

Premiers essais de détermination du coefficient de ruissellement décennal à l'aide de règles floues

F. DELCLAUX & V. HILAIRE

Laboratoire d'Hydrologie, Orstom, BP 5045, F-34032 Montpellier Cedex, France

Résumé Dans le cadre de la prédétermination des crues décennales au Sahel, Jean Rodier a proposé une méthode d'évaluation des coefficients de ruissellement décennaux. Cette méthode est basée sur l'estimation de trois paramètres explicatifs, la surface du bassin-versant, l'indice global de pente et la classe d'infiltrabilité. L'utilisation d'abaques permet la détermination du coefficient recherché: pour une surface donnée, on effectue une double interpolation sur les variables pente et infiltrabilité. Un des points délicats de la méthode réside dans la détermination et la classification des variables explicatives. En effet, ce travail met en jeu des connaissances qualitatives qu'il s'agit de traduire quantitativement. Or la théorie des ensembles flous permet le traitement numérique d'informations qualitatives ou incertaines. Nous testons un système à bases de règles floues dans lequel les variables sont appréhendées en tant que nombres flous, et où les algorithmes numériques sont remplacés par des règles d'inférence. La calibration du modèle s'effectue sur un jeu de données en faisant varier la forme des fonctions d'appartenance ainsi que la mise en oeuvre des règles. Le modèle est ensuite validé sur un autre échantillon de données.

Determination of the ten year period return runoff coefficient with a fuzzy rules-based model

Abstract Within the framework of the determination of the 10-year return floods in the Sahel, Rodier proposed an evaluation method for the runoff coefficient. This method is based on estimation of three parameters, the catchment basin area, a global slope index and an infiltrability index. The runoff coefficient is found by interpolation between these variables on predefined curves. One of the problems is to determine quantitative values for these variables, mainly for the infiltrability index which is qualitatively known. We test a fuzzy rules-based model in which the variables are considered as fuzzy numbers and where the curves are modelled by inference rules. The model is calibrated on a data set: the calibration parameters are the shape of the membership functions, the defuzzification function and the number of fuzzy classes. Then we validate the model on another data set.

INTRODUCTION

En matière de modélisation hydrologique, nous sommes parfois conduits à poser des hypothèses et classifications trop strictes que nous aimerions pouvoir moduler en

fonctions de connaissances qualitatives. Le problème de l'accès à la valeur exacte d'une variable incertaine ou imprécise est souvent résolu par une approche statistique où la variable est remplacée par une fonction de probabilité. L'approche Bayésienne permet d'intégrer une connaissance a priori sur l'incertitude. Dans ce domaine cependant, la théorie des ensembles flous propose un autre cadre de travail. Aussi nous sommes-nous intéressés, dans le cadre de la prédétermination des coefficients de ruissellement décennaux, à la modélisation à base de règles floues dans laquelle les variables s'expriment en tant que nombres flous et leurs relations par des règles d'inférence.

ENSEMBLES ET NOMBRES FLOUS

Définitions

Les ensembles flous, introduits en 1965 par Zadeh, sont basés sur le concept d'appartenance graduelle. Dans la théorie classique des ensembles, un élément appartient ou n'appartient pas à un ensemble donné. Au contraire, un ensemble flou a une frontière vague: par exemple, l'ensemble des longues rues d'une ville.

Un sous-ensemble flou A d'un ensemble X est un ensemble de couples tels que:

$$A = \{[x, \mu_A(x)], x \in X\}$$

où $\mu_A(x)$ est la fonction d'appartenance de x dans A et prend ses valeurs dans l'intervalle $[0, 1]$. Plus $\mu_A(x)$ est proche de 1, plus l'élément x appartient à A . Par contre si $\mu_A(x)$ vaut zéro, il est certain que x n'appartient pas à A . Dans le cas des ensembles ordinaires, la fonction $\mu_A(x)$ est binaire et prend les valeurs 0 ou 1.

Un nombre flou est un cas particulier des ensembles flous et se définit de la manière suivante:

- il existe au moins une valeur x telle que $\mu_A(x) = 1$; et
- la fonction $\mu_A(x)$ est convexe: $\forall a, b, c, a < c < b: \mu_A(c) \geq \min[\mu_A(a), \mu_A(b)]$.

