

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

**ASPECTS THEORIQUE ET APPLIQUE DE LA
MODELISATION EN HYDROLOGIE :**

*Amélioration d'une méthode de calage et application d'une technique
de détection de changements du comportement hydrologique des
bassins forestiers altérés*

Josep Lluís RUIZ BELLET

Directeurs de recherche :

Vazken ANDREASSIAN

Thibault MATHEVET



Cemagref
UR Hydrosystèmes et Bioprocédés
Parc de Tourvoie BP 44
92163 Antony

Septembre 2005



REMERCIEMENTS

Merci beaucoup à mes maîtres de stage : Thibault Mathevet et Vazken Andréassian pour leurs aide et disponibilité. Merci aussi à Charles Perrin, pour la résolution des doutes et pour la lecture du rapport.

Je veux remercier aussi tous ceux qui nous ont prêté leurs données, notamment Eric Prinsloo de l'Afrique du Sud et Vickie Dreher des Etats-Unis.

Merci aussi à Nico pour son aide avec le français et à Golnoush pour son aide avec le persan.

Merci à Anne pour son soutien.

I, finalment, gràcies a la Fundació Ruiz Bellet per concedir-me la substanciosa beca que porta el meu nom i que m'ha permès d'estudiar a París.

*Los computadores han llegado a ser
indispensables como herramientas para el
desarrollo y uso de muchos modelos. Pero
nótese la analogía entre modelado por
computador y poesía por máquina de escribir.*

Óscar García, 1994

*Mais, bien sûr, nous qui comprenons la vie,
nous nous moquons bien des numéros !*

Le petit prince, Antoine de Saint-Exupéry

TABLE DES MATIERES

INTRODUCTION GENERALE.....	1
PARTIE THEORIQUE : AMELIORATION D'UNE METHODE DE CALAGE.....	2
1. INTRODUCTION.....	2
1.1 Calage.....	2
1.2 Fonction objectif	3
1.3 Les caractéristiques d'une bonne méthode de calage.....	5
1.4 Amélioration des méthodes locales.....	6
1.5 Comparaison à l'aide d'un grand nombre de bassins versants.....	6
2. OBJECTIFS	7
3. MATERIAUX ET METHODES	8
3.1 Données et modèles utilisés	8
3.2 Fonction objectif	9
3.3 Split-sample test	9
3.4 Méthodes de calage préexistantes	10
3.5 Construction des nouvelles méthodes de calage	10
3.6 Comparaison des méthodes de calage	12
4. RESULTATS ET DISCUSSION.....	14
4.1 Efficacité	14
4.2 Performance	14
4.3 Robustesse.....	16
4.4 Résumé des résultats	18
5. CONCLUSIONS.....	20
PARTIE D'APPLICATION : APPLICATION D'UNE TECHNIQUE DE DETECTION DES CHANGEMENTS DU COMPORTEMENT HYDROLOGIQUE DES BASSINS FORESTIERS ALTERES.....	21
1. INTRODUCTION.....	21
1.1 Le comportement hydrologique	21
1.2 Influence de la forêt sur le comportement hydrologique	21
1.3 Techniques de détection de changements du comportement hydrologique.....	21
1.4 La technique de la Réponse hydrologique potentielle (Rhp)	23
2. OBJECTIFS	25
3. MATERIAUX ET METHODES	26
3.1 Bassins versants.....	26
3.2 Modélisation pluie-débit	26
3.3 La technique de la Réponse hydrologique potentielle (Rhp)	26
4. RESULTATS ET DISCUSSION.....	28
5. CONCLUSION	30
6. BIBLIOGRAPHIE	31
ANNEXE C : RESULTATS DE LA PARTIE THEORIQUE	36
ANNEXE D : RESULTATS DE LA PARTIE APPLIQUEE	44
ABSTRACT	52
RESUM	52
RESUME.....	52

INTRODUCTION GENERALE

En raison, d'une part, du développement des sciences informatiques et de la collecte des données, et, d'autre part, de l'augmentation des problèmes liés à la gestion de l'eau et de la vulnérabilité aux crues, la modélisation mathématique est de plus en plus utilisée en hydrologie, de sorte qu'elle en est actuellement la branche la plus dynamique.

La modélisation hydrologique pourrait être divisée, en faisant le parallèle avec la physique, en modélisation théorique et appliquée, ou autrement dit : la modélisation *sensu stricto* et son application dans les différents domaines de l'hydrologie.

Dans la modélisation *sensu stricto* (la partie « théorique »), on trouve tous les aspects liés à son développement : construction et amélioration des modèles (Michel 1983, Edijatno et Michel 1989, Perrin *et al.* 2003, Le Moine 2005, Asadian 2005), tests de comparaison des performances (Mathevet *et al.* 2005), calcul des incertitudes (Berthier 2005), couplage avec des modèles hydrogéologiques ou atmosphériques, méthodes de calage (Mathevet 2005).

Dans, la modélisation appliquée, on trouve les techniques liées à son utilisation comme outil de simulation, avec pour exemples : la gestion de barrages, les plans d'urgence en cas de crue ou l'hydrologie forestière (Andréassian 2002).

Ce stage a eu un double objectif : théorique et appliqué. Il a donc comporté deux parties : dans la première on a essayé d'améliorer une méthode de calage et, dans la deuxième, on a utilisé cette méthode de calage lors l'application d'une technique de détection des changements du comportement hydrologique des bassins forestiers altérés.

PARTIE THEORIQUE : AMELIORATION D'UNE METHODE DE CALAGE

1. INTRODUCTION

Ces dernières années, en hydrologie, la simulation (appliquée en prévision et en prédétermination) est devenue un domaine stratégique à cause de l'augmentation des problèmes liés à la gestion de l'eau et de la vulnérabilité aux crues. Malheureusement, la grande complexité des processus hydrologiques à l'échelle du bassin versant rend la simulation difficile. Malgré tout, les hydrologues ont créé un grand nombre de modèles pluie-débit, qui sont à ce jour des outils puissants, quoique encore imparfaits.

En effet, un modèle est une simplification de la réalité qui permet de travailler avec celle-ci plus facilement. Il existe différents types de modèles : graphiques, physiques (maquettes), mathématiques. Les modèles pluie-débit appartiennent dans ce dernier groupe, car ils sont formés par un nombre d'équations et de fonctions.

La variété des modèles pluie-débit est grande et on peut les classer selon diverses caractéristiques :

- i) approche : modèles de type boîte noire, à réservoirs, à base physique
- ii) division de l'espace : modèles globaux et distribués
- iii) pas de temps : mensuel, journalier, horaire
- iv) objectif : prévision des crues, gestion de la ressource eau

Cette grande diversité est la conséquence du fait que chaque type de modèle a des avantages et des inconvénients ; par exemple, pour les modèles à réservoirs, l'un des inconvénients est la détermination des paramètres.

1.1 Calage

Un modèle global pluie-débit à réservoirs est constitué typiquement d'une structure fixe et d'un nombre de paramètres libres lui permettant de s'adapter à un bassin versant déterminé. Ces paramètres ont rarement une signification physique, c'est-à-dire qu'on ne peut pas les mesurer ou les calculer à partir des caractéristiques du bassin, mais on doit les estimer dans la phase dite de calage, aussi appelée calibration ou optimisation. « L'optimisation des paramètres d'un modèle pluie-débit a pour but de trouver le jeu de paramètres qui rapproche le plus possible le comportement du modèle de celui du bassin modélisé, la similitude des comportements étant quantifiée par un critère (fonction objectif) servant à la fois à l'optimisation des paramètres et mesurant le degré de similitude » (Perrin 2000). Comme les modèles sont des systèmes d'équations non-linéaires, le calage est une étape très délicate et la qualité des paramètres (et celle des résultats qu'ils fourniront) dépend fortement de la puissance et la robustesse de la méthode de calage.

En effet, les méthodes de calage¹ sont sensibles aux problèmes de nature numérique : equifinalité (Beven 2001) (interdépendance des paramètres), existence de maxima locaux, non-différentiabilité de la fonction en certains points, faible sensibilité de la fonction objectif. De plus, les résultats sont dépendants de la fonction objectif et des séries de données utilisées. A cause de tous ces problèmes, il faut davantage demander aux méthodes de calage des jeux de paramètres stables et « conceptuellement acceptables » que des maxima mathématiques.

¹ On ne parle ici que des méthodes automatiques ; les méthodes manuelles, basées sur l'analyse graphique, sont trop subjectives et limitées pour travailler avec un grand nombre de données

Les méthodes de calage peuvent se diviser en deux grandes familles: les méthodes locales et les méthodes globales (Perrin 2000). Les méthodes locales adoptent une stratégie itérative dans laquelle on se déplace dans une direction qui améliore continûment la valeur de la fonction objectif, jusqu'à ne plus pouvoir générer d'amélioration. Les méthodes globales, par opposition aux méthodes locales, explorent une partie beaucoup plus grande de l'espace des paramètres, afin de trouver l'optimum de la fonction objectif en évitant les optima locaux. Ces deux méthodes ont leurs avantages et inconvénients : les méthodes locales convergent rapidement vers un optimum mais peuvent tomber sur un optimum local et leur résultat dépend fortement du jeu de paramètres initial, alors que les méthodes globales identifient mieux l'optimum mais ont des temps de calculs prohibitifs.

Enfin, on peut diviser les méthodes locales en :

- i) méthodes directes, qui n'utilisent comme information au cours de l'optimisation que la valeur de la fonction objectif (exemples : Rosenbrock, Simplex, Pas à pas)
- ii) méthodes de gradient, qui utilisent, avec la valeur de la fonction objective, la valeur du gradient de cette fonction (exemple : David-Fletcher-Power) ; elles sont plus rapides que les directes, mais beaucoup moins robustes.

De même, les méthodes globales peuvent se classer selon la stratégie d'exploration en:

- i) déterministes : discrétisation systématique de l'espace des paramètres en mailles qui comporte des temps de calcul prohibitifs (exemple : Exhaustive Gridding)
- ii) stochastiques : échantillonnage de l'espace des paramètres de façon aléatoire à partir de lois de probabilité des paramètres et concepts d'évolution de populations (exemple : Sce-Ua, Ga-Edf)
- iii) combinatoires : méthode stochastique d'échantillonnage pour trouver les points de départ et méthodes locales pour suivre l'optimisation (exemple : Msx)

D'après Perrin (2000), les méthodes globales sont plus efficaces que les globales sur des données synthétiques, mais sur des données réelles, cet avantage théorique n'est pas vérifié, et les méthodes locales sont aussi bonnes que les globales ; de plus, les méthodes locales sont plus rapides que les globales ; par exemple, la méthode du Simplex est trente fois plus rapide que Sce-Ua et 100 fois plus rapide que Msx.

1.2 Fonction objectif

Comme on l'a vu antérieurement, le calage trouve les valeurs des paramètres qui rapprochent le plus les résultats du modèle des observations. Or, pour faire cela, il faut un critère d'évaluation de l'écart entre les débits calculés et les observés. On peut en distinguer deux types : les qualitatifs et les quantitatifs ; les premiers sont basés sur l'analyse des graphiques pendant que les deuxièmes se servent de formules mathématiques et, sont pour cela, automatisables et plus objectifs. Dans la suite de ce travail, on ne parlera que des critères quantitatifs.

Le critère d'évaluation, aussi appelé fonction objectif, mesure l'écart entre les débits simulés et les débits observés, c'est-à-dire, l'erreur du modèle. Géométriquement, on peut définir la fonction objectif comme une hypersurface dans un hyperespace de $n+1$ dimensions, où n est le nombre de paramètres du modèle ; la Figure 1 montre le cas pour un modèle à deux paramètres. Le calage est donc l'opération de trouver un jeu de paramètres qui minimise l'erreur du modèle, c'est-à-dire, trouver le minimum (ou maximum) de la fonction objectif.

La fonction objectif joue aussi un rôle central dans la phase de validation (aussi dite de contrôle) où elle mesure la divergence entre les observations et les résultats du modèle pour des données appartenant à une période autre que celle utilisée pour le calage des paramètres.

De cette façon, le calage des paramètres est testé sur des données indépendantes et, par conséquent, la fonction objectif en validation nous informe sur la vraie qualité du calage et sur sa robustesse.

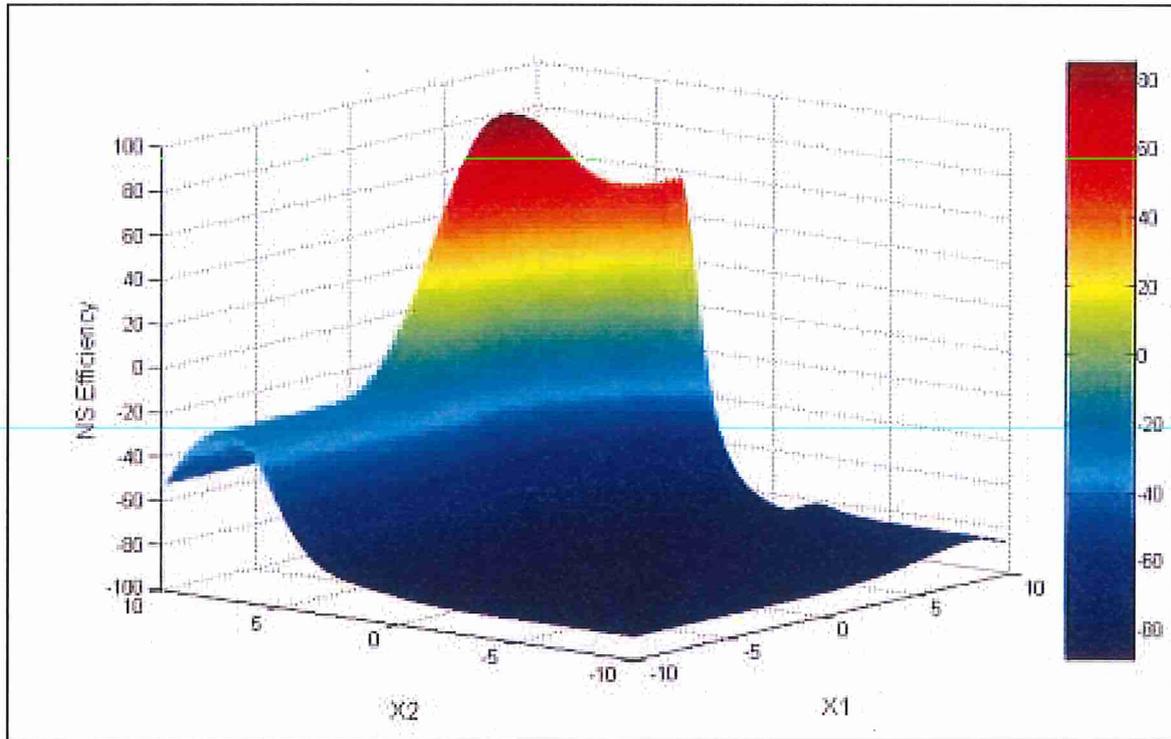


Figure 1. Fonction objectif d'un modèle a deux paramètres (Source : Mathevet 2005)

Il y a plusieurs types de fonctions objectif ; le choix dépend des objectifs de l'utilisateur : prévision des crues, gestion de ressource, simulation d'étiages. Dans cette étude on parlera de trois types de fonction objectif très utilisés, basés sur trois types d'erreur : l'erreur quadratique (Eq. 1), l'erreur absolue (Eq. 2) et l'erreur cumulée (Eq. 3).

$$F = \sum_{i=1}^n (Q_{obs,i} - Q_{cal,i})^2 \quad (\text{Eq. 1})$$

$$A = \frac{1}{n} \sum_{i=1}^n |Q_{obs,i} - Q_{calc,i}| \quad (\text{Eq. 2})$$

$$EC = \frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - Q_{calc,i}) = \overline{Q_{obs,i}} - \overline{Q_{calc,i}} \quad (\text{Eq. 3})$$

Les critères d'erreur quadratique, aussi appelés critères en moindres carrés, sont très utilisés en modélisation pluie-débit. Une des ces fonctions objectif est le critère de Nash-Sutcliffe (NS) qui compare (Eq. 5) l'erreur quadratique du modèle à l'erreur quadratique d'un modèle naïf de référence (Eq. 4) ; ce modèle naïf est la simple moyenne arithmétique des débits observés (Nash et Sutcliffe 1970).

$$F_0 = \sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2 \quad (\text{Eq. 4})$$

$$NS = 1 - \frac{F}{F_0} \quad (\text{Eq. 5})$$

NS peut prendre des valeurs entre $-\infty$ et 1 ; NS prend des valeurs négatives quand l'erreur quadratique du modèle est plus grande que l'erreur quadratique du modèle naïf (c'est-à-dire, que le modèle explique moins bien les observations que la moyenne arithmétique) et des valeurs positives quand l'erreur quadratique du modèle est plus petite, une valeur de 1 signifiant un modèle parfait, sans aucune divergence avec les observations.

Malgré sa popularité parmi les hydrologues, NS a quelques inconvénients. Un de ces inconvénients est qu'il donne de relativement grandes valeurs, même pour des modèles médiocres, puisque utiliser la moyenne arithmétique comme modèle naïf n'est pas adapté pour des bassins à comportement extrême. Un autre inconvénient est que NS n'est pas borné inférieurement (Tableau 1), fait qui crée un biais des moyennes de NS sur plusieurs bassins lors qu'on fait des comparaisons entre modèles.

Pour corriger ce dernier inconvénient, Mathevet *et al.* (2005) ont proposé C_{2M} (Eq. 6) une fonction objectif bornée inférieurement et supérieurement (Tableau 1) qui permet des comparaisons fiables entre modèles.

$$C_{2M} = \frac{1 - \frac{F}{F_0}}{1 + \frac{F}{F_0}} \quad (\text{Eq. 6})$$

Tableau 1. Valeurs de deux fonctions objectif (Nash-Sutcliffe et C_{2M}) pour différentes qualités de calage

qualité du calage	valeur de la fonction objectif	
	Nash-Sutcliffe (NS)	C_{2M}
calage parfait	100	100
égal à celle du calage avec le modèle naïf	0	0
calage très mauvais	$-\infty$	-100

Les fonctions objectif commentées (NS et C_{2M}) permettant une appréciation globale de la performance du modèle sur l'ensemble de la période ; mais, d'autres critères s'intéressent à des classes particulières de débit : les débits en crue ou en étiage, par exemple. Une solution pour donner plus de poids à une classe de débits est d'utiliser des débits transformés ; ceci permet de limiter les effets de l'hétéroscédasticité (la valeur des résidus dépend du valeur des débits). Ainsi, par exemple pour donner plus d'importance aux débits de basses eaux, on peut faire une transformation logarithmique des débits.

1.3 Les caractéristiques d'une bonne méthode de calage

Comme on a vu précédemment, les méthodes de calage des modèles pluie-débit présentent des inconvénients : soit elles sont trop lentes, soit elles sont sensibles à des problèmes numériques qui peuvent conduire à des calages de mauvaise qualité.

Donc, une bonne méthode de calage doit trouver un bon jeu de paramètres dans le moindre temps possible (efficacité). Trouver un bon jeu de paramètres implique savoir identifier l'optimum lors du calage (performance en calage) et que ce jeu de paramètres demeure bon lors de la validation (robustesse). Ainsi, les trois caractéristiques qu'on doit contrôler pour juger bonne une méthode de calage sont :

- i) Efficacité : temps de calcul dont la méthode a besoin pour trouver le maximum
- ii) Performance en calage: capacité de trouver l'optimum
- iii) Robustesse : capacité du jeu de paramètres trouvé en la phase de calage à conserver une bonne performance en la phase de validation

1.4 Amélioration des méthodes locales

Une méthode locale a besoin d'un point de départ du calage, d'un jeu de paramètres initial. Comme la méthode n'explore pas tout l'espace des paramètres mais seulement un voisinage de ce point, le choix du point de départ peut avoir une grande influence.

Pour essayer de gommer cette influence, un certain nombre d'essais ont été menés en vue de l'amélioration des méthodes de calage locales en les combinant à des méthodes stochastiques d'échantillonnage ; ainsi, les méthodes stochastiques trouvent un ensemble de points de départ à partir desquels la méthode locale va initier le calage (Johnston et Pilgrim 1976). On peut citer, par exemple, la méthode Msx (Duan *et al.* 1992).

Une autre stratégie de détermination des points de départ est l'utilisation d'information *a priori*. Par exemple, Perrin (2000), pour la méthode Pas-à-pas, a utilisé les valeurs des paramètres d'un calage préalable pour identifier deux points de départ plausibles ; malgré tout, cette modification obtient une faible amélioration de la performance en calage. De plus, l'auteur est de l'avis que, pour la méthode Pas-à-pas, les tests avec un plus grand nombre de points de départ n'obtiendraient pas d'améliorations non plus.

Une troisième stratégie a été celle de Kuzmin *et al.* (2004) qui utilisent l'information physiographique du bassin pour estimer des bons points de départ pour une méthode locale.

1.5 Comparaison à l'aide d'un grand nombre de bassins versants

Lors qu'on compare des modèles de pluie-débit (ou des méthodes de calage), pour rendre les résultats généralisables, il paraît logique d'utiliser des bassins versants avec des caractéristiques très différentes de climat, couvert végétal, taille, géologie, sols, physiographie, etc.

Une façon d'assurer cette diversité est d'utiliser un grand nombre de bassins versants. Or, l'inconvénient des grands échantillons de données est la difficulté de tester la qualité des données, vu leur grand nombre. Dans ce cas, Perrin (2000) et Mathevet *et al.* (2005) proposent de les prendre toutes en compte car, dans une comparaison on ne s'intéresse pas aux performances absolues mais aux relatives, et tous les modèles (ou les méthodes de calage) seront affectés de la même façon par le manque de qualité des données.

