
<b>Tableau 4</b> : Descripteurs Climatiques calculés à partir de la base de données SAFRAN .....	27
<b>Tableau 5</b> : Liste des descripteurs pédologiques utilisés .....	28
<b>Tableau 6</b> : Les descripteurs d'occupation de sol .....	29
<b>Tableau 7</b> : Les descripteurs Climato-physiques .....	30
<b>Tableau 8</b> : La gamme des valeurs des paramètres testées .....	35
<b>Tableau 9</b> : Récapitulatif des résultats de la régionalisation spatiale.....	36
<b>Tableau 10</b> : Résultats d'évaluation de la première et de la deuxième méthode .....	39
<b>Tableau 11</b> : Résultats de la régionalisation par similarité physique avec la troisième méthode .....	39
<b>Tableau 12</b> : Descripteurs sélectionnés comme les plus performants pour la similarité physique avec la méthode 3.....	40
<b>Tableau 13</b> : Comparaison des résultats de la proximité géographique avec les similarités physiques .....	41
<b>Tableau 14</b> : Optimisation de poids de six descripteurs physiques .....	42
<b>Tableau 15</b> : Optimisation de poids de dix descripteurs physiques pour le modèle GR4J .....	42
<b>Tableau 16</b> : Optimisation de poids de dix descripteurs physiques pour le modèle TOPMO ..	42
<b>Tableau 17</b> : Optimisation de poids de dix descripteurs physiques pour le modèle IHA .....	42
<b>Tableau 18</b> : Résultats de la troisième méthode sans et avec optimisation de poids de descripteurs physiques des bassins.....	43
<b>Tableau 19</b> : Tableau récapitulatif des résultats de la proximité spatiale et de la similarité physique .....	43
<b>Tableau 20</b> : Résultats de la deuxième approche de combinaisons de la similarité physique et de la proximité géographique .....	45

## **Abstract**

Hydrological models are in general calibrated using measured flows, for which we try to match the simulated flows.

This procedure is impossible for ungaged basins, for which there is no flow measurements observation. In this case, we use a so-called regionalization approach that allows the transfer of parameters determined on gaged basins to ungaged basins. These methods are based on the similarity of basins' behavior. The difficulty of such approaches is the determination of this similarity. In order to draw some general conclusions on this issue, we have followed a three step methodology:

First, we have prepared a synthesis of previous works focus on the regionalization of hydrological models. These works allowed us to define the general context of this study, analyze and diagnose the existing methods of regionalization and propose new perspective on this issue.

Second, we prepared, a large catchments set on the French territory. On this set, we applied an automatic procedure to detect interpolated flow. Then we collect physical descriptors of those catchments'. Most of those descriptors are internationally available, in order to overcome the limitation of data heterogeneity for further international studies. The data base encompasses a large number of descriptors to explain watersheds hydrological response.

Third, several regionalization approaches are tested and developed. Some are based on physical similarity between catchments and others are based on spatial proximity. Results show that physical similarity approach outperforms the spatial proximity approach, which is a novel conclusion on the regionalization issue.

# 1 Introduction

L'hydrologie est une science qui étudie le cycle de l'eau dans tous ses aspects, depuis l'évaporation jusqu'à la formation et la propagation des écoulements, aussi bien pour le domaine superficiel que souterrain. Cependant, cette science est aussi chargée d'apporter des réponses concrètes aux demandes des aménageurs et des responsables locaux.

Avant la réalisation de toute étude, les hydrologues commencent par la récolte des mesures et des observations de certains phénomènes physiques ayant trait à l'eau, ceci afin d'étudier les problèmes plus en amont, et tenter d'aboutir à une analyse de la dynamique qui les sous-tend. Cependant, on ne possède actuellement qu'une compréhension partielle des mécanismes physiques de circulation de l'eau à l'échelle du bassin versant.

Aux problèmes liés à l'incompréhension des processus hydrologique à l'échelle du bassin versant, s'ajoute celui de la mesure et de la collecte des données : dans la majorité des cas, les bassins étudiés sont soit non jaugés, soit partiellement jaugés, ce qui accroît encore la complexité des études hydrologiques. C'est pour cela que la plupart des hydrologues s'orientent vers des approches stochastiques de modélisation, tentant d'exploiter au mieux les séries des données collectées, et cherchant à remplir les vides causés par le manque de données tout en se basant sur diverses méthodologies. **C'est dans ce cadre s'intègre notre sujet d'étude concernant les bassins versants non-jaugés**, il porte sur le thème de la régionalisation des modèles hydrologiques, et s'appuie sur une vaste base de données de bassins versants français.

Le terme régionalisation recouvre un ensemble de méthodes qui autorisent le transfert de paramètres optimisés sur un bassin versant jaugé vers un autre non jaugé, sous hypothèse de similarité des comportements des bassins versants, qui dépendra :

- de la proximité géographique ;
- de la similarité physiographique (superficie, pente topographique, longueur de cours d'eau, occupation du sol, etc.) et climatique (pluviométrie, température de l'air, etc.).

En transférant un jeu des paramètres calés à partir d'un bassin versant jaugé, on peut simuler le débit d'un bassin versant non jaugé. Cependant, on doit tenir compte de l'unicité des comportements pour chaque bassin versant. On doit donc essayer, dans la mesure du possible, d'adapter ces paramètres pour tenir compte des différences entre le bassin jaugé 'donneur' et le bassin non jaugé 'receveur', soit par transformation de ces paramètres (facteurs multiplicatifs, coefficients de pondérations, etc.) ou en combinant plusieurs jeux de paramètres issus d'un calage sur des bassins jaugés similaires.

## **Organisation du mémoire**

Ce mémoire s'organise en trois parties : la première partie est introductive et vise à préciser le cadre général de l'étude ainsi que la base bibliographique sur laquelle nous allons nous baser durant cette étude. La deuxième partie présente l'échantillon de données utilisées pour répondre aux objectifs et aux ambitions espérées de ce travail de régionalisation des modèles hydrologiques. La troisième partie concerne l'étude de la régionalisation proprement dites avec ces différents volets ainsi que les principaux résultats trouvés (interprétation et discussion).



## 2 Synthèse bibliographique

Dans ce qui suit, nous analysons et commentons certains articles publiés sur notre sujet d'étude permettant d'exposer le contexte général de notre étude et de définir les mots clés qu'il le constitue : bassin non jaugé, modèles hydrologiques, régionalisation, etc.

### 2.1 Qu'est-ce qu'un bassin versant non jaugé ?

Un bassin versant non jaugé est un bassin dont les données hydrologiques mesurées sont considérées comme insuffisantes (autant au niveau qualité, que quantité) pour permettre le calcul des variables hydrologiques d'intérêt à des échelles spatiale et temporelle appropriées, et à une précision acceptable pour des applications pratiques (Sivapalan *et al.*, 2003). Nous ferons l'hypothèse que des données météorologiques existent toujours. Donc, un bassin versant sera considéré "non jaugé" quand les mesures débitométriques à son exutoire sont manquantes et/ou absentes.

### 2.2 Quelle approche pour les bassins non jaugés ?

Les outils développés par les hydrologues ces dernières décennies sont pour la plupart basés sur l'utilisation de données historiques à l'échelle du bassin versant. L'application de ces outils sur les bassins non jaugés suppose alors que les informations recueillies sur un bassin sont utilisables pour estimer la réponse hydrologique d'un autre bassin. Cependant, les modèles développés restent encore soit trop sophistiqués en termes d'alimentation en données observés, soit trop spécifiques, avec des hypothèses restrictives et difficiles à transposer à d'autres situations, en particulier lorsque les modèles sont élaborés et/ou paramétrés grâce au calage avec les données historiques de bassins jaugés.

En absence de données sur un site donné, les approches les plus utilisées sont basées sur l'estimation des variables hydrologiques à partir des données provenant d'autres stations, jugées "similaires" ou appartenant à une même région climatiquement et/ou hydrologiquement "homogène". L'information régionale se substitue, ainsi, à l'information locale ; on parle alors de "régionalisation". La principale difficulté pour la simulation des débits des bassins versants non jaugés est donc de définir cette notion de similarité entre les bassins.

### 2.3 Classification des techniques de régionalisation?

#### 2.3.1 Classification fondée sur l'information utilisée

Pour un bassin non jaugé, les informations hydrologiques que l'on peut exploiter sont celles de bassins versants jaugés considérés comme "similaires". Deux grandes tendances apparaissent dans la littérature :

La régionalisation basée sur des critères géographiques: (Egbuniwe et Todd, 1976; Vandewiele et al, 1991)

Cette approche est une des plus simples, elle fait l'hypothèse que des bassins proches ont de fortes chances d'avoir une réponse hydrologique semblable, car les caractéristiques physiques et climatiques sont relativement homogènes dans une même région. (Merz et Blöschl, 2004) et (Oudin *et al.*, 2008) ont comparé plusieurs méthodes pour identifier les bassins similaires et ont montré que la régionalisation basée sur la proximité géographique donnait les meilleurs résultats.

La régionalisation selon des similarités physiographiques et climatiques : (Acreman and Sinclair, 1986; McIntyre et al, 2005)

Cette approche a pour objectif de transposer des informations d'un bassin jaugé vers un bassin cible non jaugé, considéré comme suffisamment similaire physio-graphiquement (Oudin *et al.*, 2010). Les caractéristiques des bassins versants prises en compte sont, dans ce cas, considérées comme attributs représentatifs des processus hydrologiques.

Une approche combinée :

Une troisième approche consiste à combiner les critères géographiques et les attributs des bassins versants. Chaque site est considéré comme le centre de sa propre région ou zone d'influence. L'identification de cette zone d'influence se base sur la mesure de la distance euclidienne dans un espace multidimensionnel de plusieurs attributs des bassins versants. Le donneur (ou le groupe de donneurs) est celui parmi les voisins qui ressemble le plus au bassin non jaugé. Une fonction de poids peut être définie pour refléter les importances relatives de chaque site par rapport au site cible non-jaugé (Hundecha et Bardossy, 2004).

### 2.3.2 Classification fondée sur le mode d'estimation des paramètres de modèle

Après avoir répondu à la question d'identification des similarités hydrologiques et qui occupe la première partie de la méthodologie d'estimation régionale (ou régionalisation), nous nous intéressons au transfert de l'information hydrologique d'un site jaugé à un site non-jaugé. Dans la littérature, différents modes d'estimation des paramètres de modèle ont été envisagés, parmi lesquels on cite :

Méthode de régression : (Magette *et al.*, 1976) ; (Young, 2006)

La méthode la plus utilisée est la régression, c'est-à-dire le développement de relations a posteriori entre, d'une part, les caractéristiques des bassins versants et, d'autre part, la variable à expliquer (le plus souvent, il s'agira des valeurs des paramètres calés des modèles hydrologiques sur les sites jaugés).

Une fois la relation établie, elle est utilisée pour déterminer les paramètres d'un bassin non jaugé à partir de ses descripteurs physiographiques et climatiques. Cette méthode repose sur deux hypothèses sous-jacentes :

- Hypothèse n°1 : il est considéré qu'une relation simple existe entre les caractéristiques observables du bassin et les paramètres du modèle ;
- Hypothèse n°2 : il est admis que les descripteurs physiographiques et climatiques choisis pour la régression fournissent des informations pertinentes pour identifier le comportement du bassin.

### Méthode de transposition :

C'est une approche dite analogique qui consiste en une transposition de l'information obtenue sur un ou plusieurs bassins jaugés similaires. Il peut s'agir d'une similarité physique (à définir) ou d'une proximité géographique. La valeur de paramètre recherché est alors la moyenne pondérée des valeurs des paramètres des bassins similaires, le poids de chaque bassin étant relatif à son degré de similarité ou à sa proximité géographique (les bassins les plus proches ont donc plus de poids). Les informations transposées peuvent être les paramètres d'un modèle hydrologique : le jeu de paramètres d'un bassin jugé similaire est transposé tel quel au bassin non-jaugé (Parajka *et al.*, 2005).

### Méthode géostatistique :

Cette approche consiste à utiliser des méthodes d'estimation régionale basées sur la géostatistique (Merz *et al.*, 2000). Cette approche est souvent utilisée pour l'interpolation spatiale des données climatiques (les précipitations moyennes ou maximales, par exemple), mais peut également être appliquée à des variables hydrologiques : débit moyen annuel (Sauquet, 2006), les paramètres des modèles (Parajka *et al.*, 2005), les indices de basses eaux (Haberlandt *et al.*, 2000) et (Castiglioni *et al.*, 2009).

L'avantage de méthodes géostatistiques est qu'elles offrent une meilleure estimation linéaire sans biais, mais l'inconvénient demeure dans la difficulté de bien capturer la structure spatiale imposée par des conditions climatiques et géologiques des variables hydrologiques.

## **2.4 La méthode de transposition des paramètres**

### **2.4.1 Principe général**

Dans le domaine de la régionalisation des modèles hydrologiques, la plupart des modélisateurs s'appuie sur le principe de sélection des bassins versants donneurs et de transposition des paramètres de ces bassins vers le bassin non jaugé dit « receveurs ». Dans ce mémoire la sélection est basée sur le principe suivant : chaque bassin constituant notre base de données va être successivement considéré comme non jaugé, les autres bassins sont jaugés.

### **2.4.2 Choix des bassins donneurs**

Pour optimiser le choix de sélection de bassins versants donneurs, les modélisateurs ont envisagé différents approches pour cet objectif, parmi lesquels on cite :

- ✓ L'approche décrite par Kay et al en 2007 qui définit la similarité entre les bassins comme étant la distance euclidienne dans l'espace des caractéristiques physiographiques et climatiques des bassins.
- ✓ L'approche décrite par Oudin et al en 2008 qui consiste à classer les bassins donneurs (jaugés) par ordre décroissant de proximité au bassin versant receveur (non-jaugé), cette proximité dépend de l'approche de régionalisation adaptée. Donc chaque bassin doit avoir un rang qui exprime leur similarité par rapport au bassin non jaugé.

### **2.4.3 Synthèse des descripteurs utilisés**

Différents études de régionalisation à grande échelle ont été conduites, chaque étude ayant ses spécificités et les résultats de ces études sont assez disparates. Ces différences sont dus d'une

part à la variation du domaine d'étude et d'autre part à l'hétérogénéité des données disponibles utilisées pour chaque étude. Si on met l'accent sur le nombre et les types de descripteurs physiques utilisés pour chaque étude, on peut donc remplir le tableau de classification suivante.

**Tableau 1:** Classification des descripteurs suivant le domaine d'étude

Référence	Sélection		Mode de sélection		Nombre des descripteurs utilisés	Liste de descripteurs			
	Oui	Non	A priori	Optimisé		Type1	Type2	Type3	Type4
Merz et al. 2009		x	X		18	6	6	3	1
Reichl et al. 2009	x			x	25	16	3	4	2
Oudin et al. 2008	x		X		16	4	1	0	1
Parajka et al. 2007	x		X		18	9	6	2	1
McIntyre et al. 2005		X	X		6	1	3	1	1

Type de descripteur n° 1 : Descripteurs géomorphologiques

Type de descripteur n° 2 : Descripteurs climatiques

Type de descripteur n° 3 : Descripteurs pédologiques

Type de descripteur n° 4 : Descripteurs occupation de sol

Ce tableau récapitulatif des descripteurs physiques montre que, dans leurs études, les auteurs ont mis l'accent sur les données morphologiques et climatiques des bassins versants, alors qu'on trouve peu de descripteurs pédologique et d'occupation de sol même si ces descripteurs influence a priori la réponse hydrologique des bassins versants. Ce tableau montre également la disparité des études, d'un point de vu des descripteurs physiques utilisés. Plus de détails, concernant les descripteurs utilisés par ces auteurs, se trouvent dans l'annexe n°1.

#### 2.4.4 Combinaison des paramètres

Comme nous avons évoqués précédemment que l'approche de transposition des paramètres consiste à permuter les paramètres de modèle entre le bassin non jaugé et ceux jaugés, les questions qui méritent avoir des réponses à ce niveau sont : quel est le nombre optimum de bassins donneurs qui doivent être mis en jeu pour ce transfert de paramètres ? Et quel est le mode de transfert le plus approprié afin d'avoir les meilleurs résultats ?

##### ✓ Choix de nombre des bassins donneurs

Oudin et al. (2008) ont montré que le choix de nombre de bassins versants donneurs pour la régionalisation dépend de type de l'information utilisée, en effet, ils ont prouvé que ce nombre est optimisé à 5 pour la régionalisation par proximité géographique, alors que ce nombre peut atteindre les 20 bassins dans le cas de la régionalisation par proximité physique.

##### ✓ Mode de transfert des paramètres de modèle

Dans les études de la régionalisation, deux modes de transfert de paramètres ont été envisagés: soit un transfert direct avec des poids égaux entre tous les bassins donateurs, soit un transfert pondéré, c'est-à-dire le poids de chaque bassin donneur étant relatif à son degré de similarité ou à sa proximité du bassin receveur (non jaugé). Les équations suivantes expliquent ces deux modes de transferts.

Equation pour le mode de transfert direct :

$$Q_r(j) = \frac{1}{m} * \sum_{i=1}^m Q_{pi}(j) \quad \text{Équation 1}$$

Equation pour le mode de transfert pondéré :

$$Q_r(j) = \frac{1}{\sum W_k} * \sum_{i=1}^m W_k * Q_{pi}(j) \quad \text{Équation 2}$$

$$W_k = \frac{1}{D^{\delta}} \quad \text{Équation 3}$$

Avec :

$Q_d$ : Débit du bassin donneur ;

$Q_r$ : Débit du bassin receveur ;

$D$  : Distance calculé pour la régionalisation : (soit distance spatiale entre bassins, soit distance entre les descripteurs physiques des bassins);

$m$ : Nombre des bassins donateurs ;

$\delta$ : Indique la puissance appliquée sur l'inverse de la distance.