La fonction μ_A est ainsi constituée d'une partie croissante et d'une partie décroissante (voir Fig. 1).

Caractéristiques des nombres flous

Les caractéristiques principales de la fonction d'appartenance sont (voir Fig. 1):

- le support, défini comme l'intervalle $\{x; \mu_A(x) > 0\}$; et
- la forme de la fonction d'appartenance: triangulaire, trapézoïdale, ou suivant une loi puissance telle que les fonctions *LR* de Dubois et Prades (Bardossy & Duckstein, 1995).

Il est parfois utile de trouver une valeur réelle équivalente à un nombre flou. On procède alors à une opération dite de « déflouification » pour laquelle plusieurs méthodes sont utilisées:

- la moyenne des maximum:

$$M_A = \int_{-\infty}^{+\infty} \max[\mu_A(x)] dx$$

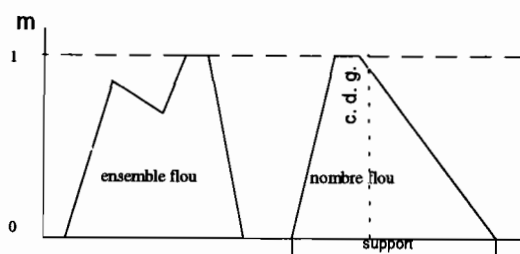


Fig. 1 Fonction d'appartenance d'un ensemble flou et d'un nombre flou.

— la médiane:

$$\int_{-\infty}^{M_A} \mu_A(x) dx = \int_{M_A}^{+\infty} \mu_A(x) dx$$

— le centre de gravité:

$$M_A = \frac{\int_{-\infty}^{+\infty} x \mu_A(x) dx}{\int_{-\infty}^{+\infty} \mu_A(x) dx}$$

Enfin, une arithmétique floue a été élaborée. C'est une généralisation de l'arithmétique courante, basée sur la notion d'intervalle. En effet, un nombre flou peut être considéré comme un ensemble discret d'intervalles. Il est alors possible d'effectuer des opérations d'addition, de soustraction, etc., ainsi que d'appliquer des opérateurs d'intersection, d'union et de calcul de distances.

MODELISATION A BASES DE REGLES FLOUES

Les ensembles flous sont utilisés dans de nombreux contextes dont les plus connus sont le contrôle de processus en automatique. Mais ils sont également mis en oeuvre en tant qu'outils d'analyse de décision et de risque et en tant qu'outils scientifiques, notamment dans le domaine de l'environnement (Droesen & Geelen, 1993; Bardossy *et al.*, 1990). Dans ce contexte, la modélisation à base de règles floues nous intéresse plus particulièrement car elle permet de modéliser quantitativement le discours qualitatif du spécialiste en respectant les caractéristiques du raisonnement humain.

Principes

Supposons qu'un système soit défini par n variables X_i explicatives d'une variable Z , que nous disposions d'un jeu de données expérimentales de m uplets (x_{ij}, z_j) et que les relations s'expriment un ensemble de règles de la forme {si X_1 et si X_2 et si X_3 et si X_n alors Z }. Ce formalisme peut exprimer, par exemple: si la pente est forte et le sol très perméable, alors le ruissellement sera très important.

On décrit chacune des variables par plusieurs nombres flous de fonction d'appartenance μ_{X_i} dont les intersections sont non vides. Ainsi, pour chaque donnée k

(x_{i_k}, z_k) , plusieurs règles pourront s'appliquer simultanément, mais avec des poids différents. On définit alors pour une donnée k un degré de satisfaction v_j^k de la règle j , v_j^k étant fonction des appartenances des X_i :

$$v_j^k = \min_i [\mu_{X_i}(x_{i_k})]$$

L'ensemble flou résultant Z_k est une pondération des ensembles Z_j^k déduits de chaque règle j , soit:

$$Z_k = \cup_j v_j^k Z_j^k$$

La valeur réelle z'_k , déduite de l'ensemble flou Z_k par une méthode de « déflouification », est alors comparée à la valeur observé z_k .

