

## **Etude de la variabilité spatiale des régimes hydrologiques: comparaison de différents modèles d'homogénéité régionale**

**TAHA B. M. J. OUARDA, BERNARD BOBEE & PETER RASMUSSEN**

*Chaire en Hydrologie Statistique, INRS-Eau, 2800 Einstein, CP 7500, Sainte-Foy, Québec, Canada G1V 4C7*

e-mail: [bernard\\_bobee@inrs-eau.quebec.ca](mailto:bernard_bobee@inrs-eau.quebec.ca)

**Résumé** Le développement de méthodologies d'étude de l'homogénéité spatiale qui sont particulièrement adaptées au type de variabilité hydrologique rencontrée en Afrique est d'une grande importance. L'objectif de ce travail est de présenter une comparaison des méthodes les plus utilisées de détermination des régions hydrologiquement homogènes. Deux types de régions sont considérés: les régions fixes qui peuvent être représentées par des régions géographiquement contiguës ou non-contiguës, et les régions du type "voisinage" qui peuvent être déterminées par la méthode des régions d'influence ou par analyse canonique des corrélations. Les résultats de l'intercomparaison démontrent la supériorité des régions du type voisinage. Des recommandations sont faites sur la base des conclusions de cette recherche, et des pistes de recherche future sont identifiées afin de développer des méthodologies qui soient particulièrement adaptées au type de variabilité hydrologique rencontrée en Afrique.

### **INTRODUCTION**

La détermination de régions hydrologiques homogènes représente un problème important en général et plus particulièrement dans le continent africain caractérisé par une assez grande hétérogénéité spatiale. Les hydrologues ont souvent recours à l'identification de régions hydrologiques homogènes pour des fins d'estimation régionale. En effet, les réseaux de stations d'acquisition de mesures de débit ne couvrent pas tous les points où il est nécessaire de disposer de données. Le problème du manque de données hydrologiques peut être particulièrement critique en Afrique, où les contraintes budgétaires freinent souvent le développement des réseaux hydrométriques. Les hydrologues doivent alors souvent recourir à des méthodologies d'estimation régionale (ou de régionalisation) qui reposent sur l'hypothèse d'existence de régions hydrologiquement homogènes à l'intérieur desquelles on peut effectuer un transfert d'information des sites où l'information est disponible aux sites où l'on dispose d'une information insuffisante. La qualité de l'estimation régionale obtenue dépend de la validité de l'hypothèse d'homogénéité régionale, et de la qualité et quantité des données disponibles. L'objectif de ce travail est de présenter une évaluation comparative des méthodes de détermination des régions hydrologiquement homogènes les plus utilisées. Le développement de méthodologies d'étude de l'homogénéité spatiale qui sont particulièrement adaptées au type de variabilité

hydrologique rencontrée en Afrique pourra découler des résultats de cette étude. Ceci est d'autant plus important dans le cadre de la sécheresse extrême et prolongée qui est observée dans certaines parties du continent africain et de son impact sur le développement économique et social de la région (Servat & Sakho, 1995).

## METHODES D'ANALYSE HYDROLOGIQUE REGIONALE

Le principe de base des méthodologies d'estimation régionale de variables hydrologiques est d'utiliser l'information spatiale provenant des bassins ayant un régime hydrologique similaire à celui du bassin cible. Ces méthodologies comprennent généralement deux étapes:

- (a) définition et détermination des régions hydrologiquement homogènes: c'est à dire regroupement des stations ayant un comportement hydrologique semblable; et
- (b) estimation régionale: transfert à l'intérieur d'une même région hydrologiquement homogène de l'information des sites jaugés à un site non-jaugé ou partiellement jaugé pour lequel on ne dispose pas d'information suffisante.

Différentes méthodologies d'estimation régionale peuvent alors être définies en considérant les couples formés par les différentes méthodes de détermination des régions homogènes et par les différentes méthodes d'estimation régionale des débits  $Q(T)$  de période de retour  $T$ . Par définition, la performance d'une méthodologie d'estimation régionale dépend des méthodes retenues durant ces deux étapes. Un projet de recherche a été mené par une équipe de scientifiques canadiens (GREHYS, 1996a, b), afin de comparer différents modèles d'estimation régionale des crues.