## 2.5 Discussions des résultats de différentes approches

Les différentes approches évoquées ci-dessus ont produit des résultats variés, cependant, on peut avancer plusieurs explications à cela : d'une part, les études ont été menées sur des ensembles de bassins versants différents, et même si les échantillons considérés sont assez grands, les conclusions peuvent encore être influencées par l'échantillon utilisé. Deuxièmement, comme l'ont souligné Merz *et al.*, (2006), les descripteurs relatifs aux bassins versants disponibles diffèrent d'une étude à une autre. Troisièmement, les méthodologies utilisées dans chaque type d'approche dépendent d'un certain nombre de choix arbitraires faits par les auteurs, qui peuvent influencer la performance du modèle régionalisé (nombre de bassins versants donateurs, la densité des bassins versants utilisés, l'utilisation de bassins versants mal modélisés comme bassins donateurs). Enfin, comme en témoignent Kay *et al.*, (2006), la performance de chaque approche de régionalisation peut dépendre de la structure et des paramètres des modèles pluie-débit utilisé.

## **2.6 Positionnement de notre étude**

Vu ces différentes raisons de contradiction entre les approches de régionalisation, notre objectif consiste à explorer de nouvelles voies d'estimation des paramètres des modèles dans un contexte de bassins versants non jaugés. En effet les travaux antérieurs sur ce sujet ont montré que les bassins apparemment physiquement similaires ne le sont pas forcément de point de vue comportemental. Ils préconisent donc, de tester les méthodes sur des échantillons plus contrastés et de développer de nouveaux indices de similarité. C'est sur ce dernier point que porte le stage. En effet, la première partie de ce stage s'est déroulée en deux étapes : la première étape a été consacrée à la constitution de l'échantillon de bassins versants, en développant notamment une technique automatique de vérification des données de débits, alors que la deuxième étape concerne la collecte et le traitement des données physico-climatiques disponible à l'échelle internationale. La deuxième partie du stage, a été consacrée à l'étude de la régionalisation des modèles pluie-débit, qui est constituée de trois axes. Le premier axe concerne la régionalisation par proximité géographique et la recherche d'une distance géographique entre les bassins versants appropriée à la régionalisation, le deuxième axe concerne la régionalisation par proximité physique ainsi que l'identification des descripteurs physiques propice pour exprimer la similarité hydrologique entre les bassins versants, alors que le troisième axe de régionalisation rassemble la proximité géographique, physique et climatique.

### 3 Les modèles hydrologiques utilisés

#### 3.1 Introduction

Les modèles hydrologiques sont aujourd'hui couramment utilisés, non seulement comme instruments de recherche mais également dans la pratique de l'ingénieur hydrologue, que ce soit pour la conception, la gestion ou les études d'impact d'un aménagement. Pratiquement, le choix d'un modèle hydrologique est toujours un compromis, qui doit prendre en compte l'objectif visé, les données disponibles et les possibilités de mise en œuvre à un coût raisonnable. Pour l'hydrologue praticien, tel modèle global peut sembler trop frustré pour l'étude qu'il envisage. Tel autre modèle maillé lui posera souvent des problèmes de données et de mise en œuvre insurmontables, ou inutilement compliqués pour son objectif précis. Donc le choix du modèle convenable à l'étude permet une optimisation des moyens à mettre en œuvre pour parvenir au but fixé. Pour cette étude nous avons fixé le choix sur les modèles conceptuels globaux pluie-débit tels que GR4J, TOPMO, HBV et IHA.

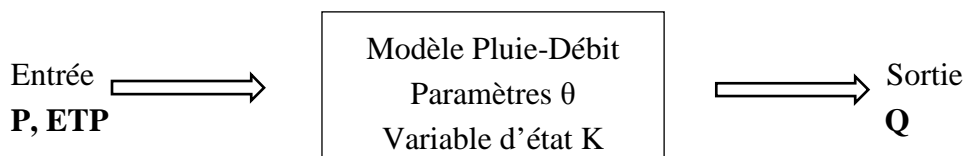
Les modèles Pluie-Débit sont des modèles hydrologiques qui permettent de simuler le débit des rivières à partir de l'observation des précipitations sur le bassin versant. Les objectifs de cette modélisation sont variés et on peut notamment citer :

- ✓ La simulation des débits dans le cas de lacunes de débits, cela permet d'étendre les chroniques à partir de celles de pluie ;
- ✓ La prévision en temps réel des débits des cours d'eau notamment pour la prévision des crues ou des étiages ;
- ✓ La prédétermination des débits de crue ou d'étiage : par exemple les valeurs de débits d'une crue de période de retour de 100 ans, notamment dans la perspective de dimensionner des ouvrages et des aménagements dans le lit des cours d'eau ;
- ✓ La détermination des conséquences de changement climatique sur l'hydrologie des rivières.

Dans cette étude, nous nous limiterons à l'étude des modèles pluie-débit dans le contexte de la simulation (premier objectif de modélisation).

#### Principe de base de la modélisation Pluie-Débit

La majorité des modèles Pluie-Débit fonctionne selon le principe suivant :



Les variables d'entrée du modèle, encore appelées variables de forçage sont le plus souvent des mesures de la pluie (P) et des estimations de l'évapotranspiration potentielle (ETP).

Les variables de sortie sont les débits (Q). Le modèle dépend de variables d'état qui traduisent les caractéristiques du système modélisé (par exemple le taux de saturation en eau du sol ; la profondeur de sol ; la pente...) et de paramètres représentant des caractéristiques du bassin qui n'ont pas d'interprétation physique directe (Gaume *et al.*, 2004).

### 3.2 Le modèle GR4J

Le modèle GR4J (Perrin *et al.*, 2003) est un modèle hydrologique conceptuel global fonctionnant au pas de temps journalier. Le nombre de paramètres est limité à quatre, ce qui en fait un modèle parcimonieux. Ce modèle a été développé sur la base d'un grand nombre de bassins versants, incluant des bassins versants à climats contrastés. La figure ci-dessous indique le schéma conceptuel du modèle ainsi que les quatre paramètres à caler.

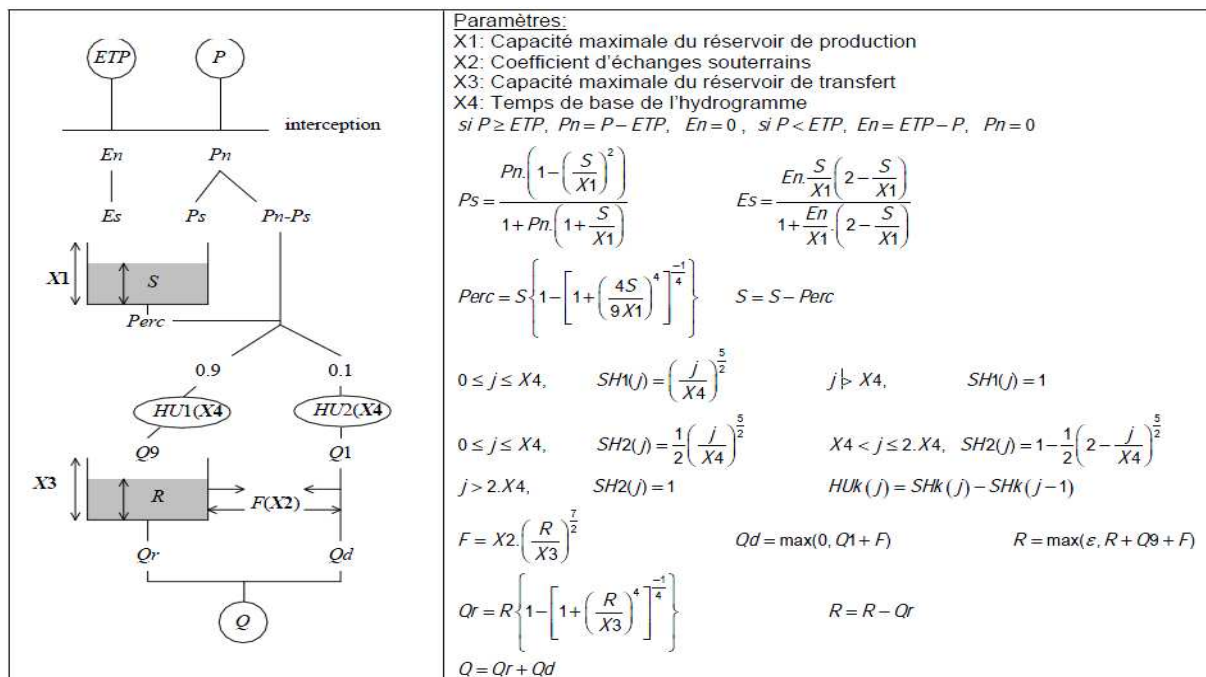


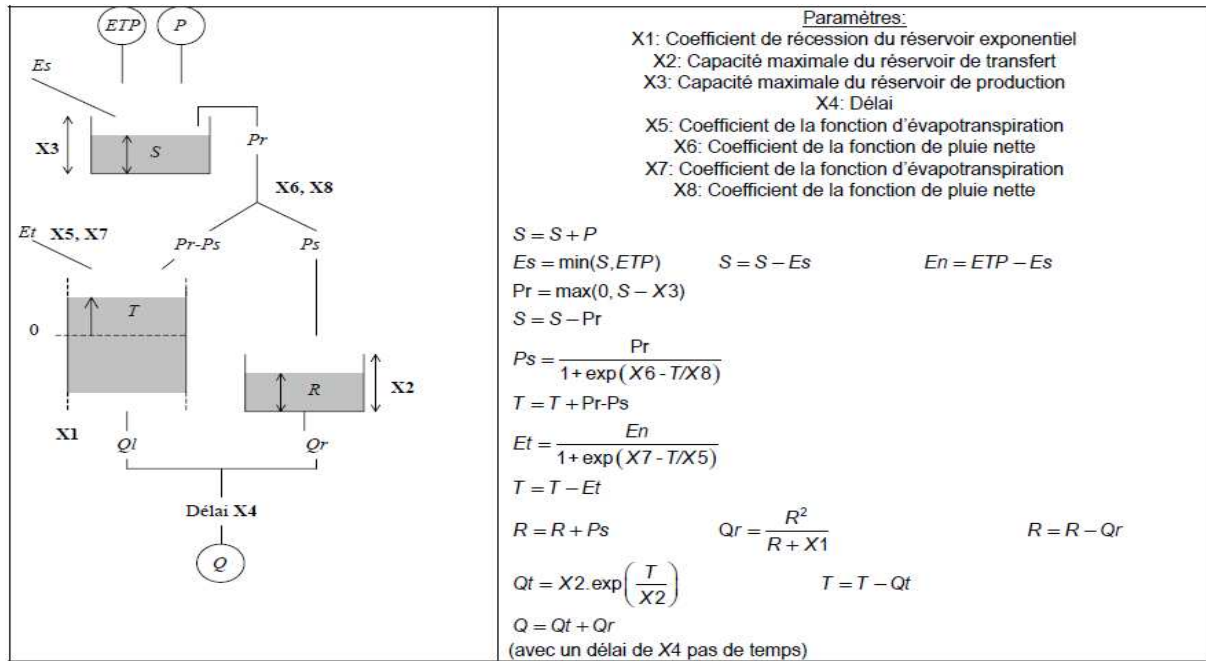
Figure 1 : Schéma conceptuel du modèle GR4J et les quatre paramètres calés.

(D'après Oudin 2004)

### 3.3 Le modèle TOPMO

Le modèle TOPMODEL (TOPography-based hydrological MODEL), développé au pas de temps journalier par Beven et Kirkby en 1979, a la particularité d'utiliser un indice topographique et la notion de zone contributive variable. Nous utilisons ici une version modifiée de ce modèle, que nous appelons TOPMO. La distribution de l'indice topographique, habituellement calculée à partir d'un modèle numérique de terrain (MNT), est approximée ici par une fonction logistique dépendant de deux paramètres à caler. Le ruissellement est attribué aux seuls écoulements sur les zones contributives. Comme GR4J, le modèle prend en entrée les pluies, l'évapotranspiration potentielle et utilise les débits observés pour le calage des paramètres. Il restitue les débits calculés. La figure suivante indique le schéma structurel du modèle TOPMO :

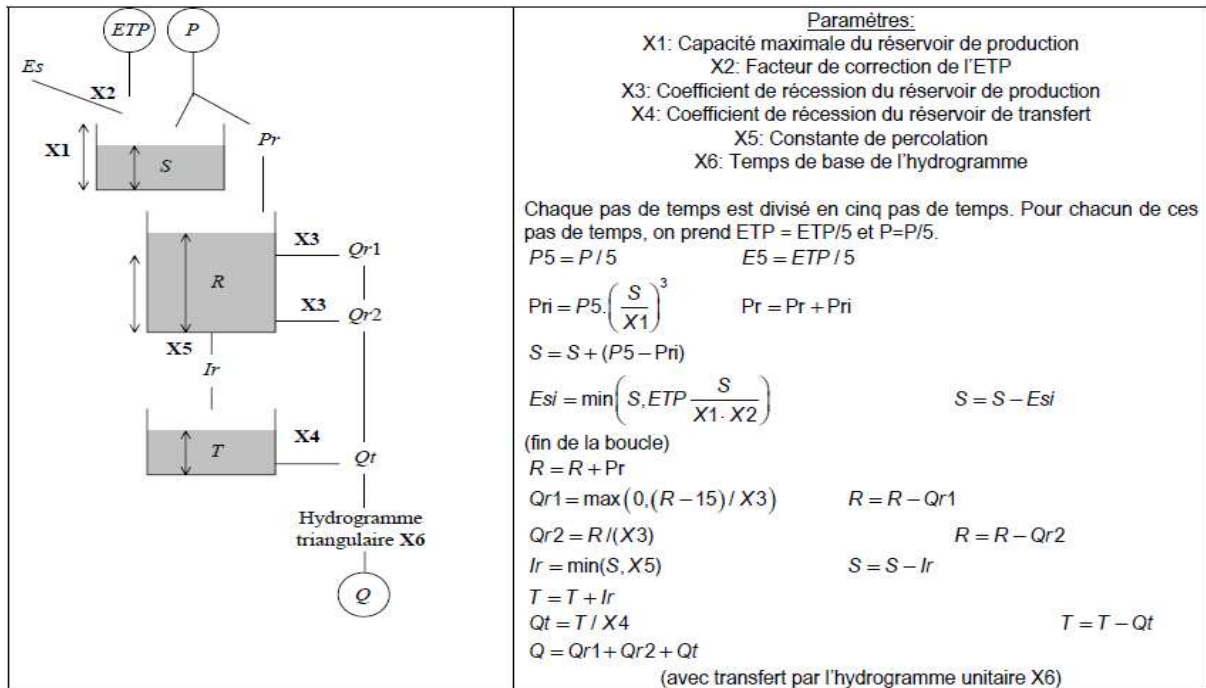




**Figure 2:** Schéma conceptuel et le calcul effectué au sein du modèle TOPMO (D'après Oudin 2004)

### 3.4 Le modèle HBV0

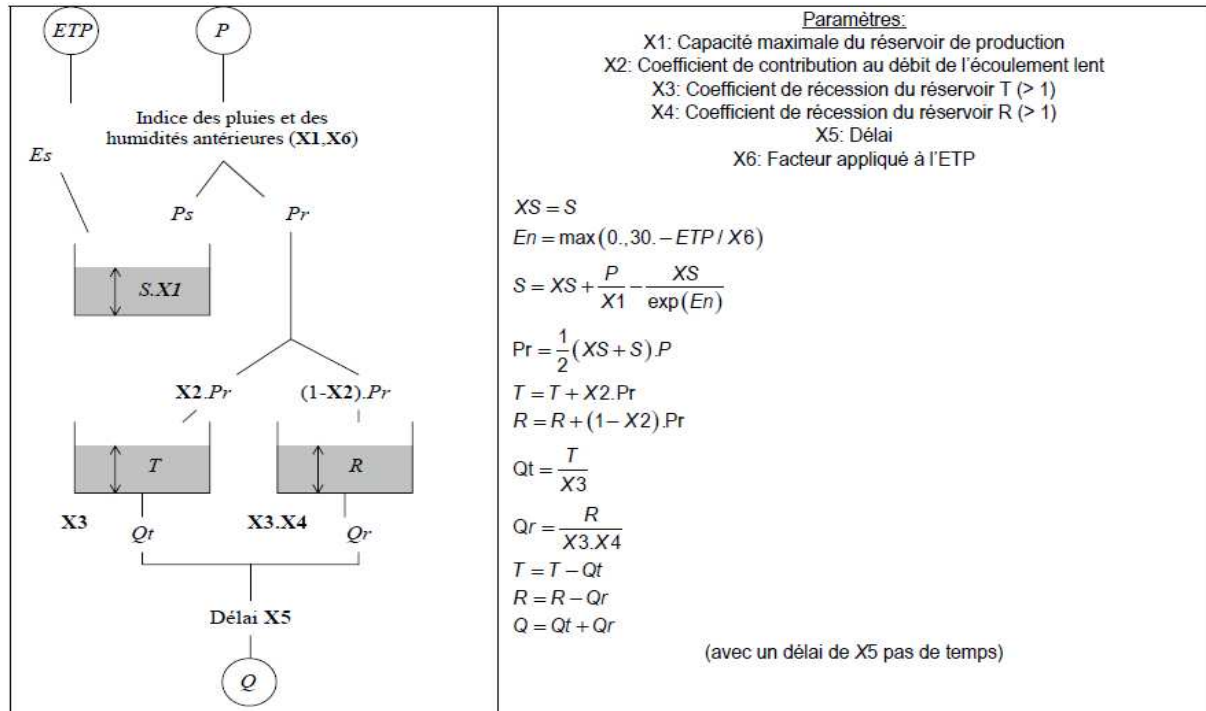
Bergström et Forman ont développé en 1973 le modèle HBV qui est un modèle conceptuel global spécialement conçu pour des conditions avec fortes influences neigeuses dans les pays nordiques. La version utilisée ici (appelée HBV0) est une version sans stock neigeux, à six paramètres.



