



Mémoire de Master 2^{ème} année

Sciences de la terre spécialité Hydrologie, Hydrogéologie et sols

Année 2004-2005

Charles-Henry Berthier

Quantification des incertitudes des débits calculés par un modèle pluie-débit empirique

Encadrant :

M. Claude MICHEL

Résumé

La transformation de la pluie en débit à l'échelle du bassin versant est simulée depuis le début des années soixante. L'accroissement des capacités de calcul permet d'accroître la qualité des modèles. Malgré cela, il est rare de trouver dans la littérature des résultats de modélisation associés à un intervalle d'erreurs. Nous avons tenté dans ce travail de DEA de quantifier les incertitudes sur les débits calculés par deux modèles pluie-débit conceptuel GR4J et TOPMO. Nous avons comparé trois méthodes d'estimation des incertitudes, une méthode dérivée de Monte-Carlo fondée sur une approximation linéaire des écarts-types des paramètres du modèle, la méthode GLUE et une dernière méthode que nous proposons pour représenter les écarts entre les débits observés et les débits calculés par TOPMO ou GR4J.

Abstract

Rainfall-runoff transformation and simulation has been carried out since the beginning of the sixties. The increase in calculation capacities enables the improvement of model quality. However, the modelling results reported in the literature are rarely presented with their error bounds. The aim of this Masters' work was to quantify the uncertainties of the runoff values calculated by two conceptual rainfall-runoff models; GR4J and TOPMO. We compared three uncertainty estimation methods. The first is a Monte-Carlo type approach based on a linear approximation of the standard deviations of the models' parameters. The second is the GLUE and the third method is a new approach we suggest to represent the standard deviations between the observed runoff values and those calculated by TOPMO or GR4J.

REMERCIEMENTS

Je tiens à remercier chaleureusement mon encadrant, Claude Michel, hydrologue au Cemagref d'Antony, pour son expérience, sa patience et son entière disponibilité qui m'ont permis de mener à bien ce stage de DEA.

J'adresse ma sincère reconnaissance à Vasken Andréassian, chef de service de l'unité Hydrologique du Cemagref d'Antony, pour ses encouragements, ses conseils et sa bonne humeur.

Je souhaite également remercier l'équipe enseignante du laboratoire d'hydrologie et de géochimie isotopique, qui nous a fait part de son expérience et de ces connaissances, et tout particulièrement les responsables Christelle Marlin et Laurent Dever.

Merci aussi à toutes les personnes de l'équipe hydrologie du Cemagref, qui ont contribué à ce travail par leur aide et leur accueil au sein du service.

SOMMAIRE

RÉSUMÉ	1
ABSTRACT.....	1
I- INTRODUCTION	5
I-1. Qu'est ce que la modélisation hydrologique.....	6
I-2. Objectif de l'étude : évaluation de l'incertitude des débits calculés	7
II- CADRE DE LA RECHERCHE : LA MODÉLISATION PLUIE-DÉBIT	9
II-1. Les Modèles : classification et exemples	9
II-1.a. TOPMO	10
II-1.b. IHACRES.....	11
II-1.c. GR4J	12
II-2. Les variables	13
II-2.a. Les précipitations (P)	13
II-2.b. L'évapotranspiration potentielle (E).....	13
II-2.c. Le débit (Q).....	14
II-3. Les Paramètres.....	14
II-4. Le calage.....	15
II-4.a. Le calage manuel	16
II-4.b. Le calage automatique.....	16
II-5. Critère d'ajustement (ou fonction objectif).....	17
II-5.a. Le critère de Nash (Nash et Sutcliffe, 1970).....	18
II-5.b. Le critère de Nash borné.....	18
II-6. Le contrôle de la performance du modèle.....	18
II-7. Méthodes d'estimation des incertitudes des débits calculés	19
II-7.a. Méthode dérivée de Monte-Carlo fondée sur une approximation linéaire des écarts-types des paramètres du modèle.....	19
II-7.b. La méthode GLUE : l'échantillonnage par ordre d'importance	21
II-7.c. La méthode proposée	23
III- MÉTHODES ET RÉSULTATS.....	25
III-1. Zones d'étude.....	25
III-2. Représentation graphique des incertitudes	26
III-2.a. Hydrogrammes de crues.....	26
III-2.b. Rapport $Q_{\text{observé}}/Q_{\text{calculé}}$ au calage sur la période de calage	27
III-2.c. Rapport $Q_{\text{observé}}/Q_{\text{calculé}}$ au calage sur la période 1974-1979 par tranche de débit calculé.....	29
III-2.d. Rapport $Q_{\text{mesuré}}/Q_{\text{calculé}}$ au contrôle sur l'ensemble de la période 1974- 1992	32
III-2.e. Qualité du modèle et fiabilité de l'estimation des incertitudes.....	35
III-3. Analyses des résultats fournis par GR4J	37
III-3.a. Estimation, des incertitudes : Méthode proposée.....	37
III-3.b. Peut-on exploiter l'information pour effectuer un réajustement a posteriori des débits calculé par le modèle ?.....	38

III-3.c.	Estimation des incertitudes : Méthode dérivée de Monte-Carlo fondée sur une approximation linéaire des écarts-types des paramètres du modèle.....	41
III-3.d.	Estimation des incertitudes : Méthode GLUE.....	42
III-4.	Analyses des résultats obtenus sur un autre modèle : TOPMO	43
III-4.a.	Estimation des incertitudes : Méthode proposée.....	43
III-4.b.	Comparaison des résultats obtenus avec TOPMO et GR4J.....	44
III-4.c.	Estimation des incertitudes : Méthode de Monte-Carlo fondée sur une approximation linéaire des écart-types des paramètre du modèle.....	45
IV-	CONCLUSION	47
V-	BIBLIOGRAPHIE	49
VI-	WEBOGRAPHIE.....	51
VII-	TABLES DES LÉGENDES	51
VII-1.	Index des figures	51
VII-2.	Index des tableaux.....	52
VIII-	ANNEXES	53
VIII-1.	Annexe 1 : Rappels sur les Quantiles.....	53
VIII-2.	Annexe 2 : Résultats avec GR4J et la méthode proposée	54

*L'eau n'oublie pas son chemin**(Proverbe russe)***I- INTRODUCTION**

L'objectif de ce travail est de quantifier les incertitudes des débits calculés par un modèle pluie-débit conceptuel. L'hydrologie est loin d'être une science précise en raison des incertitudes extrêmement fortes qui affectent mesures et modèles : c'est pour cela qu'il serait intéressant de quantifier les incertitudes afin d'avertir les utilisateurs des erreurs possibles, et de leur permettre de prendre du recul sur la signification des résultats.

Nous adoptons ici un point de vue selon lequel un modèle hydrologique n'a pas pour but de serrer au plus près la réalité physique des phénomènes qui se déroulent dans un bassin versant, mais de prévoir au mieux les débits à l'exutoire d'un bassin versant et accessoirement de constituer un outil efficace pour résoudre des problèmes pratiques sur les ressources en eau.

Le modèle est alors une représentation inévitablement simplifiée dont l'objectif est de représenter le comportement du système bassin versant. Tout modélisateur sait que son modèle est inexact, donnant des résultats plus ou moins éloignés d'une réalité. Pour utiliser son modèle et ainsi prendre des décisions, l'utilisateur doit connaître l'ordre de grandeur des inexactitudes de son modèle et ainsi de pouvoir quantifier et annoncer ces incertitudes.

Les origines des erreurs peuvent être diverses (Perrin, 2002) :

Type d'erreurs	Nature de l'erreur
Incertitudes des données d'entrée	Erreurs au moment de la collecte des données et de leurs traitements, mauvaise représentation de la variabilité spatio-temporelle
Incertitudes du modèle	Le modèle reste une représentation grossière d'un système naturel complexe
Incertitudes des paramètres du modèle	Dépend du choix de la fonction objectif, des performances d'optimisation, des échantillons des données du calage

Tableau 1 : Classification des erreurs en modélisation hydrologique selon leur nature

Les sources d'erreurs génèrent des incertitudes à chaque étape du processus de modélisation, et sont propagées tout le long de la procédure (du calage à l'application). Malgré le potentiel important des modèles, ces outils restent encore assez faiblement utilisés pour la gestion opérationnelle de l'eau. Parmi les multiples obstacles, le principal est probablement le peu de communication entre les modélisateurs et les gestionnaires. Les gestionnaires, peu familiarisés avec les méthodes de modélisation et les techniques informatiques, sont peu disposés à mettre en oeuvre des procédures qui leur semblent mystérieuses et secrètes. Un deuxième obstacle qui peut limiter l'utilisation des modèles pour la gestion de l'eau, est probablement le manque de données suffisantes pour le calage du modèle. Enfin l'absence de possibilités d'apprécier quantitativement l'incertitude des prévisions est un troisième obstacle, (Petruscu Maftai, 2002) et c'est précisément sur ce point que nous allons travailler.

I-1. Qu'est ce que la modélisation hydrologique

Un modèle hydrologique est une représentation mathématique simplifiée de tout ou partie des processus du cycle hydrologique. Un ensemble de concepts hydrologiques sont exprimés en langage mathématique pour représenter le comportement correspondant observé dans la nature. Pour l'utilisateur, l'intérêt d'un modèle réside dans sa capacité à apporter une réponse "satisfaisante" aux questions qu'il se pose à propos de l'objet modélisé. De nos jours il existe différents types de modèles : du modèle physique au modèle empirique, du distribué au global, et chaque type exprime une façon de concevoir le cycle hydrologique (Mouelhi, 2003). Une classification peut être faite sur quelques clefs de différenciation d'après des critères énoncés par Clarke (1973) & Ambroise (1998). (cf. chpt.II.1)

L'apparition et l'essor des modèles pluie-débit a commencé dans les années soixante. (Nascimento, 1995). Le souci de la modélisation pluie-débit est de mettre à la disposition des aménageurs ou bien des ingénieurs un outil «simple» permettant l'estimation ou la prévision des débits pour une étude d'aménagement désirée (barrages, lacs, ponts, etc...).

La modélisation pluie-débit a de nombreuses applications (Michel, 1989 ; Refsgaard & Abbott, 1996) parmi lesquelles on peut citer :

- Estimation des débits de rivières pour des sites non jaugés ;
- Dimensionnement des ouvrages d'art (barrages, ponts, déversoirs de sécurité...);
- Délimitation des zones inondables ;
- Simulation des débits pour des reconstitutions historiques ;
- Prévision des crues ;

- Détection de l'influence des changements d'occupation du sol ;
- Simulation de l'impact du changement climatique ;
- Pré détermination des débits de crue ou d'étiage (fréquence et durée) ;
- Prévision des étiages.

I-2. Objectif de l'étude : évaluation de l'incertitude des débits calculés

Tout scientifique sait que des résultats ne sont pleinement exploitables que s'il sont associés à une barre d'erreur, symbolisée par le signe qui suit les résultats « \pm ». On appelle cela un intervalle de confiance. Cette notion d'intervalle de confiance a un rôle essentiel dans la communication de résultats scientifiques. Pour certains, cette valeur qui borne le résultat a autant d'importance, voire plus, que le résultat lui-même. Chaque prédiction devrait être associée à une incertitude (Beven, 1996).

Les résultats sont et resteront toujours nécessairement empreints d'une marge d'incertitude, tout particulièrement en ce qui concerne l'estimation des risques à venir. La description de l'état des connaissances devrait donc inclure celle des incertitudes. Il s'agit là d'un domaine où les possibilités de malentendu entre les rédacteurs et les lecteurs sont particulièrement importantes (Petit, 2005).

Enfin, il faut reconnaître qu'avec les techniques de mesure actuellement disponibles à l'hydrologiste, il y a peu d'espoir d'obtenir des descriptions de données d'entrée très précises.

II- CADRE DE LA RECHERCHE : LA MODÉLISATION

PLUIE-DÉBIT

II-1. Les Modèles : classification et exemples

Les classifications de modèles que l'on trouve dans la littérature sont nombreuses et peuvent être fonction de plusieurs caractéristiques (Oudin, 2004). Pour ironiser, Perrin (2002) disait qu'il y a presque autant de classifications que d'hydrologues. Le tableau 2 propose de classer les différents types de modèles selon leurs caractéristiques (Clarke, 1973 ; Ambroise, 1998 ; Andréassian, 2002 ; Perrin, 2002 ; Oudin, 2004).

Caractéristique	Type de modèle
Selon la nature de la variable	Modèle déterministe (variables non aléatoires)
	Modèle stochastique (variables aléatoires)
Traitement de l'espace (Homogénéité du bassin)	Modèle global
	Modèle semi-distribué (ou semi-global)
	Modèle spatialisé
Fondement théorique et approche de développement	Modèle empirique
	Modèle conceptuel
	Modèle à base physique

Tableau 2 : Exemple de caractéristiques des modèles pluie-débit

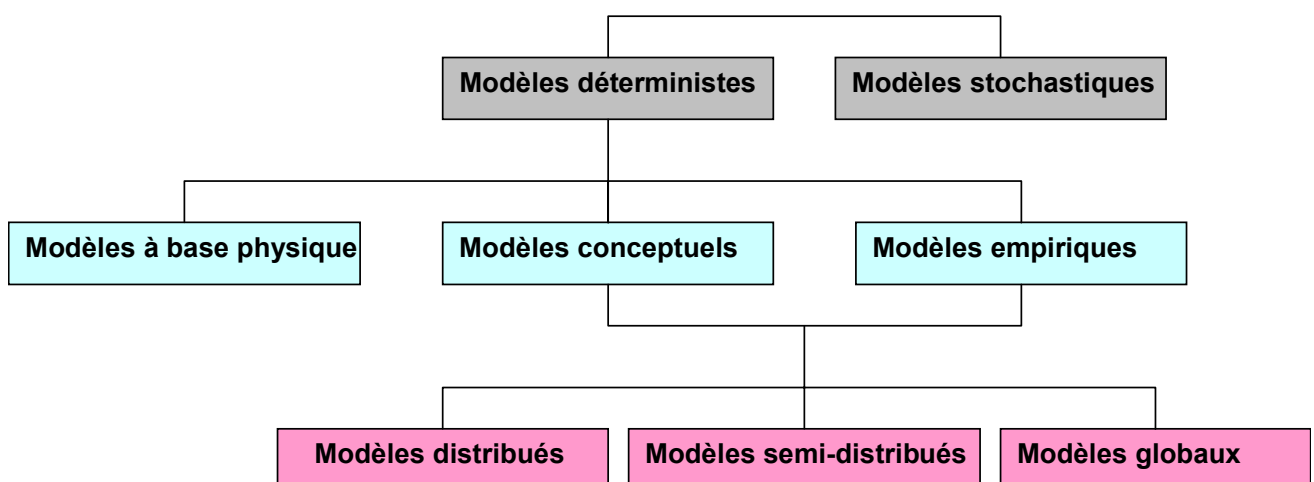


Figure 1 : Relation entre les modèles pluie-débit (Gaume, 2000)

Les données de pluie (P), de débit (Q) ainsi qu'une partie des données d'évapotranspiration potentielle (E) sont fournies au pas de temps journalier.

Nous allons présenter trois modèles fréquemment utilisés au Cemagref afin d'illustrer ce chapitre. Ces trois modèles sont tous déterministes et globaux. GR4J est empirique, alors que IHACRES et TOPMO sont conceptuels.

II-1.a. TOPMO

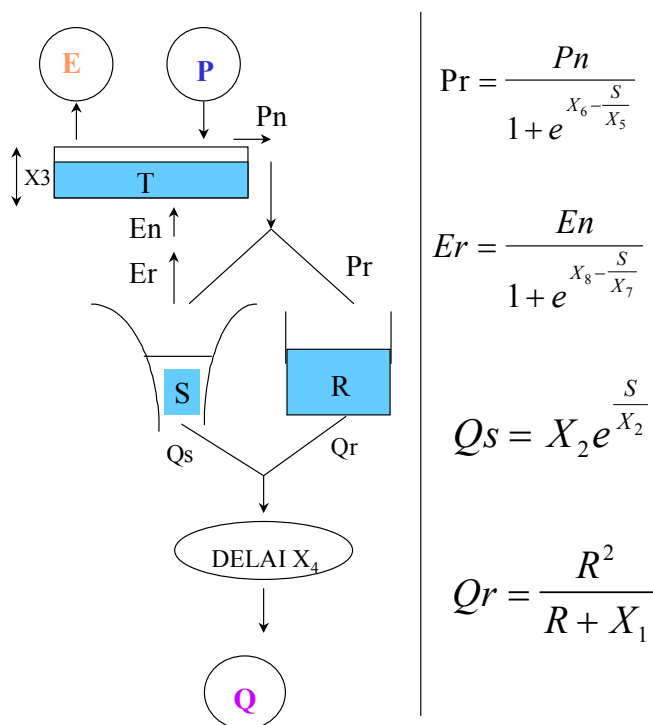


Figure 2 : Représentation schématique de TOPMO

TOPMO est une variante du modèle TOPMODEL tel qu'il est décrit par Beven et Kirkby (1979). TOPMO est un modèle ne dépassant pas 8 paramètres libres

P et E sont reçues par un réservoir d'interception dont la capacité maximale est le paramètre X3.

Si ce réservoir déborde, il s'ensuit un écoulement Pn. Si au contraire, le contenu du réservoir ne permet pas de satisfaire la demande potentielle E, la différence En est reportée sur le réservoir S qui joue un rôle central dans le modèle. L'écoulement Pn est réparti selon le niveau du réservoir S entre ce dernier et un réservoir de routage quadratique R (paramètres X5 et X6). La demande En n'est satisfaite qu'à hauteur de Es en fonction du remplissage de S (paramètres X7 et X8). S se vidange selon une loi de type exponentiel, dépendant du paramètre X2, pour donner Qs. Le réservoir de routage R se vidange selon une loi quadratique, dépendant du paramètre X1, pour donner Qr. Un retard X4 est appliqué à la somme de Qs et de Qr pour donner le débit de sortie Q.