## APERÇU SUR LES METHODES DE DETERMINATION DES REGIONS HYDROLOGIQUEMENT HOMOGENES

La détermination des *régions homogènes* consiste à regrouper, pour une variable hydrologique donnée, les stations ayant le même comportement hydrologique. La définition d'homogénéité reste cependant assez vague et subjective, et souvent les données disponibles ne permettent pas d'effectuer des tests statistiques concluants. On distingue deux types d'approche pour la détermination des régions homogènes:

- (a) **Approche basée sur des régions fixes:** ensemble de stations formant une même région homogène. Deux sous-groupes peuvent être considérés:
  - *Régions géographiquement contiguës:* Ces régions peuvent être définies par exemple à partir de la similarité des densités non-paramétriques des débits de crue (Hosking & Wallis, 1993; Adamowski et al., 1994).
  - *Régions non-contiguës:* Ces régions peuvent être déterminées par exemple par analyse factorielle des correspondances et classification ascendante hiérarchique (Lebart & Fenelon, 1975).
- (b) **Approche basée sur des régions du type voisinage:** où on associe à chaque station cible son propre voisinage. La méthode des *Régions d'influence* (Burn, 1988, 1990; Zrinji & Burn, 1994) définit la proximité de deux sites par une distance dans un espace multidimensionnel dont les axes sont des caractéristiques

hydrologiques, physiographiques et météorologiques. Les voisinages peuvent aussi être définis par *analyse des corrélations canoniques* (Cavadias, 1989).

La Fig. 1 illustre les différentes approches pour la détermination des régions homogènes. La définition des régions hydrologiques homogènes sur la base de la proximité géographique a été d'abord considérée (Dalrymple, 1960). Cette approche peut être justifiée si les facteurs qui expliquent le régime d'écoulement, tels que le type de sol ou l'altitude, présentent une cohérence spatiale. En pratique, ceci n'est généralement pas le cas pour les variables physiographiques explicatives (telles que l'aire du bassin versant ou sa forme) et des améliorations substantielles sont alors possibles en considérant des régions du type voisinage. Réalisant que la proximité géographique des bassins versants n'est pas une garantie de leur similarité hydrologique, plusieurs auteurs ont contesté le concept de régions géographiques homogènes (Cavadias, 1989). Les méthodes utilisées dans cette étude sont:

- (a) Méthode des régions d'influence de Burn ( $DRH_{11}$ ) pour le cas des stations cibles jaugées.
- (b) Méthode des régions d'influence modifiée ( $DRH_{12}$ ) pour le cas des stations cibles jaugées.
- (c) Méthode des régions d'influence pour sites non-jaugés ( $DRH_{13}$ ).
- (d) Méthode d'analyse canonique des corrélations de Cavadias ( $DRH_2$ ). Cette méthode s'applique aux sites jaugés et non-jaugés.
- (e) Analyse factorielle des correspondances et classification ascendante hiérarchique ( $DRH_3$ ). Cette méthode s'applique aux sites jaugés et non-jaugés.
- (f) Détermination des régions à partir des L-moments ( $DRH_4$ ). Cette méthode s'applique aux sites jaugés et non-jaugés.