2. OBJECTIFS

Les objectifs de cette première partie sont :

- i) Ajouter une stratégie multidépart à la méthode de calage local Pas-à-pas afin d'en améliorer les caractéristiques
- ii) Evaluer l'amélioration de cette modification en faisant une comparaison à l'aide d'un grand nombre de bassins versants et sur un ensemble de modèles de différents degrés de complexité

3. MATERIAUX ET METHODES

La Figure 2 schématise la démarche méthodologique suivie dans cette première partie

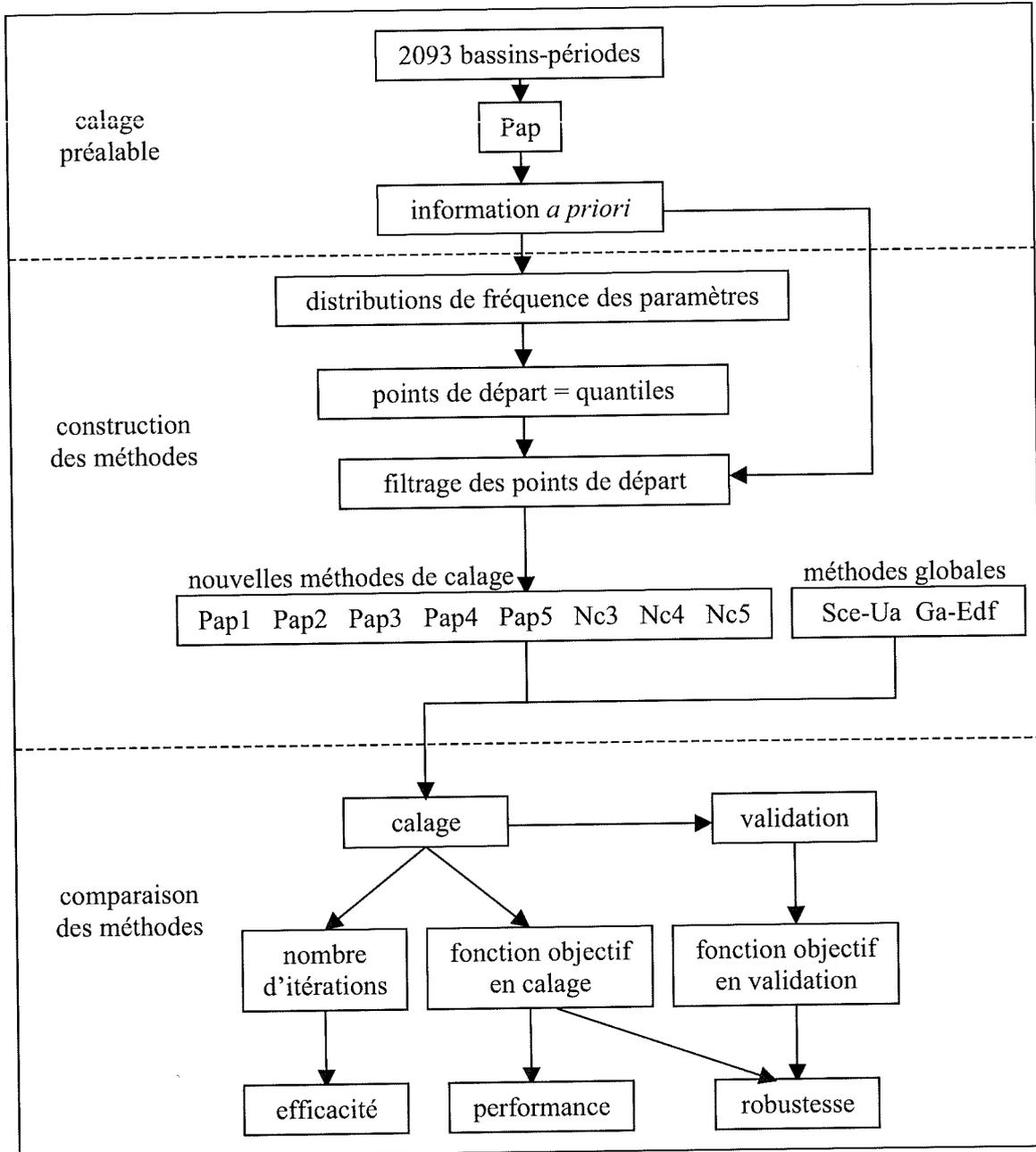


Figure 2. Schéma de la démarche méthodologique ; cette démarche a été appliquée à chaque modèle séparément (Note : les méthodes Nc 3, Nc 4 et Nc 5 n'ont été appliquées que sur Gr4h)

3.1 Données et modèles utilisés

On a utilisé les données de 313 bassins versants, comportant 6 climats différents sur trois continents, avec des conditions hydrologiques très contrastées (Tableau 2 ; pour la liste complète des bassins, voir Mathevet 2005).

Les données dont on a disposé ont été : la pluie (P) horaire, l'évapotranspiration potentielle (Etp) décadaire et le débit (Q) horaire ; l'Etp décadaire a été calculée avec la méthode

proposée par Oudin (2004). En raison du pas de temps horaire utilisé, certaines transformations des données ont été menées : ainsi, les données d'Etp décadaires ont été désagrégées pour obtenir l'Etp horaire. P et Etp sont les entrées (*inputs*) des modèles et le débit observé Q est utilisé dans le calcul de l'erreur quadratique, en le comparant à la sortie (*output* : le débit simulé) des modèles.

Tableau 2. Caractéristiques hydro-climatiques (minimum/moyenne/maximum) des 313 bassins utilisées, par situation géographique

Pays	France	Etats-Unis	Australic	Catalogne	Slovénie
Nombre de bassins	227	70	12	2	2
Climats	tempéré, méditerranéen, continental, océanique	tempéré, océanique, semi-aride	semi-aride	méditerranéen de montagne	montagnard
Surface (km ²)	1.1-280-4978	1.2-33-334	2.7-48-2538	0.56-4.71	457-1385
Durée de la chronique (ans)	3-8-34	3-11-43	6	3-4	5
Pluie annuelle moyenne (mm)	403-963-2067	193-1163-2996	569-674-1025	1183	1384
Etp annuelle moyenne (mm)	595-791-1252	1104-1545-2085	1226	639	735
Ecoulement annuel moyen (mm)	35-44-1655	0-279-1612	9-35-96	429-736	1024-1182
Coefficient d'écoulement annuel moyen (%)	5-44	0-20-81	1-5-13	36-62	74-85

Quant aux modèles utilisés, ce sont été des modèles empiriques à réservoirs et globaux (Tableau 3 et Annexe A) ; ils se différencient par le nombre de paramètres qui rend compte du degré de complexité (Tableau 3). On a utilisé ces modèles au pas de temps horaire. Il faut dire que les modèles d'Ihac, Topmo et Mordor ne sont pas les versions originales, il s'agit de versions de travail.

Tableau 3. Modèles pluie-débit utilisés

Modèle	Gr4j	Gr4h	Ihac	Topmo	Mordor
Nombre de paramètres	4	4	6	8	10

Le grand nombre et la diversité des données et des modèles assurent la généralité des résultats de la comparaison entre les méthodes de calage.

3.2 Fonction objectif

En raison des ses bonnes propriétés, la fonction objectif utilisée a été le critère C_{2M} (voir section 1.2).

3.3 Split-sample test

Le calage a été fait sur des périodes d'un an de longueur. Le *split-sample test* (Klemeš 1986) permet de faire le calage pour chaque année de la série d'un bassin successivement, pendant qu'on réserve le reste pour la validation. Aussi, cette technique, tient compte de l'initialisation des modèles. En effet, les modèles ont besoin d'un an de mise en route pour trouver les conditions initiales de l'année suivante, qui servira au calage. Ainsi, le calage a besoin de deux années : la première pour l'initialisation et la deuxième pour le calage proprement dit. De la même façon, la phase de validation a besoin d'une année de mise en route ; cette année et les deux qui ont servi au calage ne s'utilisent en validation que pour calculer les conditions initiales des années suivantes.

Le Tableau 4 schématise le fonctionnement du *split-sample test* : pour un bassin avec une série de longueur n ans, le *split-sample test* permet de faire $n-1$ couples de calage et validation. Par la suite, on appellera ces couples bassins-périodes ; alors, les données des 313 bassins versants permettent de constituer 2093 bassins-périodes.

Tableau 4. Schéma du *split-sample test* ; pour un série de n années, on a $n-1$ bassins-périodes ou périodes de calage (I : initialisation pour calage ; C : calage ; R : initialisation pour validation ; V : validation)

année \ bassin-période	1	2	3	4	5	...	n	longueur de la période de validation (ans)
1	I	C	V	V	V	...	V	n-2
2	R	I	C	V	V	...	V	n-3
3	R	V	I	C	V	...	V	n-3
4	R	V	V	I	C	...	V	n-3
5	R	V	V	V	I	...	V	n-3
...
n-1	R	V	V	V	V	...	C	n-3

3.4 Méthodes de calage préexistantes

Les méthodes de calage préexistantes qu'on a utilisés ont été :

- i) Pas-à-pas (Pap, par la suite) : méthode locale directe développée au Cemagref (Nascimento, 1995) ; son mode de fonctionnement est de faire varier en petites quantités les paramètres l'un après l'autre et s'arrêter lorsqu'elle n'est plus capable d'améliorer la valeur de la fonction objectif. Un diagramme schématique de la méthode est présenté à l'annexe B.
- ii) Shuffle Complex Evolution (Sce-Ua, par la suite) : méthode globale génétique développée par l'Université d'Arizona
- iii) Genetic algorithm (Ga-Edf, par la suite) : méthode globale génétique développée par Electricité de France

3.5 Construction des nouvelles méthodes de calage²

Notre objectif était d'améliorer la méthode de calage Pap en lui ajoutant une stratégie multidépart. Pour sélectionner les points de départ, on a décidé d'utiliser l'information *a priori*, dont on disposait grâce à un calage préalable.

Ainsi, pour chaque paramètre, on connaissait sa distribution (Figure 3); on a divisé successivement celle-ci en un, deux, trois, quatre et cinq quantiles et on a choisi les points centraux de chaque de ces quantile (Tableau 5).

Ensuite, on a combiné les quantiles de tous les paramètres pour obtenir des points de l'espace des paramètres, qui seront les points de départ proprement dits. De cette façon, on a obtenu cinq ensembles de points de départ, un pour chaque ensemble de quantiles du Tableau 5. Il faut noter que pour chaque modèle, les ensembles de points de départ n'ont pas la même taille (si les modèles ont un nombre différent de paramètres) ni les mêmes valeurs.

Tableau 5. Quantiles utilisés en la construction des points de départ

nombre de parties dont on a divisé la distribution des paramètres	quantiles
une	0.50
deux	0.25 ; 0.75
trois	0.17 ; 0.50 ; 0.83

² On fait noter que cette démarche a été répétée pour chaque modèle séparément.

quatre	0.13 ; 0.38 ; 0.63 ; 0.88
cinq	0.10 ; 0.30 ; 0.50 ; 0.70 ; 0.90

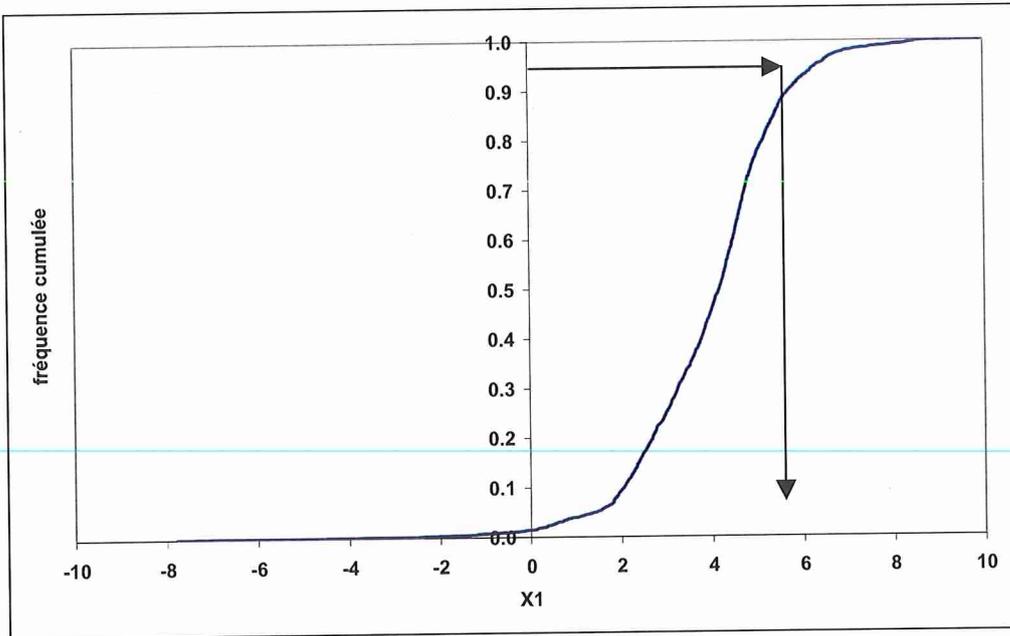


Figure 3. Distribution et quantile $p=0.9$ du paramètre X_1 du modèle Gr4h

Une façon plus simple d'obtenir un ensemble de points de départ est de diviser l'espace des paramètres avec maille régulière. Cependant, en utilisant l'information *a priori* on fait un découpage de l'espace de paramètres plus dense dans la zone où il y a plus de chances de trouver l'optimum (Figure 4).

Comme il s'agit d'une combinaison, le nombre de points de départ (np) dépend du nombre de quantiles (q) et du nombre de paramètres du modèle (p) de la suivante façon : $np = q^p$; ainsi donc, pour des modèles à un grand nombre de paramètres, le nombre de points de départ devient astronomique (presque 10 millions pour Mordor avec cinq quantiles).

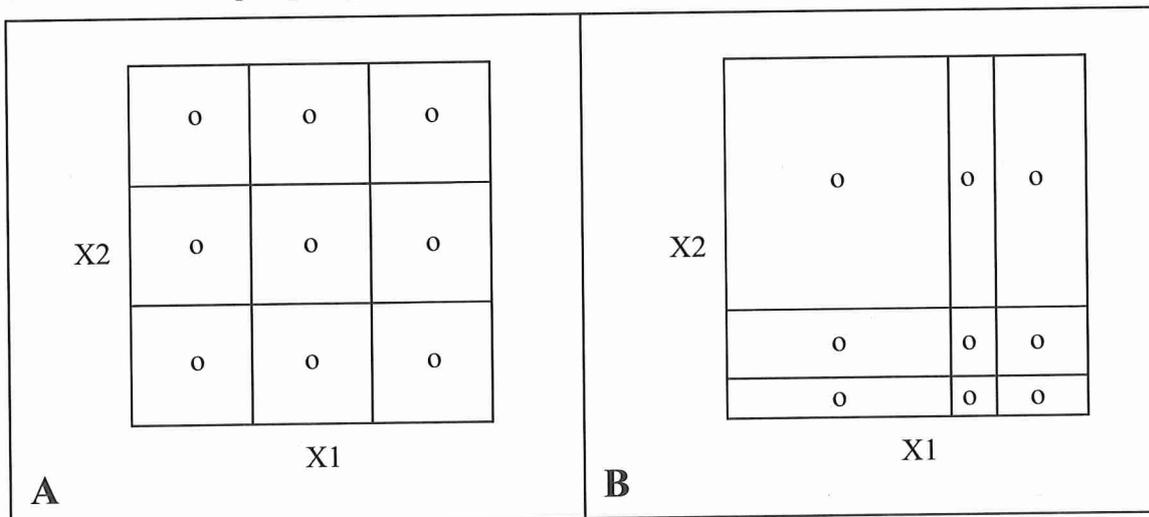


Figure 4. Découpage de l'espace des paramètres d'un modèle à deux paramètres ; A : découpage en maille régulière ; B : découpage avec l'information *a priori* (distribution des paramètres). (Note : les points de départ sont les points centraux de chaque maille)

Pour éviter cet énorme inconvénient, qui nuit à la rapidité de la méthode Pap, on a filtré les points de départ en utilisant, encore une fois, l'information *a priori*. Ainsi, on a divisé l'espace des paramètres en régions hypercubiques, dont le centre est chaque point de départ. Ensuite, on a enlevé toutes les régions où le calage préalable n'avait jamais trouvé un optimum pour les 2093 bassins-périodes. De cette façon, on arrive à réduire à 454 les 625 points de départ pour Gr4h à cinq quantiles et, à 1468 les dix millions points pour Mordor à cinq quantiles.

Enfin, on a combiné chaque ensemble de points de départ avec la méthode Pap, pour obtenir cinq nouvelles méthodes : Pap 1, Pap 2, Pap 3, Pap 4 et Pap 5 (avec l'ensemble de points de départ dérivé de l'ensemble d'un, deux, trois, quatre et cinq quantiles, respectivement). La façon dont ces méthodes fonctionnent est très simple : elles calculent la valeur de la fonction objectif pour chaque point de départ de l'ensemble et en sélectionnent celui avec la valeur la plus grande pour ensuite commencer le calage avec la méthode Pap classique.

En fait, la seule différence de ces nouvelles méthodes avec Pap est qu'elles disposent d'un ensemble de candidats à point de départ au lieu d'un seul. Pap 1 fait l'exception à tous ces commentaires puis qu'elle ne dispose que d'un seul point de départ ; la seule chose que la rend différente de Pap c'est que son point de départ a été calculé à partir d'information *a priori*. Pour cette raison et, comme le but de la comparaison est de voir l'amélioration produite par l'augmentation du nombre de points de départ, on ne différenciera plus par la suite Pap et Pap 1. Les autres variantes (Pap 2, Pap 3, Pap 4 et Pap 5) seront appelées méthodes locales multidépart, dans la suite.

Aussi, seulement pour le modèle Gr4h, et avec l'objectif d'évaluer le fonctionnement de la stratégie multidépart, on a créé les méthodes Nc 3, Nc 4 et Nc 5 (avec l'ensemble de points de départ dérivé de l'ensembles de trois, quatre et cinq quantiles, respectivement). Ces méthodes ont la particularité qu'elles s'arrêtent quand elles ont trouvé le point de départ avec la valeur la plus grande et ne continuent pas avec la méthode Pap ; c'est-à-dire qu'elles ne font pas de calage proprement dit (pour cette raison on les a appelées méthode de non-calage (Nc) ; par la suite, on les appellera aussi méthodes de calage discret, par opposition au calage continu fait par les méthodes Pap).

3.6 Comparaison des méthodes de calage

La comparaison de méthodes de calage a impliqué les méthodes affichées dans le Tableau 6, c'est-à-dire, dix méthodes pour le modèle Gr4h et sept méthodes pour les autres modèles. Comme le but de la comparaison était d'évaluer les améliorations produites sur Pap par l'adition de la stratégie multidépart, la méthode de référence de la comparaison, c'est-à-dire, celle à laquelle on compare toutes les autres, a été Pap.

Tableau 6. Méthodes de calage utilisées selon le modèle

type de méthode		méthode	modèle				
			Gr4h	Gr4j	Ihac	Topmo	Mordor
méthodes globales		Sce-Ua	x	x	x	x	x
		Ga-Edf	x	x	x	x	x
méthodes locales	un seul point de départ	Pap	x	x	x	x	x
	méthodes locales multidépart	Pap 2	x	x	x	x	x
		Pap 3	x	x	x	x	x
		Pap 4	x	x	x	x	x
		Pap 5	x	x	x	x	x
calage discret		Nc 3	x	---	---	---	---
		Nc 4	x	---	---	---	---
		Nc 5	x	---	---	---	---

La comparaison a porté sur les trois qualités qu'une bonne méthode de calage devrait avoir : efficacité, performance et robustesse (section 1.3). Pour quantifier l'efficacité on a utilisé le nombre d'itérations dont la méthode a besoin pour trouver l'optimum; pour la performance, la valeur de la fonction objectif en calage ; et pour la robustesse, la valeur de la fonction objectif en validation, toute seule et en la comparant à celle obtenue en calage. Bien entendu, les grandeurs utilisées sont les valeurs moyennes obtenues sur les 2093 bassins-périodes.

Afin de comparer le fonctionnement de chaque méthode de calage, on s'est intéressé aux jeux de paramètres calés. On a mesuré la distance que separe le jeu de paramètres calé avec chaque méthodes et le jeu de paramètres calé avec Pap. Cette distance est, en fait, une distance relative, normée, par l'écart type du paramètre (Eq. 7). De cette façon on gomme l'influence que certains paramètres avec une grande variabilité pourraient avoir sur le calcul de la distance.

$$d_j^{a \rightarrow b} = \frac{1}{n} \left[\sum_{i=1}^n \left(\frac{X_i^a - X_i^b}{\sigma_i^a} \right)^2 \right]^{1/2} \quad (\text{Eq. 7})$$

où, $d_j^{a \rightarrow b}$ est la distance qui sépare les jeux de paramètres du bassin-période j calés avec les méthodes a et b ; n est le nombre de paramètres ; X_i^a est la valeur calée du paramètre i avec la méthode a ; et σ_i^a est l'écart type calculé sur les 2093 bassins-périodes du paramètre i calé avec la méthode a . La méthode a est la méthode de référence (Pap dans notre cas).

4. RESULTATS ET DISCUSSION

4.1 Efficacité

Le Tableau 7 montre le nombre d'itérations nécessaires pour trouver l'optimum, la grandeur choisie pour évaluer l'efficacité des méthodes.

Comme attendu, les méthodes locales sont entre 10 et 50 fois plus rapides que les méthodes globales pour les modèles simples ; cependant, pour les modèles plus complexes, cette différence est réduite, les méthodes locales étant seulement quatre fois plus rapides que les globales. De leur côté, les méthodes locales multidépart sont entre une et deux fois plus lentes que Pap (cette différence est aussi très atténuée pour les modèles complexes). Logiquement, les méthodes de calage discret (Nc) sont plus rapides que leurs homologues multidépart.