METHODE ORSTOM DE PREDETERMINATION DES CRUES DE FREQUENCE DECENNALE

Historique

Dès 1965, une première note de synthèse (Rodier & Auvray, 1965) établissait les bases de la méthode Orstom à partir d'un échantillon de 65 bassins versants d'Afrique de l'Ouest, de superficie inférieure à 120 km² et situés entre les isohyètes interannuelles 150 et 1600 mm.

Le volume des observations s'est notablement accru et des analyses hydrologiques de plus en plus fines ont permis de mieux appréhender les phénomènes hydrologiques. De 1984 à 1986, un travail de révision a été réalisé pour les bassins versants de superficie inférieure à 10 km² (Rodier, 1986; 1989; Rodier & Ribstein, 1986; 1988).

A partir de 1988, Rodier a étendu cette révision à 250 petits bassins versants des zones sahélienne et tropicale sèche, pour des pluviométries annuelles comprises entre 300 et 1000 mm (Rodier, 1993; 1995; FAO, 1995).

But et fondement de la méthode

Rodier (1993) a proposé de prendre pour référence de crue de projet la crue dite « décennale ». Celle-ci est définie comme étant une crue fictive dont les principales caractéristiques ont une période de retour de 10 ans. Plus précisément, c'est la crue provoquée par une pluie journalière de hauteur décennale, toutes les autres caractéristiques correspondant aux valeurs les plus fréquentes pour les fortes précipitations observées (humectation du sol, état de la végétation, forme du hyétogramme). L'approche proposée, déterministe, est celle d'un modèle global « pluie-débit » fondé sur la théorie de l'hydrogramme unitaire. Le débit de pointe correspondant au ruissellement de la crue décennale peut être défini par la relation:

$$Q_{10} = K_{10} P_{10} Kr_{10} \alpha_{10} S / T b_{10}$$

avec: K_{10} : coefficient d'abattement; P_{10} : hauteur de pluie ponctuelle journalière de fréquence décennale; Kr_{10} : coefficient de ruissellement; α_{10} : coefficient de pointe; S :

superficie du bassin versant; et Tb_{10} : temps de base.

Le coefficient d'abattement est un coefficient de réduction qui permet de passer d'une hauteur de pluie ponctuelle à une hauteur moyenne sur une certaine superficie située dans une zone pluviométriquement homogène. Il est déterminé en Afrique de l'ouest par la formule de Vuillaume (1974).

Le coefficient de ruissellement est le rapport entre le volume de ruissellement rapide à l'exutoire du bassin et le volume précipité sur l'ensemble du bassin versant.

Le coefficient de pointe est le rapport entre le débit maximum ruisselé et le débit moyen correspondant au volume ruisselé sur le temps de base. On admet que α_{10} est voisin de 2,6 quelle que soit la superficie du bassin; cependant l'aspect du réseau hydrographique peut modifier cette valeur.

Enfin, le temps de base décrit l'intervalle de temps compris entre le début et la fin du ruissellement rapide.

Le coefficient de ruissellement Kr_{10}

Dans le cadre de notre étude, nous nous sommes limités à la détermination du coefficient de ruissellement décennal Kr_{10} . Pour un bassin-versant donné, et pour une pluie P_{10} donnée, Kr_{10} est calculé par interpolation à partir de deux coefficients Kr_{70} et Kr_{100} correspondant à des précipitations de 70 et 100 mm. Ces coefficients sont déterminés en fonction de la superficie (S) des bassins pour cinq classes d'infiltrabilité (Ci) et pour différentes valeurs de l'indice global de pente (Ig) à partir des observations de 250 bassins versants. Le coefficient Kr_{10} s'écrit donc:

$$Kr_{10} = Kr_{70}(S, Ci, Ig) + [Kr_{100}(S, Ci, Ig) - Kr_{70}(S, Ci, Ig)](P_{10} - 70)/(30)$$

L'indice global de pente Ig est déterminé après construction de la courbe hypsométrique du bassin. C'est le rapport entre la dénivelée globale et la longueur du rectangle équivalent.