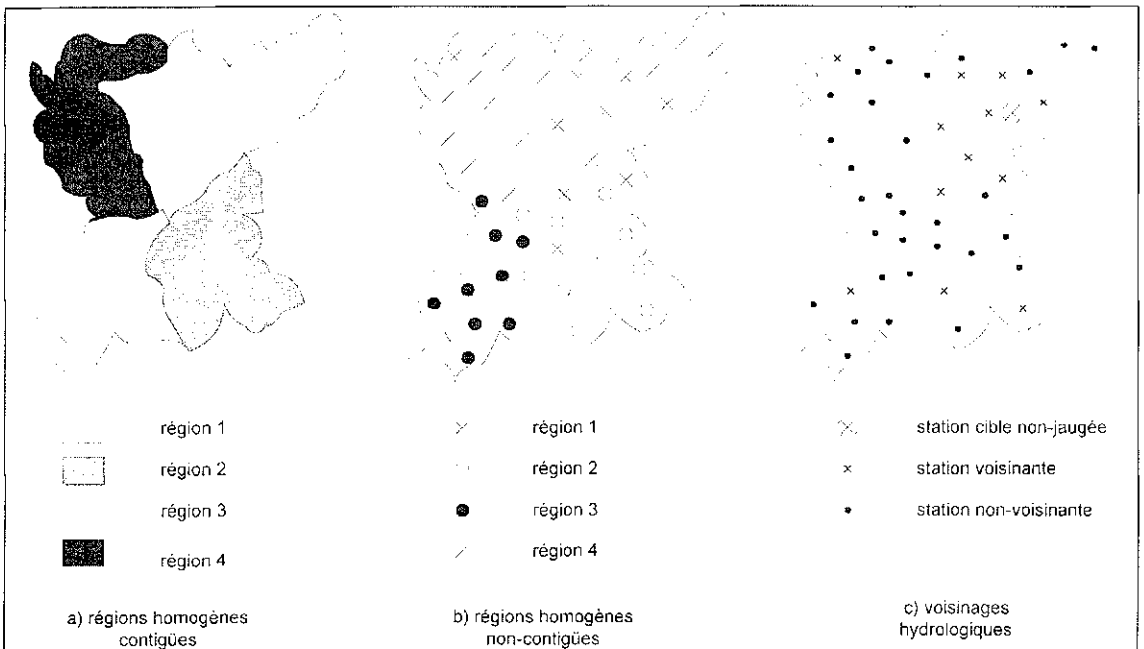


Fig. 1 Approches pour la détermination des régions homogènes.

Nous considérons donc dans le cas des sites jaugés cinq méthodes pour la détermination des régions homogènes. Dans le cas des sites non-jaugés quatre méthodes sont utilisées. Les détails de ces méthodes sont décrits dans GREHYS (1996a). La méthode des régions d'influence et la méthode d'analyse des corrélations canoniques sont présentées brièvement dans ce qui suit.

### Méthode des régions d'influence

La méthode des régions d'influence est une méthode de détermination des régions "homogènes" au sens hydrologique. Selon cette approche, chaque site peut être considéré comme le centre d'une région formée de sites dont les caractéristiques de crue sont similaires. Cette approche implique généralement que l'on dispose d'un minimum d'information hydrologique au site pour lequel on souhaite définir la région d'influence. Dans le cas contraire, les variables climatologiques et physiographiques susceptibles de pallier ce manque de données peuvent être identifiées et utilisées comme supplément d'information. On calcule ensuite les distances entre les sites dans l'univers des variables (attributs) sélectionnées. Dans l'espace des attributs physiographiques et/ou hydrologiques on définit la distance entre deux sites  $i$  et  $j$  par:

$$D_{ij} = \sqrt{\sum_{m=1}^M \omega_m (C_m^i - C_m^j)^2} \quad (1)$$

où  $M$  est le nombre d'attributs considérés,  $C_m^i$  est la valeur standardisée de l'attribut d'ordre  $m$  au site  $i$ , et  $\omega_m$  est le poids associé à l'attribut  $C_m$ . Plusieurs variantes de la méthode ont été présentées, en fonction des attributs utilisés et selon que le site cible est partiellement jaugé ou non-jaugé (Burn, 1988, 1990; Zrinji & Burn, 1994; GREHYS, 1996a). Les poids  $\omega_m$  permettent d'accorder davantage d'importance à un attribut de façon à tenir compte de l'importance relative de certaines variables descriptives dans la réponse hydrologique du bassin. La région d'influence  $ROI_i$  du site  $i$  est définie par l'ensemble des sites  $j$  satisfaisant:

$$ROI_i = \left\{ \text{site } j \mid \text{tel que } D_{ij} \leq \theta_j \right\} \quad (2)$$

où  $\theta_j$  est un seuil qui permet d'obtenir un compromis entre le nombre de stations voisines à considérer dans l'analyse régionale et l'homogénéité hydrologique du groupe de stations voisines sélectionnées. Une fonction de pondération peut aussi être calculée pour refléter la proximité relative des différents sites voisins du site cible. Une description plus détaillée de la méthode est présentée dans Burn (1990) et GREHYS (1996a).

### Méthode d'analyse des corrélations canoniques

La méthode d'analyse des corrélations canoniques permet d'identifier les sites dont le régime de crues est similaire au site cible (Cavadias, 1989; Ribeiro-Corréa et al., 1995; Ouarda et al., 1997). L'analyse des corrélations canoniques est un outil

d'analyse statistique multivariée qui permet de décrire la relation de dépendance existant entre deux ensembles de variables aléatoires. Cette méthode permet de déterminer des paires de combinaisons linéaires de chaque ensemble de variables, que l'on appelle des variables canoniques, telles que la corrélation entre les variables canoniques d'une paire est maximisée, et la corrélation entre les variables de paires différentes est nulle. On obtient ainsi un ensemble de variables canoniques pour les deux ensembles de variables aléatoires associées à des coefficients de corrélation canonique. Il est alors possible d'inférer sur les variables canoniques d'un ensemble connaissant les variables canoniques de l'autre ensemble. On peut aussi calculer une distance entre deux variables canoniques ce qui permet de déterminer des voisinages hydrologiques.