II-1.b. IHACRES

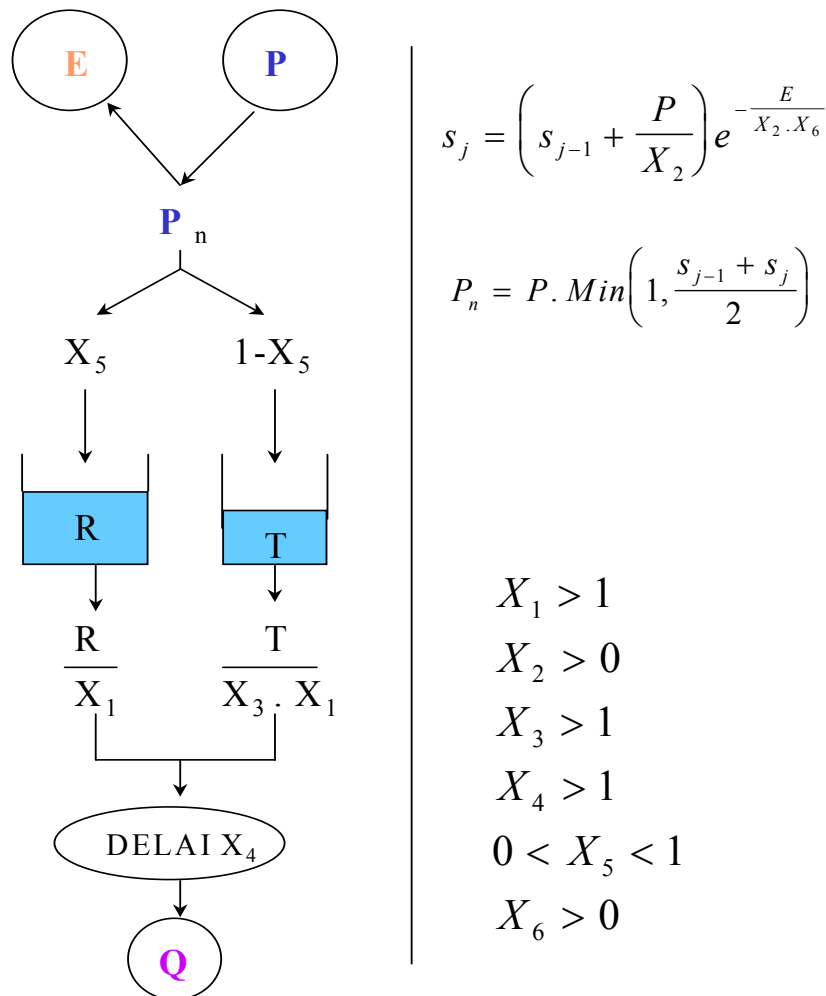


Figure 3 : Représentation schématique de IHACRES

Le modèle original (Jakeman et al.1990) n'ayant que six paramètres, ce modèle a été peu modifié par rapport au modèle initial. Un changement a toutefois été nécessaire pour pouvoir substituer l'évapotranspiration potentielle E à la température journalière utilisée dans le modèle original. Dans ces conditions on peut décrire le fonctionnement du modèle de la façon suivante : Un indice de saturation (s) définit la quantité Pr orientée vers l'écoulement, en fonction d'un paramètre X2. La demande évaporatoire est satisfaite en fonction de ce même indice s (à l'aide d'un paramètre X6). Pr est partagée en deux parties (fractions X5 et 1 - X5), chacune transitant par un réservoir linéaire, ce qui permet de différencier l'écoulement généré par le bassin en une composante lente (paramètre X1) et une composante rapide (paramètre X1 * X2, avec X2 > 1). Les débits en sortie de ces deux réservoirs subissent un retard X4 avant de former le débit à l'exutoire. Ces indications sont reprises dans la figure 3.

II-1.c. GR4J

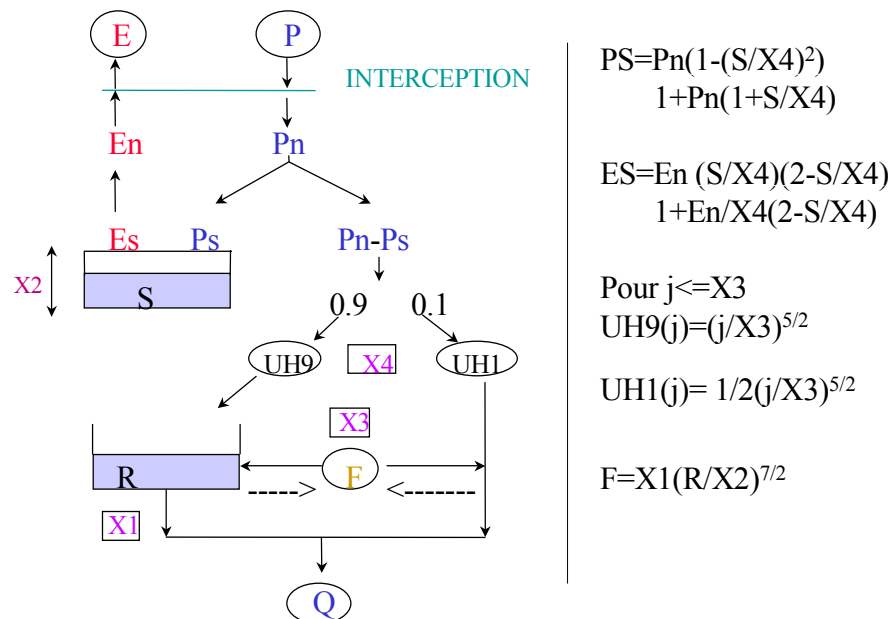


Figure 4 : Représentation schématique de GR4J

Développé au Cemagref et au Cergrene (Edijatno et al.1999), GR4J sera utilisé dans ce présent document pour quantifier les incertitudes.

Le modèle n'a que 4 paramètres à caler (X_1 à X_4). P et E agissent sur un réservoir d'interception de capacité nulle. A sa sortie, on a donc soit une pluie nette P_n , soit une évapotranspiration nette E_n . L'évaporation réelle s'effectue dans un réservoir S de X_2 mm de capacité et dépend du taux de remplissage de ce réservoir de façon quadratique. Une fraction P_r de P_n dépendant de façon presque quadratique du remplissage de S est orientée vers l'écoulement. Le complément contribue au remplissage de S . La fraction P_r est partagée de façon fixe (90% et 10%) en deux composantes qui sont transformées par deux hydrogrammes unitaires de même temps de montée (paramètre X_3). L'une des composantes (celle correspondant à 90% de P_r) traverse un réservoir R qui se vidange selon la puissance $7/2$ de son taux de remplissage et dont la capacité maximale à 1 jour est le paramètre X_1 . Le réservoir R est soumis à un flux, entrant ou sortant, qui dépend de la puissance cinq de son taux de remplissage (le paramètre X_1 en est le coefficient multiplicatif). Ce flux est également additionné algébriquement à la deuxième composante et l'on ne retient le résultat de cette addition que s'il est positif. Le flux de cette deuxième composante ainsi modifié est alors ajouté à l'écoulement de vidange du réservoir R pour donner le débit à l'exutoire. La figure 4 illustre et précise quantitativement ce qui précède. Les valeurs optimisées sont des transformées des paramètres.

II-2. Les variables

En hydrologie on a rarement la possibilité de réaliser des expériences à des échelles de temps et d'espace compatibles avec celles du bassin versant ou de reproduire en laboratoire l'ensemble des phénomènes complexes qui ont lieu. Donc les données disponibles pour la modélisation sont constituées des séries historiques des mesures de précipitation, de débit et de température (Nascimento, 1995).

Les données traitées par les modèles sont la pluie, l'évapotranspiration potentielle (ETP), et les débits. Elles proviennent de METEO-France pour les pluies (banque PLUVIO) et pour les ETP, et auprès du Ministère de l'Environnement pour les débits (banque HYDRO) (ces mêmes séries de données doivent être concomitantes et disponibles sur la même période d'observation).

Les seules données d'entrées, variables jour après jour, sont les précipitations (P en mm) et quelquefois les évapotranspirations potentielles (E en mm). Les seules données de sortie sont les débits à l'exutoire du bassin versant.

II-2.a. Les précipitations (P)

La *pluviométrie* consiste presque toujours à mesurer une hauteur de précipitations pendant un intervalle de temps. On a coutume d'exprimer les précipitations journalières, mensuelles ou annuelles en millimètres par jour, par mois ou par an, réservant généralement l'expression en intensité (mm/h) à des intervalles de temps plus courts. L'usage a également consacré la *pluviométrie* comme l'étude de la répartition et du régime des précipitations. (Hubert, 1996).

Dans notre cas, la pluviométrie (P) est exprimée en mm sur le pas de temps en question. Elle représente une hauteur d'eau tombée sur le bassin versant considéré et intégrant la distribution spatio-temporelle (Mouelhi, 2003).

II-2.b. L'évapotranspiration potentielle (E)

L'évapotranspiration Potentielle (ETP) est la quantité d'eau susceptible d'être évaporée par une surface d'eau libre ou par un couvert végétal dont l'alimentation en eau n'est pas le facteur limitant. Elle est estimée par une mesure par un évaporomètre ou déduite par des formules empiriques telles que celles de Penman (1948), Morton (1983), (Oudin, 2004). L'unité de l'évapotranspiration est le millimètre sur le pas de temps désiré (jour, mois, an...).

II-2.c. Le débit (Q)

Suite à une pluie tombée sur le bassin versant, tout un mécanisme complexe d'interactions de phénomènes hydrologiques et climatologiques se met en œuvre. Cela se traduit par une circulation, directe ou indirecte, de l'eau à travers le sol ou le sous-sol en passant par les cours d'eau secondaires puis principaux, et finissant par l'exutoire du bassin versant. A l'échelle globale (celle du bassin versant), le débit à l'exutoire traduit la réaction de tout le bassin suite à une pluie. Il est mesuré par différentes techniques hydrométriques, par exemple : le jaugeage par moulinet ou le jaugeage chimique (Mouelhi, 2003). Ce débit (Q) est exprimé en mètre cube par seconde (m^3/s). Il convient d'uniformiser les dimensions des différentes variables $\{(Q, m^3/s) ; (P, mm/pas \text{ de temps}) \text{ et } (E, mm/pas \text{ de temps})\}$.

Pour ce faire, deux possibilités sont envisageables :

✓ **Uniformiser toutes les variables en m^3/s**

Dans ce cas, il est nécessaire d'intervenir au niveau des deux variables (P et E) en multipliant par la surface du bassin. Cette façon de procéder est très peu usuelle. D'autre part, les valeurs des débits sont très variables d'un bassin à un autre, ce qui peut poser des problèmes lors du calage des modèles (Mouelhi, 2003).

✓ **Uniformiser toutes les variables en mm/pas de temps**

Seule une opération est effectuée sur la variable débit en divisant par la superficie du bassin versant considéré. De ce fait, la diversité entre les valeurs de débit d'un bassin à un autre sera plus faible. Cette méthode permet des comparaisons aisées (Mouelhi, 2003).

II-3. Les Paramètres

Les paramètres correspondent soit à des grandeurs physiques (quantifiables sur le terrain), soit à une valeur déterminée par une optimisation (non quantifiable sur le terrain) (Perrin, 2002).

La loi de Darcy est un exemple de modèle empirique global. Le processus sous-jacent, écoulement d'eau dans un réseau complexe de pores, ne peut être décrit finement. La conductivité hydraulique d'un sol n'est pas accessible directement à la mesure et doit être évaluée à partir de tests d'infiltration (Gaume, 2000).

Les paramètres des modèles globaux n'ont généralement pas de signification physique et ne peuvent pas être, par conséquent, mesurés sur le terrain. Ils doivent être déterminés mathématiquement par un processus d'optimisation ou de calage. Le choix d'une telle méthode repose sur son efficacité, c'est-à-dire sa capacité à fournir des jeux

de paramètres permettant au modèle de donner des simulations satisfaisantes. (Perrin, 2002)

Les paramètres permettent à l'utilisateur d'adapter le modèle aux caractéristiques particulières du bassin considéré, de la même façon que les dimensions d'un patron sont adaptées par le tailleur aux mensurations de la personne à habiller (Perrin, 2002).

II-4. Le calage

Le calage consiste donc à sélectionner le jeu de paramètres d'un modèle de façon à ce que celui-ci simule le comportement hydrologique du bassin versant de la meilleure façon possible (Madsen, 2000). Le calage se fait avec une année de mise en route (fig.2), c'est-à-dire une année pendant laquelle les erreurs du modèle ne sont pas comptabilisées dans le critère d'adéquation permettant de juger la performance du modèle calé.

Le critère utilisé est le critère des moindres carrés des erreurs, présenté sous la forme adimensionnelle du critère de Nash borné (cf. chapitre II-5.b.). L'opération de calage consiste à trouver des valeurs des paramètres du modèle qui minimise l'erreur de modélisation sur la période considérée. Une période de 5 ans généralement est utilisée pour caler le modèle.

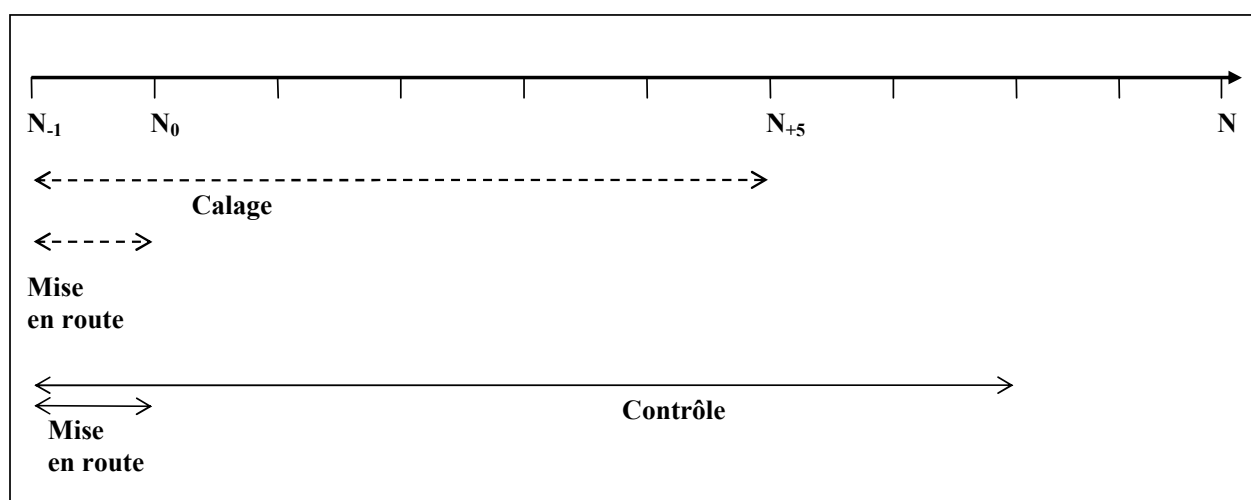


Figure 5 : Schéma de la chronologie du calage

→ Pourquoi utiliser une période de 5 ans pour le calage ?

Cette période de 5 ans représente, généralement, le délai moyen entre le projet et la réalisation d'ouvrage. En effet un des buts de la modélisation est de prédire les débits de rivières dans les années à venir sur des bassins versants non jaugés et de prédéterminer la fréquence de retour des crues. A partir du moment où il est décidé la construction d'ouvrage sur un bassin versant non jaugé, l'utilisateur du modèle (le

gestionnaire) ne possèdera que 5 années en moyenne de données pour faire fonctionner le modèle.

Les écarts entre ce que le modèle prédit et ce que l'on observe peuvent être calculés soit sur les débits soit sur les débits transformés par la fonction racine carrée ou la fonction logarithmique (dans ce dernier cas on ajoute au débit le quarantième du module, estimé sur les données observées, pour éviter le cas du logarithme de débit nul)

II-4.a. Le calage manuel

Pour cette méthode de calage, c'est l'utilisateur du modèle qui recherche la valeur des paramètres qu'il va utiliser pour faire tourner le modèle.

On pourrait comparer le calage manuel au réglage d'un oscilloscope où l'on doit superposer deux courbes (l'une correspondant au débit observé et l'autre au débit calculé) en faisant varier manuellement 4 boutons (correspondant 4 paramètres). Cette technique laisse une part de subjectivité due à l'utilisateur.

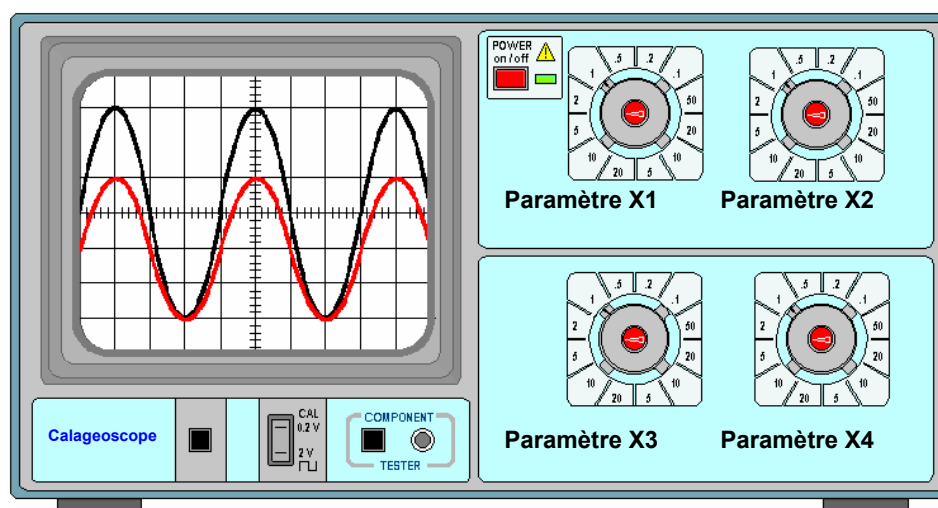


Figure 6 : Représentation schématique du calage manuel

Les différents critères de qualité qui accompagnent le calage permettent d'évaluer la capacité du modèle à bien représenter quelques caractéristiques importantes de la série des débits.