La quantification des classes d'infiltrabilité est le point le plus délicat de la méthode proposée: les éléments déterminants de l'infiltrabilité en zones sahélienne et tropicale sèche sont les états de surface qui intègrent le couvert végétal, les organisations pédologiques superficielles et les modalités de mise en culture ou d'occupation du sol. La classification qualitative adoptée (FAO, 1995) comprend cinq classes d'infiltrabilité caractérisées par des plages de lames d'eau fictives (Pr) correspondant à une pluie utile de 50 mm:

– TI	: bassin rigoureusement imperméable	$Pr > 30$ mm;
– I	: bassin imperméable	$24 > Pr > 26$ mm;
– RI	: bassin relativement imperméable	$12 > Pr > 18$ mm;
– P	: bassin perméable	$6 > Pr > 9$ mm; et
– TP	: bassin très perméable	$Pr < 5$ mm.

Des abaques fournissent les valeurs des coefficients de ruissellement Kr_{70} et Kr_{100} pour différentes classes de surfaces, d'indices de pente et d'infiltrabilité (voir Fig. 2).

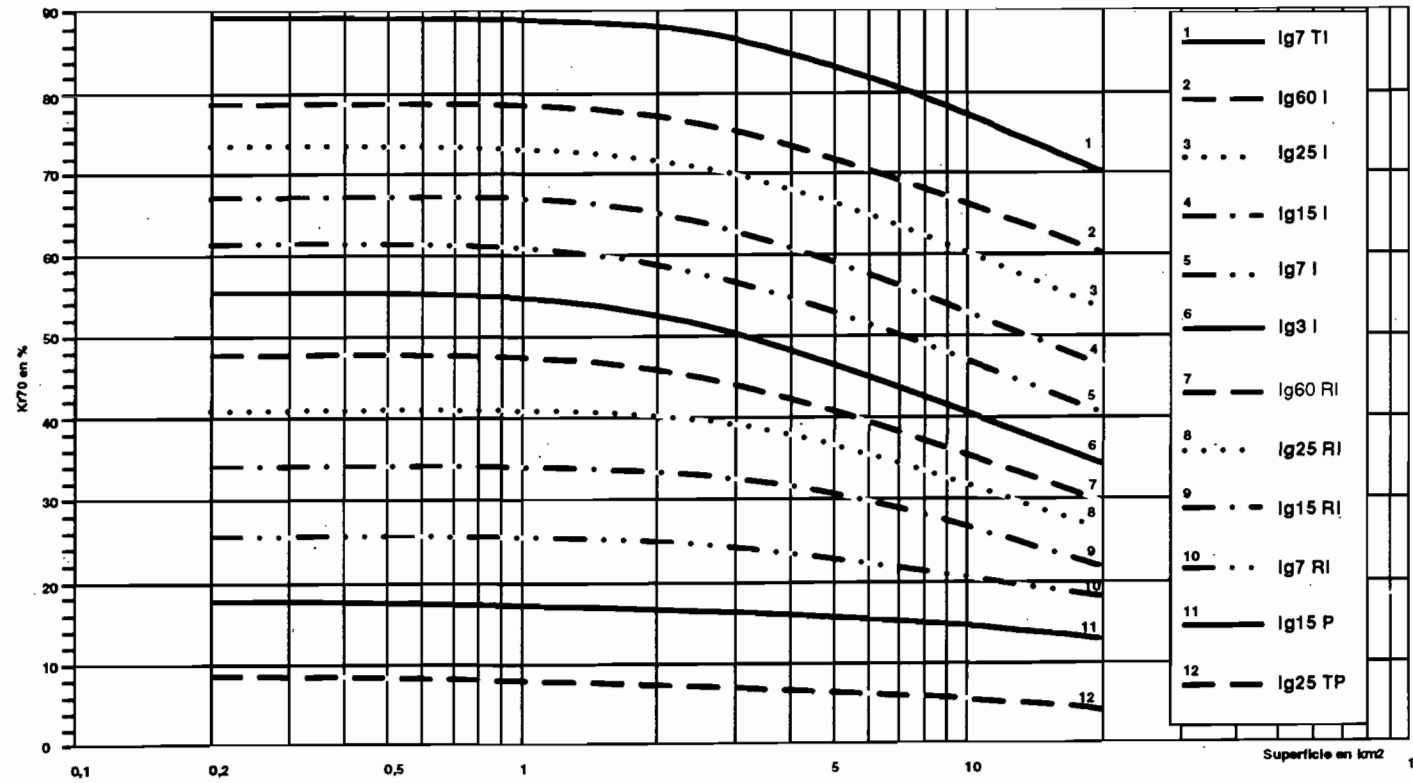


Fig. 2 Evolution du K_{70} en fonction de la surface ($S < 10 \text{ km}^2$), des indices globaux de pente et des classes d'infiltrabilité.