Tableau 7. Nombre d'itérations moyen pour les 2093 bassins-périodes et pour chaque modèle et méthode

méthode	modèle					moyenne
	Gr4h	Gr4j	Ihac	Topmo	Mordor	
Sce-Ua	7000	7000	6800	11500	15000	9460
Ga-Edf	14000	14000	16000	18000	18000	16000
Pap	337	319	948	2424	3845	1575
Pap 2	267	323	875	2148	3878	1498
Pap 3	338	373	1252	2795	4025	1757
Pap 4	483	504	1748	3157	4778	2134
Pap 5	700	718	2076	3274	4578	2269
Nc 3	81	---	---	---	---	---
Nc 4	232	---	---	---	---	---
Nc 5	454	---	---	---	---	---

4.2 Performance

La performance en calage a été mesurée avec la fonction objectif (Tableau 8 et Figures de l'Annexe C) ; si l'on prend les valeurs moyennes de fonction objectif, on peut voir que les méthodes globales sont plus performantes (de deux points en moyenne) que les méthodes locales ; cette différence est plus grande pour les modèles complexes (Topmo et Mordor).

Les méthodes locales multidépart sont plus performantes que la méthode de référence Pap (de 0.73 points) ; cette supériorité est maximale pour le cas du modèle Mordor et minimale pour le modèle Gr4j. Aussi, on peut noter que les méthodes de calage discret sont d'une dizaine de points moins performantes que leurs homologues locales multidépart.

Le fait que les méthodes les plus élaborées (globales et locales multidépart) soient supérieures à Pap pour les modèles complexes, est sans doute lié à la complexité des surfaces de réponse, avec plus de problèmes numériques et d'optima locaux.

Si on s'intéresse au pourcentage de bassins-périodes dont le calage s'est amélioré ou dégradé par rapport à celui de Pap (Tableau 17 de l'Annexe C), le bilan est légèrement négatif (-1.5%) pour les méthodes multidépart Pap 2 et Pap 3, c'est-à-dire, qu'elles dégradent le calage un peu plus souvent qu'elles ne l'améliorent. Par contre, pour Pap 4 et Pap 5 le bilan est positif (4%). Aussi, dans cet aspect, le modèle Gr4h est une exception puisque les méthodes multidépart obtiennent une assez nette amélioration par rapport à Pap. On confirme aussi les mauvaises performances en calage des méthodes de calage discret, puis que c'est presque toujours que leurs calages sont pires que leurs homologues multidépart (Tableau 19 de l'Annexe C).

Tableau 8. Fonction objectif en calage moyenne ((C_{2M})_c) pour les 2093 bassins-périodes et pourcentage de fonctions objectif positives (F(>0))

		Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Sce-Ua	(C _{2M}) _c	59.1	54.8	53.8	62.1	63.4	58.6
	F(>0)	98.8	98.3	97.6	98.9	99.2	98.6
Ga-Edf	(C _{2M}) _c	59.1	54.5	54.0	60.6	63.4	58.3
	F(>0)	98.6	98.1	97.5	97.5	99.1	98.2
Pap	(C _{2M}) _c	57.5	53.1	52.9	55.4	59.8	55.7
	F(>0)	97.2	97.4	96.2	88.7	95.3	95.0
Pap 2	(C _{2M}) _c	57.8	53.0	53.4	53.5	61.1	55.8
	F(>0)	97.7	97.3	96.8	88.2	97.3	95.5
Pap 3	(C _{2M}) _c	58.2	53.3	53.6	56.2	60.0	56.3
	F(>0)	97.8	97.4	97.3	91.9	96.3	96.1
Pap 4	(C _{2M}) _c	58.1	53.6	53.6	57.1	62.9	57.1
	F(>0)	97.7	97.6	97.3	92.7	98.9	96.8
Pap 5	(C _{2M}) _c	58.2	53.3	53.7	57.5	61.2	56.8
	F(>0)	97.9	97.4	97.4	92.6	98.3	96.7
Nc 3	(C _{2M}) _c	45.0	---	---	---	---	---
	F(>0)	91.0	---	---	---	---	---
Nc 4	(C _{2M}) _c	49.4	---	---	---	---	---
	F(>0)	93.8	---	---	---	---	---
Nc 5	(C _{2M}) _c	51.0	---	---	---	---	---
	F(>0)	95.1	---	---	---	---	---

Si on s'intéresse maintenant aux jeux de paramètres calés, on voit (Tableau 9) que dans 60% des cas, les méthodes locales multidépart trouvent le même optimum que Pap. Cette proportion est plus grande pour les modèles simples (80%) et bien plus faible pour les modèles complexes ; cela veut dire que pour des surfaces de réponse simples, la stratégie multidépart n'a pas un grand impact sur la performance en calage.

Tableau 9. Distance moyenne (dist) des jeux de paramètres calés avec les différents méthodes et avec Pap et pourcentage (F(p=)) de jeux de paramètres similaires. (Note : les jeux de paramètres sont similaires si leur distance est inférieure à ± 0.1)

méthode	Gr4h		Gr4j		Ihac		Topmo		Mordor		moyenne	
	dist	F(p=)	dist	F(p=)	dist	F(p=)	dist	F(p=)	dist	F(p=)	dist	F(p=)
Sce-Ua	0.31	59.3	0.25	75.2	0.18	68.9	0.63	4.1	0.23	42.9	0.32	50.1
Ga-Edf	0.30	60.0	0.26	75.7	0.18	69.7	0.64	3.6	0.27	36.2	0.33	49.0
Pap 2	0.12	85.0	0.09	85.4	0.13	74.3	0.27	26.0	0.22	44.8	0.17	63.1
Pap 3	0.12	83.7	0.10	85.1	0.13	75.3	0.28	22.4	0.19	45.7	0.16	62.4
Pap 4	0.12	82.7	0.09	85.1	0.14	73.0	0.29	20.8	0.21	48.6	0.14	62.0
Pap 5	0.12	83.5	0.09	83.8	0.13	73.6	0.27	24.0	0.23	41.4	0.17	61.3
Nc 3	0.29	19.1	---	---	---	---	---	---	---	---	---	---
Nc 4	0.25	30.0	---	---	---	---	---	---	---	---	---	---
Nc 5	0.23	38.9	---	---	---	---	---	---	---	---	---	---

En analysant les bassins-périodes où les méthodes multidépart ont trouvé un optimum différent que la méthode de référence Pap, étonnamment, moins de la moitié de ceux-ci sont meilleurs que ceux trouvés par Pap (Tableau 10). Encore, on peut faire des différences selon le modèle utilisé, puis que les optima différents sont plus souvent pires que ceux de Pap pour les modèles complexes (seulement 32% des cas pour le modèle Mordor) ; cependant, il faut noter que, pour le modèle Gr4h, la stratégie multidépart améliore les performances en calage de Pap dans 58% des cas.

Mais on a vu que la supériorité des performances des méthodes multidépart était plus marquée pour les modèles complexes (Tableau 8). Cela voudrait dire que les stratégies multidépart se trompent plus souvent que Pap, mais que, dans les cas où elles réussissent, l'optimum trouvé

est de très bonne qualité, c'est-à-dire que l'amélioration des performances ne serait pas générale mais restreinte à certains cas.

Tableau 10. Analyse des bassins-périodes avec un optimum différent de celui calé par Pap : pourcentage des bassins-périodes où la performance en calage est meilleure que celle de Pap

méthode	Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Pap 2	58.1	37.4	42.0	28.0	35.4	40.2
Pap 3	56.7	41.2	45.6	37.1	23.7	40.9
Pap 4	56.1	46.8	45.6	41.5	42.7	46.5
Pap 5	60.1	45.9	46.9	45.8	27.4	45.2

L'information présentée au Tableau 11 est aussi intéressante : où l'on peut voir que la distance qui sépare les optima trouvés par les méthodes calage discret et avec ceux de leurs homologues locaux multidépart est assez petite (seulement le double de ce qui sépare les méthodes multidépart et Pap). Cela nous amène à croire que les méthodes de calage discret ont réussi, dans un grand nombre de cas, à s'approcher de la zone où l'optimum se trouve.

Tableau 11. Comparaison des jeux des paramètres calés avec les méthodes de calage discret et avec les méthodes multidépart : Distance moyenne (dist) des jeux des paramètres et pourcentage (F(p=)) de jeux de paramètres similaires. (Note 1 : les jeux de paramètres sont similaires si leur distance est inférieure à ± 0.1). (Note 2 : modèle utilisé : Gr4h)

méthode	dist	F(p=)
Nc 3 comparé à Pap 3	0.25	21.6
Nc 4 comparé à Pap 4	0.22	31.3
Nc 5 comparé à Pap 5	0.2	40.2

4.3 Robustesse

On peut définir la robustesse comme la capacité du jeu de paramètres trouvé en la phase de calage pour conserver une bonne performance en la phase de validation. Ainsi, pour analyser la robustesse, on a évalué la performance en validation de deux façons : par elle-même et en relation à la performance en calage.

Si on analyse la performance en validation toute seule (Tableau 12 et Figures de l'Annexe C), on peut voir que les méthodes locales ont obtenu des meilleurs résultats (0.8 points en moyenne) que les méthodes globales; cette différence est encore plus grande pour Topmo (2.3 points).

Les méthodes locales ne se différencient guère entre elles; en général, sauf pour le modèle Gr4h, où les méthodes Pap 2 et Pap 3 sont un peu meilleures, les méthodes multidépart sont un peu moins performantes que Pap (0.2 points en moyenne) ; pour les modèles complexes, cette différence s'accroît. Une explication de cela pourrait être que, lorsque les surfaces de réponse sont très complexes, essayer de trouver le maximum trop précisément peut conduire à de mauvais résultats en validation, parce que, en raison de la surparamétrisation, des instabilités peuvent apparaître ; cela expliquerait aussi la moindre robustesse des méthodes globales.

Les méthodes de calage discret ont des performances en validation plus basses que celles de leurs homologues locaux multidépart. Cela est logique, puis que leurs performances en calage étaient mauvaises au départ; mais il faut noter que leurs performances étaient pires de 10 points en calage et seulement de 3 points en validation. Une raison de cette apparente robustesse des méthodes de calage discret pourrait être le hasard ; en effet, les calages étant très mauvais, ce n'est qu'une question de hasard que le jeu de paramètres calé ait un meilleur résultat en validation.

Tableau 12. Fonction objectif en validation moyenne (C_{2M}) pour les 2093 bassins-périodes et pourcentage de fonctions objectif positives ($F(>0)$)

		Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Sce-Ua	C_{2M}	33.5	29.5	29.3	26.4	35.4	30.8
	$F(>0)$	87.1	86.0	92.6	80.2	90.2	87.2
Ga-Edf	C_{2M}	33.4	29.7	30.2	25.9	35.2	30.9
	$F(>0)$	87.4	86.9	93.1	77.1	90.4	87.0
Pap	C_{2M}	33.8	30.2	29.6	28.4	36.2	31.6
	$F(>0)$	87.7	87.0	89.7	75.7	91.7	86.4
Pap 2	C_{2M}	34.2	29.7	29.9	26.7	35.8	31.3
	$F(>0)$	88.8	86.8	93.6	75.6	91.4	87.2
Pap 3	C_{2M}	34.0	29.7	29.9	28.7	36.2	31.7
	$F(>0)$	88.5	86.4	93.3	79.5	90.9	87.7
Pap 4	C_{2M}	33.9	29.5	29.9	28.2	35.5	31.4
	$F(>0)$	87.6	86.2	93.3	80.4	90.4	87.6
Pap 5	C_{2M}	34.0	29.7	29.9	28.4	34.8	31.4
	$F(>0)$	87.8	86.4	93.1	79.5	90.6	87.5
Nc 3	C_{2M}	29.9	---	---	---	---	---
	$F(>0)$	89.5	---	---	---	---	---
Nc 4	C_{2M}	31.7	---	---	---	---	---
	$F(>0)$	90.4	---	---	---	---	---
Nc 5	C_{2M}	32.5	---	---	---	---	---
	$F(>0)$	90.2	---	---	---	---	---

Si on s'intéresse au pourcentage de bassins-périodes avec une performance en validation améliorée ou dégradée par rapport à celle de Pap (Tableau 18 de l'Annexe C), le bilan pour les méthodes multidépart est neutre ($\pm 1.1\%$), bien que pour le modèle Gr4h, il soit clairement positif (6%). Pour les méthodes de calage discret, le bilan est très négatif : empiriquement leurs performances se dégradent en validation par rapport à celles de leurs homologues multidépart presque deux fois plus souvent qu'elles ne s'améliorent (Tableau 20 de l'Annexe C).

La comparaison de la performance en validation avec la performance en calage est présentée dans le Tableau 13. On peut voir que, pour les méthodes globales, les performances en validation tombent de plus de 27 points par rapport aux performances en calage ; pour les méthodes locales, la descente est légèrement inférieure : 24.8 points en moyenne. Cela signifie que les méthodes locales sont légèrement plus robustes que les globales ; cette supériorité est plus marquée pour les modèles complexes, surtout pour le modèle Topmo.

Mais surtout, le Tableau 13 montre que les méthodes de calage discret sont celles qui subissent la perte de performance la plus petite, c'est-à-dire, qu'elles sont les plus robustes. Mais, en fait, comme on l'a dit auparavant, cette robustesse est apparente et due au hasard, qui fait que les jeux de paramètres de mauvaise qualité puissent avoir des performances en validation moins mauvaises qu'en calage. Cet effet du hasard, peut se voir aussi dans la plus grande proportion (autour de 17%) de bassins-périodes où la fonction objectif est plus grande en validation qu'en calage (pour le reste des méthodes, cette proportion est autour de 8%).

Tableau 13. Différence moyenne entre les performances en validation et en calage $((C_{2M})_v - (C_{2M})_c)$ pour les 2093 bassins-périodes et pourcentage de bassins-périodes où $(C_{2M})_v > (C_{2M})_c$. (Note : C_{2M} est la fonction objectif)

		Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Sce-Ua	$(C_{2M})_v - (C_{2M})_c$	-25.6	-24.8	-24.5	-35.7	-28.1	-27.7
	$p((C_{2M})_v > (C_{2M})_c)$	7.2	8.2	5.2	6.1	3.8	6.1
Ga-Edf	$(C_{2M})_v - (C_{2M})_c$	-25.7	-25.3	-23.6	-34.7	-28.3	-27.5
	$p((C_{2M})_v > (C_{2M})_c)$	7.3	8.0	5.8	5.5	4.3	6.2
Pap	$(C_{2M})_v - (C_{2M})_c$	-23.7	-22.9	-23.3	-27.0	-23.6	-24.1
	$p((C_{2M})_v > (C_{2M})_c)$	8.5	9.6	6.5	9.9	8.9	8.7
Pap 2	$(C_{2M})_v - (C_{2M})_c$	-23.6	-23.3	-23.5	-26.8	-25.3	-24.5
	$p((C_{2M})_v > (C_{2M})_c)$	8.7	9.7	6.1	10.8	6.6	8.4
Pap 3	$(C_{2M})_v - (C_{2M})_c$	-24.1	-23.6	-23.6	-27.5	-23.8	-24.5
	$p((C_{2M})_v > (C_{2M})_c)$	8.8	8.8	5.5	9.0	7.4	7.9
Pap 4	$(C_{2M})_v - (C_{2M})_c$	-24.3	-23.9	-23.7	-28.9	-27.4	-25.6
	$p((C_{2M})_v > (C_{2M})_c)$	8.6	9.1	5.6	8.5	4.9	7.3
Pap 5	$(C_{2M})_v - (C_{2M})_c$	-24.1	-23.6	-23.8	-29.1	-26.	-25.3
	$p((C_{2M})_v > (C_{2M})_c)$	8.5	9.6	5.3	7.7	5.3	7.3
Nc 3	$(C_{2M})_v - (C_{2M})_c$	-14.9	---	---	---	---	---
	$p((C_{2M})_v > (C_{2M})_c)$	20.5	---	---	---	---	---
Nc 4	$(C_{2M})_v - (C_{2M})_c$	-17.5	---	---	---	---	---
	$p((C_{2M})_v > (C_{2M})_c)$	17.2	---	---	---	---	---
Nc 5	$(C_{2M})_v - (C_{2M})_c$	-18.9	---	---	---	---	---
	$p((C_{2M})_v > (C_{2M})_c)$	14.0	---	---	---	---	---

De la même façon qu'on l'a fait dans l'analyse de la performance au calage, on a comparé les performances en validation des bassins-périodes pour lesquels les méthodes multidépart ont trouvé des jeux de paramètres différents de ceux trouvés par Pap (Tableau 14). On peut voir que le Tableau 14 et le Tableau 10 sont assez semblables ; cela nous amène à confirmer l'hypothèse que, quand les méthodes multidépart trouvent un meilleur optimum que Pap, celui-ci est de bonne qualité. Cependant, dans la majorité des cas où Pap et les méthodes multidépart divergent, Pap trouve des jeux des paramètres plus performants en calage et en validation. Une exception à tout cela est le cas du modèle Gr4h, pour lequel les méthodes multidépart améliorent les performances en validation de Pap dans 54% des cas.

Tableau 14. Analyse des bassins-périodes avec un optimum différent de celui calé par Pap : pourcentage des bassins-périodes où la performance en validation est meilleure que celle de Pap

méthode	Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Pap 2	56.8	32.1	50.7	36.2	38.5	42.9
Pap 3	55.3	29.9	49.3	47.2	35.3	43.4
Pap 4	51.7	32.7	45.6	45.7	40.4	43.2
Pap 5	53.6	37.9	46.4	47.0	30.8	43.1

4.4 Résumé des résultats

Le Tableau 15 fait un résumé des résultats de cette première partie.

En général, les méthodes globales sont plus performantes mais moins efficaces et moins robustes que les méthodes locales.

Les méthodes multidépart sont, en moyenne, un peu plus performantes que Pap en calage, mais un peu moins en validation, c'est-à-dire, qu'elles sont moins robustes. Elles sont aussi moins efficaces en général. De plus, on a démontré que, même si les méthodes multidépart obtiennent des meilleurs résultats en performance moyenne au calage, quand les jeux de paramètres calés divergent de ceux trouvés avec Pap, c'est celui obtenu avec Pap qui est le plus souvent le meilleur.

Une exception à tout cela est le modèle Gr4h, pour lequel la stratégie multidépart paraît générer des améliorations au niveau des performances en calage et en validation. Apparemment, sa surface de réponse simple permet d'utiliser de façon fiable l'information *a priori*. Le fait que cette amélioration ne se reproduise pas pour le modèle Gr4j vient peut-être du fait que ce modèle n'est pas bien adapté au pas de temps horaire.

Les méthodes de calage discret sont légèrement plus efficaces que les méthodes locales, mais nettement moins performantes en calage et en validation.

Tableau 15. Résumé des résultats

méthode	efficacité	performance		robustesse
	nombre d'itérations	fonction objectif en calage	fonction objectif en validation	Différence validation – calage (perte de performance en validation)
Sce-Ua	9460	58.6	30.8	-27.7
Ga-Edf	16000	58.3	30.9	-27.5
Pap	1575	55.7	31.6	-24.1
Pap 2	1498	55.8	31.3	-24.5
Pap 3	1757	56.3	31.7	-24.5
Pap 4	2134	57.1	31.4	-25.6
Pap 5	2269	56.8	31.4	-25.3

5. CONCLUSIONS

Une stratégie multidépart a été appliquée à la méthode de calage locale Pas-à-pas avec le but d'en améliorer les caractéristiques pour le calage des modèles au pas de temps horaire. La comparaison a indiqué que, globalement, l'addition de cette stratégie n'a pas apporté l'amélioration espérée, notamment en termes de performance en validation et de robustesse.

Ces résultats sont en accord avec ceux de Perrin (2000), qui, pour le cas du pas de temps journalier, indique que l'ajout d'un deuxième point de départ à la méthode Pas-à-pas n'apporte que des améliorations limitées en calage et occasionne une perte de robustesse en validation.

Cependant, la stratégie multidépart paraît générer des améliorations pour le modèle Gr4h ; ainsi, dans le calage de ce modèle, on se permet de conseiller l'utilisation de la méthode Pap 2 (avec deux quantiles pour trouver les points de départ) en raison de ses meilleures efficacité, performance en calage et en validation que Pap. Une explication de ce dernier résultat pourrait être que, pour des surfaces de réponse pas trop complexes, la zonation de l'espace de paramètres à l'aide de l'information *a priori* permet de se rapprocher de la zone où se trouve l'optimum. Ainsi, paradoxalement, la complexification de la méthode de calage, qui voulait améliorer le calage des modèles complexes, ne profite que les modèles simples.

PARTIE D'APPLICATION : APPLICATION D'UNE TECHNIQUE DE DETECTION DES CHANGEMENTS DU COMPORTEMENT HYDROLOGIQUE DES BASSINS FORESTIERS ALTERES

1. INTRODUCTION

1.1 Le comportement hydrologique

Dans le contexte actuel, dans les pays occidentaux, de nombreux bassins versants subissent des changements importants d'occupation de sol (urbanisation, afforestation, abandon de cultures). Ces changements peuvent produire, à leur tour, un changement sur l'hydrologie des bassins, plus précisément sur leur comportement hydrologique.

Le comportement hydrologique d'un bassin versant est la façon dont le bassin transforme les entrées climatiques (précipitations et évapotranspiration) en écoulement (Andréassian 2000).

Le comportement hydrologique, aussi appelé réponse hydrologique, dépend des caractéristiques du bassin : géologie (lithologie, structure géologique, géomorphologie), physiographie (pente, superficie, réseau hydrographique), sol (type, profondeur, structure, texture) et couverture du sol (végétation, occupation du sol). Comme on peut le supposer, la relation entre le comportement hydrologique et les caractéristiques du bassin est très complexe.

1.2 Influence de la forêt sur le comportement hydrologique

Afin d'illustrer la relation entre les caractéristiques du bassin et son comportement hydrologique, on présente quelques conclusions générales de l'influence de la forêt sur l'hydrologie:

- i) impact sur le bilan annuel : la déforestation augmente l'écoulement annuel et l'afforestation le diminue
- ii) impact sur les crues : la déforestation augmente le volume et le débit des crues, mais cet effet est très variable, et n'est pas toujours vérifié ; l'afforestation a, par contre, un effet très faible
- iii) impact sur les étiages : la déforestation augmente le débit en étiage et l'afforestation le diminue

1.3 Techniques de détection de changements du comportement hydrologique

Comme on l'a dit au début, aujourd'hui, les changements d'occupation des sols sont très courants, et ont des conséquences sur la double nature de l'eau (ressource et risque), convertissant ainsi, l'étude du comportement hydrologique en un sujet capital pour la planification territoriale et la gestion de l'eau.