II-4.b. Le calage automatique

Pour enlever la part de subjectivité due à l'utilisateur pendant le calage manuel, un algorithme mathématique peut déterminer les paramètres par essais successifs, en tâtonnant. Cet algorithme est capable de déterminer les paramètres les plus appropriés pour que les débits calculés soient les plus proches possibles de la réalité. Mais cette

méthode, comme le calage manuel, reste soumis au risque de rencontrer un optimum secondaire.

Une majorité des utilisations de modèle reposent sur ces techniques d'estimations mathématiques des paramètres. Le mode de calage automatique est préféré en général à des méthodes manuelles ou semi-manuelles.

II-5. Critère d'ajustement (ou fonction objectif)

On désigne ainsi le critère numérique qui permet de juger de l'adéquation d'un modèle, c'est-à-dire de la comparaison entre la chronique réelle des débits et celle simulée par le modèle (Edijatno, 1987). Le but du calage est de minimiser cette fonction objectif.

Le critère représente la somme des erreurs quadratiques (ou somme des écarts élevés au carré), permettant ainsi de juger les performances du modèle. On élève les erreurs au carré pour avoir la valeur absolue des écarts et donner moins d'importance aux petits écarts.

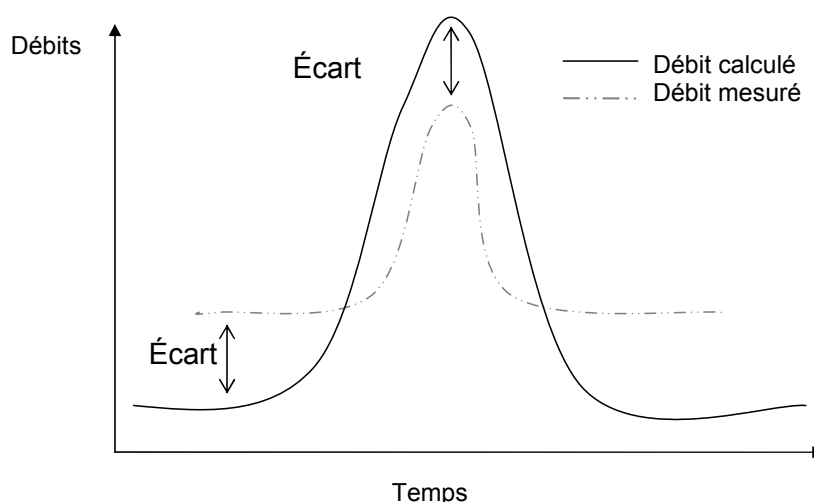


Figure 7 : Représentation schématique des écarts entre le débit observé et le débit calculé

Les critères d'ajustement peuvent être calculés (entre autres) sur :

- les débits Q *Renseigne sur les débits élevés*
- la racine carrée du débit \sqrt{Q} *Renseigne sur les débits moyens*
- le logarithme du débit $\log Q^1$ *Renseigne sur les faibles débits*

¹ Dans ce cas on ajoute au débit le quarantième du module, estimé sur les données observées, pour éviter une impossibilité en cas de débit nul

II-5.a. Le critère de Nash (Nash et Sutcliffe, 1970)

Le critère d'optimisation le plus communément utilisé, est le critère la somme des carrés des écarts, qui peut-être présenté sous la forme adimensionnelle du critère de Nash (qui est limité supérieurement à 100%, mais n'est pas borné inférieurement)

$$\text{Nash} = \left[1 - \frac{\sum (\sqrt{Q_{obs}} - \sqrt{Q_{cal}})^2}{\sum (\sqrt{Q_{obs}} - \sqrt{Q_{obs}})^2} \right] \times 100$$

Un Nash dont la valeur est nulle signifie que la qualité de la simulation équivaut à celle du modèle de référence.

\sqrt{Q} Permet de ne pas donner trop de poids aux erreurs associées aux événements de crue.

II-5.b. Le critère de Nash borné

Une légère modification du critère de Nash peut être apportée dans le but de simplifier la comparaison des critères de chaque bassin. Tout au long de ce document, nous allons utiliser le critère de Nash borné (aussi appelé le C2M développé par Mathevet et al. en 2004).

$$\text{Nash}_{\text{borné}} = \frac{\text{Nash}}{200 - \text{Nash}}$$

II-6. Le contrôle de la performance du modèle

Après l'étape du calage, il est nécessaire de vérifier la représentativité des paramètres calés. C'est l'opération de contrôle qui consiste à tester le modèle dans des conditions différentes de celles du calage (Chahinian, 2004). C'est l'opération 'vérité' qui permet de juger réellement la fiabilité et la robustesse du modèle. Les performances à retenir (au bénéfice du modèle étudié) sont essentiellement celles obtenues au cours des différentes sous-périodes de contrôle. On peut également réaliser un contrôle période par période et observer les valeurs prises par les critères présentés plus haut.

Pour tester la validité de l'ajustement obtenu sur une série de données, il est nécessaire de juger des performances du modèle sur une période différente de celle ayant servi à ajuster les paramètres du modèle, c'est ce qui est appelé le contrôle. Au contrôle on est plus proche des conditions d'utilisation opérationnelle qu'au calage.

Le degré de réalisme des paramètres ainsi identifiés peut être évalué par le niveau de performance du modèle sur la période de contrôle.

La qualité du modèle se mesure autant à la valeur du critère d'ajustement qu'à la différence de résultat entre la valeur prise par la fonction sur la période de calage et celle prise sur une période de contrôle différente de la précédente. (Edijatno, 1987)

II-7. Méthodes d'estimation des incertitudes des débits calculés

Nous allons présenter dans ce chapitre, les deux méthodes les plus classiques d'estimation des incertitudes, utilisées dans la littérature, afin de mettre en évidence leurs limites, puis dans la section III-2, nous présenterons la nouvelle méthode développée dans le cadre de ce DEA.

II-7.a. *Méthode dérivée de Monte-Carlo fondée sur une approximation linéaire des écarts-types des paramètres du modèle*

La méthode de Monte-Carlo a été développée par Stanislaw Ulam et Nicholas Metropolis, en 1949. L'appellation « Monte Carlo » est due à Metropolis, inspiré de l'intérêt de Ulam pour le poker, car Monte Carlo est un grand centre de casinos, et a pour origine les similarités avec les jeux de hasard. Le travail de Ulam et de Metropolis consistait à simuler les problèmes de dispersion et d'absorption de neutrons pour les matériaux fissibles pour la fabrication de la bombe atomique.

Cette méthode a pour but de quantifier les incertitudes des paramètres, en reconstituant la distribution des variables de sortie. Elle nécessite des simulations répétées en tirant aléatoirement la valeur de la variable d'entrée. C'est une méthode facile (mais longue) à mettre en oeuvre qui n'est pas trop complexe comparée à des méthodes mathématiques. Cependant en hydrologie, cette méthode donne des résultats avec des bornes d'erreurs très faibles et quasi inutilisables en pratique (Tuffin, 1997). Dans ce présent document, cette technique a été appliquée à l'hydrologie. Pour illustrer cette technique nous allons prendre comme exemple un modèle simple à deux paramètres connue de tous : le modèle linéaire simple.

$$y_c = ax + b \quad \text{Ceci est ce que l'on calcule}$$

$$y_0 = ax + b + e_0 \quad \text{Ceci est la réalité}$$

Où y est la donnée de sortie (Q) et x la donnée d'entrée.

Où a et b sont les paramètres du modèle.

Où e_0 est l'erreur associée aux paramètres,

Classiquement les incertitudes se calculent en faisant la différence de y_0 et de y_c :

$$e_0 = y_0 - y_c$$

On ne peut pas faire de différence car l'erreur peut être plus grande que le débit calculé et prédire un débit négatif. **Ex : $Q=10\text{mm/j} \pm 11\text{mm/j}$** . On préférera donc utiliser les rapports.

On détermine les paramètres a et b en calant ce modèle linéaire avec des valeurs d'entrée (x_1, y_1) , (x_2, y_2) , (x_i, y_i) , ..., (x_n, y_n) .

Les écarts-types des paramètres a et b sont calculés. De nouveaux jeux de paramètres a' et b' sont déterminés autour de ces écarts-types, puis de nouveaux débits sont ensuite calculés à leur tour avec les nouveaux paramètres a' et b' . Par cette méthode dérivée de Monte-Carlo nous pouvons calculer ensuite le rapport des y_c (calculés avec les paramètres a et b) avec les y'_c (calculés avec les paramètres a' et b') et obtenons des incertitudes I_c .

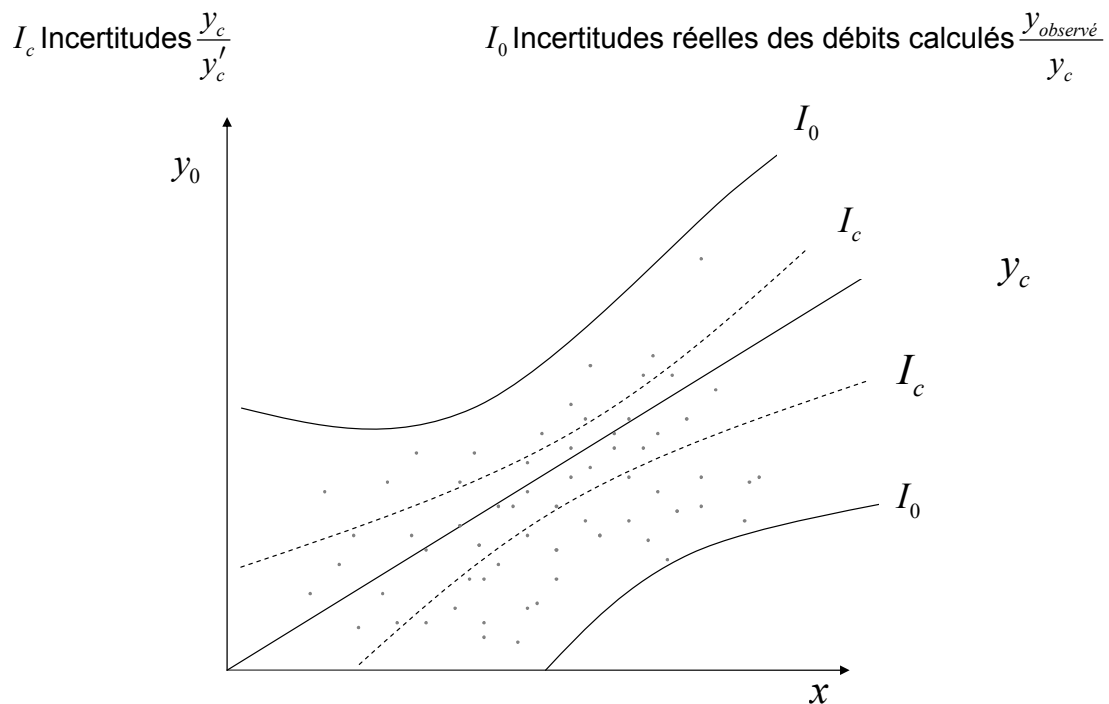


Figure 8 : Représentation des incertitudes sur les débits ayant subi une approximation linéaire des écarts-types des paramètres du modèle I_c et des incertitudes des débit calculé I_0

Comme dans le modèle linéaire, on peut calculer les écarts-types σ^0 des 4 paramètres de GR4J $[x_1^0, x_2^0, x_3^0, x_4^0]$ déterminés au calage. En supposant une loi multi-normale, on peut générer d'autres jeux de paramètres compris dans ces écarts-types :

$$\begin{array}{l}
 [x_1^0, x_2^0, x_3^0, x_4^0] \\
 [x_1^1, x_2^1, x_3^1, x_4^1] \\
 \vdots \\
 [x_1^i, x_2^i, x_3^i, x_4^i]
 \end{array}
 \left. \vphantom{\begin{array}{l} [x_1^0, x_2^0, x_3^0, x_4^0] \\ [x_1^1, x_2^1, x_3^1, x_4^1] \\ \vdots \\ [x_1^i, x_2^i, x_3^i, x_4^i] \end{array}} \right\} \begin{array}{l} \text{Jeu déterminé au calage [P0]} \\ \\ \\ \text{Jeux de paramètres [P1] à [Pi] répartis autour de [P0]} \end{array}$$

Avec le jeu de paramètre [P0] nous obtenons des débits $Q_1^0, Q_2^0, \dots, Q_n^0$

Avec le jeu de paramètre [P1] le modèle génère des débits $Q_1^1, Q_2^1, \dots, Q_n^1$

⋮

Avec le jeu de paramètre [Pi] le modèle génère des débits $Q_1^i, Q_2^i, \dots, Q_n^i$

Il va donc être généré un grand nombre de jeu de paramètres [Pi] répartie autour du jeu de paramètres déterminés au calage [P0] en fonction de σ^0 . A la différence de la méthode proposée d'estimation des erreurs (cf. chapitre précédent) la méthode fondée sur une approximation linéaire des écarts-types des paramètres du modèle ne calcule pas le rapport $\frac{Q_{\text{observé}}}{Q_{\text{calculé}}}$, mais le rapport $\frac{Q^0}{Q^i}$ c'est-à-dire le rapport des débits obtenus avec les paramètres du calage sur les débits obtenus avec les paramètres répartis autour des paramètres obtenus au calage. Le rapport que nous observons ne reflète pas la réalité car il ne compare pas le débit calculé avec le débit observé sur le terrain.

II-7.b. La méthode GLUE : l'échantillonnage par ordre d'importance

La méthode *Generalized Likelihood Uncertainty Estimation* (GLUE), développée par Beven et Binley en 1992, consiste à tirer un échantillon aléatoire de la fonction de vraisemblance et à pondérer les résultats par ordre d'importance (le meilleur résultat aura un poids maximal et le pire un poids minimal). La fonction de vraisemblance pourra être remise à jour lorsque plus de données seront disponibles (Chahinian, 2004)

Cette méthode gère la multiplicité possible des jeux de paramètres acceptables en les classant selon une mesure subjective traduisant l'ajustement aux données

observées, tel le critère de Nash et Sutcliffe (1970). Le succès de cette méthode dépend du choix de la fonction de vraisemblance et de la taille de l'échantillon (Kuczera et Parent, 1998). Elle est onéreuse en temps de calcul. GLUE rejette le principe du paramètre optimal unique. La méthode GLUE n'est pas réellement une méthode de calage à proprement dit : elle ne détermine pas un jeu de paramètres qui l'emporte sur un autre. En réalité un grand nombre de jeux de paramètres sont retenus, obtenant ainsi non pas un hydrogramme à comparer avec l'hydrogramme des débits mesurés, mais « n » hydrogrammes, ce qui permet le calcul d'un intervalle de confiance.

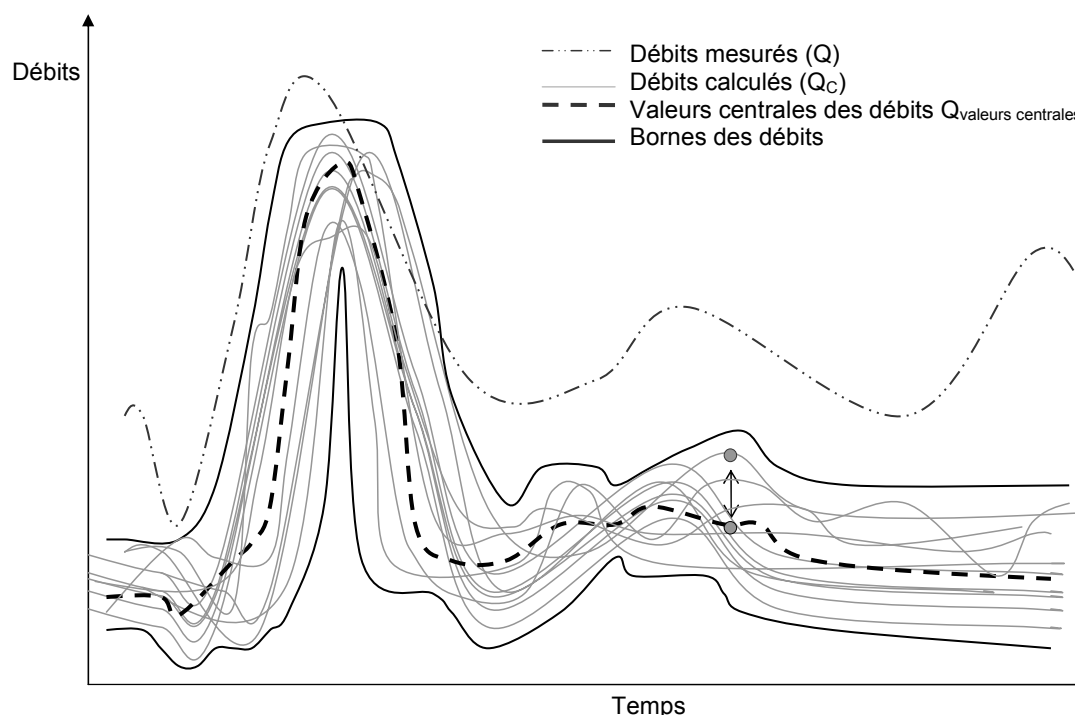


Figure 9 : Représentation schématique des incertitudes avec la méthode GLUE

Comme avec la méthode fondée sur les approximations linéaires (voir chapitre II-7.b), les incertitudes ne sont pas calculées au moyen du ratio $\frac{Q_{\text{observé}}}{Q_{\text{calculé}}}$, mais de cette manière :

$$\frac{Q_{\text{valeurs centrales}}}{Q_{\text{calculé}}}$$

Les valeurs centrales de débits (représentées en pointillés sur la figure ci-dessus) sont les valeurs médianes comprises entre la borne supérieur des débits et la borne inférieur.

Ces incertitudes sont représentées par la double flèche sur le schéma. Là aussi les débits calculés par le modèle ne sont pas comparés aux débits mesurés.

II-7.c. La méthode proposée

Les erreurs sont traditionnellement représentées comme des différences $Q_{\text{observé}} - Q_{\text{calculé}}$ comme dans le critère de Nash. Mais cette représentation n'est plus acceptable pour une utilisation pratique, car une même erreur absolue peut-être mineure pour une pointe de crue et excessive pour un étiage. Il est donc plus approprié de calculer les erreurs en faisant le rapport $\frac{Q_{\text{observé}}}{Q_{\text{calculé}}}$. Nous proposons alors de quantifier les erreurs des débits calculés par le modèle en faisant simplement le rapport $\frac{Q_{\text{observé}}}{Q_{\text{calculé}}}$.