MODELISATION DU $Kr70$ A L'AIDE DE REGLES FLOUES

Les données

Dans notre étude, nous nous limiterons à la modélisation du coefficient $Kr70$. Les trois variables explicatives sont le logarithme de la surface ($\log S$), l'indice global de pente (Ig) et la classe d'infiltrabilité (Ci). A partir des données expérimentales relevées par Rodier, concernant 74 bassins-versants, nous avons établi un échantillon de calibration et un échantillon de validation comportant chacun 37 jeux de données dont les caractéristiques statistiques sont similaires (voir Tableau 1).

Tableau 1 Caractéristiques des échantillons de calibration et de validation.

Calibration	(37 bassins versants)	Validation	(37 bassins versants)
$\log S$	moyenne	1,46	1,34
	écart type	0,55	0,63
Ig	moyenne	11,5	13,7
	écart type	12,2	10,4
Ci	moyenne	21,0	20,1
	écart type	7,2	7,5
$Kr70$	moyenne	31,7	32,8
	écart type	17,3	18,0

Méthodologie

- (a) Dans une première étape, nous calibrons le modèle flou par les opérations suivantes:
- classification floue des variables ($\log S$, Ig , Ci , $Kr70$) et définition des fonctions d'appartenance;
 - estimation, à partir des abaques, des règles traduisant les relations entre variables; et
 - application des règles et « déflouification » du $Kr70$.

Dans ce contexte, la paramétrisation du modèle est réalisée par la définition des intervalles des classes floues, par le choix des fonctions d'appartenance et par le choix d'une méthode de « déflouification ». La calibration est manuelle, de type essai/erreur et repose sur le jeu de données de calibration précédemment décrit.

- (b) Dans l'étape de validation, nous travaillons sur le deuxième échantillon de données:
- comparaison avec les « observations » de Rodier; et
 - comparaison avec les résultats de la méthode Orstom (FAO, 1995) basée sur une détermination directe du $Kr70$ à partir des abaques.

Résultats

Dans la phase de calibration, la paramétrisation du modèle a porté sur trois éléments:

- (a) **la méthode de « déflouification »**: nous avons comparé deux méthodes, la moyenne des maximum et le centre de gravité. Il s'avère que les résultats fournis par ces deux

méthodes sont équivalents. Cette relative insensibilité à la méthode de « déflouification » est également confirmée par la littérature (Bardossy & Disse, 1993; Bardossy & Duckstein, 1995);

- (b) **la forme des fonctions d'appartenance**: nous avons testé deux cas extrêmes où toutes les fonctions sont soit triangulaires soit trapézoïdales. La Fig. 3 montre la configuration retenue dans le cas de fonctions trapézoïdales. Les simulations effectuées révèlent une mauvaise adéquation de la forme triangulaire qui sous-estime le *Kr70* flou. Ceci s'explique par le fait que cette forme minimise les valeurs des fonctions d'appartenance, et donc les poids d'application des règles. En effet, une forme triangulaire traduit un caractère flou assez prononcé, ce qui n'est pas le cas pour toutes les variables. Ainsi, ce test est-il trop schématique: dans l'avenir, il faudra pouvoir estimer le caractère incertain ou imprécis de chaque variable et lui associer la fonction d'appartenance la plus cohérente; et
- (c) **le nombre de classes**: nous avons retenu deux situations, les variables *Ci* et *Ig* étant réparties en cinq classes:
- trois classes pour $\log S$ et cinq classes pour *Kr70*, soit 75 règles; et
 - deux classes pour $\log S$ et trois classes pour *Kr70*, soit 50 règles.