Or, il est rarement possible³ d'évaluer le degré d'évolution du comportement hydrologique à cause de la variabilité temporelle du forçage climatique. En effet, si, par exemple, après une coupe à blanc, il y a avait une période très humide, on ne pourrait pas savoir quelle partie de l'augmentation des débits est due au traitement et quelle partie est due à l'augmentation des précipitations.

³ Une exception est donnée par Costa *et al.* (2003) qui, à l'aide d'analyses statistiques, ont lié l'augmentation du débit du fleuve Tocantins à Porto Nacional (Brésil), à l'augmentation des pâtures : ils ont utilisé une période climatiquement stable.

Pour cette raison, il faut se munir de techniques qui permettent de tenir compte de l'influence de la météorologie. La technique la plus traditionnellement utilisée permettant de faire cela est la technique des bassins appariés, dans laquelle on choisit deux bassins très semblables et géographiquement proches, qui sont suivis ensemble pendant une période afin d'établir une relation univoque entre leurs comportements hydrologiques ; cette période doit être assez longue et assez variée pour obtenir une relation fiable. Ensuite, on applique un traitement sur le couvert du sol de l'un des bassins pendant que l'autre reste inchangé (le bassin contrôle). Après le traitement, la relation trouvée préalablement est utilisée pour reconstituer le comportement du bassin traité. En comparant les débits observés avec les reconstitués, on peut évaluer l'impact du traitement sur le comportement hydrologique. Andréassian (2004) a élaboré une liste exhaustive des études des bassins appariés publiées.

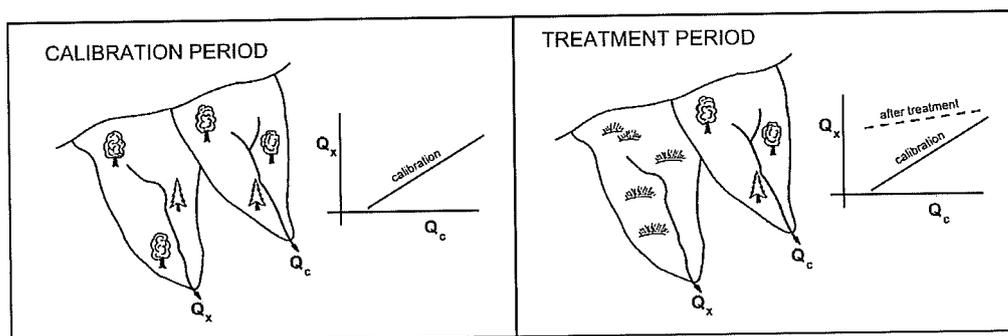


Figure 5. Schéma de la technique des bassins appariés (source : Hewlett 1982)

Or, cette technique, très coûteuse en moyens et en temps, ne peut être utilisée que dans les conditions limitées des expériences scientifiques, ne pouvant, pour les cas réels, qu'extrapoler les résultats de ces expériences. De plus, souvent il est difficile, voire impossible, de trouver un bassin de contrôle, qui doit remplir deux conditions : présenter des caractéristiques semblables à celles du bassin d'étude et être stationnaire du point de vue du comportement hydrologique. Ceci est d'autant plus vrai que le bassin est grand.

La modélisation peut être utile pour dépasser ces limitations, puisqu'on peut arriver à simuler un bassin de contrôle à l'aide d'un modèle pluie-débit. En effet, Cornish (1993) trouve que la modélisation donne des résultats similaires à ceux des bassins appariés ; cependant, il trouve aussi que les incertitudes sur la magnitude des impacts sont plus grandes.

Il y a plusieurs exemples dans la bibliographie d'utilisation des modèles dans la détection des changements du comportement hydrologique : Kuczera (1987), Brandt *et al.* (1988), Gallart et Llorens (2004). Dans la plupart des cas, la méthode était la suivante : caler le modèle sur une période préalable au traitement pour ensuite simuler les débits d'une période postérieure et les comparer aux débits observés. Ainsi, puisqu'il a été calé sur des données antérieures au traitement, le modèle calé fonctionne comme un bassin de contrôle virtuel, reconstituant les débits qu'il y aurait si le bassin n'avait pas été altéré.

Schreider *et al.* (2002) proposent d'utiliser le modèle (calé, comme dans la méthode habituelle, sur une période préalable au traitement) pour simuler les débits d'une façon continue et calculer ce qu'ils appellent la « Réponse hydrologique potentielle » (Rhp, par la suite) permettant ainsi détecter des changements graduels.

Andréassian *et al.* (2003) ont amélioré la méthode de Schreider *et al.* (2002) en développant un test non-paramétrique pour évaluer la significativité statistique des changements détectés. Pour faire cela, ils ont découpé la série en sous-périodes, sur lesquelles ils ont calé le modèle. C'est-à-dire qu'au lieu de caler le modèle au début de la série et de simuler les débits pour

chaque période postérieure, cette méthode cale le modèle pour chaque sous-période et fait les calculs avec l'ensemble des données, les mêmes pour toutes les sous-périodes. Cette approche caractérise mieux le bruit lié à la méthode ; elle a été utilisée par Andréassian (2002) pour essayer de comparer de façon graphique les changements du comportement hydrologique aux changements de couvert végétal dans des petits bassins forestiers.

1.4 La technique de la Réponse hydrologique potentielle (Rhp)

La technique de la Réponse hydrologique potentielle (Rhp) se décompose en trois étapes (Figure 6):

- i) Diviser la série de données d'un bassin en sous-périodes assez courtes rendre visible l'évolution du comportement hydrologique (normalement entre trois et cinq ans). Ces sous-périodes ne peuvent pas être trop courtes non plus, au risque d'obtenir de calages d'une très mauvaise qualité ; c'est pour cette raison qu'une méthode de calage très robuste est indispensable
- ii) Caler le modèle pluie-débit sur chaque sous-période ; on obtient ainsi un jeu de paramètres pour chaque sous-période.
- iii) Calculer pour chaque sous-période, et à l'aide du jeu de paramètres calé, une ou plusieurs variables hydrologiques avec l'ensemble des données de la série complète. En utilisant les mêmes entrées pour toutes les sous-périodes, on fait disparaître l'influence de la variabilité météorologique.

Les variables hydrologiques (ou Rhp) que l'on peut calculer sont variées : débit annuel moyen, volume d'eau des crues ou des étiages, index de débit de base (BFI), chacune informant d'un aspect du comportement hydrologique du bassin.

La Rhp d'une sous-période donne la valeur que la variable hydrologique choisie aurait eue si le bassin avait toujours eu les caractéristiques physiques de la sous-période sur laquelle le modèle a été calé. Ainsi, avec un graphique des différentes Rhp selon la sous-période, il est théoriquement possible de détecter visuellement les changements du comportement hydrologique.

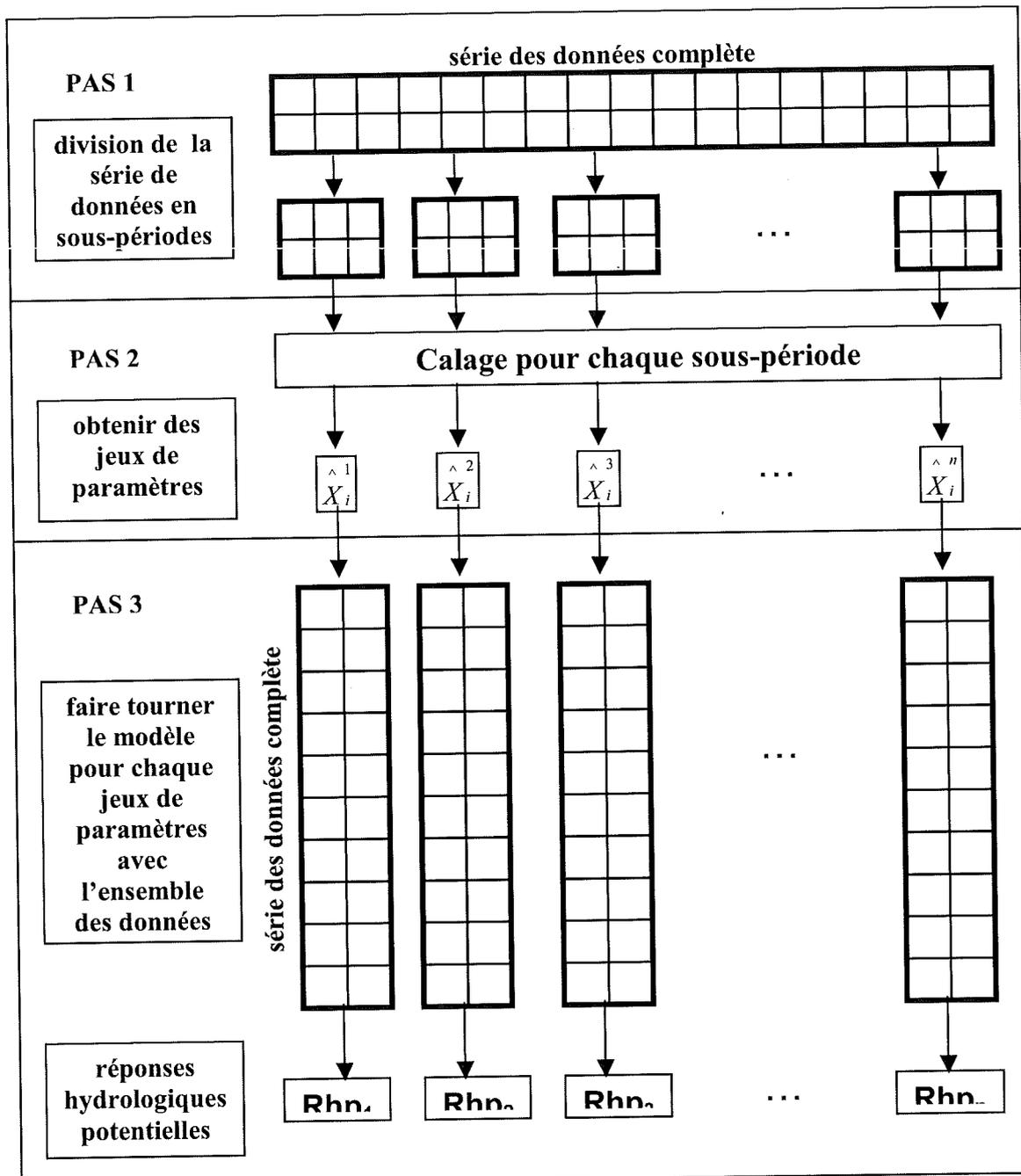


Figure 6. Schéma illustrant la technique Rhp (Note : n est le nombre de sous-périodes)

2. OBJECTIFS

L'objectif de cette deuxième partie est d'évaluer la performance d'une technique graphique de détection des changements du comportement hydrologique en l'appliquant sur un certain nombre de petits bassins forestiers.

3. MATERIAUX ET METHODES

3.1 Bassins versants

On a utilisé les données de 24 petits bassins forestiers appartenant à six ensembles de bassins expérimentaux situés aux Etats-Unis, en France et en Afrique du Sud. Ces bassins représentent différents types de climat, végétation et de traitement du couvert végétal (Tableau 16).

Les données dont on a disposé ont été : la pluie (P) journalière, l'évapotranspiration potentielle (Etp) décadaire et débit (Q) journalier. L'Etp décadaire a été calculée avec la méthode proposée par Oudin (2004). En raison du pas de temps journalier utilisé, certaines transformations des données ont été menées : ainsi, les données d'Etp décadaires ont été désagrégées pour obtenir l'Etp journalière. P et Etp sont les entrées (*inputs*) des modèles et Q est utilisé dans le calcul de l'erreur quadratique, en le comparant à la sortie (*output*) des modèles.

3.2 Modélisation pluie-débit

La technique Rhp peut s'appliquer avec tout type de modèle pluie-débit ; dans cette étude on a utilisé le modèle Gr4j, modèle empirique global à pas de temps journalier (Annexe A). Le modèle Gr4j a comme principal avantage la simplicité.

La méthode de calage utilisée a été celle qu'on a jugée la plus performante et robuste pour ce modèle Gr4j dans la première partie, c'est-à-dire, Pap.

La fonction objectif utilisée a été C_{2M} (section 1.2 de la première partie)

3.3 La technique de la Réponse hydrologique potentielle (Rhp)

Pour l'application de la technique de la Réponse hydrologique potentielle (Rhp), on a découpé les séries de données en sous-périodes de quatre ans, dont un an de mise en route et trois ans de calage. La dernière année d'une sous-période correspond à l'année de mise en route de la sous-période suivante. Exceptionnellement, quelques sous-périodes ont eu cinq ans de longueur afin d'utiliser toute la longueur de la série disponible et d'éviter avoir un traitement au milieu d'une sous-période

On a utilisé la pluie de la période la plus longue sans interruptions pour le calcul des Rhp, qui ont été : le débit annuel moyen (Q_a , en mm/an) et les quantiles 10 et 90 de la distribution des débits (Q_{10} et Q_{90} , en mm/j). Q_a donne une information sur le bilan hydrique annuel, pendant que Q_{10} renseigne sur les hautes eaux (débits en crue) et Q_{90} , sur les basses eaux (débits en étiage).

Tableau 16. Bassins versants utilisés et leurs caractéristiques. (Note : P_a est la précipitation annuelle moyenne, Q_a est l'écoulement annuel moyen avant le traitement et ΔQ^{\max} est la variation maximale d'écoulement annuel due au traitement). Sources : Scott et Lesch (1997), Andréassian (2004) et Waichler *et al.* (2004)

bassin	pays	climat	surface (ha)	P_a (mm/an)	Q_a (mm/an)	type de forêt	traitement (% du bassin)	ΔQ^{\max} (%)
Coshocton 172	Usa	tempéré	18	970	300	70% agricole	partie agricole planté avec <i>Pinus</i>	-45
Coshocton 196			123	922	407	pâtures et forêt	contrôle	---
Vaubarnier (R.Collobrier)	France	méditerranéen humide	150	1017	404	maquis	contrôle	---
Valescure (R.Collobrier)			925	1154	416		contrôle	---
Rimbaud (R.Collobrier)			153	1169	683		incendie (100%)	28
HJ Andrews 2	Usa	océanique	60	2275	1332	conifères	contrôle	---
HJ Andrews 1			96	2388	1376		coupe à blanc (100%)	34
HJ Andrews 3			101	2388	1346		coupe à blanc (25%)	22
HJ Andrews 9			9	2275	1276		contrôle	---
HJ Andrews 10			9	2330	1650		coupe à blanc (100%)	24
Three Bar B	Usa	semi-aride	19	582	11	maquis	incendie (100%) ; conversion en prairie (100%)	273
Three Bar C			39	638	58		incendie (100%) ; conversion en prairie (100%)	228
Three Bar D			33	736	85		incendie (100%) ; contrôle	---
Three Bar F			28	681	36		incendie (100%) ; conversion en prairie (100%)	225
Fernow 1	Usa	montagnard	30	1524	584	feuillus	coupes sélectives (100%)	22
Fernow 2			15	1500	660		coupes sélectives (100%)	10
Fernow 3			34	1500	607		coupes sélectives (100%)	1
Fernow 4			39	14446	636		contrôle	---
Fernow 5			36	1473	732		coupes sélectives (100%)	5
Fernow 6			22	1440	493		coupe à blanc (100%) ; plantation <i>Picea</i> (100%)	33
Fernow 7			24	1469	788		coupe à blanc (100%)	20
Mokobulaan A	Afrique du Sud	subtropical	26	1135	244	pâturage	plantation <i>Eucalyptus</i> (100%) ; coupe à blanc (100%)	-100
Mokobulaan B			35	1135	217		plantation <i>Pinus</i> (100%)	-100
Mokobulaan C			37	1186	143		contrôle	---

4. RESULTATS ET DISCUSSION

Les graphiques faits avec les Rhp pour chaque ensemble de bassins, avec les graphiques de la fonction objectif par sous-période et les graphiques de précipitation annuelle pour le bassins de contrôle sont dans l'Annexe D. Dans la suite, on analyse ces graphiques :

Coshocton (Figures 7 à 11): Le bassin de contrôle est assez stable pour Q_a et Q_{10} , mais ne l'est pas pour Q_{90} puisqu'on peut indentifier deux zones : avant et après la sous-période 1973-1976. La réponse du bassin 172 aux traitements est claire dans les trois Rhp : la plantation de *Pinus sp.* provoque une baisse de l'écoulement annuel et des débits en crue et en étiage pendant que les éclaircies génèrent un retour au comportement hydrologique avant traitement ; ces résultats sont en accord avec ceux de McGuinness et Harrold (1971).

Réal Collobrier (Figures 12 à 16) : Les bassins de contrôle n'ont un comportement stationnaire pour aucune des Rhp : le bassin du Valescure a une légère diminution graduelle des écoulements, peut-être en rapport avec la succession végétal dans le maquis (Andréassian *et al.* 2003); le bassin du Vaubarnier, bassin de structure géologique compliquée, paraît réagir violemment à l'incendie du bassin du Rimbaud, peut-être à cause des échanges souterrains entre les deux bassins. Finalement, dans Rimbaud on peut voir une réaction claire à l'incendie : une augmentation des écoulements immédiatement après celui-ci et une diminution des écoulements (en dessous, même, des écoulements avant le traitement) due à la colonisation du bassin par de végétation opportuniste avec un moindre control stomatique.

En raison du aux grand nombre de bassins, pour rendre l'analyse plus facile, on a divisé les bassins d'Andrews en deux parties : bassins 1, 3 et bassin de contrôle, et bassin 10 et bassin de contrôle.

Andrews , partie 1 (Figures 17 à 21) : Le bassin de contrôle n'est guère stable, surtout pour Q_a et Q_{10} , mais cela ne paraît pas dû à des conditions météorologiques extrêmes (Figure 17). La réaction du bassin 1 à la coupe à blanc est très bien détectée (on note que Q_{90} revient bientôt aux valeurs avant traitement, c'est-à-dire que la méthode détecte l'effet sur les débits en étiage de la végétation colonisatrice. La réponse du bassin 3 à un traitement moins agressif est aussi détectée, quoique de façon moins claire.

Andrews, partie 2 (Figures 22 à 26) : Le bassin de contrôle est assez stable (sauf pour les première et dernière sous-périodes) pour Q_a et Q_{10} ; la non-stationnarité du bassin de contrôle pour Q_{90} ne peut pas être expliquée par des années de climatologie extrême (Figure 22). Comme pour le bassin Andrews 1, la méthode détecte la réponse du bassin 10 à la coupe à blanc : le changement de courte durée de Q_{90} et les changements plus étalés de Q_a et Q_{10} .

Three Bar (Figures 27 à 31) : Comme souvent dans les cas antérieurs, le bassin de contrôle est assez stable pour Q_a et Q_{10} et un peu moins pour Q_{90} . Dans ces bassins les traitements (traitements chimiques d'élimination du maquis) sont étalés dans le temps. Le petit nombre de sous-périodes empêche de donner une grande significativité aux tendances, mais apparemment la technique a réussi à détecter les changements de comportement hydrologique provoqués par les traitements ; La magnitude des ces changements paraît être bien corrélée à la magnitude des traitements ; ainsi, la réaction des bassins C et F (traités complètement) est plus claire que la réaction de B (traité partiellement). Ces résultats sont en accord avec ceux trouvés par Hibbert (1971).

Comme avant avec Andrews, les bassins de Fernow ont été divisés en deux groupes : bassins 1, 2, 3 et bassin de contrôle, et bassins 5, 6, 7 et bassin de contrôle.

Fernow, partie 1 (Figures 32 à 36) : Le bassin de contrôle n'est stable pour aucune des Rhp. La technique ne parvient pas à détecter les changements produits dans le comportement

hydrologique par l'enlèvement de 75% de la surface terrière, mais, par contre, elle détecte les réponses des bassins 2 et 3 à des traitements parfois moins sévères.

Fernow, partie 2 (Figures 37 à 41) : Le bassin de contrôle (le même que pour la partie 1 de Fernow) n'est stable pour aucune des Rhp. Fernow 5 n'a pas assez de données pour permettre d'extraire des conclusions de l'évolution de son comportement. La technique détecte les changements de comportement dans Fernow 6 et Fernow 7 dus aux coupes sévères. Par contre, les effets de la plantation de *Picea sp.* dans Fernow 6 ne se laissent voir que dix ans après le traitement, ce qui pourrait être dû à un mauvais fonctionnement de la technique ou aux particularités de la plantation (densité, succès), information dont on ne dispose pas.

Mokobulaan (Figures 42 à 46) : Le bassin de contrôle n'est pas du tout stable, peut-être à cause, comme soutiennent Scott et Lesch (1997), des pertes en profondeur, fait qui empêcherait un bon calage du modèle pluie-débit et, donc, une application défectueuse de la technique. Néanmoins, les effets de la plantation des bassins A et B sont facilement détectables, du fait de l'importance des changements (les deux bassins se sont asséchés).

En résumé, la technique parvient à détecter, dans la plupart des cas, les changements graduels (plantation, croissance, succession végétale) aussi bien que les changements brusques (coupe, incendie) du couvert forestier. De plus, la magnitude des ces changements paraît être bien corrélée à la magnitude des traitements. Souvent, aussi, détecte des comportements non-stationnaires des bassins de contrôle.

5. CONCLUSION

Le calcul de la Réponse hydrologique potentielle (Rhp) par sous-périodes introduit par Andreassian *et al.* (2003) et sa représentation graphique permettent de développer une technique graphique de détection des changements du comportement hydrologique des bassins altérés.

On a constaté que cette technique est assez efficace dans la détection des changements brusques et graduels.