**Figure 3:** Schéma conceptuel et le calcul effectué au sein du modèle HBV0 (D'après Oudin 2004)

### 3.5 Le modèle IHA

Le modèle IHACRES (identification of unit Hydrograph and component flow from Rainfalls, Evaporation and streamflow data), est le résultat d'une collaboration entre CEH (centre of ecology and Hydrology) du Royaume-Unit et L'Université nationale de l'Australie (ANU). Ce modèle peut être appliqué pour des investigations dans des petits bassins versants équipés pour des études spéciales ou pour des études régionales dans les grands bassins. La version du modèle IHACRES (Jakeman *et al.*, 1990) retenue est un modèle à six paramètres à caler. Il comprend un module de production non linéaire et un module de transfert linéaire.



**Figure 4:** Schéma conceptuel et le calcul effectué au sein du modèle IHA (D'après Oudin 2004)

### 3.6 Critères d'évaluation

Dans le domaine de la modélisation hydrologique, il existe un grand nombre de critères d'évaluation des performances des modèles, cependant, on trouve que le critère de Nash et Sutcliffe (1970) « NSE » est le plus communément utilisé ces dernières décennies. Le critère de Nash sert à construire un index de concordance entre les débits observés et les débits simulés et qui est exprimé par l'équation suivante :

$$NSE = 100 * \frac{\sum (Q_o - Q_c)^2}{\sum (Q_o - Q_m)^2} \quad \text{Équation 4}$$

Avec

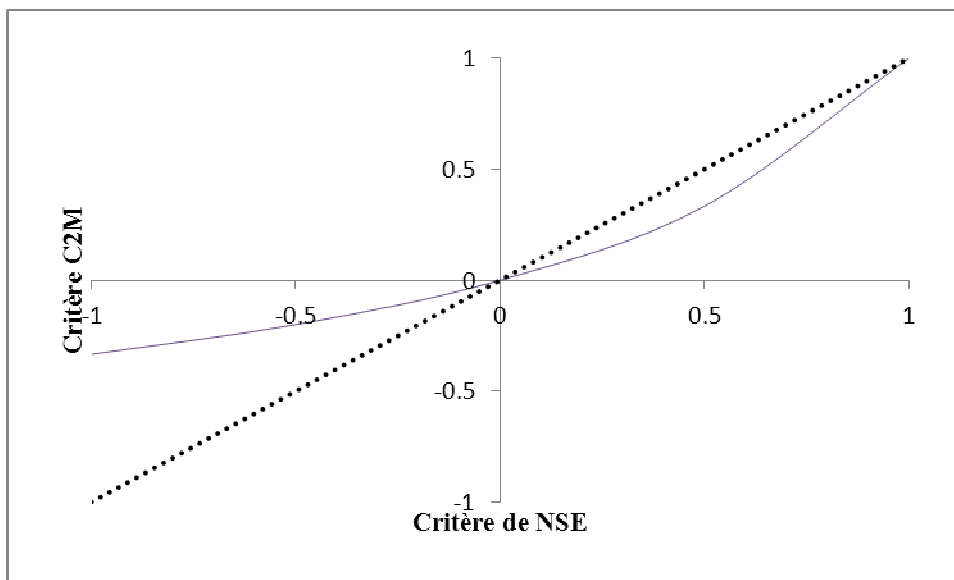
- $Q_0$  : débit observé
- $Q_c$  : débit simulé
- $Q_m$  : moyennes de débits observés

Le critère NSE varie dans l'intervalle  $]-\infty, 1]$ . Il estime l'amélioration obtenue par le modèle testé, par rapport au modèle de référence ( $Q=Q_{obs}$ ).

Pour borner le critère NSE et garder la valeur 0 comme une valeur de référence, Mathevet (2005), a proposé la formule suivante qu'il a nommé **C2M**.

$$C2M = \frac{NSE}{2 - NSE} \quad \text{Équation 5}$$

La relation entre NSE et C2M est donnée par le graphique suivant :



**Figure 5** : Relation entre le critère NSE et le critère C2M

L'intérêt de C2M est multiple :

- Le domaine de variation est compris entre -1 et 1 ;
- Le signe et le zéro du C2M coïncident avec le signe et le zéro du critère de NSE ;
- Dans sa partie positive, le critère de C2M est sensiblement inférieur au critère de NSE, ce critère donne ainsi une image moins bonne, de la qualité de simulation de modèle que le critère de NSE. Cela va dans le sens de point de vue de Garrick et al. (1978), qui faisaient remarquer que les valeurs du critère de NSE étaient généralement trop fortes.

Le fait que le C2M soit borné (Contrairement au NSE) permet d'envisager plus sereinement l'utilisation des valeurs moyenne obtenues sur plusieurs bassins. Par la suite nous utiliserons la valeur moyenne de C2M sur l'ensemble de l'échantillon, ainsi que la valeur médiane de NSE pour rendre compte de l'efficacité des différentes techniques de régionalisation envisagées.

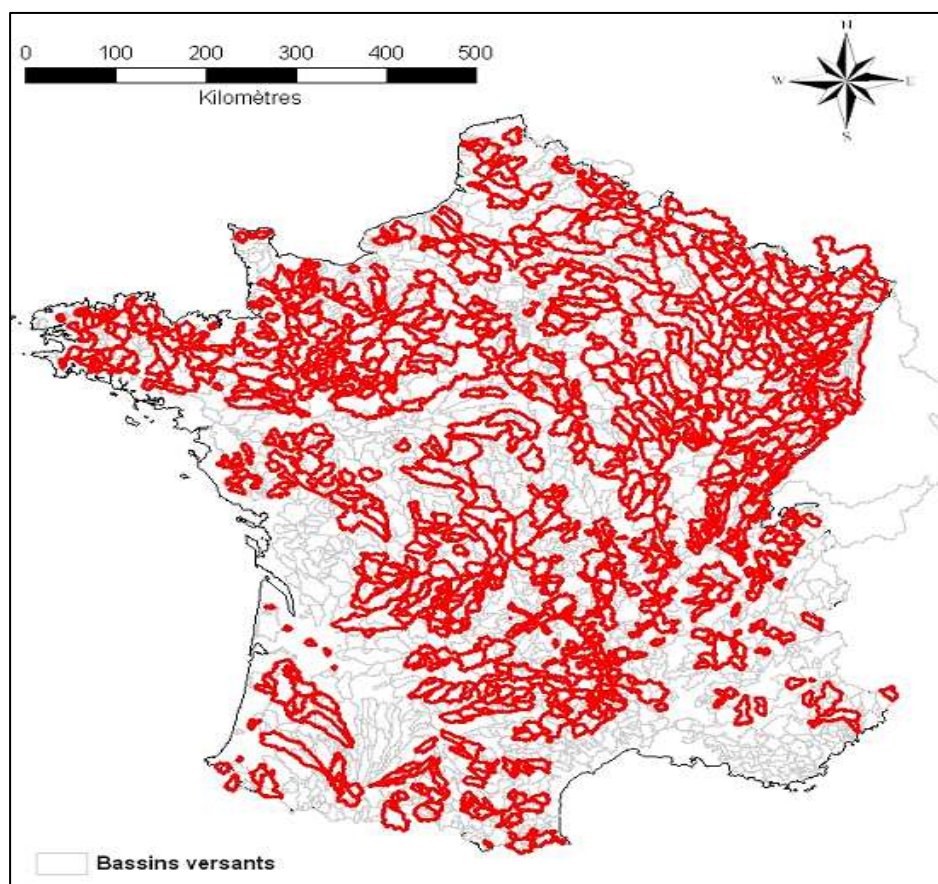
## 4 Base de données

### 4.1 Les données de modélisation

L'étude est basée sur 1056 bassins versants en France (Fig. 6). Pour chaque bassin, des séries journalières de précipitations, évapotranspiration ponctuelle et débits sont disponibles. Elles permettent le calage et la mise en route du modèle, ainsi que la simulation pluie-débit et l'évaluation des débits prévus.

Les données météorologiques proviennent de la base SAFRAN de Météo-France (Quintana-Seguí *et al.*, 2008). Les données d'évapotranspiration potentielle ont été obtenues à partir de l'équation proposée par (Oudin *et al.*, 2005), qui prend en compte la courbe de régime des températures et le rayonnement extraterrestre (obtenu via la latitude). La base de données de la Banque Hydro fournit les données de débits observés, relativement complètes, car 75 % des bassins étudiés ont plus de 27 années de données journalières.

Les données pour les descripteurs physico-climatiques provenant des différentes bases de données sont disponibles à l'échelle internationale et sont décrites par la suite dans le paragraphe traitement de données des descripteurs physico-climatiques.

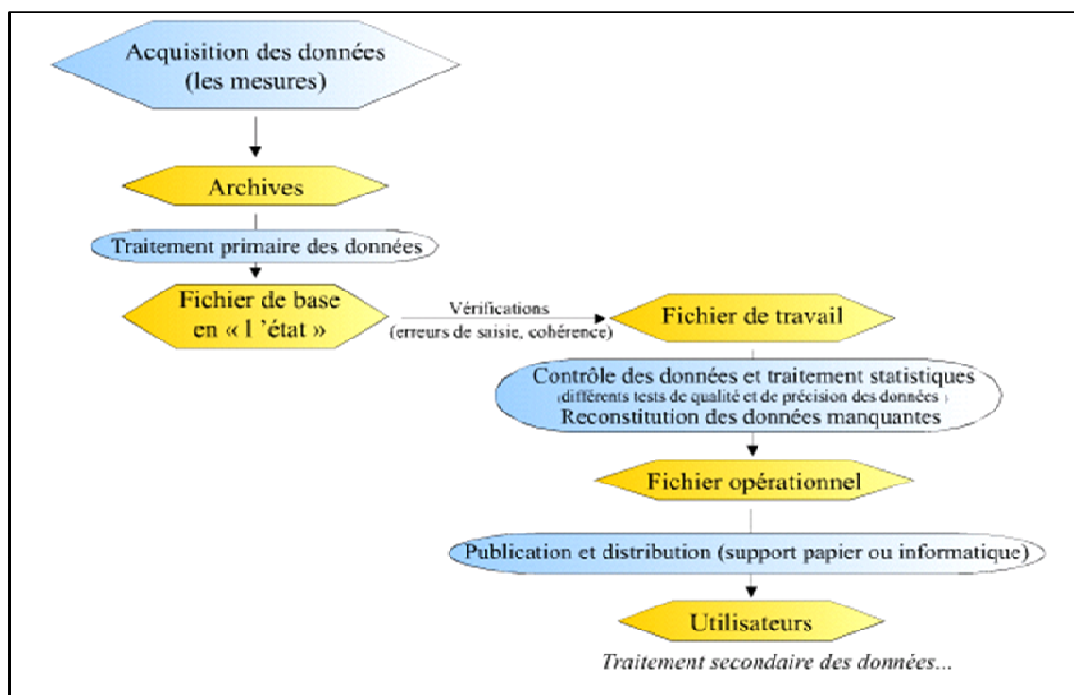


**Figure 6:** Carte représentant les 1056 bassins étudiés en France

## 4.2 Traitement des données hydrométéorologiques

La compréhension des processus intervenant dans le cycle de l'eau ainsi que l'étude de leurs variations spatiales et temporelles nécessitent de disposer de données, notamment pour le type de modèles utilisés dans ce mémoire, où les données sont utilisées pour le calage du modèle et la simulation des débits.

De façon générale, pour permettre le passage de l'acquisition des données à leur utilisation effective dans le cadre d'une analyse hydrologique, ces données passent obligatoirement par les étapes suivantes : acquisition, traitement, contrôle et validation, organisation, diffusion et publication. La figure 6 présente les différentes étapes de la chaîne d'acquisition et de traitement des données.



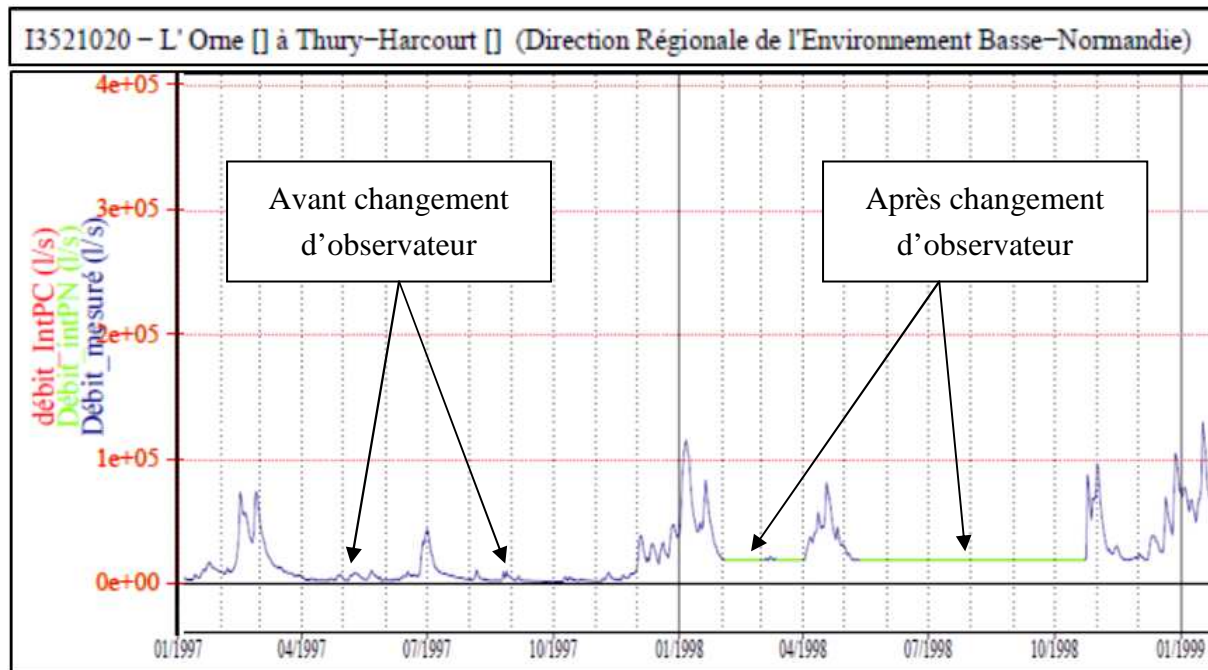
**Figure 7:** Etapes de la chaîne d'acquisition et de traitement des données  
(<http://echo2.epfl.ch/e-drologie/resumes/chapitre8/resume8.html>)

Etant donné, le nombre important d'étapes par lesquelles passent les données hydrométéorologiques, plusieurs points doivent être éclaircis : comment se fait l'acquisition de ces données ainsi que leur traitement et quels sont les différentes erreurs qui peuvent être commises lors ces étapes ?

### 4.2.1 Origine et différents types d'erreurs

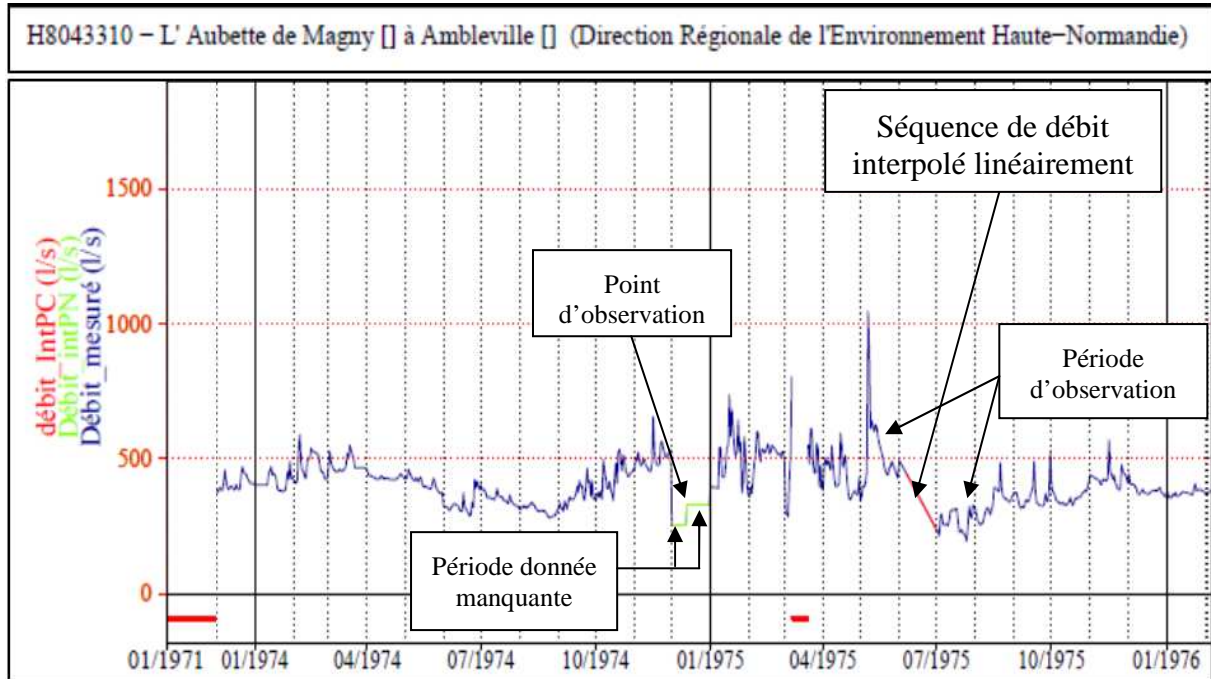
Une erreur de mesure est définie comme étant la différence entre la vraie valeur (qui est l'idéal recherché, mais qui n'est en principe et malheureusement jamais connue) et la valeur mesurée. On peut différencier deux types d'erreur : les erreurs statistiques (aléatoires) et les erreurs systématiques.

**Les erreurs systématiques :** Elles affectent la fiabilité des données et sont généralement corrélées. L'origine des erreurs systématiques est le plus souvent liée à la calibration de l'appareil de mesure qui n'est pas parfaite ou à un phénomène extérieur qui perturbe la mesure (erreur d'appareillage, changement d'observateur...). La figure 8 illustre bien un type d'erreur due à un changement d'observateur. En effet, cette élévation du niveau de débit d'étiage après le changement de l'observateur peut être expliquée par la présence d'un seuil de mesure au niveau du flotteur (ou bien l'appareil qui capte la hauteur de l'eau) : au-dessous d'une certaine hauteur d'eau, tous les mesures sont égales à ce seuil.