Les résultats (voir Fig. 4) montrent qu'un nombre de classes peu élevé induit une dispersion plus grande. Par contre, et ceci est le plus intéressant, la dispersion a lieu autour de la bissectrice, ce qui n'est pas le cas lorsque le nombre de classes est

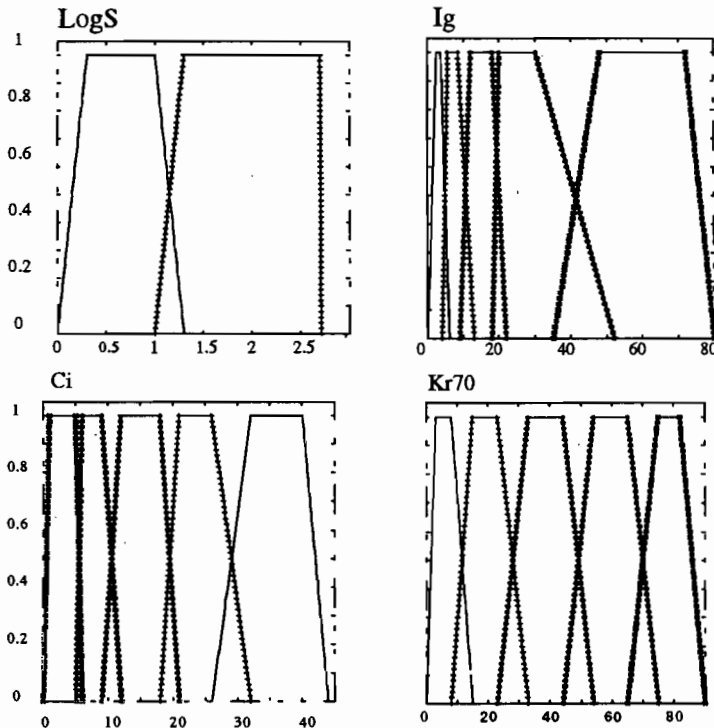


Fig. 3 Fonctions d'appartenance trapézoïdales pour les variables « logarithme de la surface » ($\log S$); indice de pente (*Ig*); classe d'infiltrabilité (*Ci*); et coefficient de ruissellement (*Kr70*).

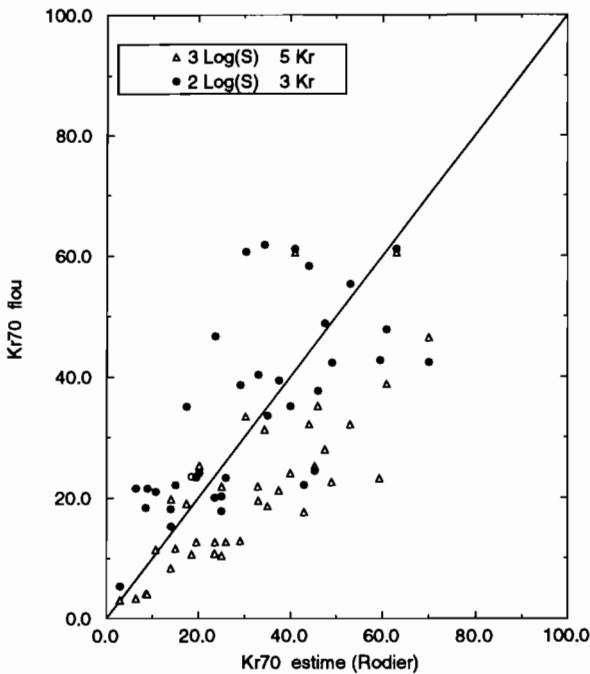


Fig. 4 Phases de calibration: influence du nombre de classes floues pour les variables $\log S$ et $Kr70$.

important. En d'autres termes, complexifier le modèle en multipliant les classes n'apporte pas de gain en précision. Par ailleurs l'accroissement correspondant du nombre de règles rend celles-ci délicates à générer.

Nous avons finalement retenu la paramétrisation suivante:

- méthode de « déflouification »: le centre de gravité;
- fonction d'appartenance: trapézoïdales; et
- nombre total de classes: 15, dont deux pour la surface, cinq pour l'indice de pente et la l'infiltrabilité, et trois pour le coefficient de ruissellement.

Les Figs 5 et 6 montrent les résultats finaux des phases de calibration et de validation du modèle flou par rapport aux données estimées par Rodier. Par ailleurs, nous mettons en parallèle les résultats obtenus avec la méthode Orstom où le $Kr70$ est déterminé par les abaques.