Cependant, elle manque d'un outil lui permettant d'évaluer les erreurs d'application; par conséquent, le bassin de contrôle sera toujours nécessaire. Mais, cette technique permet aussi de détecter les comportements non-stationnaires des bassins de contrôle, dus à des facteurs difficilement quantifiables, comme l'évolution de la formation végétale ou les transferts entre bassins. Pour ces deux raisons, une utilisation conjointe des techniques des bassins appariés et de la Rhp est recommandée, afin de contrôler les erreurs inhérentes aux deux techniques.

Enfin, la longueur des sous-périodes de calage de trois ans paraît assez judicieuse au vu des résultats. Il faut cependant adapter cette longueur (quand cela est possible) pour placer les traitements entre deux sous-périodes afin qu'elles soient les plus stables possible.

6. BIBLIOGRAPHIE

- Andréassian V (2002). *Impact de l'évolution du couvert forestier sur le comportement hydrologique des bassins versants*. Thèse de doctorat, Université Pierre et Marie Curie et Cemagref, 781 pp
- Andréassian V (2004) *Waters and forest : from historical controversy to scientific debate*. Journal of Hydrology, 291, 1-27
- Andréassian V, Parent E et Michel C (2003) *A distribution free-test t detect gradual changes in watershed behaviour*. Water Resources Research, 39(9), 1252
- Asadian G (2005) *Peut-on bénéficier de la connaissance de la pression atmosphérique dans la modélisation pluie-débit?* Mémoire de Master, Université Pierre et Marie Curie et Cemagref
- Berthier CH (2005) *Quantification des incertitudes des débits calculés par un modèle pluie-débit empirique*. Mémoire de Master, Université Paris-Sud 11 et Cemagref
- Beven K (2001) *How far can we go in hydrological modelling ?* Hydrology and Earth System Sciences, 5(1), 1-12
- Brandt M, Bergström S et Gardelin M (1988) *Modelling the effects of clear-cutting in runoff – Examples from central Sweden*. Ambio, 17, 303-313
- Costa MH, Botta A, Cardille JA (2003) *Effects of large-scale changes in land cover on the discharge of the Tocantins River, Southeastern Amazonia*. Journal of Hydrology, 283, 206-217
- Duan Q, Sorooshian S et Gupta VK (1992) *Effective and efficient global optimization for conceptual rainfall-runoff models*. Water Resources Research, 28(4), 1015-1031
- Edijatno et Michel C (1989) *Un modèle pluie-débit journalier à trois paramètres*. La Houille Blanche, 2, 113-121
- Gallart F et Llorens P (2004) *Observations on land cover changes and water resources in the headwaters of the Ebro catchment, Iberian Peninsula*. Physics and Chemistry of the Earth, 29, 769-773
- García O (1994) *The state-space approach in growth modelling*. Canadian Journal of Forest Research 24: 1894-1903
- Hewlett JD (1982) *Principles of forest hydrology*. The University of Georgia Press, 183 pp.
- Hibbert AR (1971) *Increases in streamflow after converting chaparral to grass*. Water Resources Research, 7(1), 71-81
- Johnston PR et Pilgrim DH (1976) *Parameter optimization for watershed models*. Water Resources Research, 12(3), 477-486
- Klemeš V (1986) *Operational testing of hydrological simulation models*. Hydrology Science Journal, 31(1), 13-24
- Kuczera G (1987) *Prediction of water yield reductions following a bushfire in ash-mixed species eucalyptus forest*. Journal of Hydrology, 94, 215-236
- Kuzmin V, Seo DJ et Koren V (2004) *Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search*. Soumis à Journal of Hydrology

Le Moine N (2005) *Déterminants du bilan en eau à l'échelle du bassin versant : Vers une estimation réaliste des pertes des bassins de surface*. Mémoire de Master, Université Pierre et Marie Curie et Cemagref

Mathevet, T (2005) *Modélisation Pluie-Débit au pas de temps horaire : bilan des performances des modèles existants et développement de modèles au pas de temps horaire*. Thèse de doctorat, Ecole Doctorale Géosciences et Ressources Naturelles (Engref) et Cemagref, en cours

Mathevet T, Perrin C, Andréassian V et Michel C (2005) *Large samples of watersheds : a necessary condition to compare the efficiencies of rainfall-runoff models ?* Soumis à Water Resources Research

McGuinness JL et Harrold L (1971) *Reforestation influences on small watershed streamflow*. Water Resource Research, 7(4), 845-852

Michel C (1983) *Que peut-on faire en hydrologie avec un modèle conceptuel à un seul paramètre ?* La Houille Blanche, 1, 39-44

Nascimento N (1995) *Appréciation à l'aide d'un modèle empirique des effets d'actions anthropiques sur la relation pluie-débit à l'échelle d'un bassin versant*. Thèse de doctorat, cole Nationale des Ponts et Chaussées de Paris, 447 pp.

Nash JE et Sutcliffe JV (1970) *River flow forecasting through conceptual models. Part I: a discussion of principles*. Journal of Hydrology, 10, 282-290

Oudin L (2004) *Recherche d'un modèle d'évapotranspiration potentielle comme entrée d'un modèle pluie débit global*. Thèse de doctorat, Ecole Doctorale Géosciences et Ressources Naturelles (Engref) et Cemagref, 495 pp.

Perrin C (2000) *Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative*. Thèse de doctorat, Institut National Polytechnique de Grenoble et Cemagref, 530 pp.

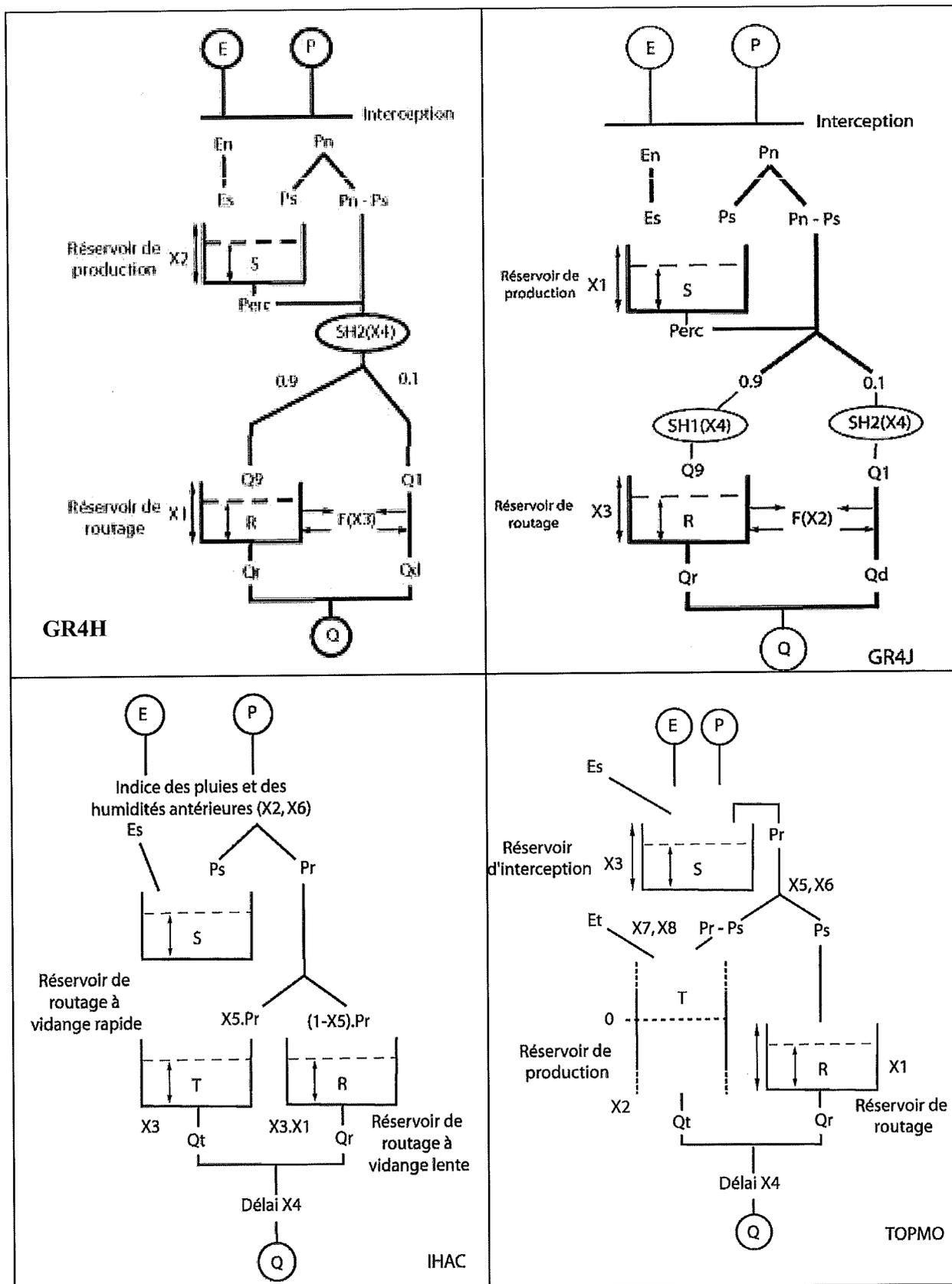
Perrin C, Michel C et Andréassian V (2003) *Improvement of a parsimonious model for streamflow simulation* Journal of Hydrology, 279, 275-289

Schreider SY, Jakeman AJ, Lechter RA, Nathan RJ, Neal BP et Davis SG (2002) *Detecting changes in streamflow response to changes in non-climatic catchment conditions: Farm dam development in the Murray-Darling basin, Australia*. Journal of Hydrology, 262, 84-98

Scott DF et Lesch W (1997) *Streamflow response to afforestation with Eucalyptus grandis and Pinus patula and to felling in the Mokobulaan experimental catchments, South Africa*. Journal of Hydrology, 199, 360-377

Waichler SR, Wemple BC et Wigmosta MS (2004) *Simulation of water balance and forest treatment effects at the HJ Andrews Experimental Forest*. Soumis à Hydrological Processes

ANNEXE A : MODELES PLUIE-DEBIT UTILISES



Schémas des modèles pluie-débit Gr4h, Gr4j, Ihac et Topmo

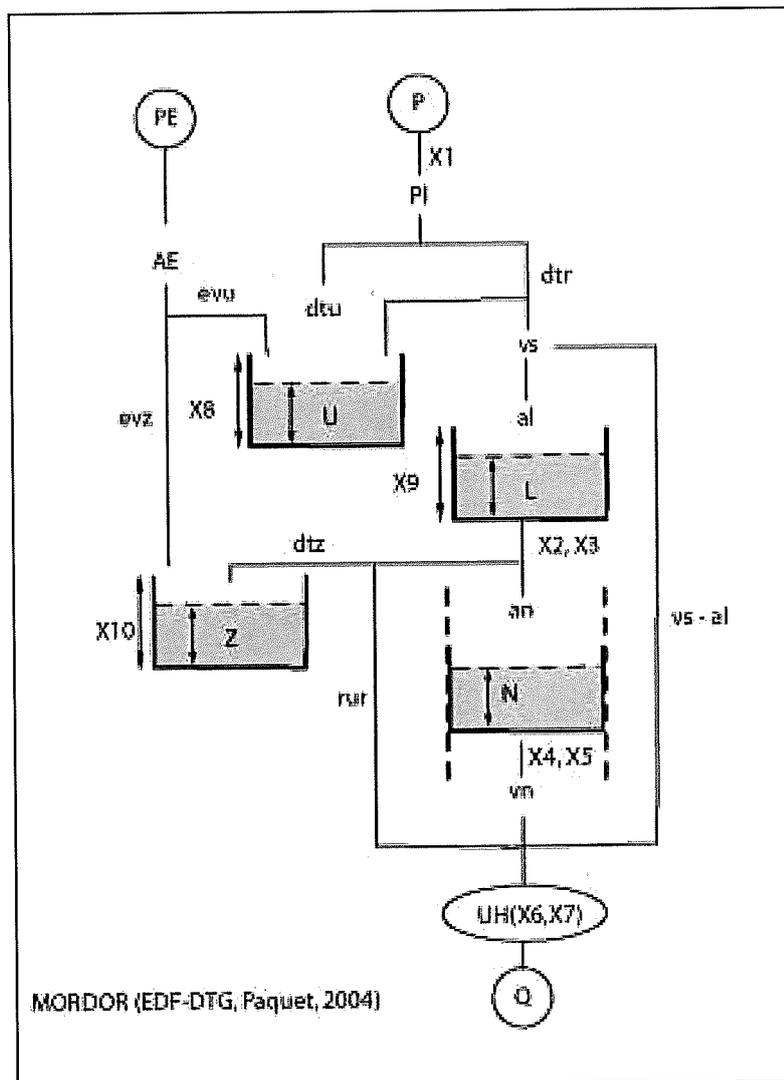


Schéma du modèle pluie-débit Mordor

ANNEXE B: LA METHODE PAS-A-PAS

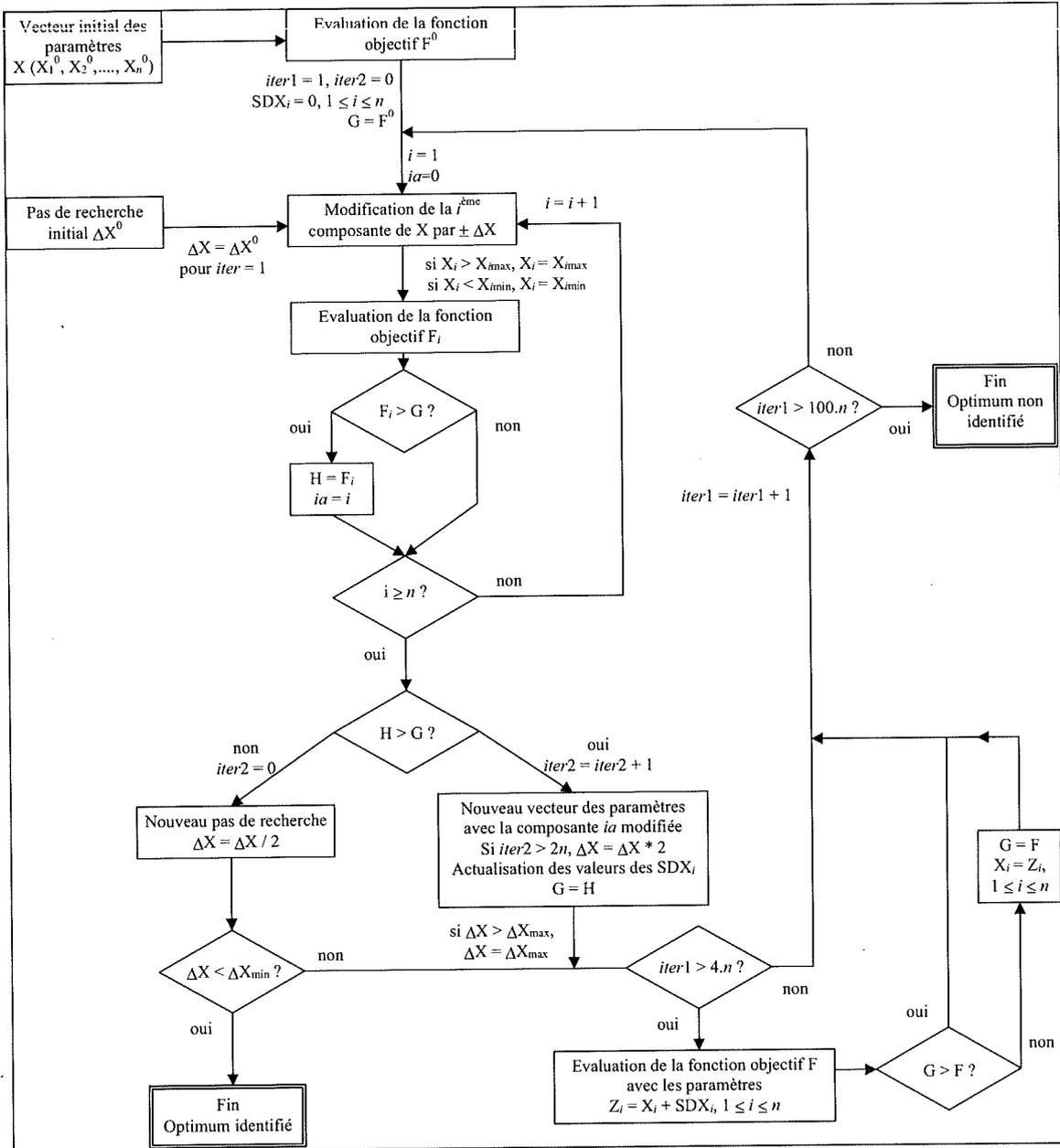
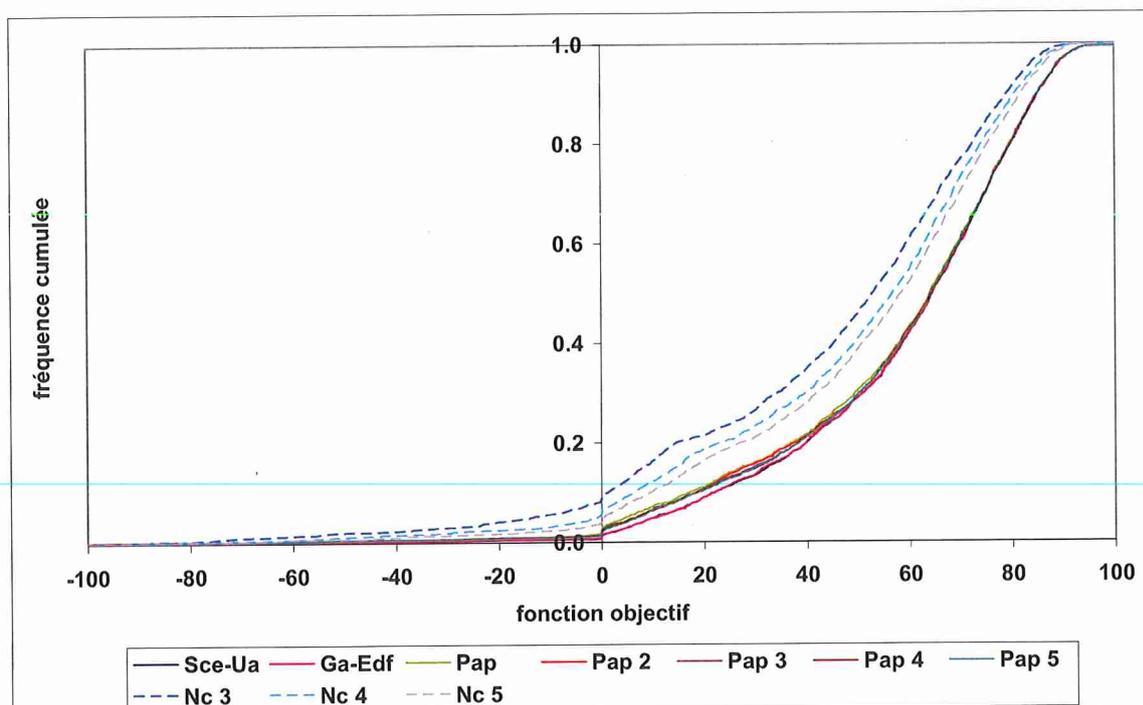
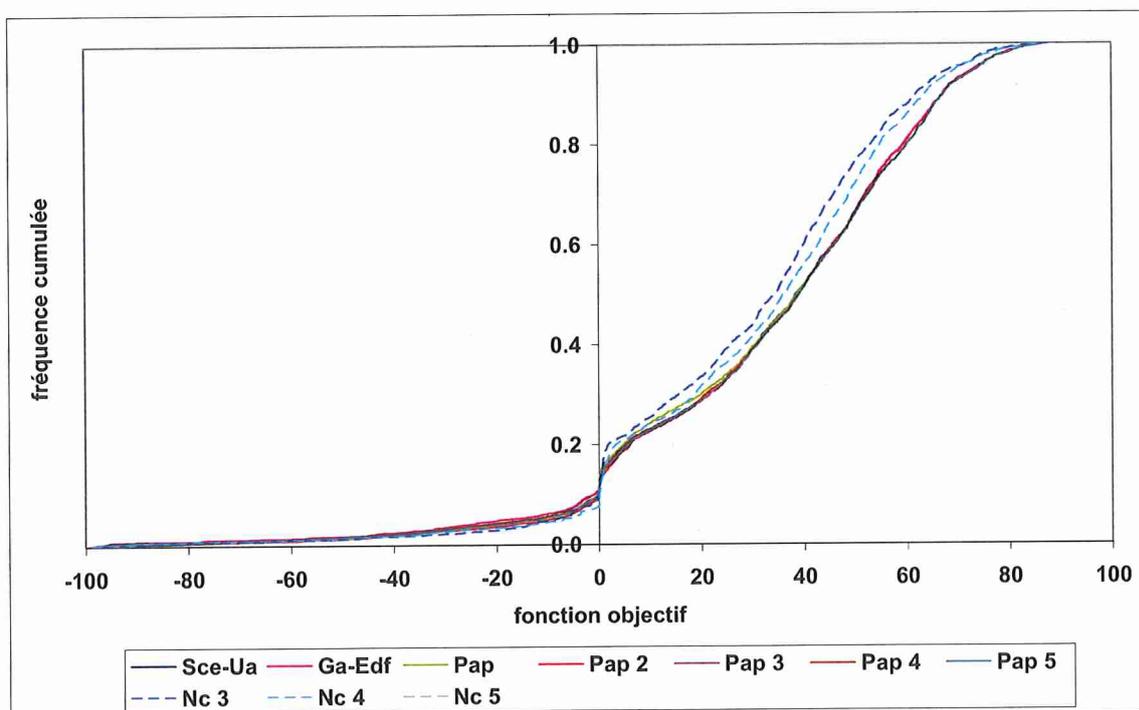


Diagramme schématique de la méthode de calage 'pas-à-pas' (source : Perrin, 2000)