**Figure 8:** Erreur due au changement d'observateur

**Les erreurs statistiques :** La constitution d'une série de valeurs, composant un échantillon au sens statistique du terme, est un processus long au cours duquel de nombreuses erreurs, de nature statistique, sont susceptibles d'être commises. Ces erreurs sont généralement commises lors de la phase de traitement. La figure 9 montre deux exemples d'interpolation de données probablement non observées : la première séquence en couleur rouge sur la courbe peut être expliquée par la présence de deux périodes d'observation intercalées par une période de données manquantes, et avec une interpolation linéaire entre ces deux périodes d'observations on aboutit à ce genre de séquence de débit interpolé. L'autre séquence en couleur verte sur la courbe (pente constante nulle) peut être expliquée par la présence de deux périodes de données manquantes qui sont intercalées par une valeur observée, ces séquences seront donc interpolées linéairement vis-à-vis de cette valeur observée, ce qui produit une séquence de débit à pente constante nulle.



**Figure 9:** Erreur due à l'interpolation des données de débits

Vu le nombre d'erreurs qui peuvent être commises lors des phases de déroulement classique de constitution d'une série de données, il est donc indispensable, avant d'utiliser ces séries de données, de se préoccuper de leur qualité et de leur représentativité, ceci en utilisant diverses techniques, en général de type statistique ou graphiques.

Ce qui nous intéresse nous, comme utilisateurs des modèles pluie-débit, c'est de mettre l'accent sur la nature même des données : S'agit-il de données observées mesurées ou bien de données manquantes interpolées. La meilleure méthode pour bien visualiser et distinguer entre ces deux types de données c'est la méthode graphique : il s'agit de représenter les séries des débits et essayer de distinguer visuellement les séquences des débits interpolés.

## 4.2.2 Traitement des données de débit

### 4.2.2.1 Motivation

Dans le domaine de la régionalisation des modèles hydrologiques, la plupart des modélisateurs s'appuient sur le principe de la sélection des bassins donneurs de l'information hydrologique. Pour que l'information transmise soit utile il faut s'assurer que :

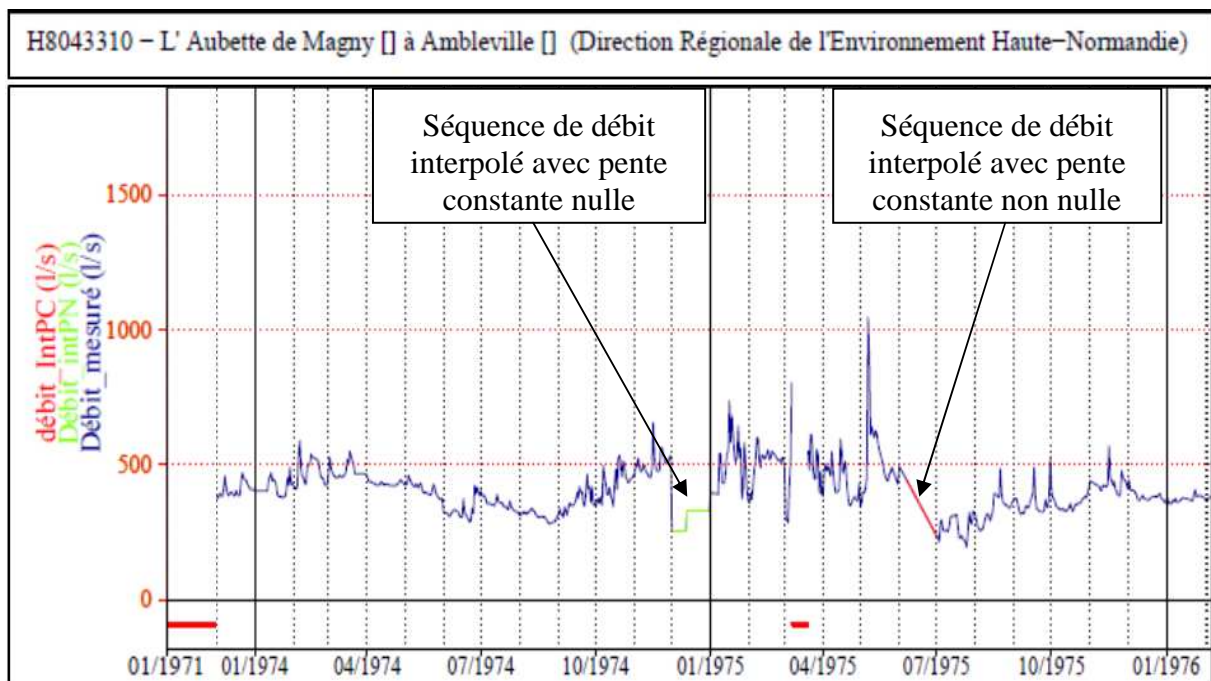
- ✓ le bassin soit non influencé par les ouvrages hydrauliques que ce soit à l'amont ou à l'aval ;
- ✓ le bassin présente un série de données complète, i.e. avec un faible pourcentage de données manquantes ;
- ✓ les données soient réellement observées et non interpolées.

Afin de remplir ces critères nous avons appliqué une pratique de traitement des séries de débit qui consiste à éliminer les données interpolées pour les remplacer par des données manquantes.

En effet, notre objectif d'étude portant sur la régionalisation des modèles hydrologiques, il n'est pas logique d'appliquer la régionalisation sur des séries de données elles même interpolées et/ou régionalisées. La difficulté est de pouvoir détecter ces données interpolées à partir d'une base de données hétérogène constituée de différentes sources de données, différentes conditions ainsi que différentes méthodes et relations d'interpolation.

Pour répondre à cette question, nous avons procédé dans notre traitement à une analyse générale des séries des données afin d'avoir une vision globale sur la ou les méthodes d'interpolation majoritairement utilisée(s). De ce fait, nous avons commencé par une représentation graphique des toutes les chroniques de débits de tous les bassins versants constituant notre base de données (1056 Bassins français).

L'analyse de certaines chroniques montre bien qu'une simple interpolation linéaire est parfois utilisée pour combler les lacunes de débit. La figure 10 illustre certaines séquences de débits interpolés :



**Figure 10 :** Les séquences des débits interpolés sont en rouge et en vert

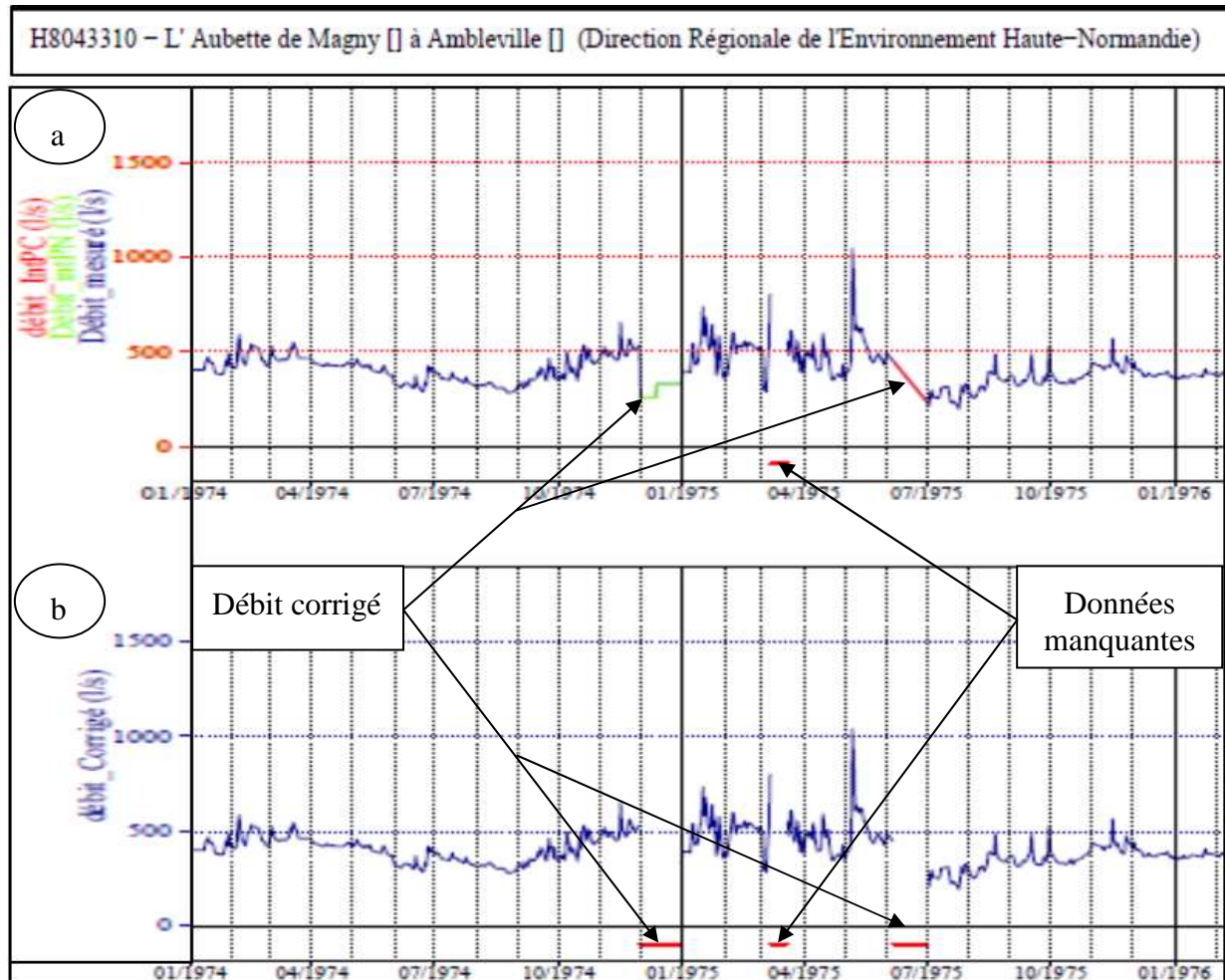
La figure 10 met en évidence les séquences de débits interpolés ainsi que la méthode d'interpolation. Cette figure envisage aussi la présence de deux types de séquences de débits interpolés, ceux qui ont une pente constante nulle (en couleur verte) et d'autres ayant une pente constante non nulle (en couleur rouge sur la courbe).

#### 4.2.2.2 Automatisation de la détection des données interpolées

Pour détecter ces séquences de débits interpolés de façon plus rapide nous avons automatisé l'analyse. Le problème majeur rencontré au début de traitement a été les séquences de débit d'étiage étaient considérées probablement, comme des séquences interpolées, c'est pour cela qu'une condition a été rajouté lors de la détection des débits interpolés : cette condition

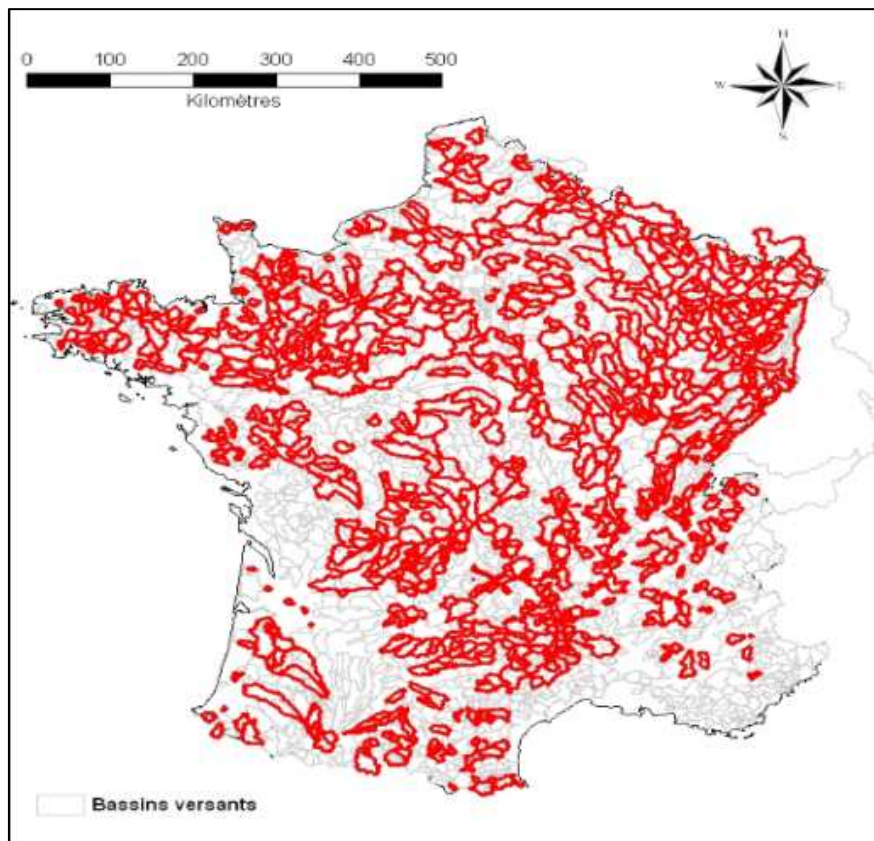


consiste à ne pas détecter les séquences pour lesquelles les débits sont inférieurs au troisième décile ( $Q_{30}$ ) de la distribution des débits observés. Le choix de cette valeur de fréquence a été guidé par une série de tests de gamme de valeurs comprises entre  $Q_{15}$  et  $Q_{40}$ , ceci en contrôlant visuellement à chaque fois les résultats sur certaines chroniques. Les figures 10-a et 10-b illustrent bien l'objectif de correction et met en évidence les interventions faites sur les séries de débit.



**Figure 11:** Exemple de traitement des séries de débit : a) chronique avant traitement, b) chronique après traitement

Une fois cette analyse réalisée sur l'ensemble des bassins, chaque bassin est finalement retenue dans l'échantillon si le pourcentage des débits interpolés ou lacunaires est inférieur à 10%. À partir d'une base de données préalable constituée de 1056 bassins versants, cette sélection a conduit à un échantillon de 973 bassins bien répartis sur le territoire français (cf. figure 12).



**Figure 12:** Répartition de 973 bassins

### 4.3 Influence de la procédure de traitement sur les résultats de calage

Le calage d'un modèle hydrologique peut se définir comme étant l'ensemble des processus d'ajustement des paramètres du modèle, dans l'objectif de reproduire certains aspects de la réalité en comparant des variables de sortie (débits simulés) à des données observées (débits observés).

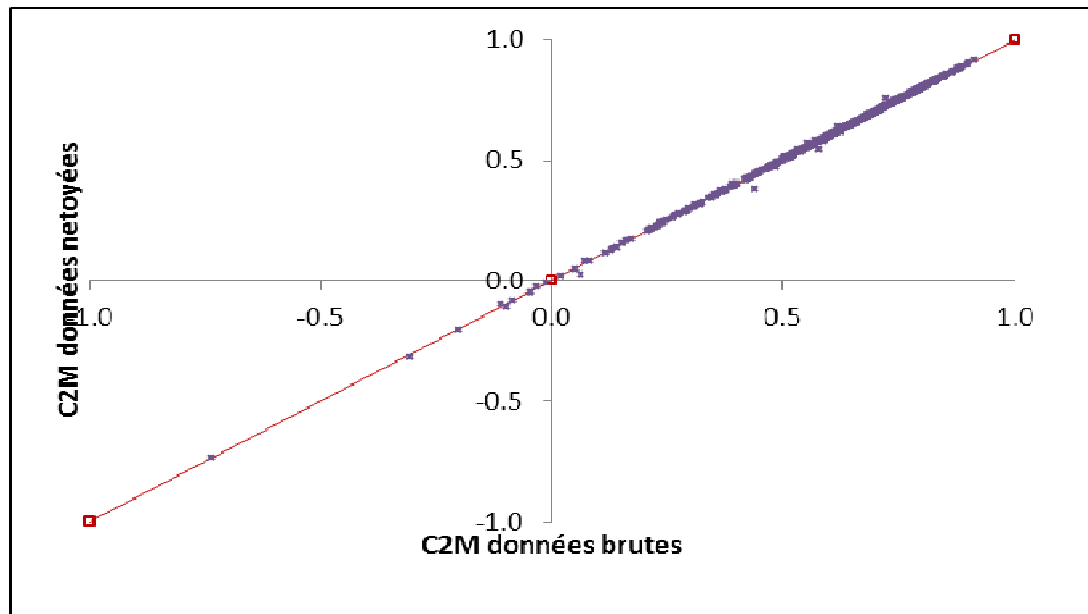
Le calage des modèles hydrologiques IHA, HBV, TOPMO et GR4J a été effectué avec d'une part les données brutes et d'autre part avec les données traitées. Ce test consiste à déterminer l'effet de la procédure de traitement sur la simulation des débits et de comparer les résultats avec ceux obtenus avec les données brutes.

**Tableau 2:** Calage des modèles avec les données traitées et les données brutes

Modèle Pluie-Débit	C2M moyen	
	Avec données brutes	Avec données nettoyées
GR4J	0.650	0.650
TOPMO	0.586	0.583
HBV	0.491	0.491
IHA	0.428	0.428

La figure ci-dessous illustre bien les résultats obtenus en faisant la comparaison entre critères de performance C2M avec données brutes et critères de performance C2M avec données traitées.

Les résultats de ce calage reportée en figure 12 montrent que la procédure de nettoyage hydrologique n'a pas une grande influence sur les critères de performance en calage.



**Figure 13:** Résultats de calage avec le modèle GR4J en considérant les données brutes et traités

Les autres modèles présentent presque les mêmes résultats que ce modèle GR4J et les figures ne sont pas reportées ici.

#### 4.4 Collecte et traitement des descripteurs physico-climatique

Il été mentionné précédemment dans le chapitre d'analyse comparée des articles que parmi les raisons de contradiction entre les approches de régionalisation, Merz *et al.*, en 2006 soulignent que les descripteurs relatifs aux bassins versants disponibles diffèrent d'une étude à une autre. Donc, afin d'homogénéiser les descripteurs et d'éviter ce problème de contradiction, nous avons collecté et traité un grand nombre de descripteurs disponibles en générale à l'échelle internationale. Pour cela, quatre sources de données ont été utilisées : SRTM, SAFRAN, European Soil Database (ESDB) et ECOCLIMAP.