Malgré une dispersion plus importante des résultats dans la phase de validation, la paramétrisation issue de la calibration est confirmée. Cette dispersion, également observable pour la méthode Orstom, semble caractéristique de la répartition des données de l'échantillon.

Par ailleurs, la méthode Orstom donne des résultats relativement meilleurs que le modèle flou. Dans l'état actuel des travaux, il est difficile d'expliquer l'origine de ces écarts.

Le calcul du $Kr70$ pour des valeurs extrêmes pose problème du fait de l'apparition de seuils: à ce niveau, il conviendra de caractériser plus finement les fonctions d'appartenance. De plus, le système d'établissement des règles pose problème: en effet,

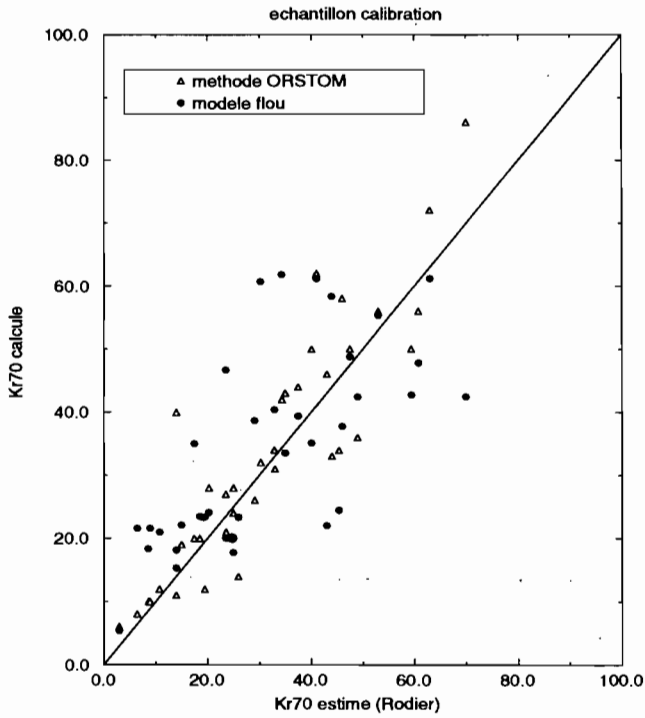


Fig. 5 Phases de calibration: comparaison des résultats issus du modèle flou et de la méthode Orstom avec les valeurs « estimées » par Rodier.

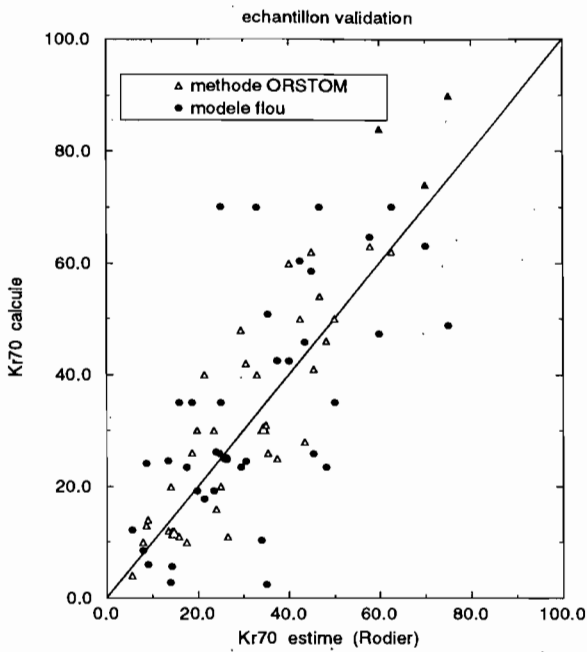


Fig. 6 Phases de validation: comparaison des résultats issus du modèle flou et de la méthode Orstom avec les valeurs « estimées » par Rodier.

il a été établi à partir des abaques utilisés dans la méthode Orstom, et, de ce fait, créée un biais méthodologique dans la comparaison des deux modèles.