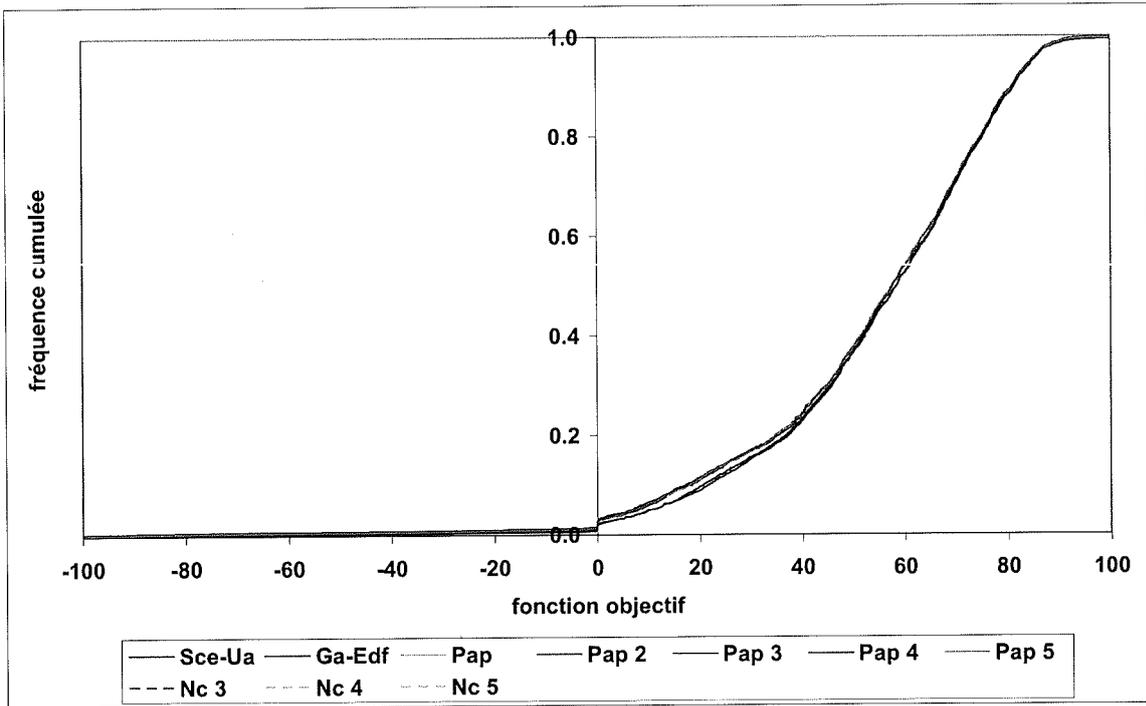
ANNEXE C : RESULTATS DE LA PARTIE THEORIQUE



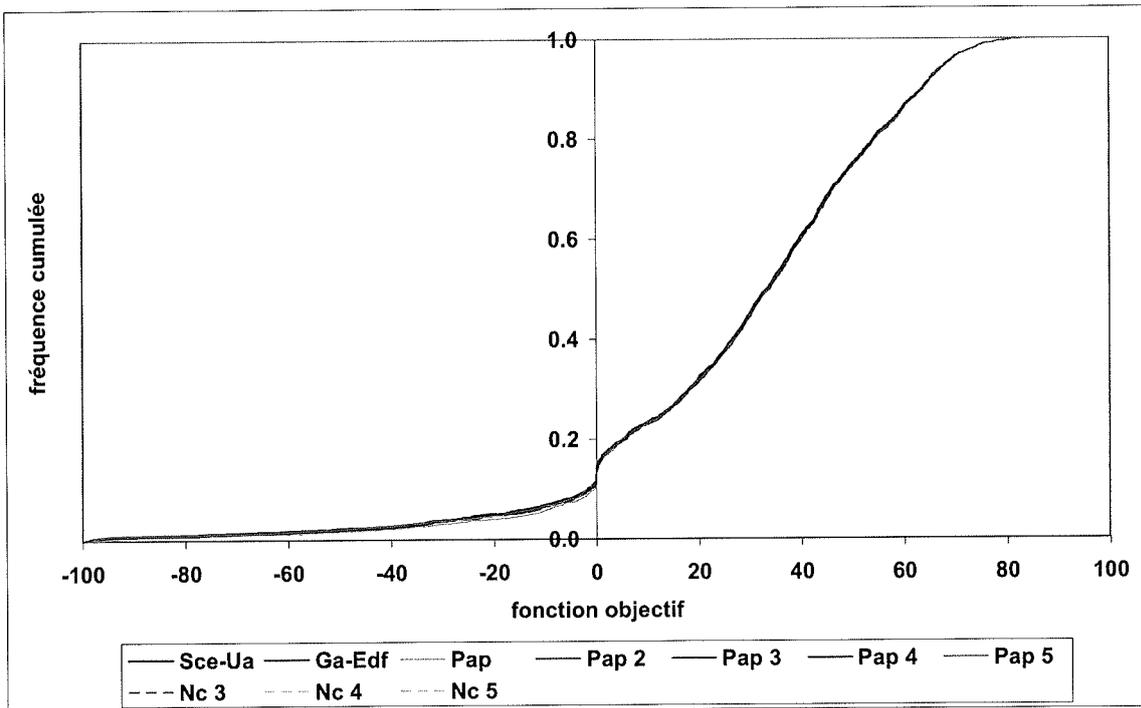
Distributions de la fonction objectif en calage pour les différentes méthodes avec le modèle Gr4h



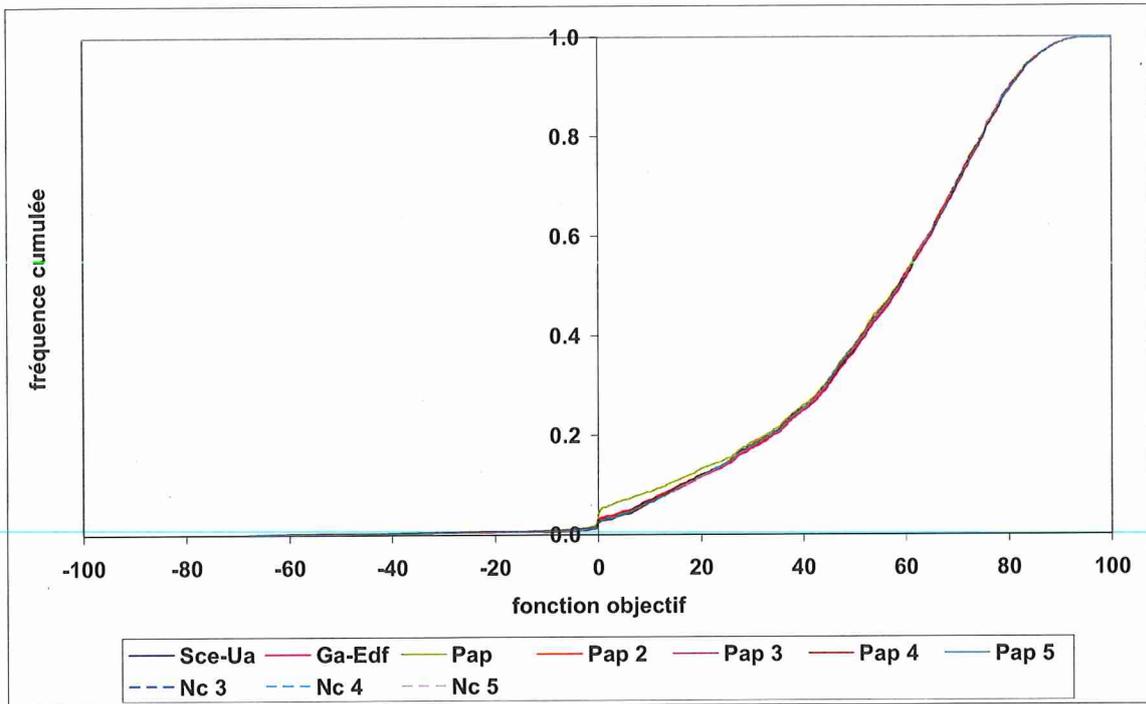
Distributions de la fonction objectif en validation pour les différentes méthodes avec le modèle Gr4h



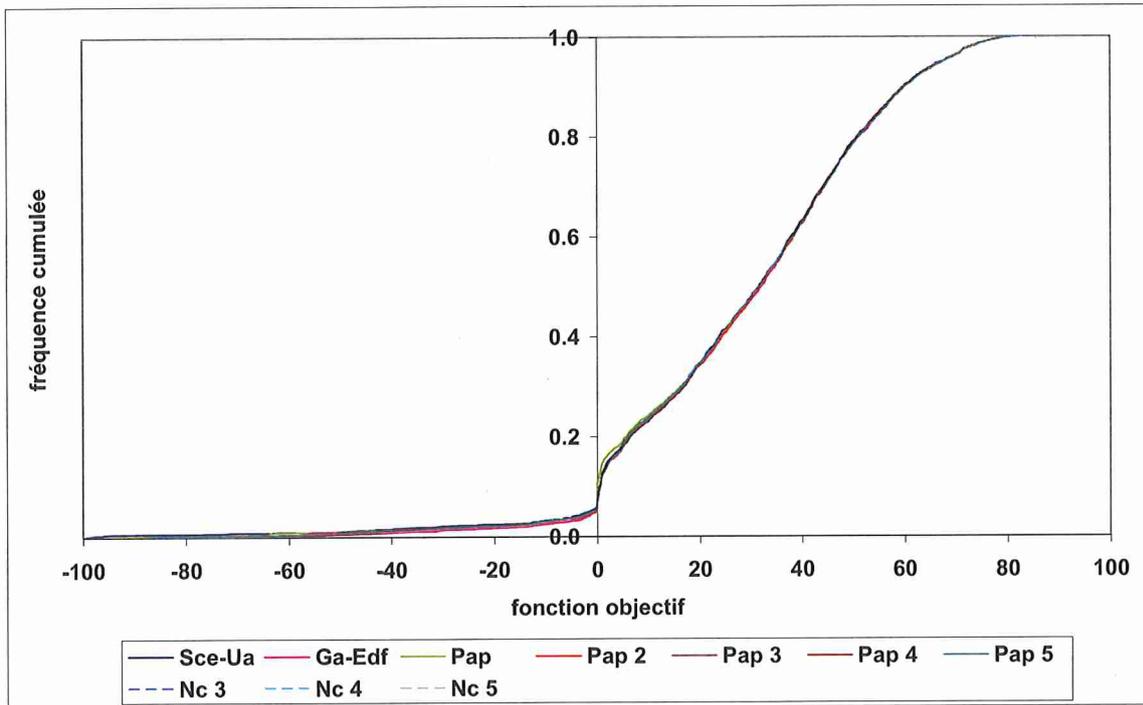
Distributions de la fonction objectif en calage pour les différentes méthodes avec le modèle Gr4j



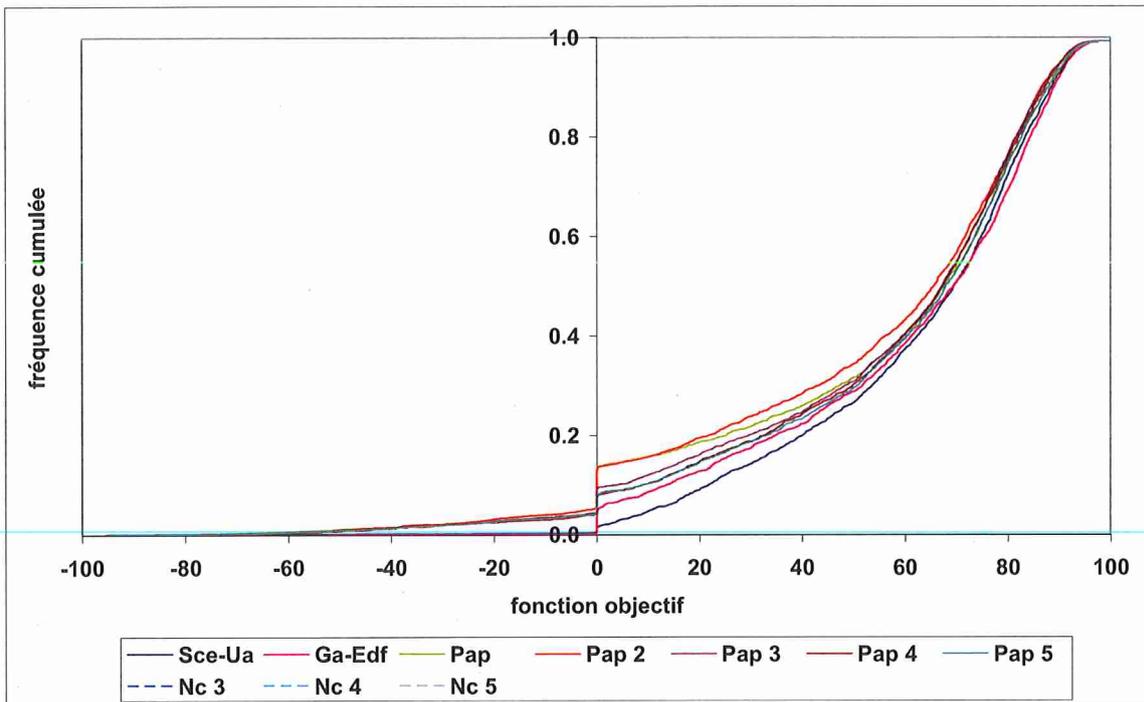
Distributions de la fonction objectif en validation pour les différentes méthodes avec le modèle Gr4j



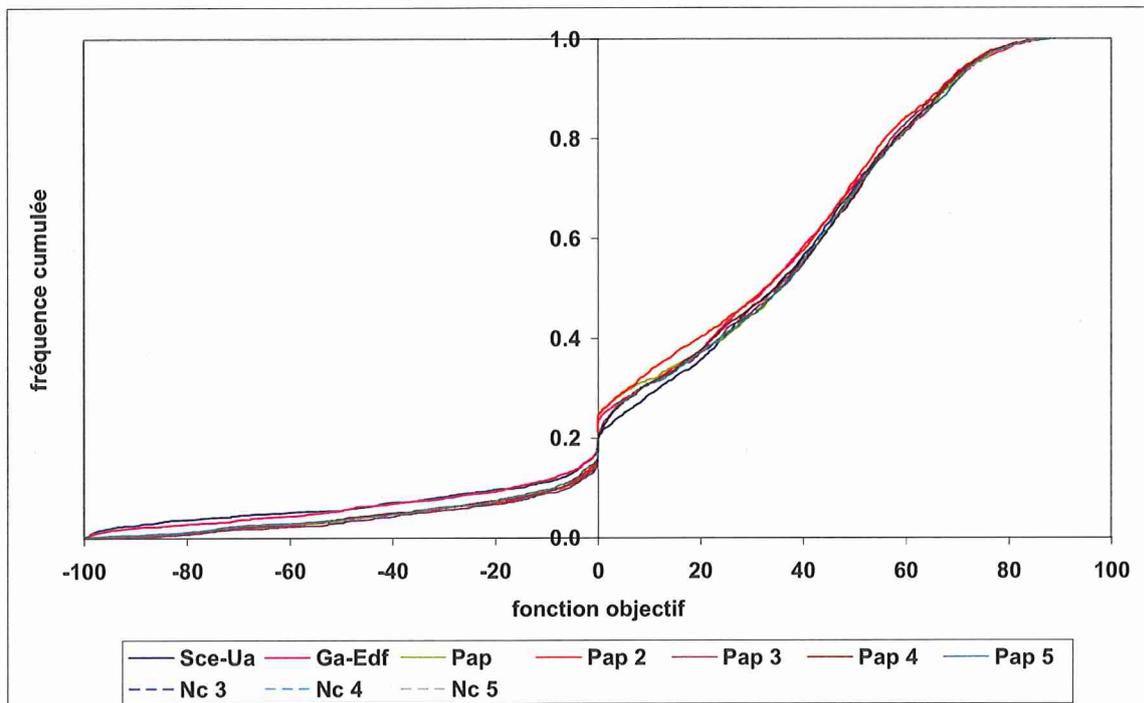
Distributions de la fonction objectif en calage pour les différentes méthodes avec le modèle Ihac



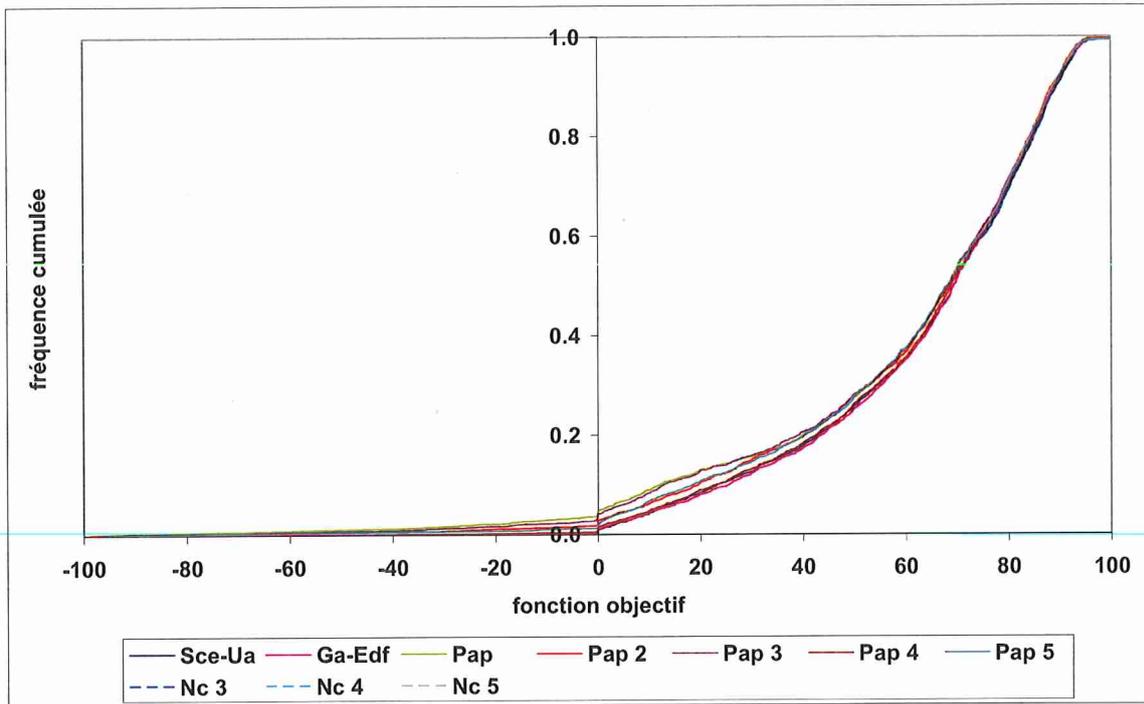
Distributions de la fonction objectif en validation pour les différentes méthodes avec le modèle Ihac



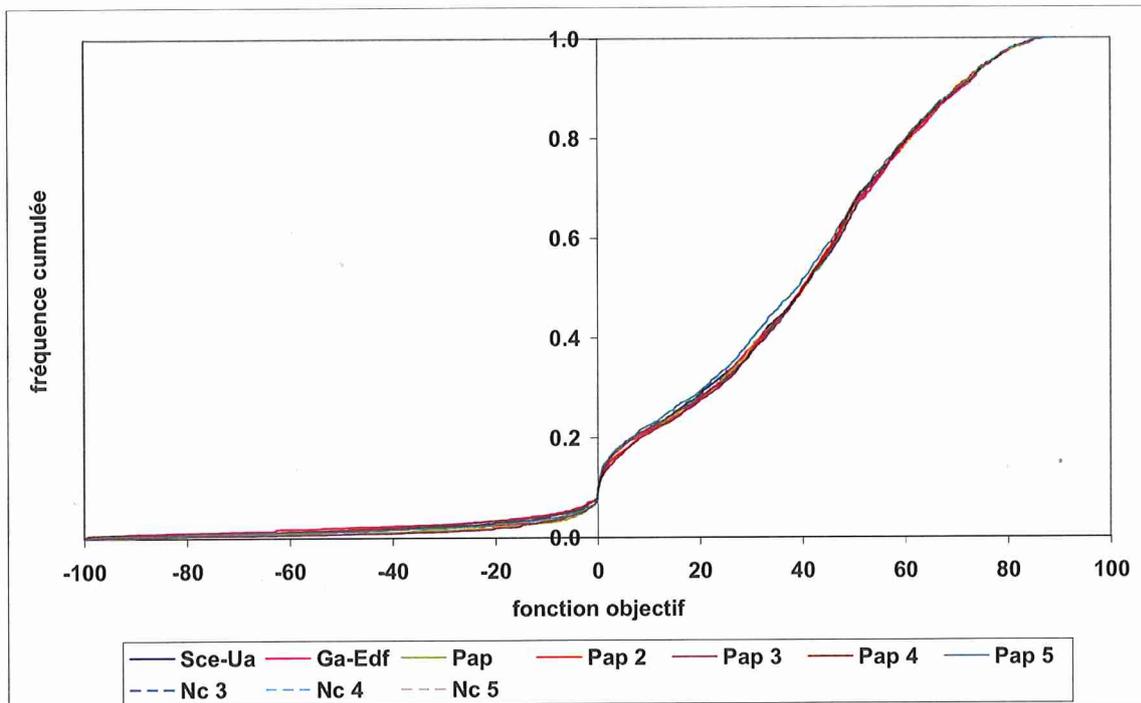
Distributions de la fonction objectif en calage pour les différentes méthodes avec le modèle Topmo



Distributions de la fonction objectif en validation pour les différentes méthodes avec le modèle Topmo



Distributions de la fonction objectif en calage pour les différentes méthodes avec le modèle Mordor



Distributions de la fonction objectif en validation pour les différentes méthodes avec le modèle Mordor

Tableau 17. Différences (maximale, moyenne et minimale) de fonction objectif en calage entre chaque méthode et Pap pour les 2093 bassins-périodes et pourcentage de bassins-périodes avec une différence positive (n1 (↑)), nulle (n2 (=)) et négative (n3 (↓)). (Note : la différence est nulle quand est inférieur à ±0.1)

		Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Sce-Ua	maximale	200.0	200.0	58.6	125.6	199.1	156.7
	moyenne	1.7	1.7	0.9	6.6	3.6	2.9
	minimale	-200.0	-24.2	-26.0	-44.1	-32.2	-65.3
	n1 (↑)	51.6	23.9	20.1	10.1	39.9	29.1
	n2 (=)	8.8	73.1	77.2	61.2	53.5	54.8
n3 (↓)	39.6	3.0	2.7	28.8	6.6	16.1	
Ga-Edf	maximale	200.0	200.0	58.5	147.8	199.1	161.1
	moyenne	1.6	1.4	1.1	5,18	3.7	2.0
	minimale	-199.8	-200.0	-29.5	-83.7	-99.7	-122.5
	n1 (↑)	51.6	23.2	20.2	71.3	38.7	41.0
	n2 (=)	8.8	73.4	78.3	11.6	42.6	42.9
n3 (↓)	39.6	3.4	1.5	17.1	18.7	16.1	
Pap2	maximale	100.0	62.3	57.0	90.0	100.0	81.9
	moyenne	0.3	-0.1	0.5	-1.93	1.3	0.0
	minimale	-200.0	-200.0	-67.5	-118.5	-37.8	-124.8
	n1 (↑)	9.6	8.2	12.4	30.0	24.5	16.9
	n2 (=)	84.7	83.1	79.7	22.3	52.3	64.4
n3 (↓)	5.8	8.7	7.9	47.7	23.2	18.7	
Pap 3	maximale	200.0	200.0	58.6	88.0	199.1	149.1
	moyenne	0.7	0.1	0.7	0,8	0.3	0.5
	minimale	-200.0	-199.3	-117.2	-97.0	-126.5	-148.0
	n1 (↑)	10.4	8.6	13.2	36.5	16.6	17.1
	n2 (=)	83.5	82.6	80.7	19.5	57.7	64.8
n3 (↓)	6.1	8.7	6.1	44.0	25.7	18.1	
Pap 4	maximale	200.0	200.0	58.6	95.8	199.1	150.7
	moyenne	0.6	0.4	0.8	1.6	3.1	1.3
	minimale	-200.0	-200.0	-109.5	-76.8	-34.0	-124.1
	n1 (↑)	10.6	9.2	13.8	40.4	28.3	20.5
	n2 (=)	82.8	82.2	79.7	17.2	54.9	63.4
n3 (↓)	6.6	8.6	6.5	42.5	16.8	16.2	
Pap 5	maximale	200.0	200.0	58.6	96.1	199.1	150.8
	moyenne	0.7	0.2	0.8	2.11	1.4	1.0
	minimale	-200.0	-182.8	-28.1	-97.0	-90.1	-119.6
	n1 (↑)	11.4	9.9	14.4	43.0	20.0	21.1
	n2 (=)	82.9	80.6	80.3	19.1	52.4	63.1
n3 (↓)	5.7	9.5	5.3	38.0	27.6	17.2	
Nc 3	maximale	18.2	---	---	---	---	---
	moyenne	-12.5	---	---	---	---	---
	minimale	-123.8	---	---	---	---	---
	n1 (↑)	1.4	---	---	---	---	---
	n2 (=)	1.4	---	---	---	---	---
n3 (↓)	97.2	---	---	---	---	---	
Nc 4	maximale	25.0	---	---	---	---	---
	moyenne	-8.1	---	---	---	---	---
	minimale	-98.5	---	---	---	---	---
	n1 (↑)	2.6	---	---	---	---	---
	n2 (=)	1.5	---	---	---	---	---
n3 (↓)	95.8	---	---	---	---	---	
Nc 5	maximale	35.5	---	---	---	---	---
	moyenne	-6.5	---	---	---	---	---
	minimale	-200.0	---	---	---	---	---
	n1 (↑)	3.5	---	---	---	---	---
	n2 (=)	1.3	---	---	---	---	---
n3 (↓)	95.2	---	---	---	---	---	

Tableau 18. Différences (maximale, moyenne et minimale) de fonction objectif en validation entre chaque méthode et Pap pour les 2093 bassins-périodes et pourcentage de bassins-périodes avec une différence positive (n1 (↑)), nulle (n2 (=)) et négative (n3 (↓)). (Note : la différence est nulle quand est inférieur à ±0.1)

		Gr4h	Gr4j	Ihac	Topmo	Mordor	moyenne
Sce-Ua	maximale	84.9	96.7	52.0	106.5	89.7	86.0
	moyenne	-0.2	0.8	-0.2	-2.5	-0.8	-0.6
	minimale	-106.3	-100.4	-64.1	-138.5	-101.2	-102.1
	n1 (↑)	47.8	23.0	24.3	41.0	33.8	34.0
	n2 (=)	7.7	50.1	57.3	8,2	28.4	35.9
n3 (↓)	44.5	26.9	18.4	50.8	37.8	35.7	
Ga-Edf	maximale	87.9	91.1	76.5	110.5	62.9	85.8
	moyenne	-0.4	-0.6	0.6	-0	-1.0	-0.3
	minimale	-106.2	-99.5	-56.7	-138.3	-102.3	-100.6
	n1 (↑)	47.6	31.8	35.0	44.0	42.6	40.2
	n2 (=)	8.3	34.9	37.4	4.6	16.6	20.4
n3 (↓)	44.1	33.3	27.6	51.4	40.8	39.4	
Pap2	maximale	87.3	81.3	88.1	104.9	73.5	87.0
	moyenne	0.5	-0.6	0.4	-1.7	-0.4	-0.4
	minimale	-99.4	-131.1	-46.9	-100.6	-102.0	-96.0
	n1 (↑)	23.6	21.8	26.4	36.7	30.1	27.7
	n2 (=)	59.1	53.8	55.3	13.2	36.3	43.5
n3 (↓)	17.3	24.4	18.3	50.1	33.7	28.8	
Pap 3	maximale	90.4	92.6	77.8	112.1	75.8	89.7
	moyenne	0.3	-0.6	0.4	0.35	0.0	0.1
	minimale	-100.0	-120.0	-56.6	-123.6	-96.9	-99.4
	n1 (↑)	23.7	21.9	24.7	46.1	28.5	29.0
	n2 (=)	58.2	53.9	55.9	11.5	38.9	43.7
n3 (↓)	18.0	24.2	19.4	42.4	32.6	27.3	
Pap 4	maximale	90.6	81.2	88.7	117.1	85.1	92.5
	moyenne	0.1	-0.7	0.4	0.2	-0.7	-0.1
	minimale	-99.5	-131.5	-40.7	-135.3	-102.3	-101.9
	n1 (↑)	23.6	22.0	24.6	44.7	32.0	29.4
	n2 (=)	58.9	52.9	55.4	10.0	35.5	42.5
n3 (↓)	17.5	25.0	20.0	45.3	32.6	28.1	
Pap 5	maximale	85.2	91.8	88.5	116.0	90.6	94.4
	moyenne	0.3	-0.6	0.3	0.0	-1.4	-0.3
	minimale	-94.8	-100.3	-56.7	140.2	-101.9	-42.7
	n1 (↑)	25.8	22.9	25.7	47.1	26.1	29.5
	n2 (=)	57.3	52.2	54.9	9.2	34.1	41.5
n3 (↓)	16.8	24.8	19.4	43.8	39.8	28.9	
Nc 3	maximale	93.3	---	---	---	---	---
	moyenne	-3.9	---	---	---	---	---
	minimale	-107.8	---	---	---	---	---
	n1 (↑)	33.2	---	---	---	---	---
	n2 (=)	3.2	---	---	---	---	---
n3 (↓)	63.6	---	---	---	---	---	
Nc 4	maximale	93.9	---	---	---	---	---
	moyenne	-2.1	---	---	---	---	---
	minimale	-95.8	---	---	---	---	---
	n1 (↑)	35.8	---	---	---	---	---
	n2 (=)	3.3	---	---	---	---	---
n3 (↓)	60.9	---	---	---	---	---	
Nc 5	maximale	95.6	---	---	---	---	---
	moyenne	-1.3	---	---	---	---	---
	minimale	-92.0	---	---	---	---	---
	n1 (↑)	38.5	---	---	---	---	---
	n2 (=)	3.6	---	---	---	---	---
n3 (↓)	57.9	---	---	---	---	---	

Tableau 19. Différences (maximale, moyenne et minimale) de fonction objectif en calage entre chaque méthode de calage discret et son homologue locale multidépart pour les 2093 bassins-périodes et pourcentage de bassins-périodes avec une différence positive (n1 (↑)), nulle (n2 (=)) et négative (n3 (↓)). (Note 1 : la différence est nulle quand est inférieur à ±0.1). (Note 2 : modèle utilisé : Gr4h)

		Gr4h
Nc 3 comparé à Pap 3	maximale	200.0
	moyenne	-13.2
	minimale	-200.0
	n1 (↑)	0.1
	n2 (=)	1.1
	n3 (↓)	98.8
Nc 4 comparé à Pap 4	maximale	200.0
	moyenne	-8.8
	minimale	-200.0
	n1 (↑)	0.1
	n2 (=)	1.1
	n3 (↓)	98.8
Nc 5 comparé à Pap 5	maximale	0.0
	moyenne	-7.2
	minimale	-200.0
	n1 (↑)	0.0
	n2 (=)	1.2
	n3 (↓)	98.8

Tableau 20. Différences (maximale, moyenne et minimale) de fonction objectif en calage entre chaque méthode de calage discret et son homologue locale multidépart pour les 2093 bassins-périodes et pourcentage de bassins-périodes avec une différence positive (n1 (↑)), nulle (n2 (=)) et négative (n3 (↓)). (Note 1 : la différence est nulle quand est inférieur à ±0.1). (Note 2 : modèle utilisé : Gr4h)

		Gr4h
Nc 3 comparé à Pap 3	maximum	100.2
	moyenne	-4.1
	minimum	108.8
	n1 (↑)	30.9
	n2 (=)	3.8
	n3 (↓)	65.3
Nc 4 comparé à Pap 4	maximum	83.1
	moyenne	-2.2
	minimum	-82.8
	n1 (↑)	34.5
	n2 (=)	4.3
	n3 (↓)	61.3
Nc 5 comparé à Pap 5	maximum	92.6
	moyenne	-1.6
	minimum	-81.2
	n1 (↑)	36.8
	n2 (=)	4.4
	n3 (↓)	58.8

ANNEXE D : RESULTATS DE LA PARTIE APPLIQUEE

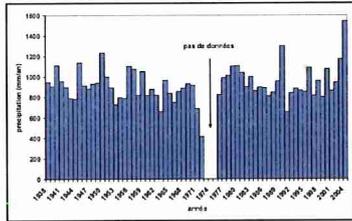


Figure 7. Précipitation annuelle à Coshocton 196

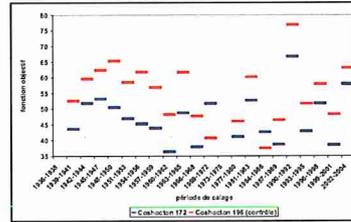


Figure 8. Fonction objectif au calage pour Coshocton

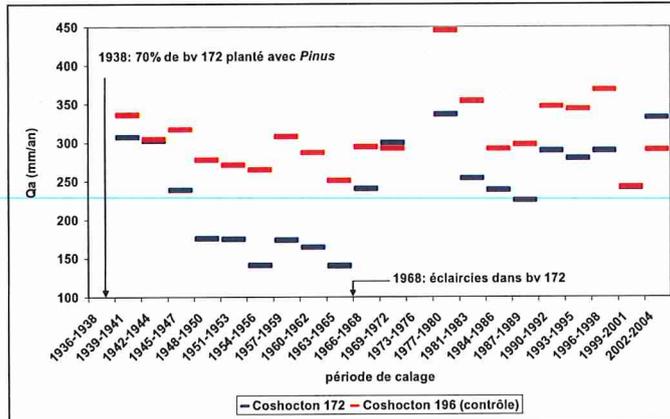


Figure 9. Q_a pour Coshocton

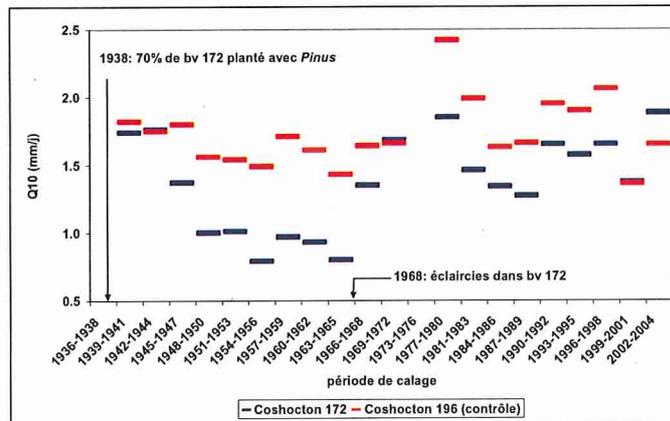


Figure 10. Q_{10} pour Coshocton

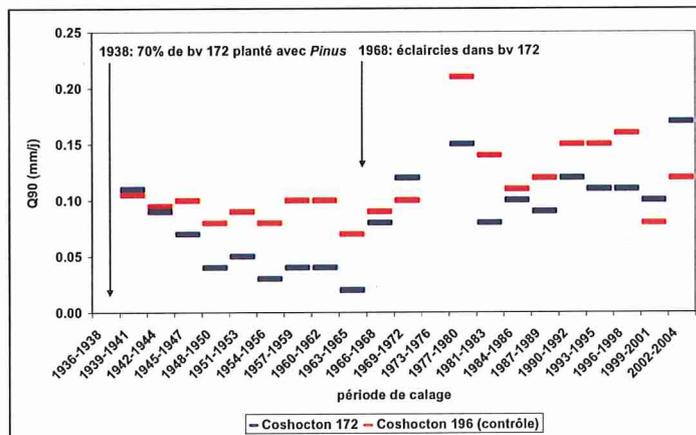


Figure 11. Q_{90} pour Coshocton

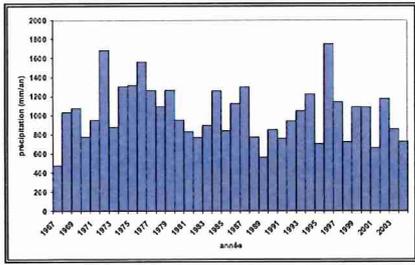


Figure 12. Précipitation annuelle à Vaubarnier

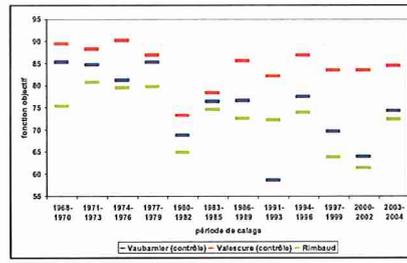


Figure 13. Fonction objectif au calage pour Réal Collobrier

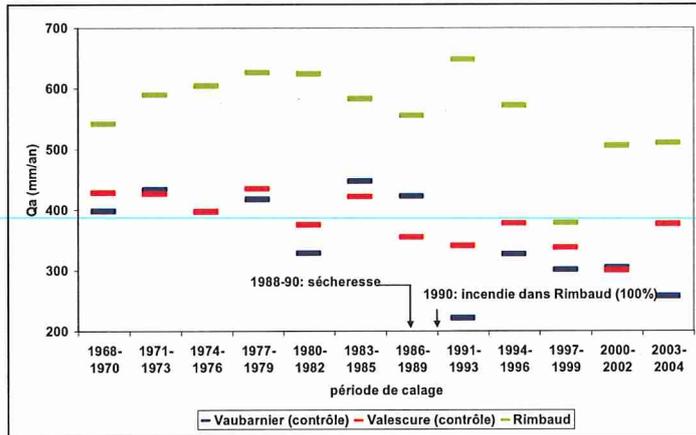


Figure 14. Q_a pour Réal Collobrier

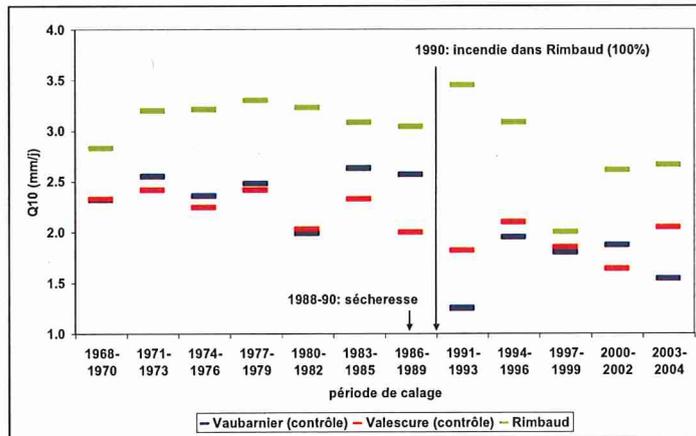


Figure 15. Q_{10} pour Réal Collobrier

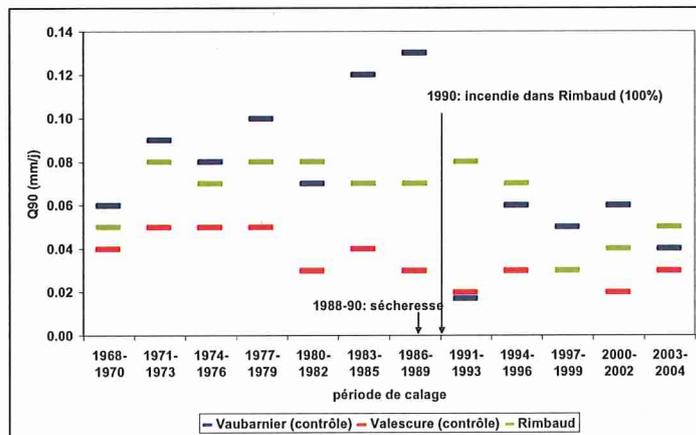


Figure 16. Q_{90} pour Réal Collobrier

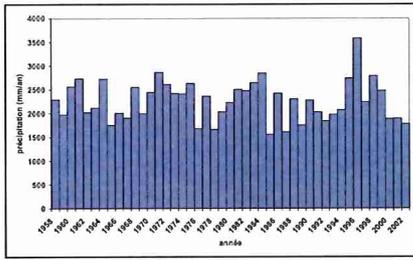


Figure 17. Précipitation annuelle à Andrews 2

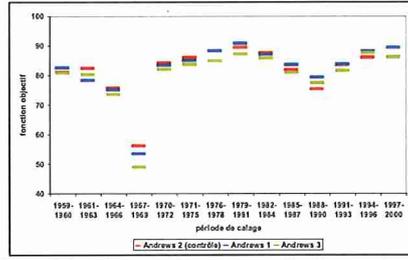


Figure 18. Fonction objectif au calage pour Andrews (partie 1)

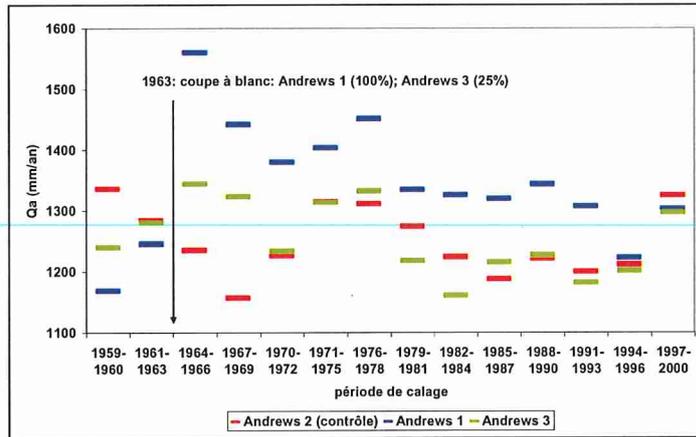


Figure 19. Q_a pour Andrews (partie 1)

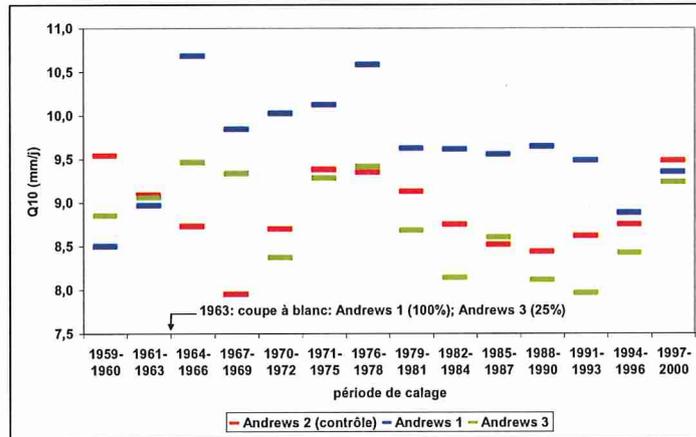


Figure 20. Q_{10} pour Andrews (partie 1)

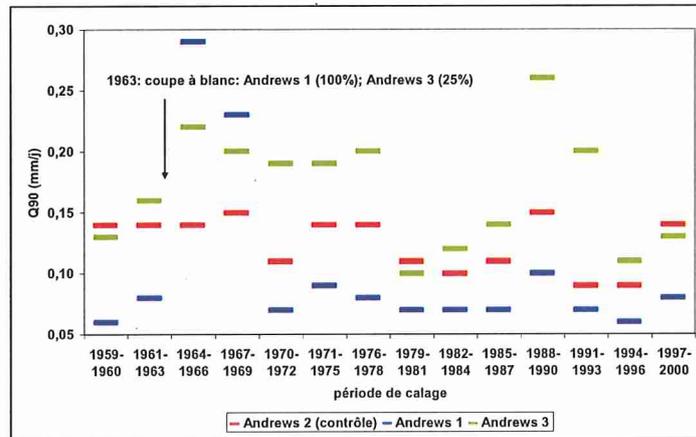


Figure 21. Q_{90} pour Andrews (partie 1)