##### 4.4.1 Les descripteurs morphologiques (base de données SRTM)

La base de données utilisée pour les descripteurs morphologique est le SRTM « *Shuttle Radar Topography Mission* » qui fait référence à des fichiers matriciels et vectoriels topographiques fournis par deux agences américaines : la NASA et la NGA (ex-NIMA). Ces données altimétriques ont été recueillies au cours d'une mission de 11 jours en février 2000 par la navette spatiale Endeavour (STS-99) à une altitude de 233 km en utilisant l'interférométrie radar. Les descripteurs morphologiques choisis pour la présente étude sont les suivants.

**Tableau 3:** Descripteurs morphologiques calculés à partir de la base de données SRTM

Notation	Descripteur	Unité
<b>S</b>	Superficie du bassin versant	km <sup>2</sup>
<b>Zmoy</b>	Altitude moyenne du bassin versant	m
<b>Z5%</b>	Altitude soupasée pour 5% de la superficie	m
<b>Z95%</b>	Altitude soupassée pour 95% de la superficie	m
<b>Pente</b>	Pente moyenne du bassin versant	m/m
<b>DD</b>	Densité de drainage	km/km <sup>2</sup>
<b>Ig</b>	Indice global de pente	-
<b>Intop</b>	Indice topographique moyen	-

#### 4.4.2 Les descripteurs Climatiques (Base de données SAFRAN)

Les descripteurs climatiques ont été déterminés à partir du système d'analyse SAFRAN (Quintana- Seguí *et al*, 2008). SAFRAN analyse huit paramètres: la vitesse du vent, l'humidité relative, la température de l'air, la nébulosité, les radiations solaires et atmosphériques, les chutes de neige et la pluie.

L'hypothèse principale de SAFRAN est que les variables atmosphériques sont considérées comme homogènes sur certaines zones bien définies (résolution de 64 km<sup>2</sup>), au sein desquelles elles ne peuvent varier qu'en fonction de la topographie. Les analyses sont calculées toutes les 6 h, et les données sont interpolées à un pas de temps horaire.

Les descripteurs climatiques extraits de cette base de données pour l'étude de régionalisation des modèles hydrologiques sont les suivants :

**Tableau 4 :** Descripteurs Climatiques calculés à partir de la base de données SAFRAN

Notation	Descripteur	Unité
<b>P</b>	Pluie moyenne annuelle	m
<b>CP</b>	Coefficient d'irrégularité de pluie	-
<b>ETP</b>	Évapotranspiration moyenne annuel	m
<b>CETP</b>	Coefficient d'irrégularité d'ETP	-
<b>PMH</b>	Pluie moyenne hivernale	m
<b>PME</b>	Pluie moyenne estivale	m
<b>PME/PMH</b>	Ratio entre la pluie moyenne estivale et hivernale	-

**C** : Coefficient d'irrégularité temporelle = (max-moyenne-mensuel–min-moyenne-mensuel) / moy-mensuel

#### 4.4.3 Les descripteurs pédologiques : (Base de données ESDB)

Cette base de données caractérise les propriétés de sol à une échelle fine. Deux bases de données interpolées sont utilisées

- ✓ Base des données géographiques des sols **SGDBE** ;

- ✓ La base de données des règles pédo-transfert *PTRDB* : Une règle pédo-transfert est conçue pour étendre la gamme des paramètres du sol à des propriétés non observées ou mesurées.

À partir d'une liste préliminaire de 70 descripteurs de sol, 29 ont finalement été retenus. Ces derniers ont été choisis après élimination de tous les descripteurs non informatifs et le rassemblement de certaines classes sous une seule.

Pour certains descripteurs, les informations qualitatives ont été traduites en information quantitatives afin de pouvoir effectuer des moyennes spatiales à l'échelle du bassin versant. Les valeurs attribuées à chaque classe sont reportées au tableau 3 et proviennent des auteurs de la base de données (CCR, 2003)

**Tableau 5:** Liste des descripteurs pédologiques utilisés

Les Descripteurs Sols											
1	La texture du sol dominante (TEXT)										
<b>Classe</b>	<b>Grossière</b>		<b>Moyenne</b>		<b>Moyenne fine</b>		<b>Fine</b>		<b>Très fine</b>		
<b>Définition des classes</b>	%Sable>65 %Argile<18		%Sable>15 18<%argile<35		%Sable<15 %Argile<35		35<%Argile<65		%Argile>60		
2	Type des matériaux d'origine hydrologique (PMH)										
<b>Classe</b>	<b>R</b>		<b>C</b>		<b>S</b>		<b>L</b>		<b>H</b>		<b>M</b>
<b>Définition des classes</b>	Poreux (-) Stockage (±) Perméabilité(+)		Poreux (++) Stockage (±) Perméabilité(+)		Poreux (+) Stockage (+) Perméabilité(+)		Stockage (-) Perméabilité (-)		Compacté Stockage (--) Perméabilité (--)		souple Stockage (--) Perméabilité (--)
3	Profondeur du sol tourbière (DGH)										
<b>Classe</b>	<b>Peu profonde (S)</b>			<b>Moyenne (M)</b>			<b>Profonde (D)</b>			<b>Très profonde (VD)</b>	
<b>Définition des classes</b>	Prof<40 cm <b>20</b>			40<Prof<80 cm <b>60</b>			80<Prof<120 cm <b>100</b>			Prof>120 cm <b>150</b>	
4	Teneur du sol en carbone organique (OC-TP)										
<b>Classe</b>	<b>Elevé (H)</b>			<b>Moyen (M)</b>			<b>Faible (L)</b>			<b>Très faible (V)</b>	
<b>Définition des classes</b>	%OC>6 <b>6</b>			2<%OC<6 <b>4</b>			1<%OC<2 <b>1.6</b>			%OC<1 <b>0.5</b>	
5	Profondeur de la couche imperméable DIMP										
<b>Classe</b>	<b>Peu profonde (S)</b>						<b>Profonde (D)</b>				
<b>Définition des classes</b>	Profondeur < 80 cm						Profondeur > 80 cm				
6	La densité apparente de surface (PD-TP)										
<b>Classe</b>	<b>Faible (L)</b>			<b>Moyenne (M)</b>				<b>Elevé (H)</b>			
<b>Définition des classes</b>	Densité<1.10 g/cm <sup>-3</sup> <b>1.10</b>			1.10<Densité<1.55 g/cm <sup>-3</sup> <b>1.55</b>				Densité>1.85 g/cm <sup>-3</sup> <b>1.85</b>			
7	Classe hydrogéologique de sol										
<b>Classe</b>	<b>1R</b>	<b>1C</b>	<b>1S</b>	<b>1L</b>	<b>1H</b>	<b>1M</b>	<b>2</b>	<b>3</b>	<b>4W</b>	<b>4D</b>	
<b>Définition des classes</b>											
8	Capacité de teneur en eau du sol (AWCTP)										
<b>Classe</b>	<b>Faible (L)</b>			<b>Moyenne (M)</b>			<b>Elevé (H)</b>			<b>Très élevé (VH)</b>	

Définition des classes	Capacité < 100 mm/m <b>50</b>	100 < Capacité < 140 mm/m <b>120</b>	140 < Capacité < 190 mm/m <b>165</b>	Capacité > 190 mm/m <b>190</b>
9	Capacité du sol en RFU (EAWCT)			
Classe	<b>Faible (L)</b>	<b>Moyenne (M)</b>	<b>Elevé (H)</b>	<b>Très élevé (VH)</b>
Définition des classes	Capacité < 100 mm/m <b>50</b>	100 < Capacité < 140 mm/m <b>120</b>	140 < Capacité < 190 mm/m <b>165</b>	Capacité > 190 mm/m <b>190</b>
10	Teneur en eau à la saturation ( $\theta_{sat}$ )			
11	Conductivité hydraulique à saturation ( $K_{sat}$ )			

#### 4.4.4 Les descripteurs d'occupation de sol (base de données ECOCLIMAP)

Pour les besoins de la modélisation en météorologie et en hydrologie, le CNRM a développé une base de données numérique mondiale des écosystèmes observés sur les surfaces continentales. Cette base de données, appelée ECOCLIMAP, a une résolution spatiale de 1 km<sup>2</sup>, et repose dans une grande mesure sur des observations satellitaires. Sur l'Europe et le bassin méditerranéen, ECOCLIMAP comporte 273 types de paysages. Ces types résultent de la fusion de bases de données existantes, CORINE-2000, qui a une résolution spatiale de 100 m mais ne concerne que les pays de l'Union Européenne, et GLC-2000, qui est une base de données mondiale. La fusion de CORINE-2000 et GLC-2000 produit une carte comprenant 76 classes. Afin d'affiner la représentation de certains écosystèmes, des séries temporelles d'indices de végétation issus des données satellitaires SPOT/VEGETATION, de 1999 à 2005, sont associées aux classes précédentes. Par classification automatique, on parvient à identifier 273 classes. Sur la France, la base de données comporte 168 classes. (<http://www.cnrm-game.fr>)

À partir de cette base de données française, une liste préliminaire constituée de 108 classes a été extraite. Sur cette liste préliminaire nous avons procédé à deux traitements. La première étape de sélection consiste à éliminer tous les classes qui n'ont pas a priori une relation directe avec la réponse hydrologique des bassins versants. La deuxième étape consiste à attribuer une valeur moyenne et un coefficient d'irrégularité pour toutes les séries temporelles (moyennes mensuelles pluriannuelles). Les descripteurs qui ont été identifiés suite à ces deux étapes sont indiqués dans le tableau n°4 ci-dessous.

**Tableau 6 :** Les descripteurs d'occupation de sol

Indice	Description	Unité
<b>F-sand</b>	Fraction du bassin versant couvert par du sable	-
<b>F-clay</b>	Fraction du bassin versant couvert par de l'argile	-
<b>d-soil</b>	Profondeur du sol	m
<b>d-root</b>	Profondeur racinaire	m
<b>rsmin</b>	Résistance stomatique	s/m
<b>F-tree</b>	Fraction du bassin versant couvert par des arbres	-
<b>F-lowv</b>	Fraction du bassin versant couvert par des arbustes	-
<b>F-bare</b>	Fraction du bassin versant couvert par du sol nu	-
<b>f-C3_crops</b>	Fraction du bassin versant couvert par des plantes de type C3	-
<b>f-C4_crops</b>	Fraction du bassin versant couvert par des plantes de type C4	-

<b>f-irr-corps</b>	Fraction du bassin versant couvert par des zones irriguées	-
<b>f-garsmland</b>	Fraction du bassin versant couvert par des prairies	-
<b>f-broadleaf</b>	Fraction du bassin versant couvert par des feuillus	-
<b>f-conif</b>	Fraction du bassin versant couvert par des conifères	-
<b>Z0h</b>	Moyenne annuelle de la rugosité au-dessus de la végétation	m
<b>CZ0h</b>	Coefficient d'irrégularité de cette rugosité	-
<b>Z0</b>	Moyenne annuelle de la rugosité	m
<b>CZ0</b>	Coefficient d'irrégularité de la rugosité	-
<b>veg</b>	Moyenne annuelle de taux de couverture végétale	-
<b>Cveg</b>	Coefficient d'irrégularité temporelle de la couverture végétale	-
<b>Lai</b>	Moyenne annuelle de l'indice foliaire sur le bassin	-
<b>Clai</b>	Coefficient d'irrégularité temporelle de l'indice foliaire	-
<b>greeness</b>	Moyenne annuelle de couvert vert	-
<b>Cgreeness</b>	Coefficient d'irrégularité temporelle de couvert vert	-
<b>glob-alb</b>	Moyenne annuelle d'albédo	-
<b>Cglob-alb</b>	Coefficient d'irrégularité temporelle d'albédo	-

#### 4.4.5 Étude de la corrélation entre les différents descripteurs

L'objectif est d'étudier le croisement entre les différents descripteurs et d'identifier ceux qui ont une forte corrélation entre eux (supérieur à 0.9), ceci afin d'éviter de redondance et de sélectionner les descripteurs les plus propices pour l'étude de la régionalisation des bassins versants par similarité. Les résultats de cette étude de corrélation sont donnés dans l'annexe n°2.

Soixante descripteurs ont été finalement retenus et sont représentés dans le tableau n° 5.

**Tableau 7:** Les descripteurs Climato-physiques

N°	Indice	Description	Unité
<b>Les Descripteurs morphologiques</b>			
1	<b>S</b>	Superficie du bassin versant	km <sup>2</sup>
2	<b>Zmoy</b>	Altitude moyenne du bassin versant	m
3	<b>Pente</b>	Pente moyenne du bassin versant	°
4	<b>DD</b>	Densité de drainage	km/km <sup>2</sup>
5	<b>Ig</b>	Indice global de pente	-
6	<b>Intop</b>	Indice topographique moyen	-
<b>Les Descripteurs Climatiques</b>			
7	<b>P</b>	Pluie moyenne annuelle	mm/an
8	<b>CP</b>	Coefficient d'irrégularité de pluie	-
9	<b>ETP</b>	ETP moyenne annuelle	mm/an
10	<b>CETP</b>	Coefficient d'irrégularité d'ETP	-
11	<b>ETP/P</b>	Indice d'aridité	-

12	<b>PMH</b>	Pluie moyenne hivernal	mm				
<b>Les descripteurs d'occupation de sol</b>							
13	<b>F-sand</b>	Fraction du bassin versant couvert par du sable	-				
14	<b>F-clay</b>	Fraction du bassin versant couvert par de l'argile	-				
15	<b>d-soil</b>	Profondeur du sol	m				
16	<b>d-root</b>	Profondeur racinaire	m				
17	<b>rsmin</b>	Résistance stomatique	s/m				
18	<b>F-tree</b>	Fraction du bassin versant couvert par des arbres	-				
19	<b>F-lowv</b>	Fraction du bassin versant couvert par des arbustes	-				
20	<b>F-bare</b>	Fraction du bassin versant couvert par du sol nu	-				
21	<b>f-C3_corps</b>	Fraction du bassin versant couvert par des plantes de type C3	-				
22	<b>f-C4_corps</b>	Fraction du bassin versant couvert par des plantes de type C4	-				
23	<b>f-garssland</b>	Fraction du bassin versant couvert par des prairies	-				
24	<b>f-broadleaf</b>	Fraction du bassin versant couvert par des feuillus	-				
25	<b>f-conif</b>	Fraction du bassin versant couvert par des conifères	-				
26	<b>Z0</b>	Moyenne annuelle de la rugosité	m				
27	<b>CZ0</b>	Coefficient d'irrégularité de la rugosité	-				
28	<b>veg</b>	Moyenne annuelle de taux de couverture végétale	m				
29	<b>Cveg</b>	Coefficient d'irrégularité temporelle de la couverture végétale	-				
30	<b>Lai</b>	Moyenne annuelle de l'indice foliaire sur le bassin	-				
31	<b>Clai</b>	Coefficient d'irrégularité temporelle de l'indice foliaire	-				
32	<b>greeness</b>	Moyenne annuelle de couvert vert	-				
33	<b>Cgreeness</b>	Coefficient d'irrégularité temporelle de couvert vert	-				
34	<b>glob-alb</b>	Moyenne annuelle d'albédo	-				
35	<b>Cglob-alb</b>	Coefficient d'irrégularité d'albédo	-				
<b>Les Descripteurs pédologiques</b>							
36	<b>Text</b>	Texture de sol dominante				-	
39		<b>Grossière</b>	<b>Moyenne</b>	<b>Moyenne fine</b>	<b>Fine</b>		
40	<b>PMH</b>	Type de matériaux d'origine hydrologique				-	
45		<b>R</b>	<b>C</b>	<b>S</b>	<b>L</b>		<b>H</b>
46	<b>DGH</b>	Profondeur du sol tourbière				m	
47	<b>OC-TP</b>	Teneur du sol en carbone organique				-	
48	<b>DIMP</b>	Profondeur de la couche imperméable				m	
49	<b>PD-TP</b>	La densité apparente de surface				-	
50	<b>HG</b>	Classe hydrogéologique de sol				-	
56		<b>HG-1C</b>	<b>HG-1S</b>	<b>HG-1L</b>	<b>HG-1H</b>		<b>HG-1M</b>
57	<b>AWCTP</b>	Capacité de teneur en eau du sol				-	
58	<b>EAWCT</b>	Capacité du sol en RFU				-	
59	<b>θsat</b>	Teneur en eau à la saturation				-	
60	<b>Ksat</b>	Conductivité hydraulique à saturation				m/s	



## 5 Régionalisation des modèles hydrologiques

### 5.1 La régionalisation par proximité géographique

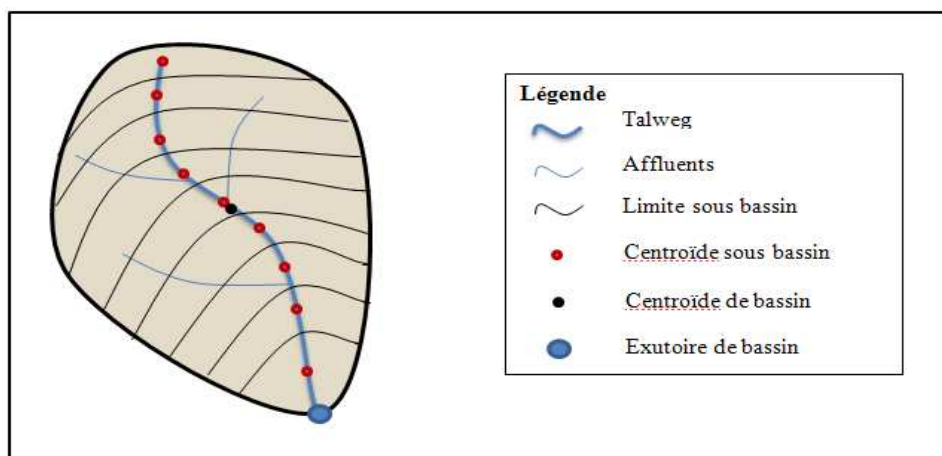
Cette approche est basée sur la proximité spatiale entre bassins versants et fait partie des premières tentatives de régionalisation pour la modélisation de bassins versants non-jaugés. Elle utilise la valeur de paramètres calés pour les voisins géographiques, ceci est justifié du fait que les caractéristiques physiques et climatiques sont relativement homogènes dans une région, de sorte que les voisins doivent se comporter généralement de façon similaire.