CONCLUSION

Cette première étude concernant la modélisation floue a permis de mettre en oeuvre les théories du flou dans la problématique hydrologique du calcul du coefficient de ruissellement décennal en Afrique de l'Ouest. Les premiers résultats sont encourageants: la facilité de mise en oeuvre des algorithmes, la non-nécessité de multiplier le nombre de classes et l'utilisation de techniques élémentaires de « déflouification » contribue à la simplicité du système.

Certes, un certain nombre d'améliorations sont à apporter. La forme des fonctions d'appartenance doit être reliée à la nature des variables, la définition des règles doit être affinée, soit par une approche statistique, soit par une approche de type expertise avec un hydrologue ayant une bonne connaissance du milieu. Il reste également à compléter cette étude par le calcul du Kr_{100} et du Kr_{10} .

Néanmoins l'approche floue est bien adaptée à la résolution du problème hydrologique posé: par sa facilité de mise en oeuvre, et par sa capacité à s'adapter au raisonnement qualitatif de l'hydrologue-expert tout en intégrant les aspects déterministes et quantitatifs de la méthode Orstom de prédétermination du coefficient de ruissellement décennal.

REFERENCES

- Bardossy, A., Bogardi, I. & Duckstein, L. (1990) Fuzzy regression in hydrology. *Wat. Resour. Res.* **26**(7), 1497-1508.
- Bardossy, A. & Disse, M. (1993) Fuzzy rule-based models for infiltration. *Wat. Resour. Res.* **29**(2), 373-382.
- Bardossy, A. & Duckstein, L. (1995) *Fuzzy Rule-based Modeling with Applications to Geophysical, Biological and Engineering Systems*. CRC Press.
- Droesen, W. J. & Geelen, L. H. W. T. (1993) Application of fuzzy sets in ecohydrological expert modelling. In: *Application of Geographic Information Systems in Hydrology and Water Resources Management* (Proc. HydroGIS 93 Conf., Vienna, April 1993) (ed. par K. Kovar & H. P. Nachtnebel), 3-11. IAHS Publ no. 211.
- FAO (1995) Manuel de prédétermination des apports annuels et des crues décennales sur les petits bassins versants en Afrique sahélienne et tropicale sèche. CIEH/Orstom/Cemagref, Rome, (à paraître). Logique Floue, Observatoire français des techniques avancées, Paris.
- Ribstein, P. & Rodier, J. A. (1989) Fortes crues des petits cours d'eau du Sahel: forme de l'hydrogramme. *Hydrol. Continent.* **4**(1), 33-47.
- Rodier, J. A. & Auvray, C. (1965) *Estimation des débits de crue décennale pour les petits bassins versants de superficie inférieure à 200 km² en Afrique occidentale*. CIEH/Orstom, Paris.
- Rodier, J. A. (1986) Caractéristiques des crues des petits bassins versants au Sahel. *Cah. Orstom, Sér. Hydrol.* **21**(2), 1-25.
- Rodier, J. A. & Ribstein, P. (1986) Transposition des données hydrologiques, utilisation des bassins représentatifs pour la prédétermination des crues et l'estimation des apports au Sahel. *Colloque International sur la Révision des Normes Hydrologiques*. CIEH, Ouagadougou.
- Rodier, J. A. & Ribstein, P. (1988) Estimation des caractéristiques de la crue décennale pour les petits bassins versants du Sahel couvrant 1 à 10 km². Rapport interne, Orstom, Montpellier.
- Rodier, J. A. (1992) *Du simulateur de pluie aux bassins représentatifs sahéliens*. Orstom, Coll. Etudes et Thèses.
- Rodier, J. A. (1993) Paramètres caractéristiques des fortes crues dans les régions tropicales sèches. Première partie. Coefficient de ruissellement. *Hydrol. Continent.* **8**(2), 139-160.
- Rodier, J. A. (1995) Paramètres caractéristiques des fortes crues dans les régions tropicales sèches. Deuxième partie. Fonction de transfert (temps de base, coefficient de pointe, temps caractéristique de base). *Hydrol. Continent.* **9**(1), 33-68.

- Vuillaume, G. (1974) L'abattement des précipitations journalières en Afrique intertropicale, variabilité et précision du calcul. *Cah. Orstom, sér. Hydrol.* **11**(3), 205-240.
- Zadeh, L. A. (1965) Fuzzy sets. *Inform. Control* **8**, 338-353.