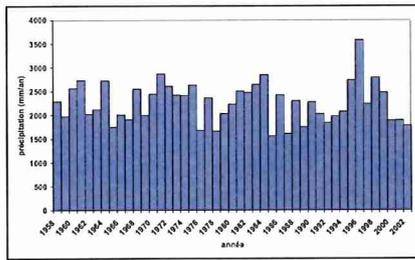


Figure 22. Précipitation annuelle à Andrews 9

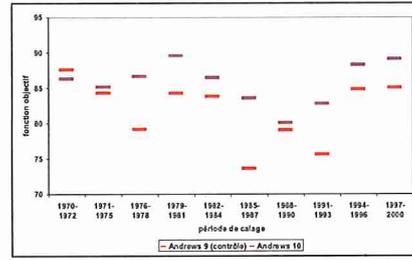


Figure 23 . Fonction objectif au calage pour Andrews (partie 2)

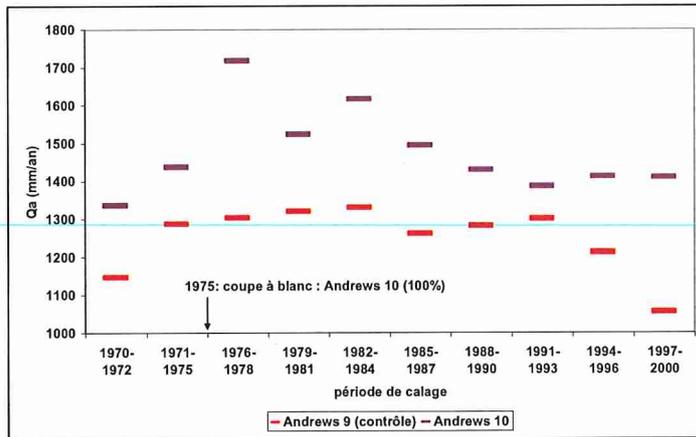


Figure 24. Q_a pour Andrews (partie 2)

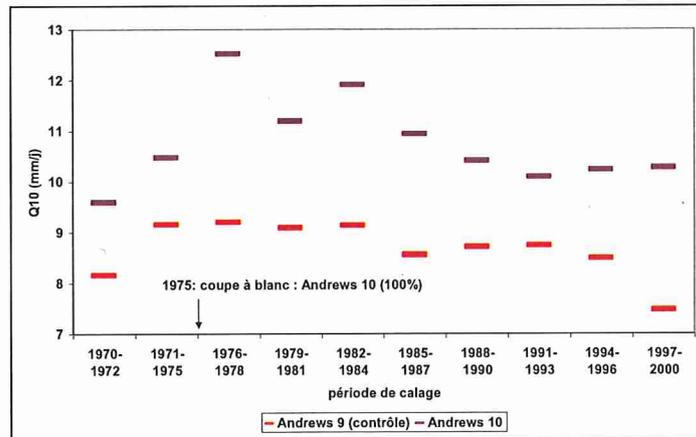


Figure 25. Q_{10} pour Andrews (partie 2)

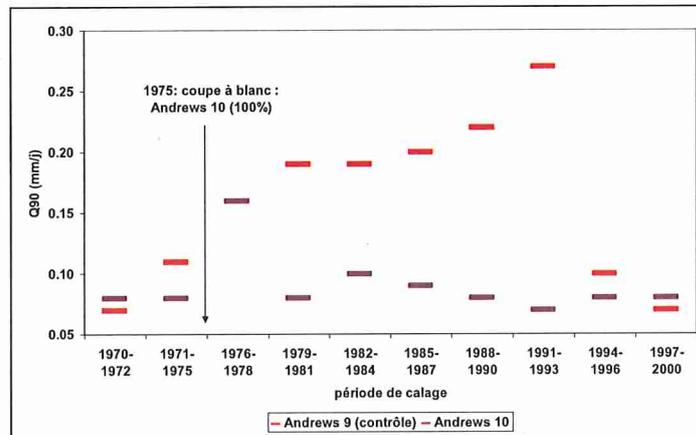


Figure 26. Q_{90} pour Andrews (partie 2)

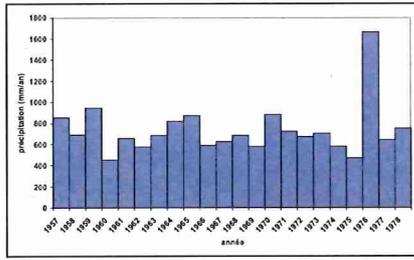


Figure 27. Précipitation annuelle à Three Bar D

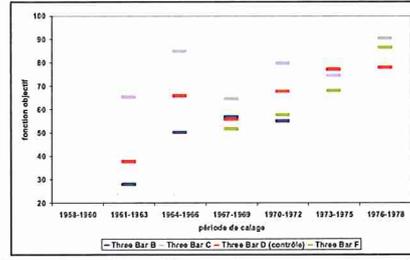


Figure 28. Fonction objectif au calage pour Three Bar

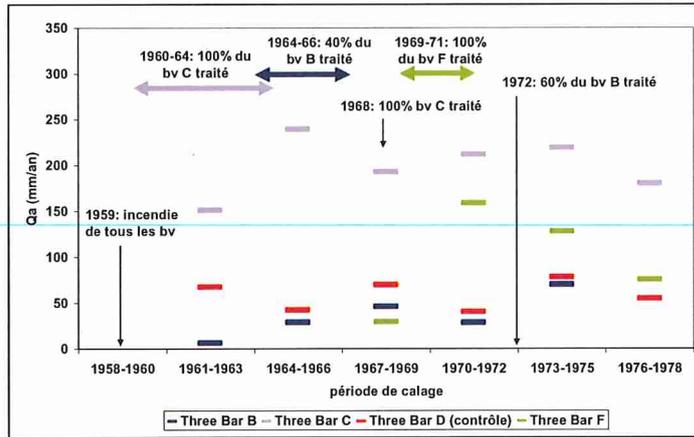


Figure 29. Q_a pour Three Bar

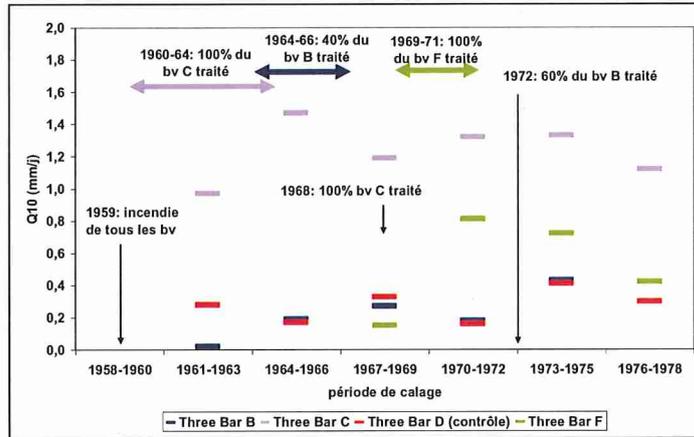


Figure 30. Q_{10} pour Three Bar

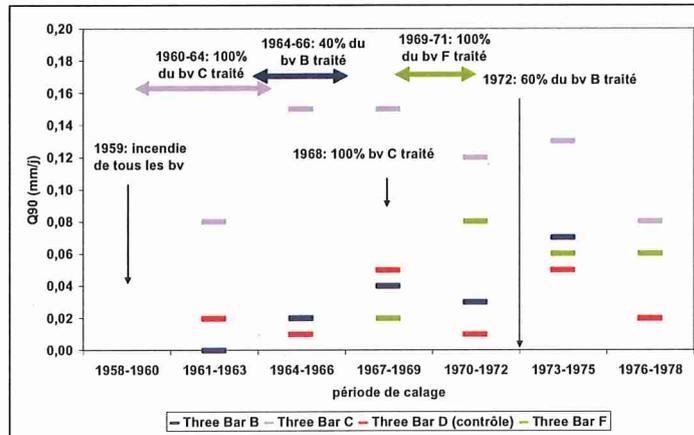


Figure 31. Q_{90} pour Three Bar

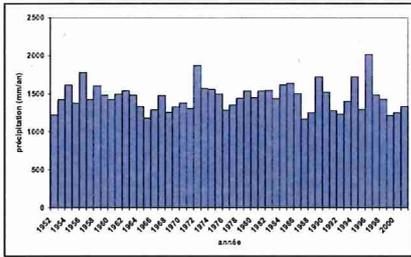


Figure 32. Précipitation annuelle à Fernow 4

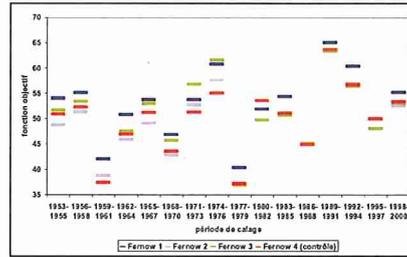


Figure 33. Fonction objectif au calage pour Fernow (partie 1)

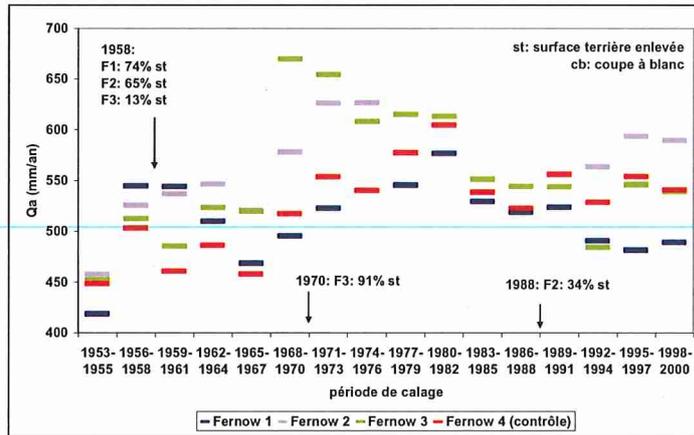


Figure 34. Q_a pour Fernow (partie 1)

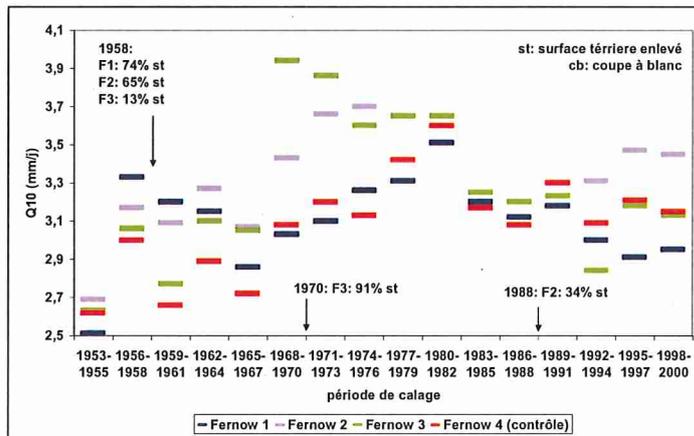


Figure 35. Q_{10} pour Fernow (partie 1)

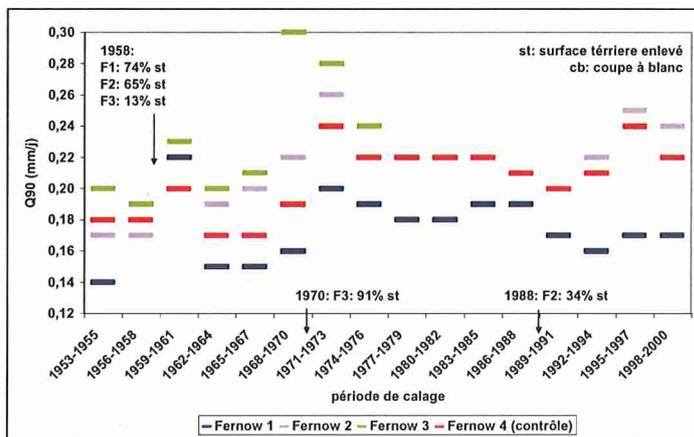


Figure 36. Q_{90} pour Fernow (partie 1)

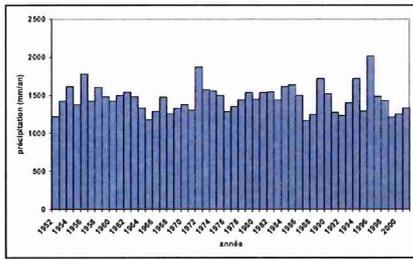


Figure 37. Précipitation annuelle à Fernow 4

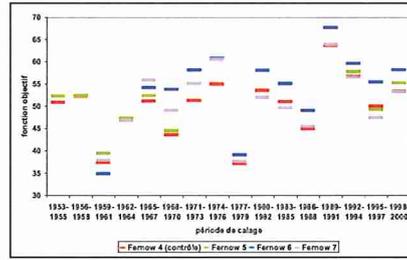


Figure 38. Fonction objectif au calage pour Fernow (partie 2)

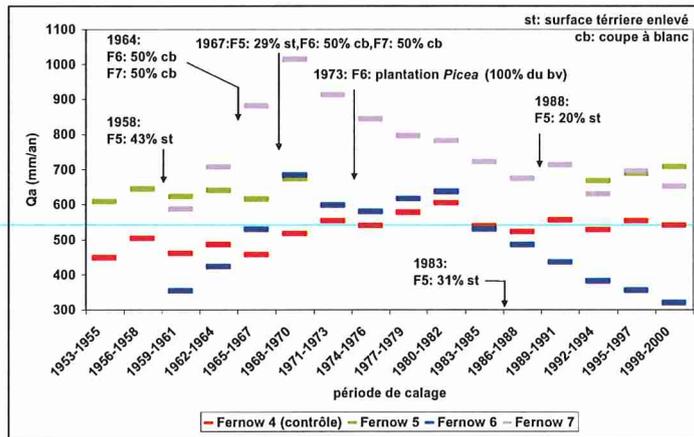


Figure 39. Q_a pour Fernow (partie 2)

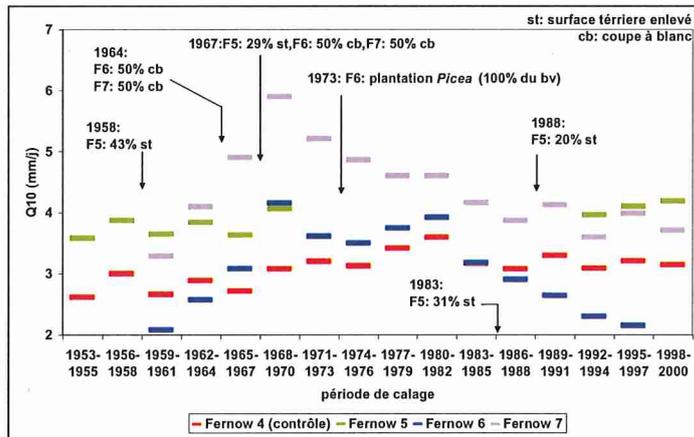


Figure 40. Q_{10} pour Fernow (partie 2)

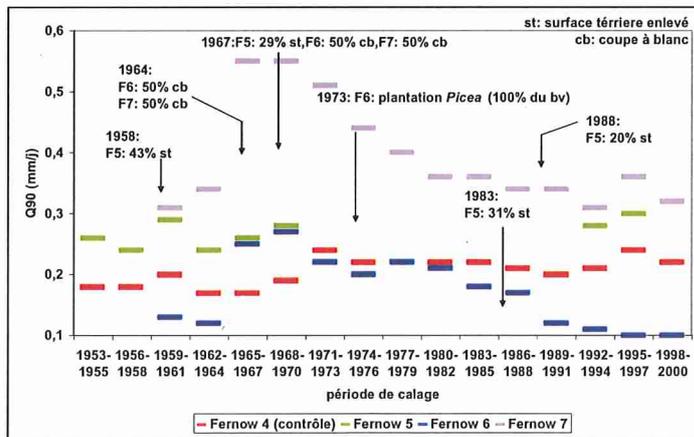


Figure 41. Q_{90} pour Fernow (partie 2)

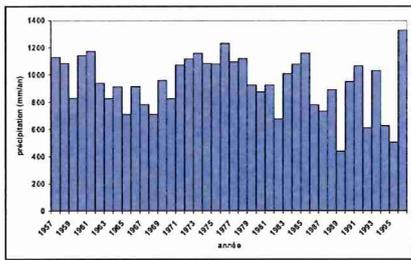


Figure 42. Précipitation annuelle à Mokobulaan C

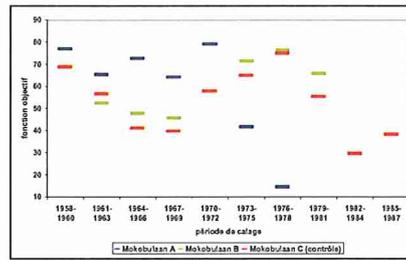


Figure 43 . Fonction objectif au calage pour Mokobulaan

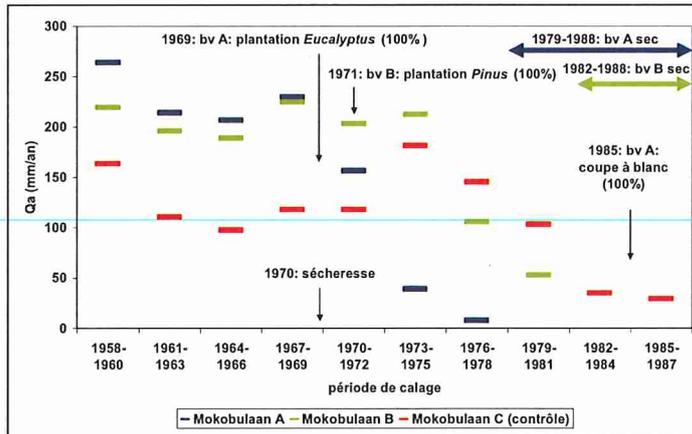


Figure 44. Q_a pour Mokobulaan

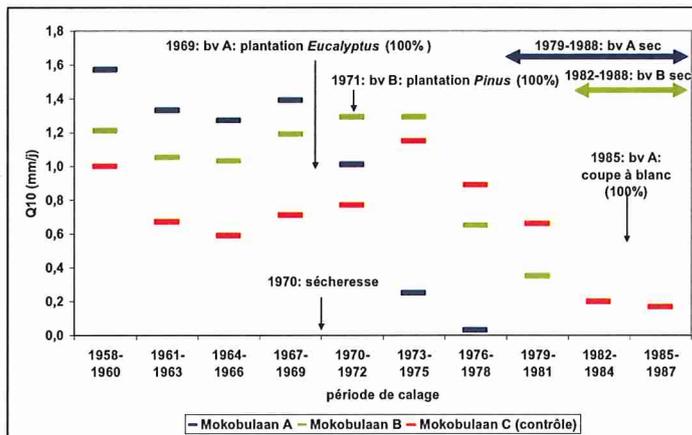


Figure 45. Q_{10} pour Mokobulaan

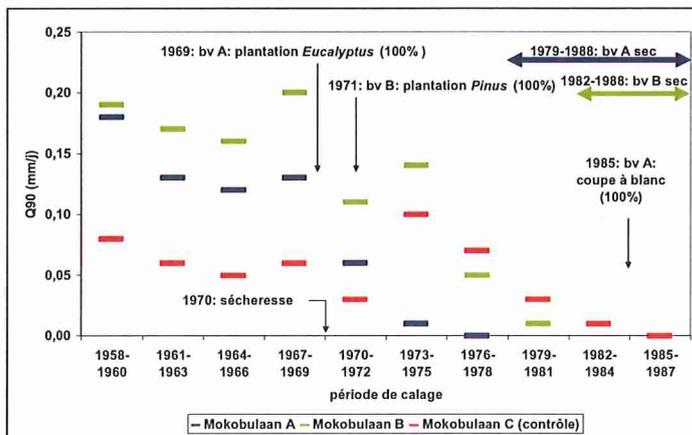


Figure 46. Q_{90} pour Mokobulaan

ABSTRACT

This study was divided into two parts: a “theoretical” modelling study and an applied modelling study.

In the theoretical study, we tried to improve a local method for the calibration of rainfall-runoff models by the addition of a multistart strategy using *a priori* information. We found that a multistart strategy did not improve the results of the original calibration method in terms of efficiency, performance in calibration, performance in validation and robustness.

In the applied study, we assessed the usefulness of a model-based technique to detect changes in treated basins streamflow behaviour. We found that this technique detects the behaviour changes but we found as well that it is not robust enough to substitute the paired-watershed technique. However, it can be helpful in detecting non-stationary behaviours in control basins. Thus, we recommend a joint use of both techniques.

RESUM

Aquest projecte de fi de carrera consta de dues parts: un estudi de modelització “teòrica” i un estudi de modelització aplicada.

En la part teòrica, hem provat de millorar un mètode de calibratge de models de pluja-escolament afegint-li una estratègia multi-inici a partir d’informació *a priori*. Hem trobat que aquest estratègia no millora els resultats del mètode de calibratge original pel que fa a eficiència, rendiment en calibratge, rendiment en validació i robustesa.

En la part d’aplicació, hem avaluat la utilitat d’una tècnica de detecció de canvis de comportament hidrològic basada en la modelització. Hem trobat que aquesta tècnica detecta els canvis de comportament però que no és prou robusta com per substituir la tècnica de les conques aparellades. Tot i això, pot ser útil en la detecció de comportaments no estacionaris de les conques control. Per això, recomanem la utilització conjunta de les dues tècniques.

RESUME

Cette étude a été divisée en deux parties : un étude de modélisation « théorique » et une étude de modélisation appliquée.

Dans la partie théorique, on a essayé d’améliorer une méthode de calage des modèles pluie-débit en lui ajoutant une stratégie multidépart à partir d’information *a priori*. On a trouvé que cette stratégie n’améliore pas les résultats de la méthode de calage originale en termes d’efficacité, performance en calage, performance en validation et robustesse.

Dans la partie d’application, on a évalué l’utilité d’une technique de détection des changements du comportement hydrologique basée sur la modélisation. On a trouvé que cette technique parvient à détecter les changements du comportement mais qu’elle n’est pas assez robuste pour se substituer à la technique des bassins appariés. Cependant, elle peut être utile pour la détection de comportements non-stationnaires des bassins de contrôle. Pour cette raison, on recommande l’utilisation conjointe des deux techniques.