Au niveau de cette approche deux options étaient possibles : l'utilisation d'un jeu de paramètres unique (obtenu en faisant la moyenne des paramètres des voisins retenus) ou l'utilisation de plusieurs jeux de paramètres (un par voisin), sachant que l'on fera par la suite la moyenne des sorties (débits simulés). Pour cette étude uniquement la deuxième option a été retenue.

#### 5.1.1 Comment définit-on la similarité géographique?

L'objectif est de définir une distance géographique entre bassins appropriée à la régionalisation. Pour cela, différentes considérations de distance ont été mis en place :

- ✓ **Distance 1** : la distance euclidienne entre les centroïdes des bassins versants ;
- ✓ **Distance 2** : la distance euclidienne entre les exutoires des bassins versants ;
- ✓ **Distance 3** : la moyenne des deux distances précédentes ;
- ✓ **Distance 4** : la distance moyenne entre les centroïdes et les exutoires des deux bassins versants. Cette distance diffère de la précédente puisque les bassins n'ayant pas la même orientation seront considérés comme plus similaire ;
- ✓ **Distance 5** : la valeur minimale entre la distance 3 et la distance 4, cette distance permet de s'affranchir du biais lié à l'orientation des distances 3 et 4.
- ✓ **Distance 6** : la distance entre dix points caractéristiques de chaque bassin versant. Pour cela nous avons pris 10 points dans chaque bassin, chaque point est le centroïde d'une partie du bassin constituant 10% de la superficie totale du bassin (figure n°13).



**Figure 14:** Répartition des 10 points dans le bassin utilisé par la distance géographique

Ces distances sont utilisées pour déterminer pour chaque bassin considéré comme non jaugé un groupe de bassins versants jaugés proches, ce que nous appelons les « bassins donneurs ». Les paramètres calés sur les bassins proches sont ensuite transférés au bassin non jaugé. La performance des modèles avec ces jeux de paramètres permet de juger de l'intérêt de chacune des distances envisagées.

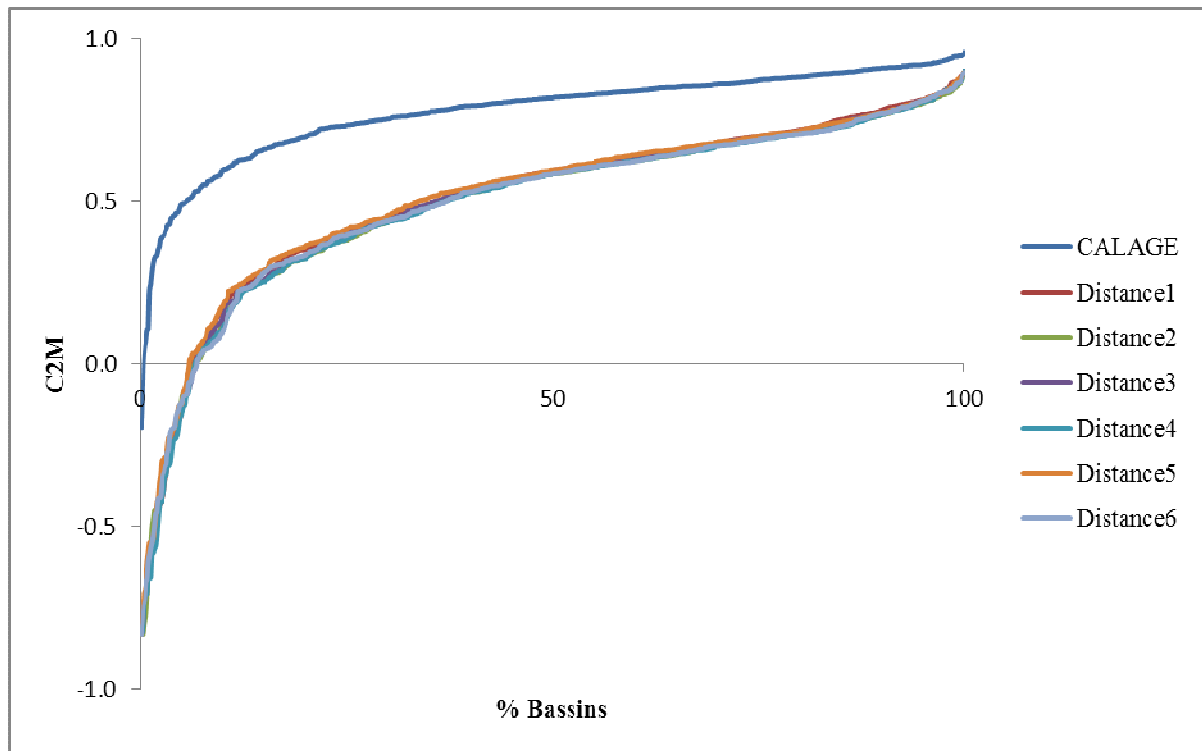
Le choix du nombre de bassins donneurs pour l'étude de la proximité spatiale a été basé sur les résultats de l'étude menée par (Oudin *et al.*, 2008). Le nombre de bassins donneurs est limité à 5 quel que soit le modèle hydrologique utilisé.

### 5.1.2 Méthode d'évaluation

Pour évaluer la performance relative des méthodes d'estimation des débits dans les bassins versants non-jaugés, chacun des 973 bassins versants a été successivement considéré comme non-jaugé, tous les autres bassins étant jaugés et donc des donneurs potentiels pour le bassin non jaugé. Pour chaque bassin pseudo-non jaugé, les 5 bassins donneurs sont identifiés à partir de la distance géographique considérée. Les paramètres de ces 5 bassins donneurs sont utilisés sur le bassin non jaugé, ce qui conduit à cinq chroniques de débit simulé. La moyenne des cinq chroniques est ensuite comparée aux observations et un calcul de la performance effectué via le C2M (Mathevet, 2005).

### 5.1.3 Résultats de la régionalisation spatiale

La figure 15 montrent, les distributions des performances (C2M) obtenues avec le modèle GR4J, les autres modèles ayant été testés et présentent les même résultats généraux.



**Figure 15:** Résultats de régionalisation par proximité géographique avec le modèle GR4J

Ces résultats indiquent que la cinquième distance présente des résultats meilleurs que les autres, même si toutes les distances testées sont très proches en terme de performances.

Le grand écart entre la distribution de performances C2M des méthodes de régionalisation avec celle obtenues en calage montre la marge de progression sur ce thème et pousse à réaliser d'autres recherches afin d'améliorer les simulations sur les bassins non jaugés. Pour cela, vu la forte ressemblance entre les résultats de la troisième et de la cinquième approche, une analyse plus poussée a été effectuée. De ce fait, une nouvelle configuration de choix de sélection des bassins donneurs a été envisagée et qui consiste à affiner et à optimiser les résultats de ces deux approches en suivant la démarche suivante.

#### - Pour la troisième approche : (Distance3)

- ✓ Nous avons appliqué la moyenne pondérée lors de la simulation de débit et qui exprimé par l'équation n° 2 (cf. équation 2 et 3 au paragraphe 2.4.4) :

$$Q_r(j) = \frac{1}{\sum W_k} * \sum_{j=1}^m W_k * Q_{pi}(j)$$

$$W_k = \frac{1}{D^\delta}$$

Avec :

- $Q_d$ : Débit du bassin donneur ;
  - $Q_r$ : Débit du bassin receveur ;
  - $D$  : Distance calculé pour la régionalisation : (soit distance spatiale entre bassins, soit distance entre les descripteurs physiques des bassins);
  - $m$ : Nombre des bassins donneurs ;
  - $\delta$ : Indique la puissance appliquée sur l'inverse de la distance.
- ✓ Nous faisons varier le nombre des bassins donneurs (m)

➡ Après ces corrections, cette distance est nommée **Distance Optimisée 1**

#### - Pour la cinquième approche : (Distance 5)

- ✓ Appliquer la moyenne pondérée des distances 1 et 2 pour la sélection des bassins donneurs, ceci est exprimée par l'équation suivante :

$$\text{Distance D} = \alpha * \text{distance1} + (1-\alpha) * \text{distance2} \quad \text{Équation 6}$$

Avec :

$\alpha$ : Indique le poids appliqué sur la distance

- ✓ Appliquer la moyenne pondérée par la distance pour la simulation de débit et qui exprimé par l'équation n° 2 (cf. équation au paragraphe 4.4.4).
- ✓ Faire varier le nombre  $m$  des bassins donneurs.

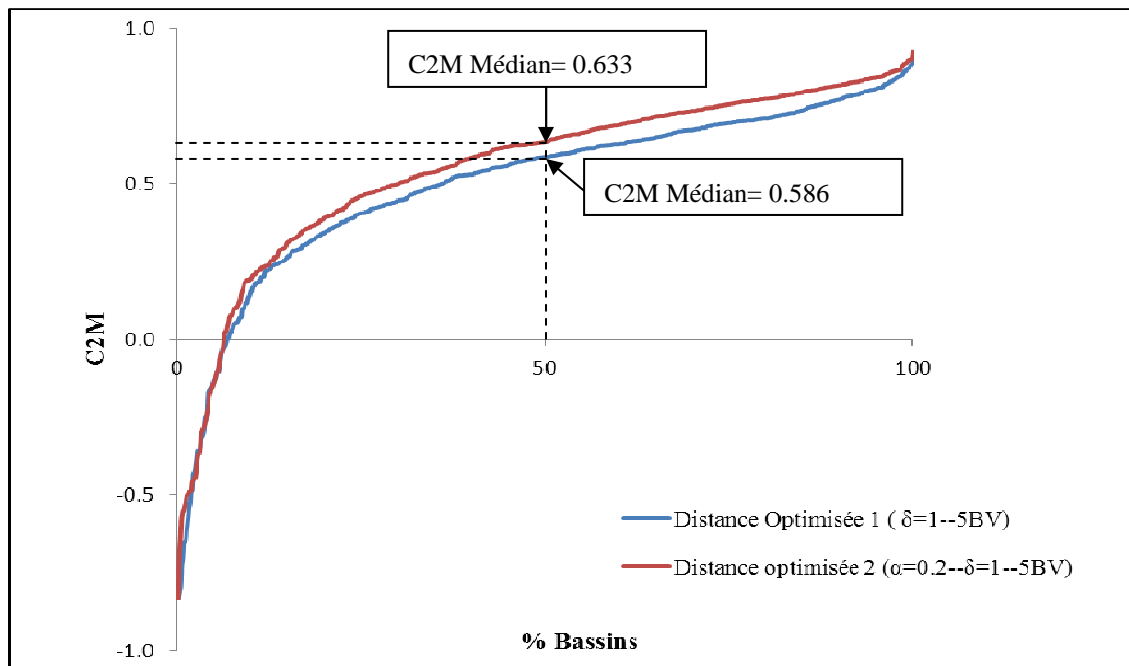
➡ Après ces corrections, cette distance est nommée **Distance Optimisée 2**

Afin d'optimiser l'évaluation de la proximité géographique, trois paramètres doivent être calés, ces trois paramètres ont été ajustés en maximisant le C2M moyen sur l'ensemble de 973 bassins versants. Le tableau suivant renseigne la gamme des valeurs testées.

**Tableau 8:** La gamme des valeurs des paramètres testées

Paramètres	Valeurs max	Valeurs min	Incrément
Poids sur la distance des centroide $\alpha$	0.9	0.1	0.1
Puissance appliquée sur l'inverse de la distance $\delta$	5	1	1
Nombre de bassins donneurs $m$	20	5	5

Les itérations avec toutes ces gammes des valeurs ont données les meilleurs résultats suivants :

**Figure 16:** Résultats d'optimisation de la proximité géographiques

La figure 16 ainsi que le tableau récapitulatif n°8 montrent qu'il y a un important progrès dans la distribution de critère de performance C2M pour **la distance optimisée 2**. Cette amélioration de performance peut être expliquée d'une part par l'optimisation de poids entre les distances des centroïdes et les distances des exutoires, et d'autre part, par l'optimisation de transfert des paramètres de modèle : transfert en appliquant une moyenne pondérée par la distance expliquée ci-dessus par l'équation n°2.

Le coefficient d'optimisation de poids  $\alpha$  est égal à **0.8**. Ce coefficient met en évidence la forte signification qu'apporte la distance entre les bassins versants, avec un poids moindre pour les distances entre les exutoires exprimées par la valeur **0.2 (1-  $\alpha$ )**.

Cette nouvelle formule de distance géographique (**distance optimisée 2**) a mis en évidence l'effet de l'orientation des bassins versants et/ou de la distance des surfaces entre les bassins. En effet, tous les bassins versants qui n'ont pas la même orientation que les bassins receveurs seront exclus de la liste des bassins donneurs.

Donc c'est **la distance optimisée n° 2** que l'on va retenir et considérer comme la plus appropriée pour la régionalisation par proximité spatiale, avec  $\alpha = 0.8$ ,  $\delta = 1$  et  $m = 5$ .

**Tableau 9** : Récapitulatif des résultats de la régionalisation spatiale

Modèle	C2M moyen					
	Calage	Distance1 (référence)	Distance3	Distance5	Distance Optimisée 1	Distance Optimisée 2
IHA	0.428	0.373	0.368	0.378	0.369	0.438
HBV	0.491	0.441	0.442	0.451	0.442	0.495
TOPMO	0.583	0.511	0.512	0.522	0.512	0.554
GR4J	0.650	0.507	0.502	0.514	0.502	0.553

### 5.1.4 La régionalisation dépend-elle de type de modèle adopté

L'analyse des résultats pour les quatre modèles hydrologiques adoptés dans cette étude montre que les valeurs de **C2M** de la régionalisation par proximité spatiale et pour les mêmes bassins versants diffèrent d'un modèle à un autre. Le modèle GR4J et TOPMO que sont très semblables en performance et apportent les meilleures garanties avec un C2M moyen de 0,552 pour GR4J et 0,554 pour TOPMO. Alors que le C2M moyen pour le modèle HBV est de 0.495 et pour le modèle IHA est de 0.438.

### 5.1.5 Conclusion relative

Les résultats de ce premier axe de régionalisation par proximité spatiale nous ont permis de définir une nouvelle distance appropriée pour la régionalisation. En effet, cette nouvelle approche (**Distance Optimisée 2**), a abouti à des résultats nettement meilleurs que ceux obtenus avec l'approche classique (**Distance 1 : distance entre les centroides**) qui a été adoptée par le passé (voir par exemple Oudin et al. (2008) et Reichl et al. (2009)).

## 5.2 La régionalisation par similarité physique

### 5.2.1 Comment définit-on la similarité physique?

Pour déterminer la similarité physique entre les bassins versants, nous nous sommes basés sur l'approche décrite par Kay et al. (2007) et qui définissent la similarité physique entre les bassins comme étant la distance euclidienne dans l'espace des caractéristiques physiographiques et climatiques des bassins. Les valeurs des caractéristiques sont normalisées par l'écart-type des caractéristiques, calculées sur l'ensemble des bassins (équation n°7).

$$dist_{a,b} = \sqrt{\sum_{j=1,J} W_j * \left( \frac{X_{aj} - X_{bj}}{\sigma_{Xj}} \right)^2} \quad \text{Équation 7}$$

- $X_{a,j}$ : Valeurs du descripteur  $X_j$  pour le bassin A ;
- $X_{b,j}$ : Valeurs du descripteur  $X_j$  pour le bassin B ;
- $\sigma_{Xj}$ : Écart-type de descripteur  $X_j$  pour tous les bassins ;
- $W_j$ : Poids attribué au descripteur  $X_j$  ;
- $J$ : Le nombre de descripteurs pris en compte.

On doit noter que les caractéristiques des bassins versant ne présentent pas les mêmes formes de distribution, ce qui peut potentiellement influencer le calcul de la distance par l'équation définie par Kay et al. (2007) (la transformation en log ne résout pas toujours le problème).

#### Le nombre des bassins donneurs $m$ :

Dans la première partie de la régionalisation spatiale, le nombre optimum des bassins donneurs trouvé dans notre étude coïncide avec celui qu'ont trouvé Oudin et al en 2008 et qui est égal à 5. Pour cette deuxième partie de régionalisation par proximité physique, nous allons suivre la même démarche et nous allons tester la même gamme de valeurs des bassins donneurs allant de 5 jusqu'au 20 avec un pas d'incrémentation égal à 5.

### **5.2.2 Différentes approches envisagées pour l'étude la similarité physiques**

Dans cette étude, trois méthodologies ont été envisagées pour définir la similarité physique entre bassins versants :

#### **5.2.2.1 Méthodologie d'Oudin et al. (2008) (Méthode 1)**

La sélection des bassins versants donneurs est basée sur la similarité des bassins non-jaugés à ceux jaugés en termes de descripteurs physiques. Par la suite, pour chacun des descripteurs, les bassins versants donneurs ont été classée par ordre décroissant de proximité au bassin versant non-jaugé. Lorsque plusieurs descripteurs ont été utilisés, nous avons utilisé les rangs des bassins donneurs pour chaque descripteur. Enfin pour chaque bassin un rang moyen a été utilisé pour classer les bassins versants donneurs. Ainsi chaque descripteur avait le même poids dans le calcul de proximité.