III- MÉTHODES ET RÉSULTATS

III-1. Zones d'étude

La présente étude va porter sur un échantillon représentatif de 50 bassins versant français (49 en métropole et 1 en Corse) choisis en fonction de la longue durée d'observation des débits et repartis sur tout le territoire.

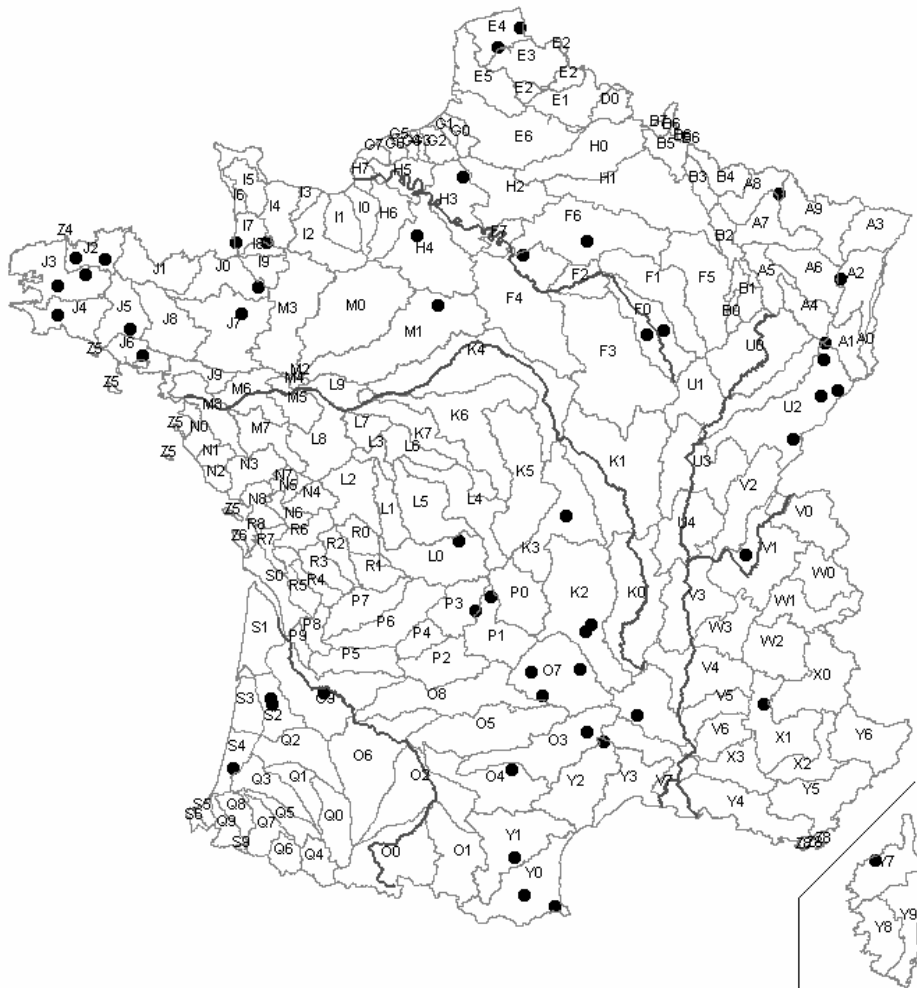


Figure 10 : Répartition des 50 bassins versants utilisés lors de l'étude

Code Hydro	n°	Code Hydro	n°	Code Hydro	n°	Code Hydro	n°	Code Hydro	n°	Code Hydro	n°
A1202020	1	H8012010	10	J4224010	19	O3165010	28	S2235610	37	Y0115410	46
A2702010	2	H9402010	11	J5613010	20	O4704030	29	S4214010	38	Y0466010	47
A8712010	3	I7001010	12	J6407120	21	O7145220	30	U2022010	39	Y1514010	48
E4035710	4	I7913610	13	J7060620	22	O7265010	31	U2142010	40	Y2002010	49
E4905710	5	J0014010	14	K2316210	23	O7625210	32	U2215020	41	Y7615010	50
H0203020	6	J2233010	15	K2365510	24	O9196210	33	U2345030	42		
H0321030	7	J2614010	16	K3374710	25	P1114010	34	V1454310	43		
H4333410	8	J3601810	17	L0244510	26	P3614010	35	V5425210	44		
H5412010	9	J3834010	18	M1024810	27	S2224610	36	X1045820	45		

III-2. Représentation graphique des incertitudes

Nous allons dans ce chapitre décrire étape par étape le processus, qui va nous emmener à une représentation graphique des incertitudes.

III-2.a. Hydrogrammes de crues

Dans cette section, nous allons utiliser comme exemple le bassin versant n°8 (H4333410) où coule le Réveillon à La Jonchère. Nous allons décrire les différentes étapes qui vont nous amener à une représentation simple et pratique des résultats issus des modèles pluie-débit afin de quantifier et de visualiser graphiquement les incertitudes des débits calculés. Après l'étape de calage où les 4 paramètres du modèle ont été déterminés, GR4J calcule les débits, qui sont comparés aux débits mesurés sur la même période. Nous utilisons pour cet exemple la période de calage 1974-1979 du bassin H4333410.

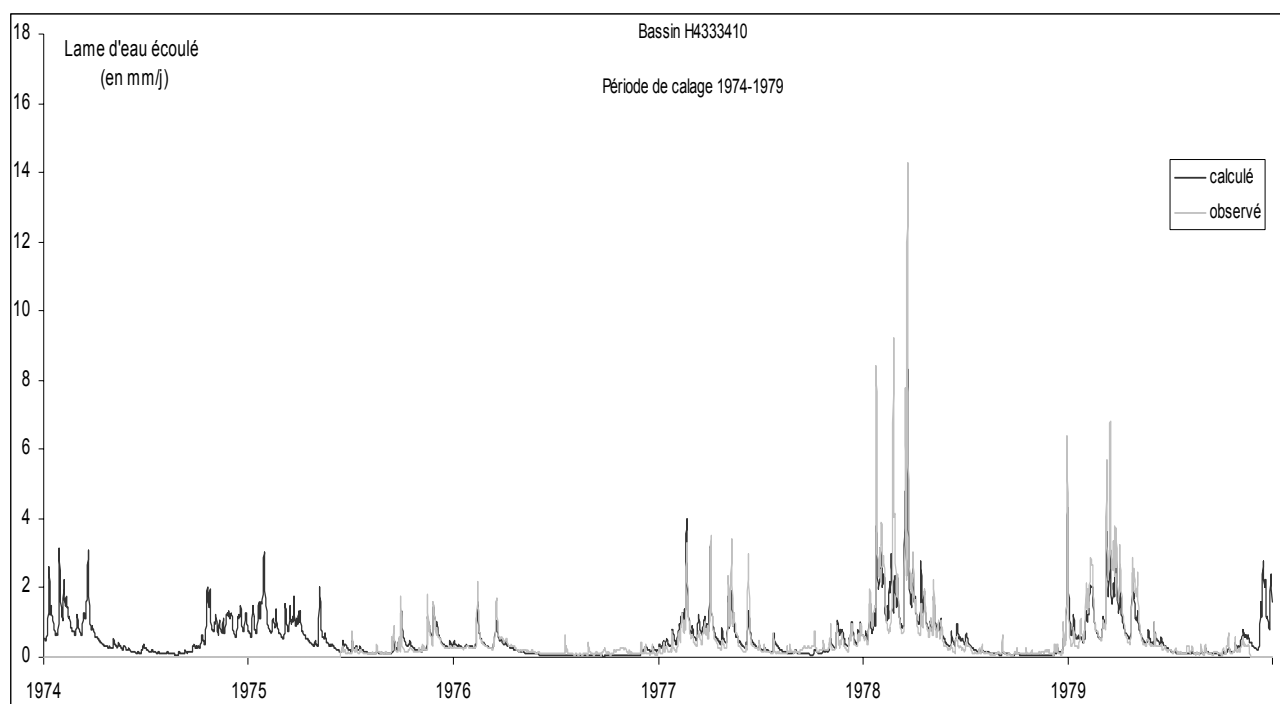


Figure 11 : Hydrogrammes de crues du bassin H4333410 sur la période de calage de 1974 à 1979 et hydrogramme calculé par GR4J

Le calage commence en 1974, avec l'année de mise en route en 1973. De 1974 à mi-1975 la banque Hydro ne dispose pas de mesures de débit sur ce bassin, ce qui explique l'absence de la courbe de débit observé sur cette courte période. Dans ce cas là le calage commence en mi-1975

III-2.b. Rapport $Q_{\text{observé}}/Q_{\text{calculé}}$ au calage sur la période de calage

Pour les raisons que nous avons développées au chapitre II-5.a, nous nous intéressons ici aux erreurs relatives représentées par les rapports $\frac{Q_{\text{observé}}}{Q_{\text{calculé}}}$ (ou $\frac{Q_{\text{obs}}}{Q_{\text{cal}}}$) et non à la différence : $Q_{\text{observé}} - Q_{\text{calculé}}$. A chaque observation (une par jour), le rapport du débit mesuré par le débit calculé de la même journée est réalisé. On obtient alors un graphique comme celui-ci.

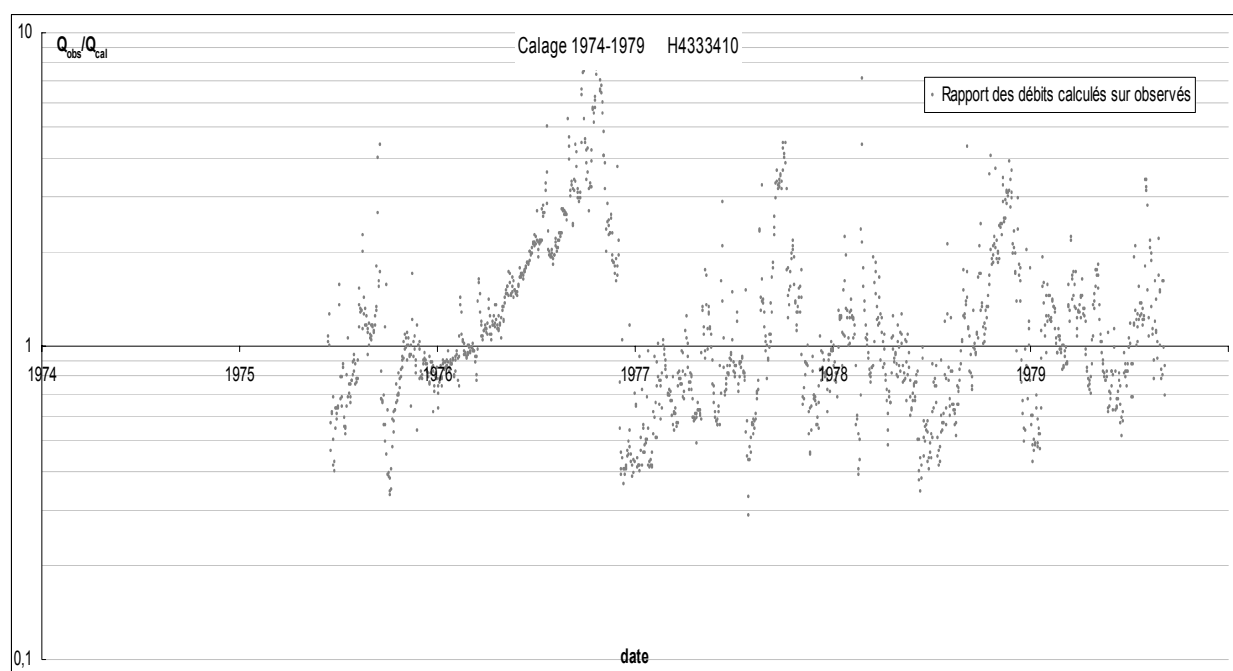


Figure 12 : Rapport des débits observés sur débits calculés pour chaque observation en fonction de la date sur la période de calage 1974-1979

La figure 12 illustre les écarts entre les deux courbes de la figure 11 du chapitre III-2.a. Nous pouvons constater en comparant ces deux figures (figure 13, page suivante) que les faibles débits sont souvent fortement sous-estimés et que les débits de pointe sont surestimés sauf pour les débits de crues exceptionnelles qui sont eux aussi sous-estimés.

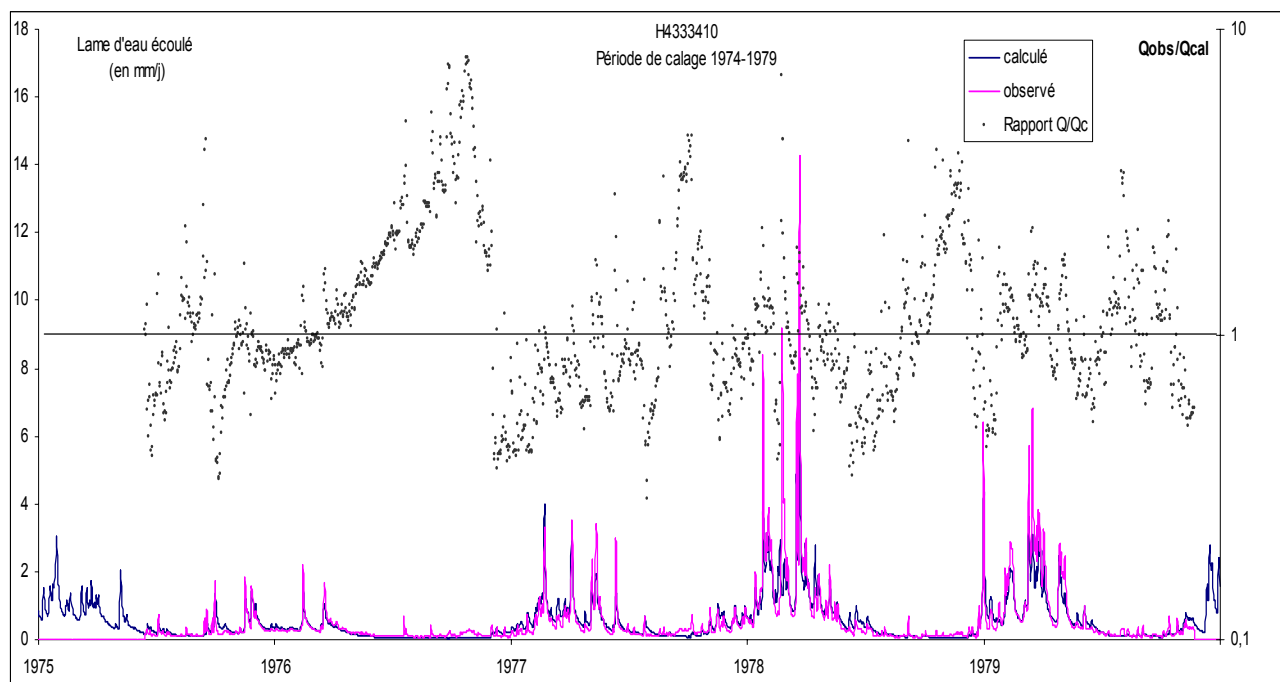


Figure 13 : Comparaison des figures 11 et 12



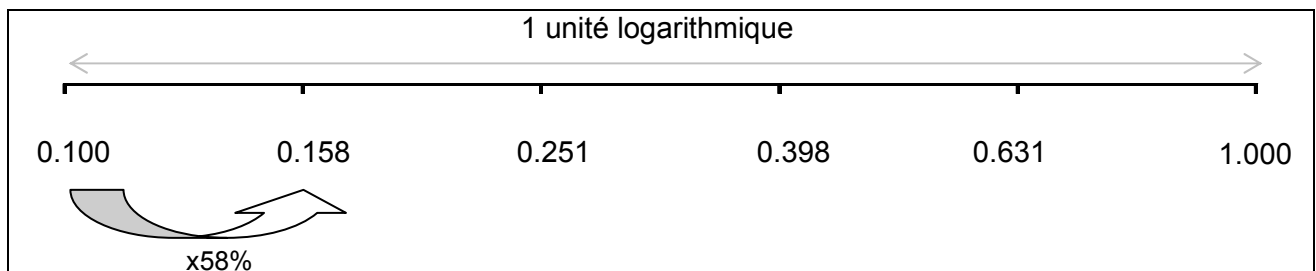
Figure 14 : Rapport Q_{obs}/Q_{cal} pour chaque observation en fonction de Q_{cal}

La figure 14 regroupe sur un fond logarithmique l'ensemble des rapports Q_{obs}/Q_{cal} en fonction des débits calculés (en abscisse).

De la même manière que les figures 12 et 13, nous pouvons effectivement observer sur cette figure que les faibles débits sont jusqu'à dix fois sous-estimés. Toutefois ce nuage de points reste difficile à interpréter. En effet les valeurs extrêmes des rapports sont peu nombreuses. Nous nous sommes donc intéressés à la dispersion des points. Pour mesurer cette dispersion nous avons utilisé les quantiles 0,15 et 0,85 ; (voir annexe au chapitre VIII-1) qui englobent 70% des effectifs autour des valeurs centrales.

III-2.c. Rapport $Q_{\text{observé}}/Q_{\text{calculé}}$ au calage sur la période 1974-1979 par tranche de débit calculé

Afin de calculer les quantiles nous avons regroupé les rapports Q_o/Q_c par tranches de débits calculés. Chaque fourchette de débit calculé est équivalente à 1/5 d'unité logarithmique. Soit $10^{0.2}$, ce qui représente une augmentation de 58% pour passer de la borne inférieure d'un intervalle à sa borne supérieure (identique à la borne inférieure du suivant).



Les résultats sont présentés sous forme d'un tableau comme ci-dessous (Tableau 2). Les débits allant de 0.01 à 100 mm/j sont donc découpés en 21 intervalles. Pour chacune de ces 21 fourchettes les quantiles (0.15 ; 0.50 ; 0.85) sont calculés s'il y a plus de 12 valeurs de débit sur la période de calcul.