Six descripteurs physiques uniquement ont été mentionnés pour évaluer cette méthode et qui sont jugés les plus pertinents dans l'étude de Oudin et al (2008) pour l'étude de la similarité physique :

- ✓ La superficie du bassin versant
- ✓ L'altitude moyenne
- ✓ La pente moyenne
- ✓ La densité de drainage
- ✓ L'indice de couvert forestier
- ✓ L'indice d'aridité

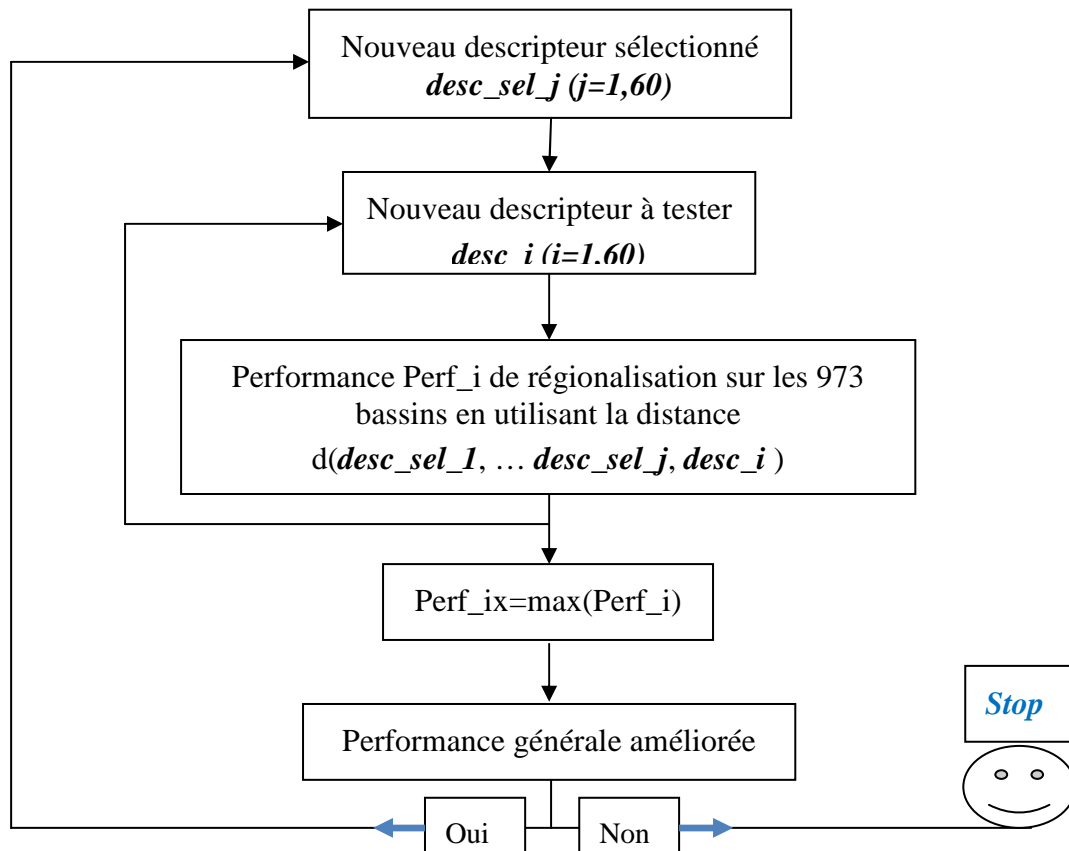
#### **5.2.2.2 Méthodologie utilisant de la distance euclidienne avec les six descripteurs d'Oudin et al. (2008) (Méthode 2)**

Avec cette méthode, la sélection des bassins donneurs se base sur le calcul de la distance euclidienne (cf. équation au paragraphe 7.2.1) entre les descripteurs physiques des bassins versants. En effet le calcul de la distance commence avec un seul descripteur et à chaque fois un autre descripteur s'ajoute dans le calcul de cette distance suivant leur proximité et leur similarité aux descripteurs du bassin non jaugé. L'évaluation de cette méthode a commencé avec les six descripteurs cités dans le paragraphe précédent avec des poids égaux.

#### **5.2.2.3 Sélection des descripteurs les plus performants (Méthode 3)**

Cette méthode se base sur le même principe que la précédente, la seule différence c'est que cette fois-ci, soixante descripteurs physiques (qui ont été sélectionnés à la partie 2) sont disponibles pour le calcul de la distance et pour l'évaluation de la similarité physique entre les

basins. L'objectif était d'élargir la base de sélection de descripteurs, de sélectionner les descripteurs les plus pertinents pour la réponse hydrologique des bassins versants et ainsi d'augmenter la performance des modèles sur les bassins non jaugés avec cette technique de régionalisation par la similarité physique. Le digramme suivant explique la procédure de fonctionnement de cette méthode.



La procédure de cette approche se résume comme suit : tout d'abord le modèle sélectionne un descripteur physique parmi la liste de 60 descripteurs disponibles, puis le modèle teste la performance de régionalisation de ce descripteur sur notre échantillon constitué par 973 bassins, par la suite le modèle détermine la valeur maximum de la performance pour ce descripteur, et il va faire un test sur ce dernier, si sa valeur est supérieure au seuil de performance préalablement fixé, le modèle reprend le calcul dès le début mais cette fois-ci, un autre descripteur qui s'ajoute au premier et le même démarche qui se répète, le calcul s'arrête le temps où le modèle trouve une valeur de performance de descripteur à l'étape  $i$  ( $Perf_{ix}$ ), inférieure au seuil de performance trouvé à l'étape  $i-1$  ( $perf_{i-1x}$ ).

### 5.2.3 Résultats et discussion

Les quatre modèles hydrologiques constituant notre support technique ont été utilisés pour évaluer les résultats de ces différentes méthodologies. Au début, nous avons commencé par le test des deux premières méthodes, sachant que ces dernières ont les mêmes descripteurs physiques (6 descripteurs décrites ci-dessus). Ceci, pour comparer ces résultats avant de passer à la troisième méthode qui prend en considération les soixante descripteurs. La comparaison des résultats des performances C2M entre les deux premières méthodes est illustrée dans le tableau suivant :

**Tableau 10:** Résultats d'évaluation de la première et de la deuxième méthode

Modèle	C2M moyen			
	Avec 5 BV donneurs		Avec 10 BV donneurs	
	Méthode 1	Méthode 2	Méthode 1	Méthode 2
<b>IHA</b>	0.394	0.437	0.393	<b>0.436</b>
<b>HBV</b>	0.406	0.486	0.398	<b>0.487</b>
<b>TOPMO</b>	0.481	0.537	0.480	<b>0.543</b>
<b>GR4J</b>	0.491	0.540	0.490	<b>0.544</b>
Modèle	C2M moyen			
	Avec 15 BV donneurs		Avec 20 BV donneurs	
	Méthode 1	Méthode 2	Méthode 1	Méthode 2
<b>IHA</b>	0.391	0.432	0.392	0.436
<b>HBV</b>	0.395	0.491	0.400	0.485
<b>TOPMO</b>	0.477	0.540	0.480	0.534
<b>GR4J</b>	0.489	0.542	0.487	0.538

Quel que soit le nombre de bassins donneurs testés, les résultats du tableau ci-dessus montrent que la deuxième méthode donne des résultats plus performants que ceux obtenus avec la première méthode. Ceci, justifie l'approche décrite par Kay et al en 2007 et qui définissent la similarité physique entre les bassins comme étant la distance euclidienne dans l'espace des descripteurs des bassins. Les résultats les plus pertinents, pour cette deuxième méthode, sont obtenus avec un nombre optimum de bassins versants donneurs égal à 10.

Ces résultats poussent les analyses vers la troisième méthode qui prend en compte les soixante descripteurs de bassins et se base sur le même principe que la deuxième méthode. Ceci afin de sélectionner les descripteurs les plus pertinents qui expriment la similarité physique entre les bassins. Le test d'évaluation de cette troisième méthode, en variant le nombre de bassins donneurs entre 5 et 20, donne les résultats des performances C2M illustrés dans le tableau suivant :

**Tableau 11 :** Résultats de la régionalisation par similarité physique avec la troisième méthode

Modèle	C2M Moyen			
	5 BV donneurs	10 BV donneurs	15 BV donneurs	20 BV donneurs
<b>IHA</b>	0.461	0.458	0.456	0.460
<b>HBV</b>	-	-	-	-
<b>TOPMO</b>	0.568	0.562	0.567	0.564
<b>GR4J</b>	0.575	0.572	0.570	0.571

Les résultats ci-dessus montrent, qu'avec la troisième méthode et avec un nombre de bassins versants donneurs égal à 5, nous avons abouti à des résultats meilleurs que ceux trouvés avec les deux premières méthodes et même plus performants que les résultats obtenus avec la régionalisation par proximité spatiale. Cette amélioration des résultats peut être expliquée par l'ajout des descripteurs pédologique qui ont été exclus avec les deux premières méthodes. En



effet, ces descripteurs sol ont apportés un plus pour l'identification de la similarité physique des bassins. Les descripteurs sélectionnés et considérés comme les plus performants, avec cette troisième approche, sont illustrés dans le tableau suivant.

**Tableau 12:** Descripteurs sélectionnés comme les plus performants pour la similarité physique avec la méthode 3

Modèle	IHA	TOPMO	GR4J	Nature d'utilisation de descripteurs
Nombre de descripteurs sélectionnés	11	13	13	
C2M moyen	<b>0.461</b>	<b>0.568</b>	<b>0.575</b>	
<b>Rang et descripteurs retenus pour chaque modèle</b>				
Superficie (km <sup>2</sup> )	8	10	-	Morphologique
Altitude moyenne (m)	7	2	-	Morphologique
Pente moyenne (°)	-	6	11	Morphologique
Densité de drainage (km/km <sup>2</sup> )	2	4	3	Morphologique
Pluie moyenne annuelle (mm)	1	7	-	Climatique
Coefficient d'irrégularité de pluie (-)	6	9	10	Climatique
ETP (mm)	4	5	5	Climatique
Coefficient d'irrégularité d'ETP (-)	3	3	6	Climatique
Indice d'aridité : ETP/P (-)	-	-	4	Climatique
K <sub>sat</sub> : perméabilité	5	8	2	Pédologique
Classe de texture de sol grossière	-	13	-	Pédologique
Classe PMH-C	10	-	1	Pédologique
Classe HG-1C	-	-	9	Pédologique
Classe HG-1L	-	1	7	Pédologique
Classe HG-1H	-	11	12	Pédologique
Classe HG-4W	9	-	13	Pédologique
Classe HG-4D	11	12	8	Pédologique

Ce tableau met en évidence les descripteurs sélectionnés comme les plus performants pour la similarité physique des bassins versants et pour chaque modèle. En plus ce tableau visualise le rang de sélection de chaque descripteur. Ce rang nous permet d'avoir une idée sur la pertinence de chaque descripteur dans l'identification de cette similarité. Si on met l'accent sur le modèle GR4J qui donne les meilleurs résultats, on trouve que ce dernier sélectionne 9 descripteurs parmi les 13 sélectionnés. Par ailleurs, aucun descripteur d'occupation de sol n'a été retenu (quel que soit le modèle hydrologique).

Cette troisième méthode nous a permis aussi d'identifier parmi la liste préliminaire de descripteurs (60 descripteurs) ceux qui sont pertinents pour la régionalisation par similarité physique. En effet, quel que soit le type de modèle hydrologique adopté 17 descripteurs uniquement sont conservés, alors que les 43 autres sont éliminés.

## Synthèse

Jusqu'au cette étape, nous constatons qu'au fils des tests, des améliorations des performances C2M sont apportées aux résultats des méthodes de régionalisation. Ceci est récapitulé dans le tableau suivant :

**Tableau 13:** Comparaison des résultats de la proximité géographique avec les similarités physiques (Les résultats indiqués pour la régionalisation avec la proximité géographique sont obtenus avec la distance optimisée 2 : moyenne pondérée entre les distances des centroides et les distances des exutoires avec optimisation des poids ( $\alpha = 0.8$ ,  $\delta = 1$  et  $m = 5$ ))

Modèle	C2M moyen			
	Proximité géographique (Distance optimisée 2)	Similarité physique		
		5 BV donneurs	10 BV donneurs	5 BV donneurs
		Méthode 1	Méthode 2	Méthode 3
<b>IHA</b>	0.438	0.394	0.436	<b>0.461</b>
<b>HBV</b>	0.495	0.406	0.487	-
<b>TOPMO</b>	0.554	0.481	0.543	<b>0.568</b>
<b>GR4J</b>	0.553	0.491	0.544	<b>0.575</b>

Ces résultats montrent que la similarité physique (troisième méthode) conduit à des meilleurs simulations de débit que la proximité géographique. Ce résultat est une nouveauté pour la régionalisation des modèles hydrologiques, car dans la littérature, on trouve que c'est la proximité spatiale qui à chaque fois donne les résultats les plus pertinents. Cette amélioration des résultats peut être rapportée à la nouvelle approche adoptée pour la sélection des descripteurs physiques ainsi qu'à la nouvelle liste de descripteurs utilisés notamment les descripteurs pédologiques.

Avec la régionalisation par proximité spatiale, nous avons fait le test d'optimisation de poids entre les distances et cela nous a apporté des bons résultats, ceci nous a poussé pour tester cette même procédure d'optimisation de poids des descripteurs physiques pris en compte pour chaque modèle hydrologique.

### 5.2.4 Optimisation des poids des descripteurs pris en compte

Les résultats obtenus précédemment avec la deuxième et la troisième méthode de similarité physique sont obtenus avec un poids égaux entre tous les descripteurs. Afin de rendre l'évaluation de ces résultats plus poussée et plus fine, une étude est faite à ce niveau et qui consiste à appliquer un poids relatif à chaque descripteur pris en compte. Pour cela, pour la deuxième méthode, les mêmes descripteurs physiques qui ont été testés (les six descripteurs sélectionnés par Oudin et al en 2008), alors que pour la troisième méthode, uniquement les descripteurs qui ont été mentionnés comme les plus performant ont été gardé pour l'étude (tableau n°11).

#### ✓ Optimisation des poids des descripteurs avec la deuxième méthode

Les résultats d'optimisation avec cette méthode ainsi que les valeurs des poids «  $\alpha_i$  » sont illustrés dans le tableau suivant.

**Tableau 14:** Optimisation de poids de six descripteurs physiques

Descripteurs	Les valeurs des poids « $\alpha_i$ »						C2M moyen
	Superficie	Altitude moyenne	Pente	Densité de drainage	Indice d'aridité	Indice de couvert végétal	
Modèle	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	
<b>IHA</b>	0.136	0.224	0.197	0.183	0.205	0.055	<b>0.446</b>
<b>HBV</b>	0.167	0.219	0.150	0.297	0.133	0.035	<b>0.497</b>
<b>TOPMO</b>	0.188	0.271	0.181	0.222	0.098	0.060	<b>0.554</b>
<b>GR4J</b>	0.167	0.219	0.193	0.219	0.147	0.055	<b>0.558</b>

✓ Optimisation des poids des descripteurs avec la troisième méthode

Les résultats d'optimisation avec cette méthode ainsi que les valeurs des poids «  $\alpha_i$  » sont illustrés dans le tableau suivant. Pour des raisons de calcul nous avons limité le nombre de descripteurs prise en compte à 10.

**Tableau 15:** Optimisation de poids de dix descripteurs physiques pour le modèle GR4J

Descripteurs	Les valeurs des poids « $\alpha_i$ »										C2M moyen
	PMH-C	$K_{sat}$	CETP	ETP/P	DD	CP	$Z_{moy}$	Pente	Pluie	HG-4D	
Modèle	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$	
<b>GR4J</b>	0.071	0.067	0.161	0.004	0.104	0.112	0.139	0.180	0.102	0.062	<b>0.580</b>

**Tableau 16:** Optimisation de poids de dix descripteurs physiques pour le modèle TOPMO

Descripteurs	Les valeurs des poids « $\alpha_i$ »										C2M moyen
	HG-1L	HG-1H	$Z_{moy}$	CETP	DD	CP	S	ETP	ETP/P	$K_{sat}$	
Modèle	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$	
<b>TOPMO</b>	0.039	0.057	0.090	0.153	0.112	0.059	0.050	0.167	0.105	0.167	<b>0.574</b>

**Tableau 17:** Optimisation de poids de dix descripteurs physiques pour le modèle IHA

Descripteurs	Les valeurs des poids « $\alpha_i$ »										C2M moyen
	DD	CETP	ETP/P	CP	$Z_{moy}$	S	HG-4W	$K_{sat}$	PMH-C	HG-4D	
Modèle	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$	
<b>IHA</b>	0.154	0.067	0.171	0.006	0.112	0.0083	0.149	0.180	0.102	0.057	<b>0.472</b>

Les résultats ci-dessus montrent que l'optimisation de poids avec la troisième méthode améliore les résultats de simulation des modèles hydrologiques. Cette amélioration est faible quel que soit le modèle hydrologique.

**Tableau 18:** Résultats de la troisième méthode sans et avec optimisation de poids de descripteurs physiques des bassins

Modèle		Nb des descripteurs	C2M moyen	Temps de calcul
<b>IHA</b>	<i>SOp</i>	11	0.461	Moyen (24 h)
	<i>AOp</i>	10	<b>0.472</b>	Élevé (11 jours)
<b>TOPMO</b>	<i>SOp</i>	13	0.568	Moyen (24 h)
	<i>AOp</i>	10	<b>0.574</b>	Élevé (11 jours)
<b>GR4J</b>	<i>SOp</i>	13	0.575	Moyen (24 h)
	<i>AOp</i>	10	<b>0.580</b>	Élevé (11 jours)
<i>SOp</i> : Sans optimisation de poids de descripteurs				
<i>AOp</i> : Avec optimisation de poids de descripteurs				

### 5.2.5 Synthèse

L'étude de ce deuxième axe de la régionalisation par proximité physique donne des résultats nettement meilleurs que la proximité spatiale (tableau n°19). Ces résultats peuvent être considérés comme des nouveautés dans les études de la régionalisation des modèles hydrologique. Cette amélioration peut être rapportée comme nous avons mentionné précédemment à l'intervention des descripteurs sol au niveau de l'identification de la similarité physique, et aussi à la nouvelle approche utilisée pour le choix et la sélection des descripteurs les plus pertinents.