Intervalle (en mm/j)	effectifs.	Quant 0,15	Quant 0,5	Quant 0,85
0.000	0.010	0		
0.010	0.016	0		
0.016	0.025	11		
0.025	0.040	94	2.273	3.266
0.040	0.063	113	1.937	2.254
0.063	0.100	184	1.049	1.376
0.100	0.158	168	0.700	1.182
0.158	0.251	180	0.532	0.854
0.251	0.398	244	0.538	0.815
0.398	0.631	235	0.545	0.735
0.631	1.000	177	0.600	0.836
1.000	1.585	105	0.727	1.031
1.585	2.512	69	0.809	1.164
2.512	3.981	34	0.862	1.110
3.981	6.310	7		
6.310	10.000	1		
10.000	15.849	0		
15.849	25.119	0		
25.119	39.811	0		
39.811	63.096	0		
63.096	100.000	0		

Tableau 3 : Extrait des résultats du bassin versant n°8 pour la période de calage 1974-1979

Le Tableau 3, contient les valeurs des rapports pour les quantiles 0,15 ; 0,50 et 0,85 pour chaque tranche de débit calculé pour la période de calage 1974-1979.

Description du Tableau 3 :

- Pour la tranche de débit comprise entre 0,016 et 0,025 mm/j, il n'y a eu que 11 valeurs mesurées de débit sur la période de calage ; les quantiles ne sont donc pas calculés.
- Pour la fourchette de débit comprise entre 0,025 et 0,040 mm/j, il y a 94 valeurs de débits sur la période de calage, les quantiles ont donc pu être calculés.
- Nous nous sommes intéressés aux quantiles 0,15 et 0,85 pour avoir une idée de la dispersion du rapport $\frac{Q_{\text{observé}}}{Q_{\text{calculé}}}$.
- 70% des 94 débits mesurés (compris entre 0,025 et 0,040 mm/j) sont compris entre 2,273 et 6,699 fois le débit calculé.

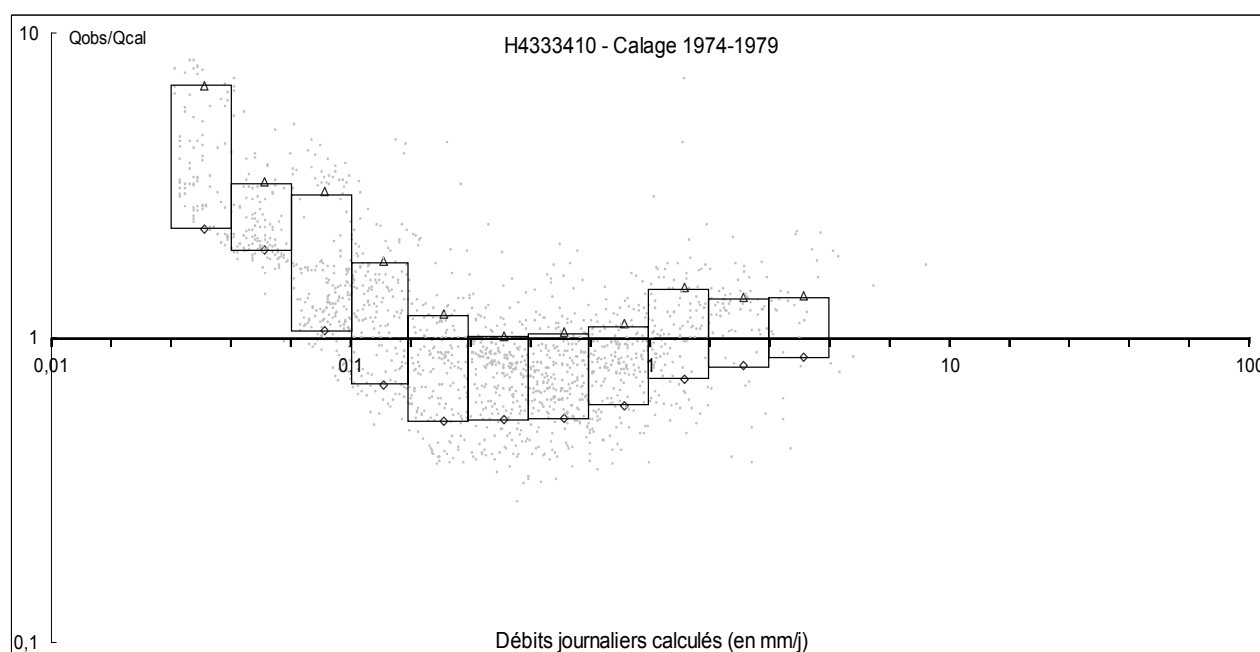


Figure 15 : Encadrement de 70% des rapports $Q_{\text{obs}}/Q_{\text{cal}}$ pour chaque fourchette de débit calculé au calage

Dans chaque case on a 70% des effectifs autour des valeurs centrales pour chacun des intervalles. La borne supérieure du rectangle correspond au quantile 0,85 et la borne inférieure correspond au quantile 0,15.

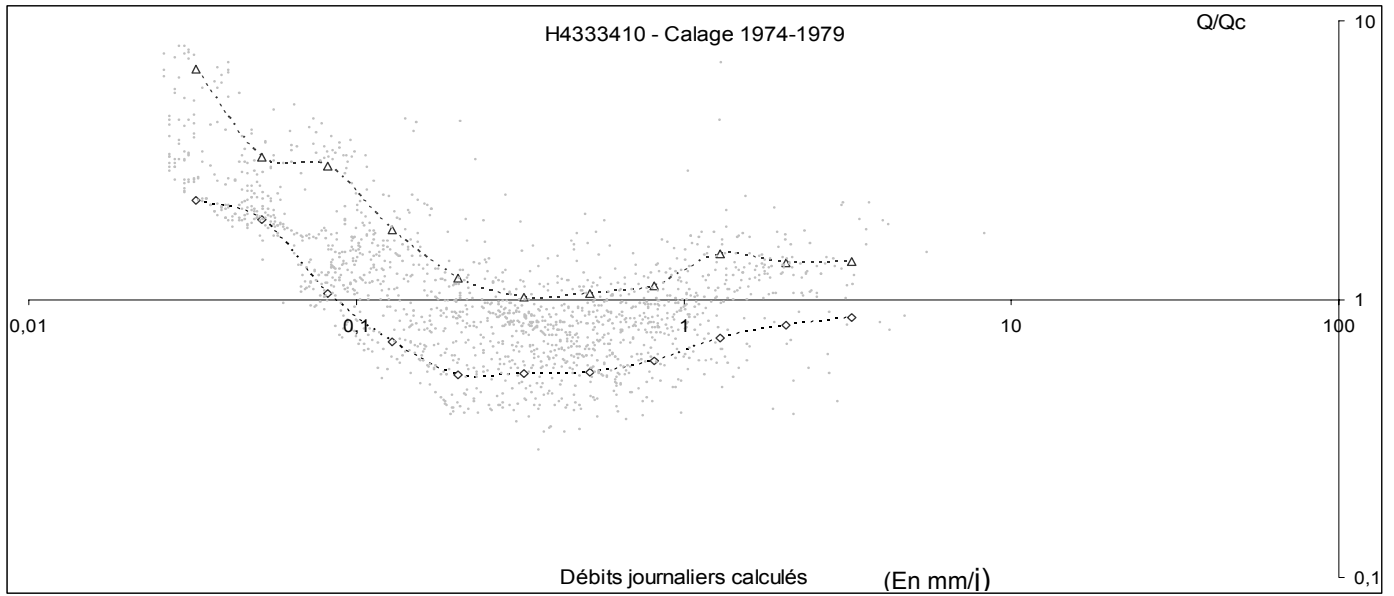


Figure 16 : Rapport Q_o/Q_c par fourchette de débit calculé par GR4J sur la période de calage 1974-1979

Enfin pour clarifier cette représentation des incertitudes, nous avons représenté la répartition des rapports Q_o/Q_c entre les deux courbes en pointillés.

III-2.d. Rapport $Q_{\text{mesuré}}/Q_{\text{calculé}}$ au contrôle sur l'ensemble de la période 1974-1992

De la même manière que dans la section III-2.c nous pouvons représenter les incertitudes des débits calculés par GR4J du bassin H4333410 sur l'ensemble de la période de 1974 à 1992 avec les paramètres déterminés au calage sur la période 1974-1979. Une note sera attribuée aux quantiles afin de mettre en évidence l'écart entre les valeurs des quantiles déterminés au calage et les valeurs des quantiles déterminés au contrôle. Cette note renseignera sur la capacité du modèle à annoncer des débits calculés sur une période de calage et sur la période de contrôle avec les mêmes incertitudes.

Intervalle (en mm/j)	effectifs.	Quant 0,15	Quant 0.50	Quant 0.85	
0.000	0.010	0			
0.010	0.016	0			
0.016	0.025	11			
0.025	0.040	118	2.267	2.896	6.089
0.040	0.063	210	1.599	2.150	3.045
0.063	0.100	533	1.067	1.483	2.119
0.100	0.158	666	0.807	1.168	1.602
0.158	0.251	732	0.603	0.794	1.134
0.251	0.398	920	0.497	0.681	0.967
0.398	0.631	990	0.460	0.654	0.920
0.631	1.000	820	0.481	0.713	0.971
1.000	1.585	613	0.560	0.826	1.207
1.585	2.512	333	0.592	0.890	1.275
2.512	3.981	126	0.547	0.953	1.343
3.981	6.310	25	0.485	1.029	1.762
6.310	10.000	4			
10.000	15.849	1			
15.849	25.119	0			
25.119	39.811	0			
39.811	63.096	0			
63.096	100.000				

Distorsion entre calage et ensemble des périodes:

Quantile 0.85 : 86.191

Quantile 0.50 : 86.376

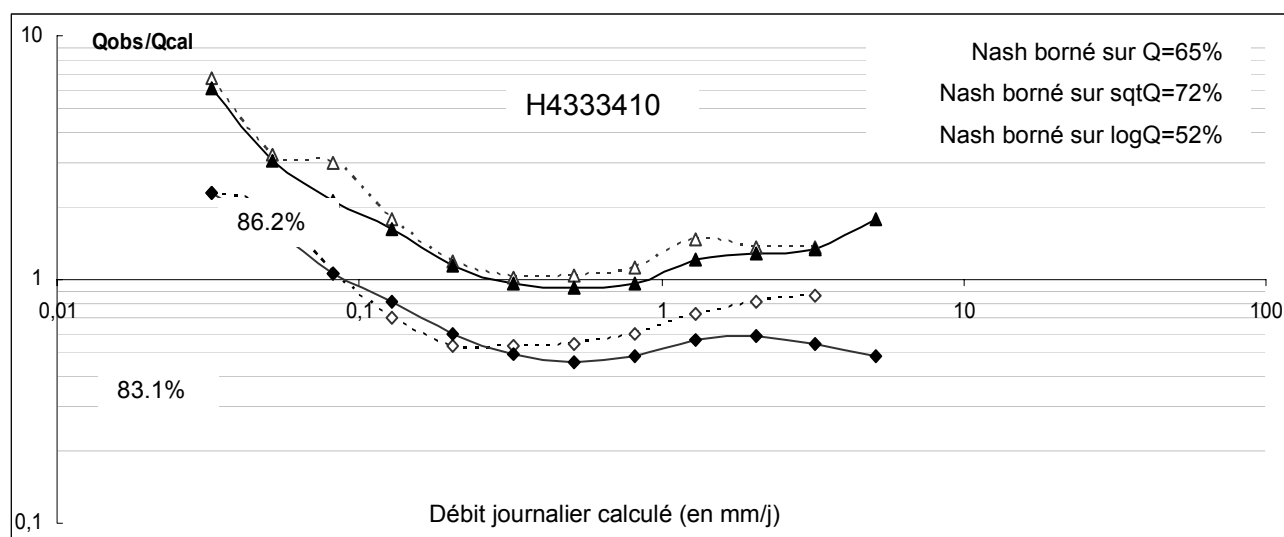
Quantile 0.15 : 83.100

Tableau 4 : Extrait des résultats du bassin versant n°25 sur l'ensemble de la période

Le Tableau 4, contient les valeurs des rapports pour les quantiles 0,15 ; 0,50 et 0,85 pour chaque tranche de débit calculé pour la période totale 1974-1992.

Les débits calculés sont comparés aux débits mesurés sur l'ensemble de la période. C'est l'opération vérité, car l'ensemble de la période est largement plus grande que la période de calage. C'est à cette étape que l'on vérifie si sur une longue période les erreurs estimées sont de même grandeur que les erreurs déterminées au calage.

Pour débits le rapport Q_{obs}/Q_{cal} du calage est comparé au rapport Q_{obs}/Q_{cal} calculé sur de l'ensemble de la période. Pour cela, nous avons reporté ces rapports sur un même graphique.



Quantiles calculés sur :

- La période de calage
- L'ensemble de la période

Figure 17 : Rapport Q_{obs}/Q_{cal} par fourchette de débit calculé par GR4J sur la période de calage 1974-1979 et sur la période totale 1974-1992

Le Nash borné donne une note sur le biais de Q_{obs} par rapport à Q_{cal} , graphiquement cela correspond à l'écart des courbes par rapport à la droite d'équation $y=1$, sur le graphique ci-dessus. Une note pour chaque quantile calculé est donnée à la distance entre la courbe au calage (---) et la courbe de contrôle sur l'ensemble des périodes (—). Cette note est de 86.2% pour les quantiles 0.85 et de 83.1% pour les quantiles 0.15.

Cette note sera égale à 100% quand les deux courbes coïncideront. Elle nous permet de mesurer la fiabilité de notre estimation des incertitudes sur la base de l'opération de calage du modèle

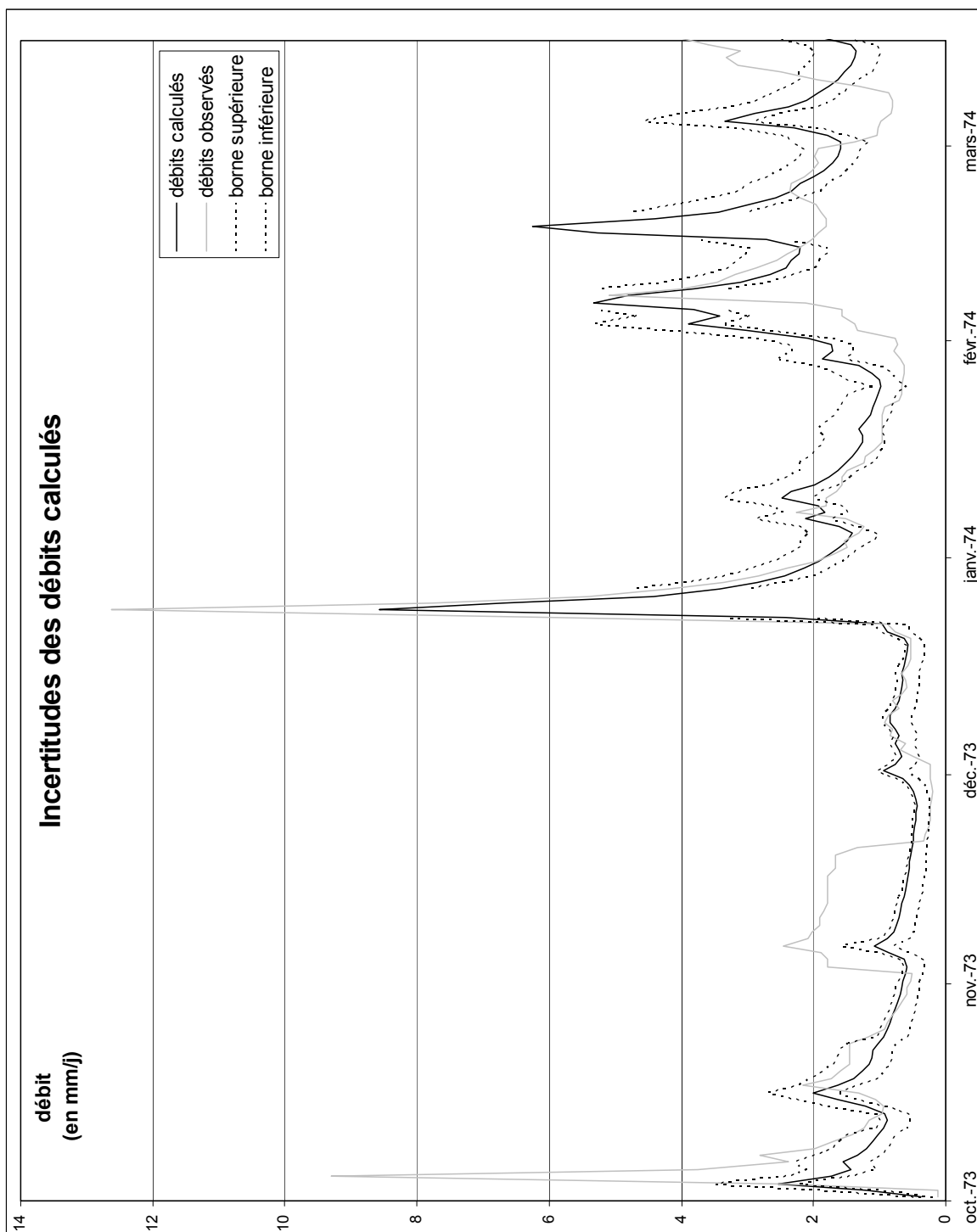
Ainsi grâce à cette représentation graphique du rapport entre Q_{obs} et Q_{cal} , nous pouvons annoncer une incertitude du débit annoncé ; incertitude qui sera d'autant plus fiable que celle annoncée au calage est la même que celle à la validation ; ce qui ne sera pas forcément valable pour tous les bassins versants étudiés.

D'après le tableau 3 et la figure 17, on a pour 70% des débits calculés compris entre 3,981 et 6,310 mm/j :

$$0,5Q_c < Q_o < 1,8Q_c$$

Enfin ajoutons que les résultats d'incertitude des débits annoncés que nous obtenons ici, sont indépendants de la méthode utilisée.

Il est ainsi possible de représenter les incertitudes pour chaque débit calculé (graphique 1).



Graphique 1 : Incertitudes des débits calculés sur un hydrogramme de crue

III-2.e. Qualité du modèle et fiabilité de l'estimation des incertitudes

La question que l'on pourrait se poser est la suivante : Existe-t-il un lien entre la qualité du modèle et la fiabilité de l'estimation des incertitudes ?

Chacune des notes données aux quantiles va être comparée au critère de Nash borné sur \sqrt{Q} de toute les périodes calages des 50 bassins versant étudiés, c'est ce qui nous intéresse et qui intéressera l'utilisateur du modèle) Rappelons que \sqrt{Q} renseigne sur les débits moyens sans donner trop de poids aux erreurs associées aux événements extrêmes de crue et d'étiage. Les figures 18 et 19 présentées ci dessous illustrent la relation entre le biais du débit calculé par rapport au débit mesuré (C2M) et la distorsion entre l'estimation des incertitudes faite au calage et au contrôle.

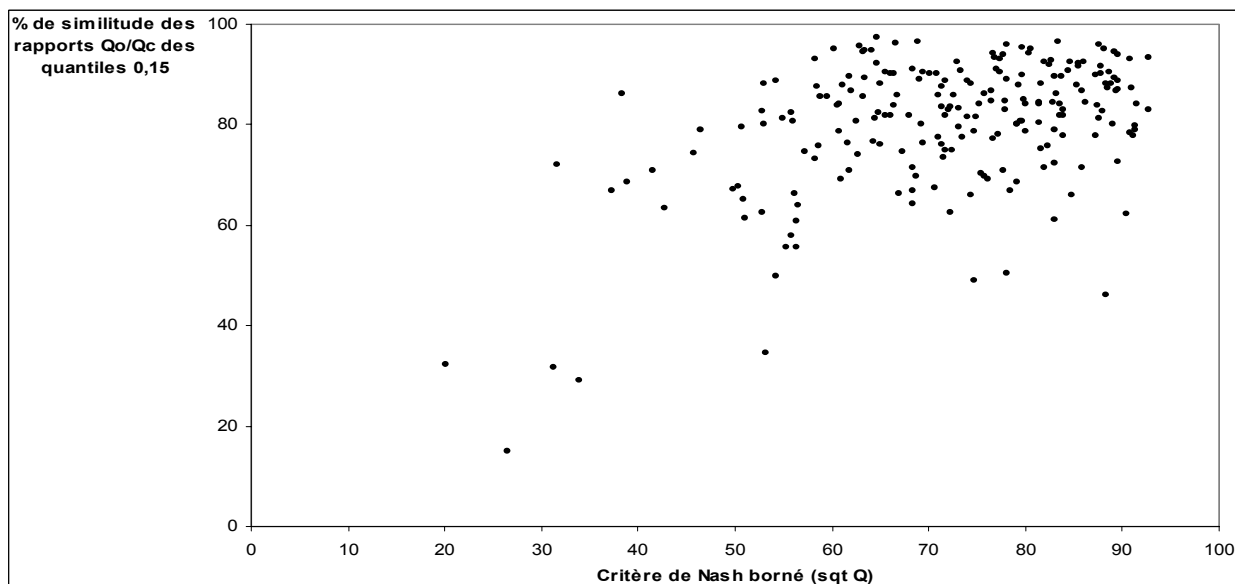


Figure 18 : Similitude des incertitudes au calage et au contrôle pour le quantile 0.15 en fonction du Nash borné (\sqrt{Q})

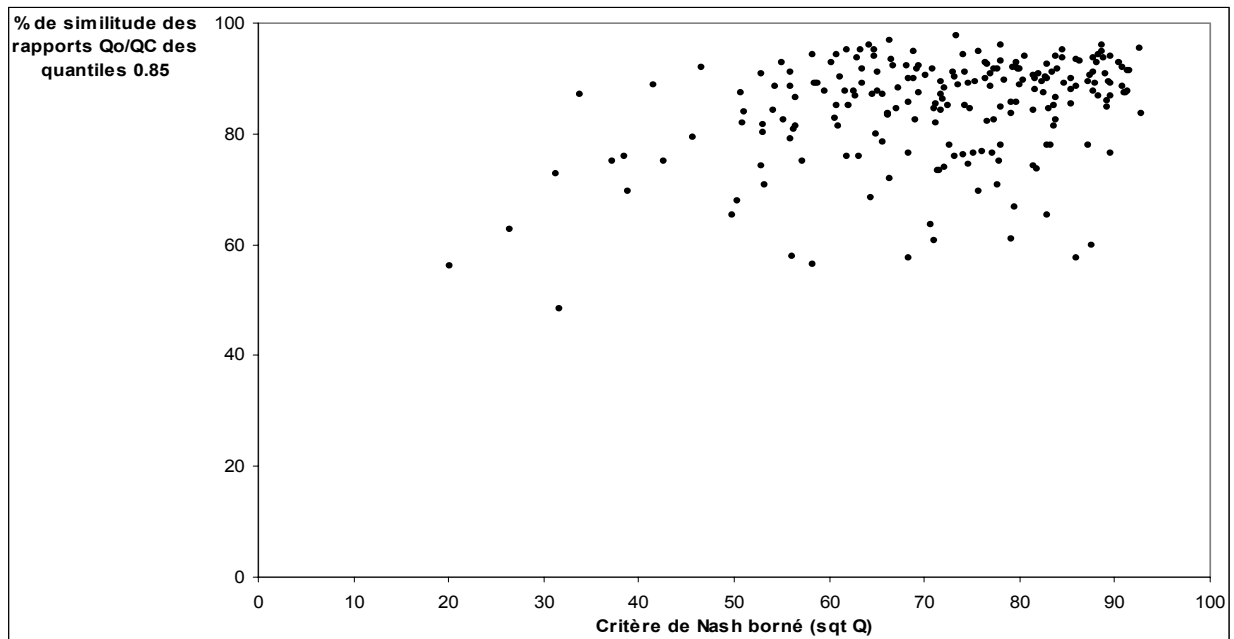


Figure 19 : Similitude des incertitudes au calage et au contrôle pour le quantile 0.85 en fonction du Nash borné (\sqrt{Q})

Nous pouvons remarquer qu'il n'y a pas de relation entre le Nash borné et la qualité de notre estimation des incertitudes. De plus la qualité de l'estimation des incertitudes est assez élevée avec peu de valeurs faibles. On peut donc dire que notre méthode d'évaluation des incertitudes sur les débits calculés par GR4J est relativement fiable et stable quelle que soit la qualité du modèle lui-même.

III-3. Analyses des résultats fournis par GR4J

III-3.a. Estimation, des incertitudes : Méthode proposée

Nous avons illustré au chapitre II, les incertitudes des débits calculés pour le bassin n°8 (H4333410), ce bassin servira aussi à illustrer les autres méthodes d'estimations d'incertitudes abordées dans ce document. Toutefois nous allons vous présenter un autre bassin pour illustrer ce chapitre, le bassin n°17 (J360810) : l'Aulne au Goasc-en-Scrignac.

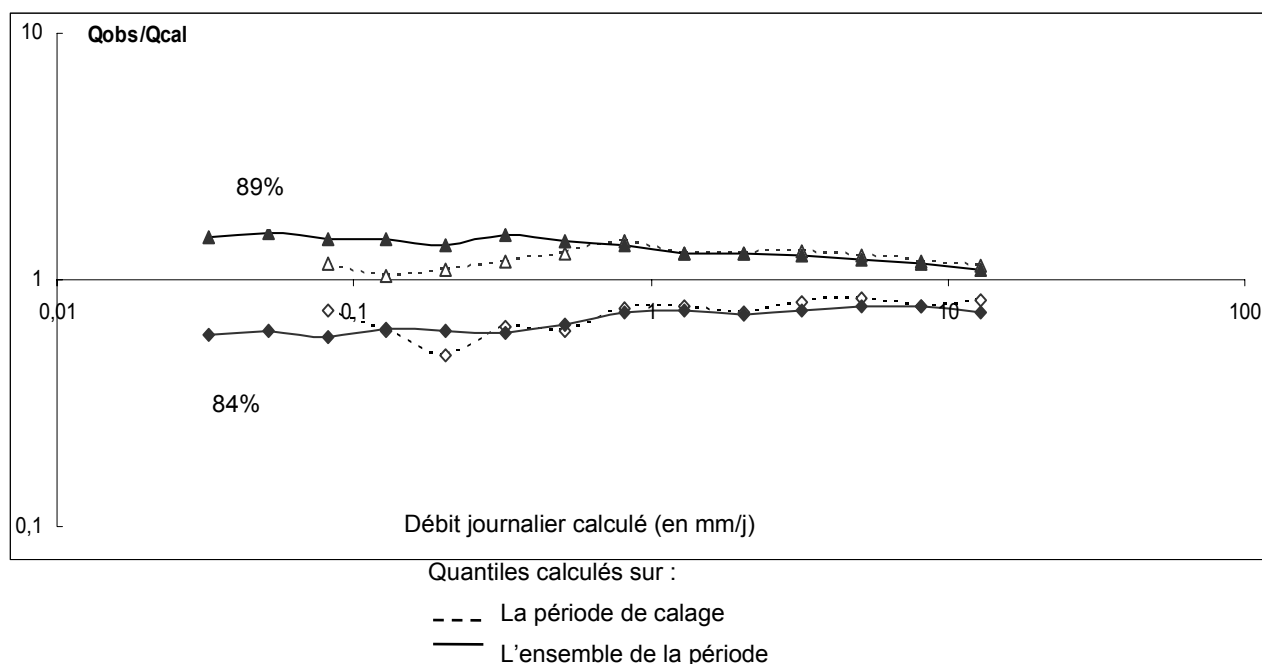


Figure 20 : Rapport Q_o/Q_c par fourchette de débit calculé par GR4J sur la période de calage 1983-1987 et sur la période totale 1974-1991 (J360810)

Sur cet exemple, on constate une forte similitude des écarts entre les débits observés et calculés ; sur la courte période de calage ou sur la période plus longue du contrôle (89% de similitude pour quantiles 0,15 et 84% pour les quantiles 0,85). Pour les débits élevés compris entre 10 et 15 mm/j le rapport Q_o/Q_c au contrôle est compris entre 0,7 et 1,1

$$0,7 < \frac{Q_o}{Q_c} < 1,1 \quad \text{soit} \quad 0,7Q_c < Q_o < 1,1Q_c$$

D'ailleurs, nous nous apercevons que ces erreurs sont relativement homogènes quelque soit les valeurs des débits calculés, et donc que les incertitudes sont globalement comprises entre 0.6 fois et 1.6 fois le débit calculé.

$$0,6 < \frac{Q_o}{Q_c} < 1,5 \quad \text{Soit} \quad 0,6Q_c < Q_o < 1,5Q_c$$

III-3.b. Peut-on exploiter l'information pour effectuer un réajustement a posteriori des débits calculé par le modèle ?

Les figures 21 à 23 montrent trois exemples des renseignements obtenus suite à un calage de GR4J. Alors que pour les figures (21) et (22), on ne peut pas améliorer la qualité de l'estimation, il semble que dans le cas de la figure (23) on peut proposer un réajustement des débits calculés par le modèle GR4J puisqu'un biais apparaît pour les faibles débits.

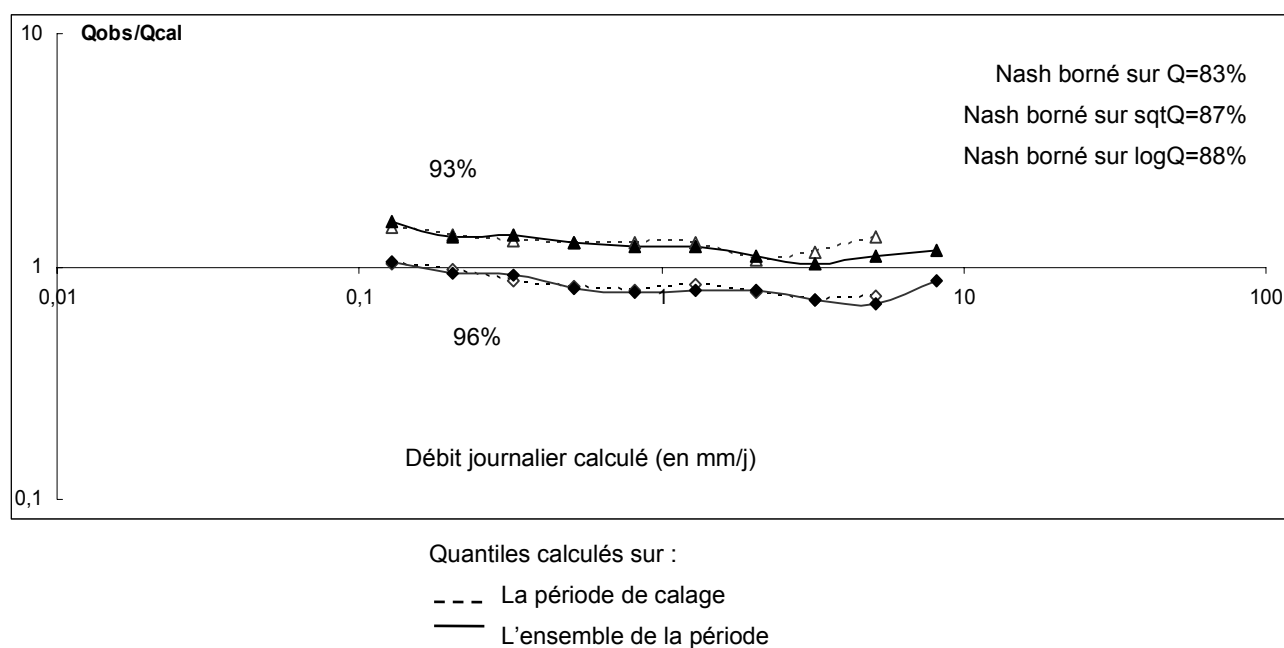
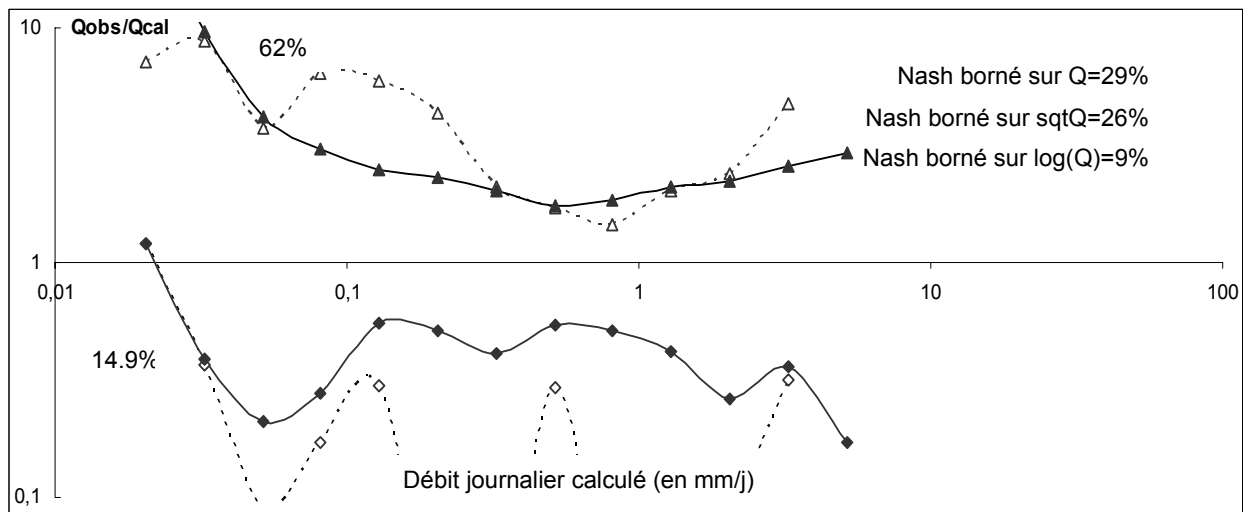


Figure 21 : Le Queffleuth à Lesquifiou Plourin-lès-Mor - J2614010 - BV n° 16 (calage 1972-1977)

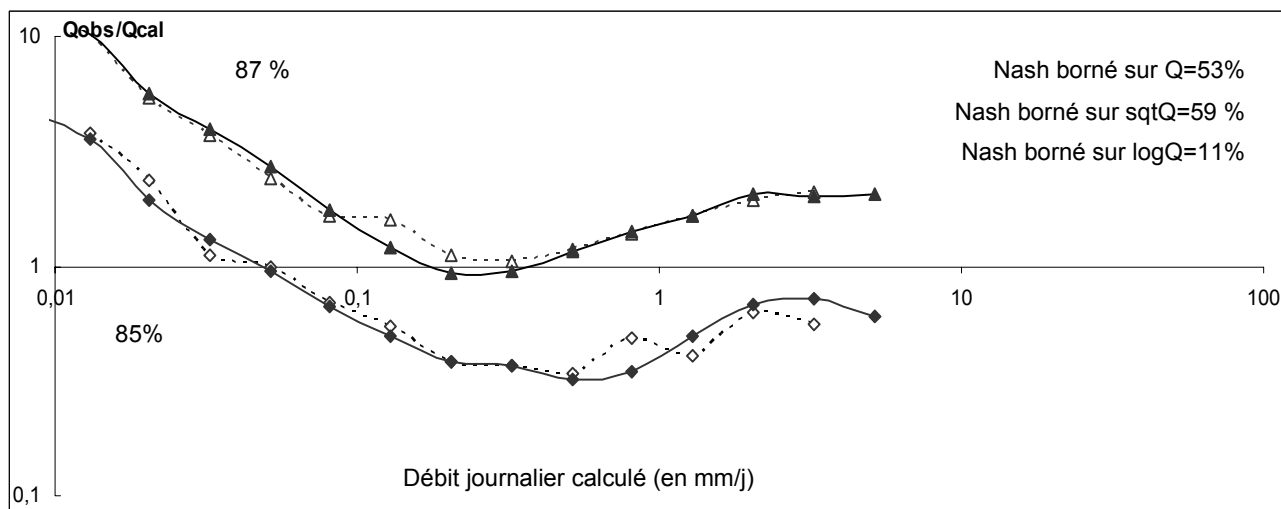
Les informations mises en évidence par la figure 21 nous indiquent qu'il y a un bon C2M (Nash borné) et une bonne similitude entre calage et période totale (distorsion faible). Les prévisions que nous apporte le modèle sur ce bassin sont très fiables.