**Tableau 19 :** Tableau récapitulatif des résultats de la proximité spatiale et de la similarité physique

Modèle	C2M moyen				
	Proximité géographique	Similarité physique			
		5 BV donneurs	10 BV donneurs	5 BV donneurs	5 BV donneurs
		Méthode 1	Méthode 2	Méthode 3 <i>SOp</i>	Méthode 3 <i>AOp</i>
<b>IHA</b>	0.438	0.394	0.436	0.461	<b>0.472</b>
<b>HBV</b>	0.495	0.406	0.487	-	-
<b>TOPMO</b>	0.554	0.481	0.543	0.568	<b>0.574</b>
<b>GR4J</b>	0.553	0.491	0.544	0.575	<b>0.580</b>

### 5.3 Combinaison de la similarité physique et de la proximité spatiale

L'étude des articles qui traitent le sujet de régionalisation des modèles hydrologiques montre qu'à chaque fois c'est la régionalisation par proximité spatiale qui donne les meilleurs résultats, alors qu'avec cette étude on a dépassé cette limite et on a prouvé que la régionalisation par similarité physique donne aussi de bons résultats. Ces résultats laissent espérer d'autres améliorations en combinant la proximité spatiale et la similarité physique des

bassins versants. Pour cela, trois approches ont été mises en évidence pour faire pousser l'étude de la régionalisation combinée.

**Approche 1 :** cette approche consiste à faire la combinaison entre les deux approches par **simple addition** de deux distances euclidiennes explicatives de chaque approche.

Vue qu'il est impossible d'additionner deux distances qui n'ont pas les mêmes dimensions, nous avons recours à la normalisation de ces distances en s'appuyant sur les équations suivantes :

$$D_s^* = \frac{D_s - \bar{D}_s}{\sigma_{D_s}} \quad \text{Équation 8}$$

$$D_p^* = \frac{D_p - \bar{D}_p}{\sigma_{D_p}} \quad \text{Équation 9}$$

$$D_m^* = D_s^* + D_p^* \quad \text{Équation 10}$$

Avec :

$D_s$  : Distance spatiale optée pour la régionalisation géographique

$D_p$  : Distance physique optée pour la régionalisation physique

$D_s^*$  : Distance spatiale normalisée

$D_p^*$  : Distance physique normalisée

$D_m$  : Distance mixte (pour la régionalisation combinée)

**Approche 2 :** combinaison de la proximité spatiale et de la similarité physique par **simple multiplication** des distances euclidiennes ( $D_s$  et  $D_p$ ) explicatives de chacune des approches. Ceci est expliqué par l'équation suivante :

$$D_m = D_s \times D_p \quad \text{Équation 11}$$

**Approche 3 :** combinaison de la proximité spatiale et de la similarité physique en se basant sur la méthode du rang moyen d'Oudin et al. (2008). Cette distance consiste à affecter à chaque bassin un rang relatif à la distance spatiale «  $R_s$  » et un autre rang relatif à la distance physique «  $R_p$  ». Par la suite un rang moyen est calculé pour chaque bassin «  $R_m$  » qui est le rang moyen des deux rangs ( $R_s$  et  $R_p$ ). Par la suite, on fait le tri croissant de ce rang moyen et on sélectionne les 5 premiers bassins qui correspondent aux 5 rangs les plus faibles.

$$R_m = R_s + R_p \quad \text{Équation 12}$$

### Avec ou sans optimisation de poids ?

L'étude de ces différentes approches a été réalisée en optimisant le poids, ceci en le faisant varier entre la valeur 0.1 et 0.9 avec un pas de 0.1. Pour chacune des approches précédentes on aura donc les équations d'optimisation de poids suivantes :

$$D_m = \alpha \times D_s^* + (1 - \alpha) \times D_p^* \quad \text{Équation 13}$$

$$D_m = D_s^\alpha \times D_p^{(1-\alpha)} \quad \text{Équation 14}$$

$$R_m = \alpha \times R_s + (1 - \alpha) R_p \quad \text{Équation 15}$$

### Qu'elles sont les distances adoptées pour l'étude de la régionalisation mixte ?

Pour cette étude de combinaison on va tenir compte uniquement des distances considérées comme les plus pertinents pour exprimer à la fois la proximité géographique et la similarité physique. Pour cela on va considérer :

#### La distance optée pour déterminer la proximité géographique $D_s$ est :

$$D_s = 0.8 * \text{distance (cent - cent)} + 0.2 * \text{distance (ex - ex)}$$

#### La distance optée pour déterminer la similarité physique $D_p$ est :

$$D_p = \alpha_1 * \text{descripteur1} + \alpha_2 * \text{descripteur2} + \dots + \alpha_{10} * \text{descripteur10} \quad \text{Équation 16}$$

Vu la contrainte de temps à la fin de ce stage nous n'avons pas pu tester les trois approches décrites ci-dessus, seule la deuxième approche a été testée. Les résultats de cette simulation sont illustrés dans tableau suivant.

**Tableau 20:** Résultats de la deuxième approche de combinaisons de la similarité physique et de la proximité géographique

Proximité Physique	C2M moyen										Proximité Géogra	
	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$	$\alpha_7$	$\alpha_8$	$\alpha_9$	$\alpha_{10}$	$\alpha_{11}$	
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	C2M- Max
Approche2	0.575	0.576	0.581	0.582	0.578	0.577	0.572	0.567	0.561	0.557	0.553	

Ce tableau montre que la combinaison des approches physique et géographique avec multiplication de ses deux distances n'améliore que très peu les résultats de simulation de débit par le modèle hydrologique GR4J. De ce fait, nous allons garder les résultats obtenus par la proximité physique comme les plus pertinents pour répondre au besoin de la régionalisation des modèles hydrologiques. Ces résultats ne défavorisent pas l'approche de combinaison de la proximité spatiale et de la similarité physique, par contre il nous laisse espérer d'autres améliorations peut-être en appliquant les deux autres approches décrites dans ce mémoire.

## 6 Conclusion

Le but de ce mémoire était d'améliorer les techniques de régionalisation des modèles hydrologiques, ceci en s'inspirant des expériences des études précédentes (les points forts et les points faibles de chaque étude) ainsi que des différentes approches adoptées. Avant de projeter la lumière sur les résultats de cette étude, nous voulons juste rappeler que dès les années soixante jusqu'à l'année 2010, différentes études ont été lancées pour atteindre cet objectif. Malgré le fait que chaque étude ait ses propres particularités, la conclusion de beaucoup d'entre elles était que la proximité spatiale représentait l'approche la plus pertinente pour la régionalisation des modèles hydrologiques. Les résultats que nous avons trouvés avec cette étude nous ont permis de :

- Proposer une nouvelle expression pour la distance géographique entre bassins (**Distance optimisée 2 : paragraphe 5.1.3**) qui a amélioré le critère de performance C2M de 10.9% (**C2M distance 1= 0.507, C2M distance optimisée 2= 0.553**).
- Démontrer que l'approche de la similarité physique était supérieure à celle de proximité géographique lorsque les descripteurs de bassins utilisés étaient sélectionnés.

Plusieurs modélisateurs ont espéré que la combinaison des approches spatiales et physiques améliore les résultats de la régionalisation, alors qu'avec notre première expérience de combinaison et avec les trois approches adoptées nous n'avons pu trouver des résultats convainquants. Ceci n'exclut pas cette combinaison, mais nous invite à pousser encore plus les analyses et les recherches vers d'autres approches appropriées à cette combinaison.

Un autre point de ce mémoire sur lequel nous pouvons projeter la lumière, c'est la base de données utilisée pour identifier les descripteurs physiques. Il s'agit d'une base de données importante, les descripteurs peuvent être calculés sur d'autres bassins en Europe puisque les bases de données géomorphologiques, pédologiques et occupation de sol sont disponibles à cette échelle et accessibles à tout le monde ce qui fait que les approches adoptées dans cette étude peuvent être appliquées à n'importe quelle base de données dans le monde, ce qui n'a été pas le cas avec les études précédentes. Ceci ouvre des nouvelles perspectives : que peuvent être les résultats de ces approches des bassins versants d'autres pays ? Peut-on uniformiser une approche de régionalisation des modèles hydrologiques à l'échelle internationale ?

## 7 Référence Bibliographique

Acreman, M. C., and C. D. Sinclair (1986), Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland, *J. Hydrol.*, 84, 365–380.

Bergström, S. et Forman, A., 1973. Development of a conceptual deterministic rainfall-runoff model. *Nordic Hydrology*, 4: 147-170.

Beven, K.J. et Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1): 43-69.

Castiglioni, S. Castellarin, A., Laaha, G., Skøien, J., Blöschl, G., 2009. "Physiographic spacebased interpolation vs. TOPKRIGING for predicting low-flow indices". Poster: EGU General Assembly 2009, in: "Geophysical Research Abstracts", 11, Paper ID EGU2009-12332.

Egbuniwe, N., and D. K. Todd (1976), Application of the stanford watershed model to Nigerian watersheds, *Water Resour. Bull.*, 12, 449–460.

Garrick, M., C. Caunnane, and J.E., 1978. Nash, A criterion of efficiency for Ranfall-runoff models., *Journal of Hydrology*, 38, 375.

Gaume, E., Livet, M., Desbordes, M., Villeneuve, J-P., 2002. Hydrological analysis of the Avene river (france) extraordinary flood 6 and 7 October 1997. *Physics and Chemistry of the Earth* 28 (6–7), 263.

Haberlandt, U., Klöcking, B., Krysanova V., Becker, A., 2001. Regionalisation of the base flow index from dynamically simulated flow components - a case study in the Elbe River Basin. *Journal of Hydrology*, 248: 35-53.

Hundecha, Y., and A. Bárdossy (2004), Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, *J. Hydrol.*, 292, 281–295.

Jakeman, A.J., Littlewood, I.G. et Whitehead, P.G., 1990. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology*, 117: 275-300.

Kay, A. L., Jones, D. A., Crooks, S. M., Calver A., Reynard, N. S., 2006. A comparison of three approaches to spatial generalization of rainfall-runoff models, *Hydrology Processessings*, 20: 3953– 3973.

Magette, W. L., V. O. Shanholtz, and J. C. Carr (1976), Estimating selected parameters for the Kentucky watershed model from watershed characteristics, *Water Resour. Res.*, 12, 472–476.

Mathevet, T., 2005. Quel modèle pluie débit Globaux au pas de temps horaire ? Développement empirique et comparaison de modèle sur un large échantillon de bassin versant. 463, Cemagref/Engref.



- McIntyre, N., H. Lee, H. Wheeler, A. Young, and T. Wagener (2005), Ensemble predictions of runoff in ungauged catchments, *Water Resour. Res.*, 41, W12434, doi:10.1029/2005WR004289.
- Merz, R., Blöschl, G., Piock-Ellena, U., Rieger, W., 2000. Regionalisierung von Bemessungshochwässern mit geostatistischen Verfahren (Regional estimation of design floods by geostatistical techniques). *Interpraevent* (1): 71-84.
- Merz, R., Blöschl, G., 2004. Regionalisation of catchment model parameters. *Journal of Hydrology*, 287: 95-123.
- Merz, R., Blöschl, G., Humer, G., 2008. National flood discharge mapping in Austria. *Natural Hazards* 46:53-72.
- Oudin, L., 2004. Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global. 495, Cemagref/Engref
- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005. Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. *Journal of Hydrology*, 303: 290-306.
- Oudin, L., Andreassian, V., Perrin, C., Michel, C., Lemoine, N., 2008. Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research* 44: 15 p.
- Oudin, L., Kay, A. L., Andreassian, V., Perrin, C., 2010. Are seemingly physically similar catchments truly hydrologically similar? *Water Resources Research*, doi: 10.1029/2009WR008887.
- Parajka, J., Merz, R., Blöschl, G., 2005. A comparison of regionalisation methods for catchment model parameters. *Hydrology and Earth System Sciences*, 9: 157–171.
- Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., Scipal, K., 2006. Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale. *Hydrology and Earth System Sciences* 10 (3): 353-368.
- Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279: 275-289.
- Quintana segui, P. Habets, F. Boone, A. Champeaux, J. L. Etchevers, P. Franchiste´guy, L. Leblois, E. Ledoux, E. Le Moigne, P. Martin, E. Morel, S. Noilhan, J. Rousset-Regimbeau, F. and Viennot, P., 2008. The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. *Journal of geophysical research*, 113, doi: 10.1029/2007JD008548.
- Sauquet, E., 2006. Mapping mean annual river discharges: Geostatistical developments for incorporating river network dependencies. *Journal of Hydrology* 331: 300- 314.
- Sivapalan, M. (2003). "Prediction in ungauged basins: a grand challenge for theoretical hydrology." *Hydrological Processes* 17(15): 3163-3170.

Vandewiele, G. L., C. Y. Xu, and W. Huybrecht (1991), Regionalisation of physically-based water balance models in Belgium: Application to ungauged catchments, *Water Resour. Manage.*, 5, 199–208.

Young, A. R. (2006), Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model, *J. Hydrol.*, 320, 155–172.

# Annexes

## Annexe n°1 : Les descripteurs utilisés par les auteurs

### Merz et al. 2009 :

- ✓ Pluie moyenne annuelle
- ✓ Evaporation moyenne annuelle
- ✓ Indice d'aridité
- ✓ Evapotranspiration réelle
- ✓ Evaporation potentielle
- ✓ Débit de base de ruissellement
- ✓ Altitude moyenne
- ✓ Pente topographique moyenne
- ✓ Densité de réseau hydrographique
- ✓ Longueur hydraulique de la rivière
- ✓ Longueur de cours d'eau principale à partir du centre de gravité
- ✓ La pente moyenne du cours d'eau principale
- ✓ Pourcentage de dépôt de sédiment quaternaire dans le bassin versant (calcaire, dolomie, argile, marne, grès, phyllite, schiste, granit et de gneiss)
- ✓ Pourcentage de la surface agricole et forestière dans le bassin versant
- ✓ Pourcentage de fluvisol, Lithosol, rendzine, phaeozem, tchernoziom, cambisol, luvisol, podzol, et la zone histosol dans le bassin versant.
- ✓ Numéro de courbe SCS (selon le type de sol et l'utilisation des terres)

### Reichl 2009 :

- ✓ COS moyen
- ✓ Ecart type de COS
- ✓ Elévation minimal du relief dans le bassin
- ✓ Altitude maximal du relief dans le bassin
- ✓ Altitude moyenne du relief dans le bassin
- ✓ Ratio d'élévation de relief  $[(\text{élévation moy} - \text{élévation min}) / (\text{élévation max} - \text{élévation min})]$
- ✓ Pente topographique moyenne
- ✓ Ecart type de la pente
- ✓ Moyenne des X
- ✓ Ecart type des X
- ✓ Moyenne des Y
- ✓ Ecart type des Y
- ✓ Distance moyenne dans l'espace de (X-Y)
- ✓ Nombre d'affluents
- ✓ Surface du bassin versant
- ✓ Densité de drainage
- ✓ La moyenne des précipitations hivernales
- ✓ La moyenne des précipitations estivales
- ✓ ETP moyenne annuelle

- ✓ Profondeur moyenne du sol
- ✓ Capacité de rétention en eau des plantes
- ✓ La perméabilité du sol
- ✓ La transmissivité latérale
- ✓ Fraction du bassin versant couvert par la végétation ligneuse indigène
- ✓ Fraction du bassin versant couvert par la végétation ligneuse

**Oudin et al. 2008 :**

- ✓ La superficie du bassin versant
- ✓ L'altitude moyenne
- ✓ La pente moyenne
- ✓ La densité de drainage
- ✓ L'indice de couvert forestier
- ✓ L'indice d'aridité

**Parajka et al 2007 :**

- ✓ Moyenne annuelle de précipitation
- ✓ La moyenne des précipitations hivernales
- ✓ La moyenne des précipitations estivales
- ✓ Hauteur de la neige
- ✓ Température de l'air
- ✓ Evapotranspiration potentiel
- ✓ Surface du bassin versant
- ✓ Altitude moyenne
- ✓ Coefficient de variation de l'altitude
- ✓ Pente moyenne topographique
- ✓ Coefficient de variation de la pente
- ✓ Minimum d'indice d'humidité topographique
- ✓ Maximum d'indice d'humidité topographique
- ✓ Densité de réseau hydrographique
- ✓ Coefficient d'atténuation des inondations par les barrages et les lacs
- ✓ La portion des aquifères poreux
- ✓ Le type de sol régional
- ✓ Les formations principales géologiques

**McIntyre et al 2005 :**

- ✓ Précipitation moyenne annuelle
- ✓ Indice de débit de base
- ✓ Densité de drainage
- ✓ Altitude de bassin versant
- ✓ Fraction de zone urbaine dans le bassin versant
- ✓ Coefficient de ruissèlement

**Annexe n° 2: Etude de corrélation entre les 60 descripteurs physiques**

## Résumé

Les modèles hydrologiques sont classiquement calés en utilisant des mesures de débit, auxquelles on essaie de faire correspondre au mieux les simulations. Cette procédure est inadaptée pour les bassins non jaugés, sur lesquels on ne dispose d'aucune observation de débit. Dans ce cas, on utilise des méthodes dites de régionalisation, qui permettent de transposer des paramètres déterminés sur des bassins jaugés vers des bassins non jaugés. Ces méthodes se basent sur la similarité de comportement des bassins. Toute la difficulté de telles approches réside dans la détermination de cette similarité. Afin d'aboutir à des conclusions robustes pour cet objet nous avons suivis le démarche suivant :

Dans la première partie du mémoire nous avons fait une synthèse des travaux antérieurs qui porte sur la régionalisation des modèles hydrologiques. Ces travaux nous a permis de définir le contexte générale de l'étude, d'analyser et de diagnostiquer les méthodes existantes de la régionalisation et de proposer de nouveau pistes pour améliorer les résultats.

Durant la deuxième partie, nous avons préparé en premier lieu un échantillon de bassins versants bien répartis sur le territoire français. Sur cet échantillon nous avons appliqué une procédure de traitement automatique des données aberrantes (débit interpolé). Et en second lieu nous avons entamé une phase de collecte et de traitement des descripteurs physiques qui sont disponibles à l'échelle internationale, ceci afin de dépasser la contrainte d'hétérogénéité des données entre les études et de rassembler un nombre de descripteurs suffisant pour expliquer la réponse hydrologiques des bassins versants.

Concernant la troisième partie, des nombreuses approches de régionalisation ont été développées, certaines exploitant la ressemblance physique des bassins, d'autres la proximité géographique. Avec ces nouvelles approches testées nous avons pu trouver des résultats qui favorisent la similarité physique, ce qui est nouveau pour la régionalisation des modèles hydrologiques.