Quantiles calculés sur :
 - - - La période de calage
 — L'ensemble de la période

Figure 22 : Lidennes à Couteuges - K2365510 - BV n°24 (calage 1980-1985)

En revanche, la figure 22 nous montre pour un bassin où le Nash borné est mauvais au contrôle, que la similitude des erreurs entre le calage et la période totale est elle aussi mauvaise (distorsion élevée).



Quantiles calculés sur :
 - - - La période de calage
 — L'ensemble de la période

Figure 23 : La Cadanne à Pondaurat - O9196210 - BV n°33 (calage 1980-1985)

La figure 23 nous montre l'exemple d'un critère médiocre mais d'une bonne similitude entre calage et période totale.

Afin de corriger le biais observé sur les faibles débits calculés (sous-estimation) et d'améliorer la performance (critère de Nash borné) au contrôle, le débit calculé Q_{cal} a été multiplié systématiquement par le rapport médian $(Q_{obs}/Q_{cal})_{0.5}$ de chaque classe de débits.

$$Qc_{\text{modifié}} = Qc \times \left[\left(\frac{Qo}{Qc} \right)_{0.5} \right]^x \quad \text{Eq. III-3-2.b}$$

L'exposant x est là pour tester dans quelle mesure une correction amoindrie ($x < 1$) ne serait pas plus efficace qu'une correction totale ($x = 1$). Nous allons donc voir si cette correction se traduit bien par une amélioration des simulations c'est-à-dire si elle permet d'améliorer la performance au contrôle. Graphiquement les courbes devraient tendre à se rapprocher de l'axe des abscisses. La moyenne des critères de Nash borné de toutes les périodes de calage des 50 bassins étudiés a été calculé avec le débit calculé modifié $Qc_{\text{modifié}}$. Nous avons ensuite calculé le Nash borné moyen en faisant varier l'exposant x et nous les avons représentés sur la figure 24.

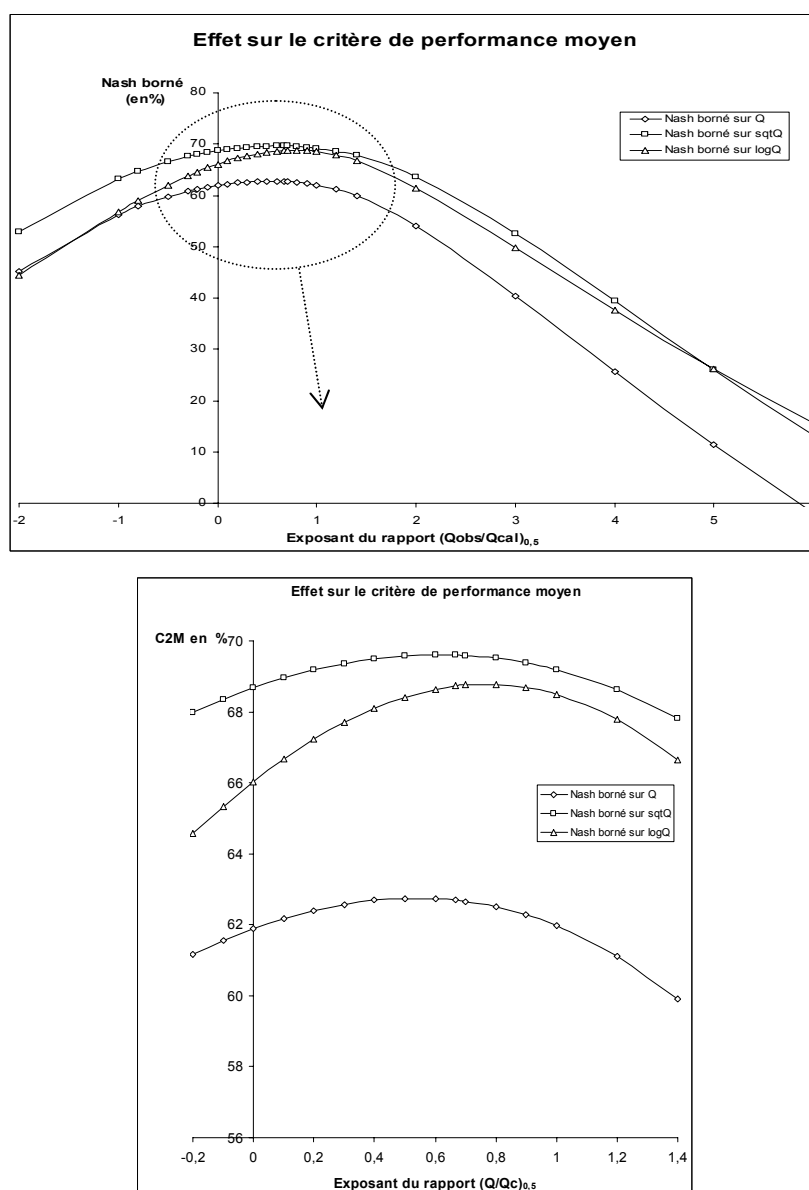


Figure 24 : Effet sur le Nash borné moyen d'une correction systématique des débits fournis par GR4J sur les 50 bassins étudié

En faisant varier l'exposant du rapport médian $(Q_{\text{obs}}/Q_{\text{cal}})_{0,5}$ nous pouvons remarquer une amélioration du Nash borné moyen au contrôle (de toutes les périodes de calage des 50 bassins versants étudiés). Ce critère de Nash borné est calculé avec les débits corrigés (Eq.III-3-2.b). Si l'exposant est nul alors cela correspond à une absence de correction. On voit qu'une correction complète ($x > 1$) n'améliore pas les simulations, cependant :

- ✓ Le critère sur les débits Q s'améliore de 0,85% pour un exposant de 0,5
- ✓ Le critère sur la racine carrée du débit \sqrt{Q} (noté sqrtQ) s'améliore 0,93% pour un exposant de 0,6
- ✓ Le critère sur le logarithme du débit s'améliore 2,74% pour un exposant de 0,8

Cette amélioration n'est pas négligeable car elle correspond grossièrement au gain que l'on peut obtenir de l'introduction d'un paramètre supplémentaire dans le modèle.

III-3.c. Estimation des incertitudes : Méthode dérivée de Monte-Carlo fondée sur une approximation linéaire des écarts-types des paramètres du modèle

Reprenons le bassin n°8 (H4333410), qui nous a servi à illustrer le chapitre II. Nous allons illustrer dans les sections suivantes, avec GR4J et TOPMO, les résultats des trois méthodes d'estimation des incertitudes (notre méthode proposée, GLUE et la méthode dérivée Monte-Carlo).

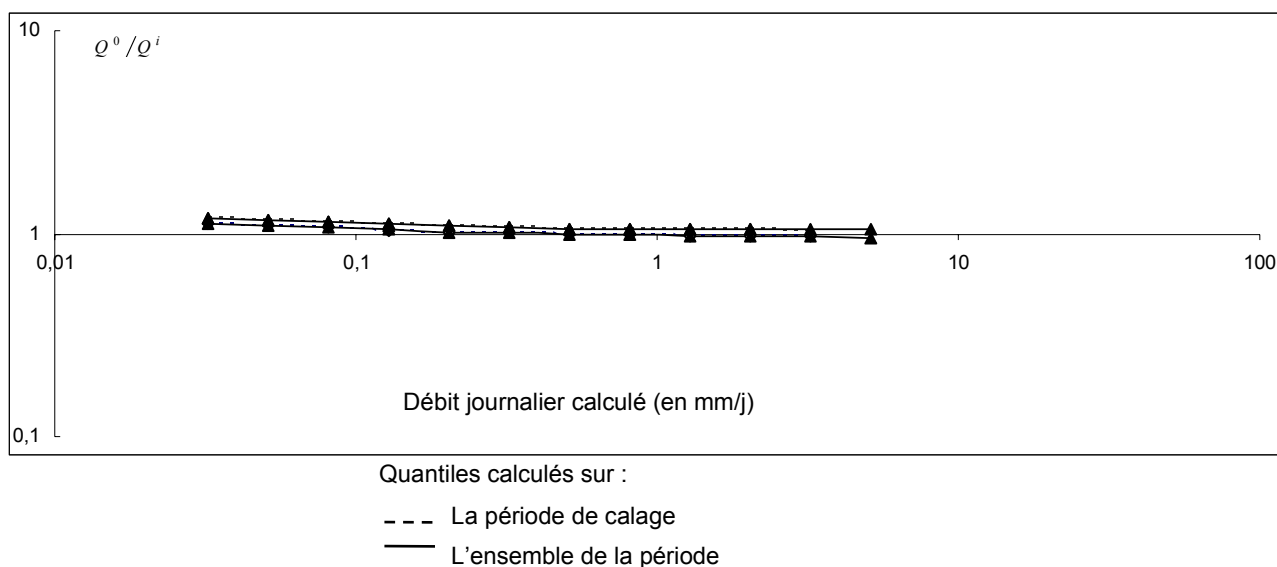


Figure 25 : Rapport Q^0/Q^i par fourchette de débits calculés par GR4J sur la période de calage 1974-1979 et sur la période totale 1974-1992 (H4333410)

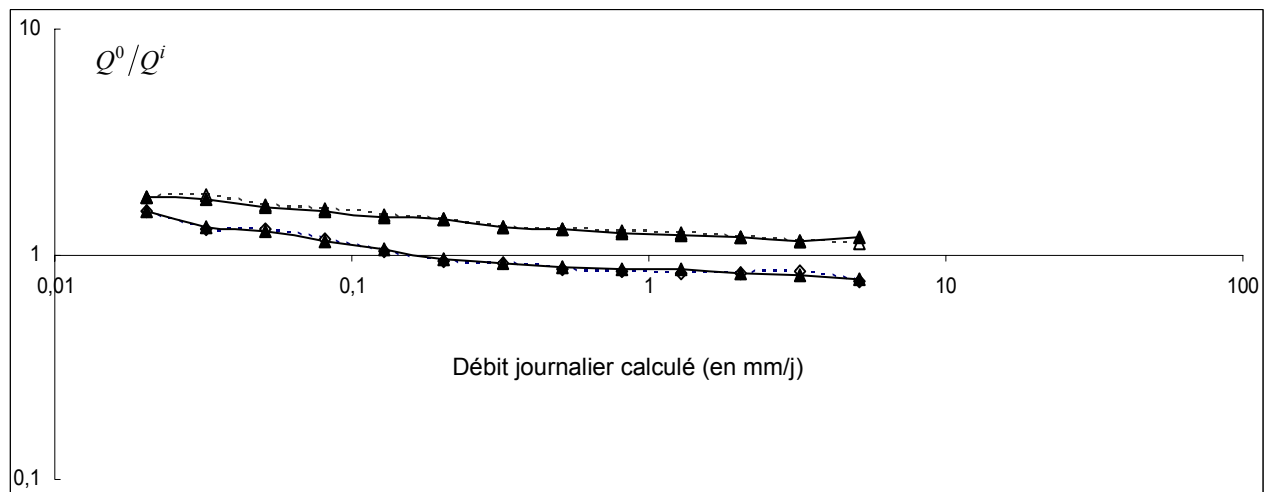
On observe que cette méthode donne une estimation des incertitudes sur les débits calculés trop optimiste, ne reflétant aucunement la réalité. En effet les débits calculés ne sont pas comparés aux débits mesurés, les erreurs obtenues ne sont donc pas exploitables. D'après la figure 25 pour la gamme complète de débits calculés on a environ :

$$1,1 < \frac{Q^o}{Q^i} < 1,2 \text{ et donc } 0,85Q^o < Q^i < Q^o 1,05$$

En réalité cela représente l'écart de ce qui est calculé avec une autre valeur de calculée. Ceci explique l'absence de distorsion entre les erreurs déterminées sur la période de calage et sur la période totale de contrôle. Un tel graphique et de telles estimations d'incertitudes pourraient laisser que le modèle est parfait annonçant des débits extrêmement proches de la réalité. Cette méthode rencontrée dans la littérature ne renseigne donc pas des incertitudes des débits calculés et donne un sentiment trompeur de sécurité.

III-3.d. Estimation des incertitudes : Méthode GLUE

De la même façon que nous avons étudié la méthode de Monte-Carlo, nous allons dans cette section présenter la méthode GLUE.



Quantiles calculés sur :

--- La période de calage

— L'ensemble de la période

Figure 26 : Rapport Q^o/Q^i par intervalle de débit calculé par GR4J sur la période de calage 1974-1979 et sur la période totale 1974-1992 (H4333410)

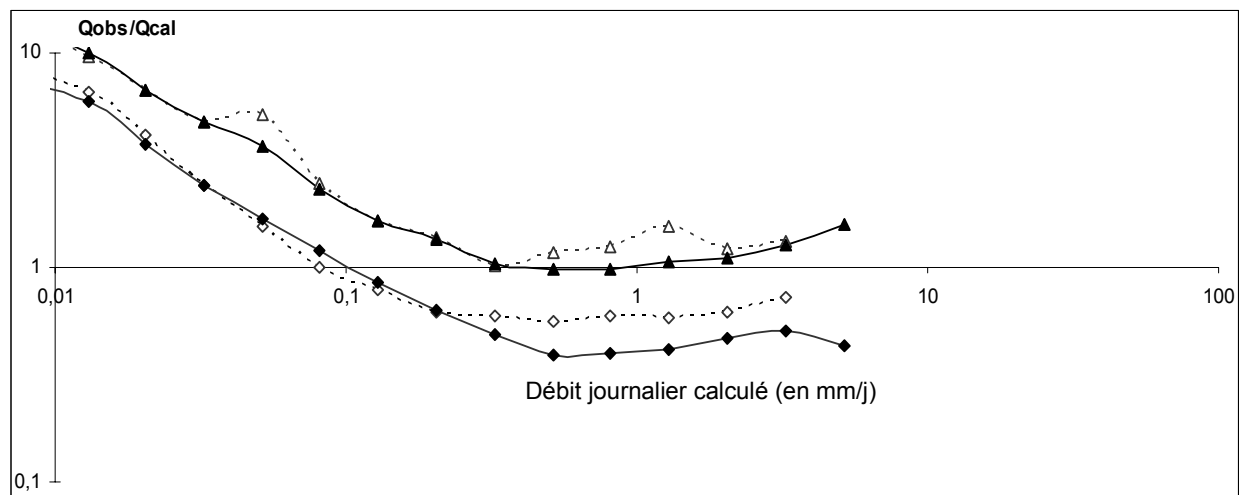
De la même manière que la méthode d'approximation linéaire des écarts-types des paramètres du modèle, la méthode GLUE n'informe pas sur les erreurs que le modèle peut faire sur les débits calculés ne révélant pas les erreurs que peuvent faire le

modèle. Car Q_c n'est pas comparé à Q_o , mais à un autre débit obtenu par le calcul. Ces observations sont aussi valables pour les 49 autres bassins étudiés lors de cette étude.

III-4. Analyses des résultats obtenus sur un autre modèle : TOPMO

En utilisant la même méthode de représentation graphique (chapitre III-2), nous avons représenté les incertitudes sur les débits calculés en utilisant un autre modèle pluie débit conceptuel : TOPMO. Nous avons choisi d'utiliser un autre modèle que GR4J afin de juger la généralité de la méthode proposée.

III-4.a. Estimation des incertitudes : Méthode proposée



Quantiles calculés sur :

--- La période de calage

— L'ensemble de la période

Figure 27 : Rapport Q_{obs}/Q_{cal} par fourchette de débit calculé par TOPMO sur la période de calage 1974-1979 et sur la période total 1974-1992 (bassin n°8 - H4333410)

Aux vues des résultats que l'on obtient sur les 50 bassins versants étudiés et sur celui illustré sur la figure 27, nous pouvons dire que les incertitudes estimées en utilisant le modèle GR4J et TOPMO sont à peu près identiques. Ci-dessus sont illustrées les incertitudes des débits calculés sur le bassin n°8 (H4333410) sur la période de calage 1974-1979. Comme avec GR4J, sur ce bassin il y a une sous estimation des faibles débits, et une bonne estimation des incertitudes au calage et au contrôle.

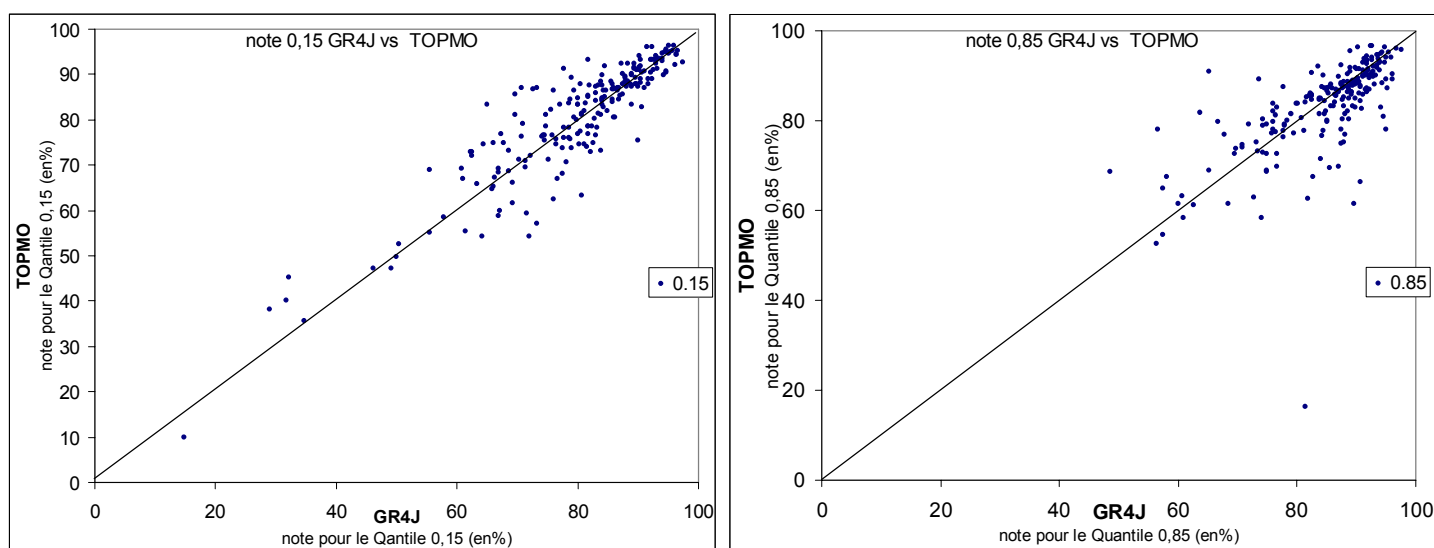
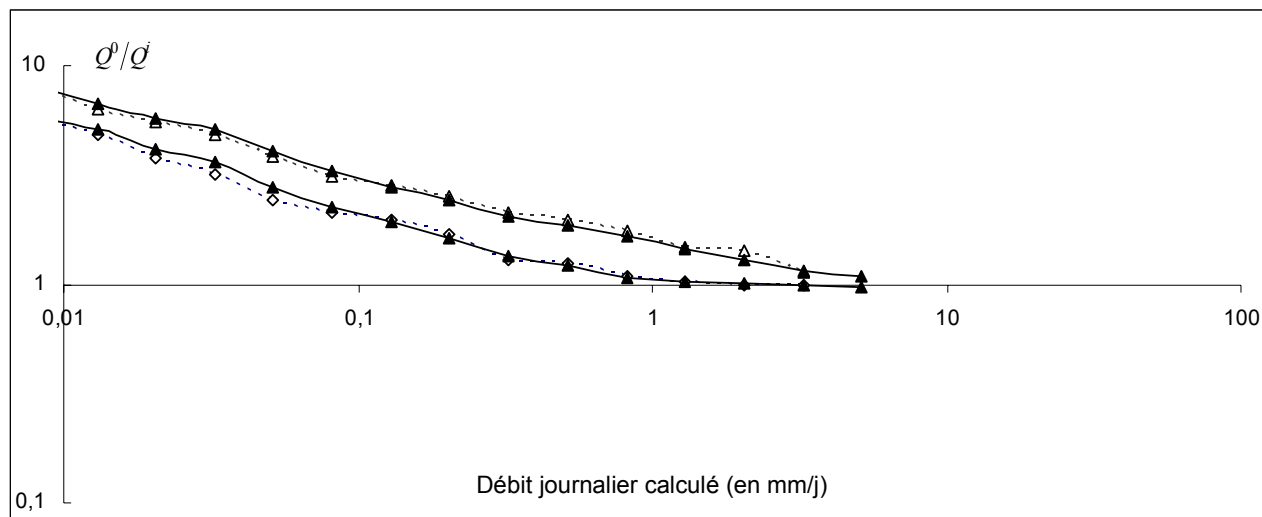
III-4.b. Comparaison des résultats obtenus avec TOPMO et GR4J

Figure 28 : Comparaison des notes (% de similitude) des quantiles 0,15 et 0,85 des 50 bassins étudiés

En comparant chaque note attribuée aux quantiles, obtenue avec GR4J et TOPMO pour les mêmes périodes de calages de chaque bassin, nous obtenons les graphiques ci-dessus.

Cette comparaison n'a pas pour but de comparer la qualité des deux modèles, mais seulement voir si de l'un des deux modèles se prête mieux à l'estimation des incertitudes sur les débits calculés que l'autre. En étudiant la figure 28, nous pouvons constater que les incertitudes annoncées avec TOPMO ou GR4J sont très proches

III-4.c. Estimation des incertitudes : Méthode de Monte-Carlo
fondée sur une approximation linéaire des écart-types des
paramètre du modèle



Quantiles calculés sur :

--- La période de calage

— L'ensemble de la période

Figure 29 : Rapport σ^0/σ^d par fourchette de débit calculé par TOPMO sur la période de calage 1974-1979 et sur la période totale 1974-1992

Nous pouvons observer sur cet exemple une forte sous-estimation des débits calculés par le modèle TOPMO, ce qui diffère des résultats obtenus avec GR4J. Cependant malgré cette sous-estimation, les incertitudes évaluées par la méthode dite de Monte-Carlo restent trop éloigné de la réalité pour des raisons que nous avons développées au chapitre II-7.b.

IV- CONCLUSION

Au fil de cette étude nous avons comparé différentes méthodes d'estimation des incertitudes des débits calculés. Nous avons proposé une méthode qui s'est avérée efficace en l'appliquant à deux modèles de structures différentes : TOPMO et GR4J. La quantification des incertitudes sur les débits calculés sur les cinquante bassins avec la méthode de GLUE ou la méthode basée sur une approximation linéaire des écarts-types des paramètres du modèle, ne renseignent pas sur les erreurs des débits calculés auxquelles il faut s'attendre lorsqu'un modèle annonce ces débits. L'utilisation de ces méthodes classiques qui sous-estiment les incertitudes, est un échec.

En revanche pour une majorité des cinquante bassins étudiés, les incertitudes des débits calculés (avec TOPMO et GR4J) estimées avec la méthode que nous proposons sur les périodes de calage sont généralement très proche de celles qui sont estimées au contrôle avec les jeux de paramètres déterminés à la période de calage correspondante. Démontrant qu'il était possible de réaliser de tels estimations sur les cinquante bassins (choisis pour leur longue chronique de débits et de pluies) ; il est envisageable d'étendre ce genre d'estimations sur d'autres bassins jaugés et ainsi donner encore plus de crédibilité au résultats annoncés.

Il est ensuite facile de simuler des précipitations sur de longues périodes avec un modèle météorologique, sur des bassins sur lesquelles il a été possible de quantifier les erreurs des débits calculés grâce aux chroniques de pluies et de débits connues. Avec les précipitations simulées le modèle peut prévoir des débits auxquelles nous pouvons joindre une incertitude en fonction de leurs valeurs.

V- BIBLIOGRAPHIE

- AMBROISE, B. (1998)** - La dynamique du cycle de l'eau dans un bassin versant - Processus, Facteurs, Modèles. *H*G*A* (Ed.), Bucarest, 200 p.
- 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 Paris VI, Cemagref.
- BEVEN K (1996)** - The limits of splitting: Hydrology-The sciences of the total environment 183 pp.89-97
- BEVEN, K, BINLEY, A. (1992)** - The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes **6**: pp.279-289.
- BEVEN, K.J.; KIRKBY, M.J. (1979)** - A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin, **24**(1), 43-69.
- CHAHINIAN, N (2004)** - Paramétrisation multi-critère et multi-échelle d'un modèle hydrologique spatialisé de crue en milieu agricole. Thèse de doctorat à l'Université Montpellier II, 264p.
- CLARKE, R.T. (1973)** - Review of some mathematical models used in hydrology, with observations on their calibration and use. Journal of hydrology **19**, pp.1-20
- EDIJATNO (1987)** – Amélioration es modèles simples de transformations pluie-débit au pas de temps journalier sur de petits bassins versants. Mémoire de DEA « Sciences et Techniques de L'Eau » Université Louis Pasteur (Strasbourg), Cemagref.242p.
- EDIJATNO (1991)** – Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier. Thèse de doctorat, Université Louis Pasteur (Strasbourg), Cemagref.242p.
- EDIJATNO ; NASCIMENTO, N.O. ;YANG, X. ; MAKLHOUF, Z., MICHEL, C. (1999)** - GR3J : a daily watershed model with three free parameters. Hydrological Sciences Journal **44**, pp.1-20
- HUBERT, P. (1984)** – Eaupuscule (Une introduction à la gestion de l'eau) ; ed. Marketing(Paris) ; 192p.
- HUBERT, P. (1996)** - "Dictionnaire Français d'Hydrologie", Comité Nationale Français des Sciences Hydrologiques, Commission terminologie.
- JACQUET, J. (1971)** – Aspects généraux de la représentation mathématique des phénomènes en hydrologie de surface. La Houille Blanche, **n°2**, pp.105-110
- JAKEMAN, A.J.; LITTLEWOOD, I.G.; WHITEHEAD, P.G. (1990)** - Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. Journal of Hydrology, **117**, pp.275-300.

- MADSEN, H. (2000)** - Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology* **235** pp.276-288.
- MATHEVET, T.; PERRIN, C.; ANDRÉASSIAN, V.; MICHEL, C. (2004)** – Large samples of watersheds: necessary condition to compare the efficiency of rainfall-runoff models? Soumis à *Water Resources Research*.
- METROPOLIS, N.; ULAM, S. (1949)** - The Monte Carlo method, *Journal of the American Statistical Association*, **44** (247), pp.335-341.
- MICHEL, C. (1989)**. Hydrologie appliquée aux petits bassins versants ruraux, Cemagref, Antony.
- MONTANARI A.; BRATH A. (2004)** - A stochastic approach for assessing the uncertainty of rainfall-runoff simulations *Water Resources Research* **40** n°1
- MORTON, F.I. (1983)**. "Operational estimates of actual évapotranspiration and their significance to the science and practice of hydrology." *Journal of Hydrology* **66** pp.1-76.
- MOUELHI, S. (2003)** - Vers une chaîne cohérente de modèles pluie-débit conceptuels globaux aux pas de temps pluriannuel, annuel mensuel et journalier. Thèse de doctorat, Cemagref, ENGREF, 323p.
- NASCIMENTO, N.O. (1995)** - Appréciation à l'aide d'un modèle empirique des effets d'actions anthropiques sur la relation pluie-débit à l'échelle du bassin versant. Thèse de Doctorat, CERGRENE / ENPC, Paris, France, 550 p.
- NASH, J.; SUTCLIFFE J. (1970)** - River flow forecasting through conceptual model. Part I: a discussion of principles. *Journal of Hydrology* **10** pp.282-290
- LOUDIN, L. (2004)** – Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global Thèse de doctorat, Cemagref, ENGREF, 495p.
- PENMAN, H.L. (1948)** - Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. London A193*: 120-145.
- PERRIN, C. (2000)** - Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Thèse de Doctorat, Cemagref. Antony, Institut National Polytechnique de Grenoble: 530 p.
- PETIT, M. (2005)** - Incertitudes scientifiques et risques climatiques – EDITORIAL. *Comptes Rendus Geosciences*, 337 pp. 393-398
- PETRUSCU MAFTEI, C (2002)** - Etude concernant les écoulements superficiels, Modélisation spatialisée de l'écoulement sur le bassin versant de Voinesti (Roumanie). Thèse de doctorat, Université Montpellier II, Université Ovidius de Constanta (Roumanie)

REFSGAARD, J.C.; ABBOTT, M.B. (1996) - The role of distributed hydrological modelling in water resources management - Chapter 1. In: Distributed Hydrological modelling, Kluwer Academic Publishers, Netherlands, 1-16.

TUFFIN, B. (1997) - Simulation accélérée par les méthodes de Monte Carlo et Quasi-Monte-Carlo : théorie et applications. Thèse de Doctorat, Université de Rennes 1.

VI- WEBOGRAPHIE

GAUME, E. (2000) - Hydrologie : cycle de l'eau et modélisation pluie-débit. Cours ENPC

<http://www.enpc.fr/cereve/HomePages/gaume/courshydro/courshydro.html>

<http://www.cig.ensmp.fr/~hubert/res/res98-10.htm>

VII- TABLES DES LÉGENDES

VII-1. Index des figures

Figure 1 : Relation entre les modèles pluie-débit (Gaume, 2000)	9
Figure 2 : Représentation schématique de TOPMO	10
Figure 3 : Représentation schématique de IHACRES	11
Figure 4 : Représentation schématique de GR4J	12
Figure 5 : Schéma de la chronologie du calage.....	15
Figure 6 : Représentation schématique du calage manuel	16
Figure 7 : Représentation schématique des écarts entre le débit observé et le débit calculé.....	17
Figure 8 : Représentation des incertitudes sur les débits ayant subi une approximation linéaire des écarts-types des paramètres du modèle I_c et des incertitudes des débit calculé I_0	20
Figure 9 : Représentation schématique des incertitudes avec la méthode GLUE.....	22
Figure 10 : Répartition des 50 bassins versants utilisés lors de l'étude	25
Figure 11 : Hydrogrammes de crues du bassin H4333410 sur la période de calage de 1974 à 1979 et hydrogramme calculé par GR4J	26
Figure 12 : Rapport des débits observés sur débits calculés pour chaque observation en fonction de la date sur la période de calage 1974-1979	27
Figure 13 : Comparaison des figures 11 et 12.....	28
Figure 14 : Rapport Q_{obs}/Q_{cal} pour chaque observation en fonction de Q_{cal}	28
Figure 15 : Encadrement de 70% des rapports Q_{obs}/Q_{cal} pour chaque fourchette de débit calculé au calage.....	30
Figure 16 : Rapport Q_o/Q_c par fourchette de débit calculé par GR4J sur la période de calage 1974-1979	31
Figure 17 : Rapport Q_{obs}/Q_{cal} par fourchette de débit calculé par GR4J sur la période de calage 1974-1979 et sur la période totale 1974-1992.....	33
Figure 18 : Similitude des incertitudes au calage et au contrôle pour le quantile 0.15 en fonction du Nash borné (\sqrt{Q}).....	35

Figure 19 : Similitude des incertitudes au calage et au contrôle pour le quantile 0.85 en fonction du Nash borné (\sqrt{Q}).....	36
Figure 20 : Rapport Q_o/Q_c par fourchette de débit calculé par GR4J sur la période de calage 1983-1987 et sur la période totale 1974-1991(J360810).....	37
Figure 21 : Le Queffleuth à Lesquifiou Plourin-lès-Mor - J2614010 - BV n° 16 (calage 1972-1977).....	38
Figure 22 : Lidennes à Couteuges - K2365510 - BV n°24 (calage 1980-1985).....	39
Figure 23 : La Cadanne à Pondaurat - O9196210 - BV n°33 (calage 1980-1985).....	39
Figure 24 : Effet sur le Nash borné moyen d'une correction systématique des débits fournis par GR4J sur les 50 bassins étudié	40
Figure 25 : Rapport Q^o/Q^i par fourchette de débits calculés par GR4J sur la période de calage 1974-1979 et sur la période totale 1974-1992 (H4333410).....	41
Figure 26 : Rapport Q^o/Q^i par intervalle de débit calculé par GR4J sur la période de calage 1974-1979 et sur la période totale 1974-1992 (H4333410).....	42
Figure 27 : Rapport Q_{obs}/Q_{cal} par fourchette de débit calculé par TOPMO sur la période de calage 1974-1979 et sur la période total 1974-1992 (bassin n°8 - H4333410).....	43
Figure 28 : Comparaison des notes (% de similitude) des quantiles 0,15 et 0,85 des 50 bassins étudiés	44
Figure 29 : Rapport Q^o/Q^i par fourchette de débit calculé par TOPMO sur la période de calage 1974-1979 et sur la période totale 1974-1992.....	45
Figure 30 : Schéma d'une boîte à moustaches.....	53

VII-2. Index des tableaux

Tableau 1 : Classification des erreurs en modélisation hydrologique selon leur nature ...	5
Tableau 2 : Exemple de caractéristiques des modèles pluie-débit.....	9
Tableau 3 : Extrait des résultats du bassin versant n°8 pour la période de calage 1974-1979.....	29
Tableau 4 : Extrait des résultats du bassin versant n°25 sur l'ensemble de la période ..	32

VIII- ANNEXES

VIII-1. Annexe 1 : Rappels sur les Quantiles

Les quantiles sont des *indicateurs de dispersion* d'une série statistique ; ils nous fournissent des informations sur la distribution étudiée. Ils définissent des découpages standard des séries et permettent de *mesurer la dispersion* (par exemple grâce à l'intervalle interquartile relatif). On appelle intervalle inter-quantile, l'intervalle qui contient les valeurs centrales de l'échantillon.

$Q_{cal}(0,15) = 0,3$ signifie que 15% des rapports $\frac{Q_{observé}}{Q_{calculé}}$ sont inférieurs à 0,3

$Q_{cal}(0,85) = 3,3$ signifie que 85% des rapports $\frac{Q_{observé}}{Q_{calculé}}$ sont inférieurs à 3,3

C'est-à-dire que 70% des valeurs se trouvent dans l'intervalle inter-quantile

$$[Q_{cal}(0,15), Q_{cal}(0,85)]$$

Ces intervalles sont à la base d'une représentation très compacte de la distribution empirique : le diagramme en boîte (ou boîte à moustaches). Elle consiste en une boîte rectangulaire, dont les deux extrémités sont les quantiles. Ces extrémités se prolongent par des traits terminés par des segments orthogonaux (les moustaches). La longueur de ces segments varie selon les auteurs. Nous proposons de la fixer valeurs extrêmes.

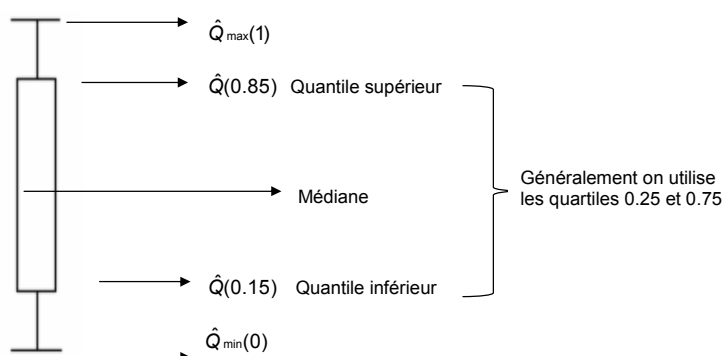


Figure 30 : Schéma d'une boîte à moustaches

Même sur de très gros échantillons, les quantiles sont peu coûteux à calculer puisqu'il suffit de trier l'échantillon par ordre croissant pour calculer les statistiques d'ordre, et donc tous les quantiles simultanément. Ils fournissent une visualisation facile de la distribution empirique. Nous avons vu que la médiane est une valeur centrale. Pour mesurer la dispersion, on peut calculer l'étendue, qui est la différence entre la plus

petite et la plus grande valeur. Mais cette étendue reflète plus les valeurs extrêmes que la localisation de la majorité des données.

VIII-2. Annexe 2 : Résultats avec GR4J et la méthode proposée