On note  $\mathbf{Y}'=(Y_1, \dots, Y_p)$  et  $\mathbf{X}'=(X_1, \dots, X_q)$  respectivement l'ensemble des variables hydrologiques et l'ensemble des variables caractérisant la géomorphologie et la météorologie des bassins versants. On note aussi  $\mathbf{W}$  le vecteur de variables canoniques hydrologiques,  $\mathbf{V}$  le vecteur de variables canoniques physiographiques,  $(\lambda_1, \dots, \lambda_p)$  les coefficients de corrélation canonique, et  $\Lambda$  la matrice diagonale formée par les coefficients de corrélation canonique. Pour les bassins non-jaugés, la position moyenne dans l'espace canonique hydrologique est déterminée, en utilisant une estimation de  $\Lambda$ , par  $\Lambda \mathbf{V}_0$ , où  $\mathbf{V}_0$  est le vecteur canonique physiographique connu. On peut ainsi définir le voisinage d'un bassin non-jaugé à un niveau de confiance  $(1 - \alpha)$  par l'ensemble de bassins dont la position  $\mathbf{W}$  dans l'espace canonique hydrologique vérifie la relation:

$$(\mathbf{W} - \hat{\Lambda} \mathbf{V}_0)' (I_p - \hat{\Lambda} \hat{\Lambda}')^{-1} (\mathbf{W} - \hat{\Lambda} \mathbf{V}_0) \leq \chi_{\alpha, p}^2 \quad (3)$$

où  $I_p$  est la matrice identité d'ordre  $p$ . Ouarda *et al.* (1997) ont indiqué que la méthode d'analyse des corrélations canoniques n'est pas sensible à des facteurs tels que la méthode d'estimation au site ou la configuration du réseau hydrométrique.

## METHODOLOGIE DE COMPARAISON ET RESULTATS

Les différents modèles sont comparés en utilisant comme référence les quantiles de crue estimés localement pour 33 sites cibles qui ont 35 années de données ou plus et pour lesquels toutes les variables physiographiques et météorologiques sont disponibles. Les sites cibles ont été choisis, parmi les 179 stations de l'étude (provinces du Québec et de l'Ontario). La méthodologie d'intercomparaison a été appliquées aux deux cas d'estimation régionale: le cas de stations cibles non-jaugées et le cas de stations cibles jaugées avec peu de données. Pour chaque cas, nous avons calculé l'indice de performance pour les quantiles de période de retour 10 et 100 ans, estimés à l'aide d'un modèle régional. Les résultats présentés dans cet article sont basés sur l'utilisation de la méthode de l'indice de crue pour l'estimation régionale. D'autres méthodes d'estimation régionale ont aussi été considérées. Une synthèse de tous les résultats obtenus est disponible dans GREHYS (1996b).

L'approche bayésienne pour la comparaison de la performance des différentes méthodes est basée sur les concepts de "super population" et d'échangeabilité. Dans l'approche bayésienne on calcule la distribution a posteriori du quantile de crue

considéré. Cette distribution tient compte à la fois de l'information contenue dans l'estimation locale et dans l'estimation régionale. L'introduction d'une fonction de coût permet alors de calculer l'espérance du coût d'une estimation régionale donnée. En faisant la moyenne sur tous les sites cibles considérés, on obtient l'indice de performance suivant:

$$AEPL = \frac{1}{M} \sum_{i=1}^M \frac{\text{Var}\{Q^i | \hat{Q}_s^i\} + [E\{Q^i | \hat{Q}_s^i\} - \hat{Q}_r^i]^2}{[E\{Q^i | \hat{Q}_s^i\}]^2} \tag{4}$$

où

$$E\{Q^i | \hat{Q}_s^i\} = \hat{Q}_s^i - \frac{\sigma_i^2}{\sigma_i^2 + V_i^2} (\hat{Q}_s^i - \underline{X}_i \underline{\alpha}) \tag{5}$$

$$\text{Var}\{Q^i | \hat{Q}_s^i\} = \frac{\sigma_i^2 V_i^2}{\sigma_i^2 + V_i^2} \tag{6}$$

et:

$Q^i$  quantile correspondant à une période de retour  $T = 10$  ou  $100$  ans. Cette quantité est considérée une variable aléatoire dans un contexte bayésien

$\hat{Q}_s^i$  estimation du quantile à partir de l'information locale

$\underline{\alpha}$  paramètres de régression qui mesurent le biais de l'estimateur régional

$V_i^2$  variance de l'estimation régionale à site  $i$

$\sigma_i^2$  variance de l'estimation locale à site  $i$

$\underline{X}$  vecteur qui contient les éléments  $(1, \hat{Q}_r^i)$

La Fig. 2 illustre la valeur de l'indice de performance  $AEPL$  pour les différentes méthodes dans le cas des sites cibles jaugés. Ces résultats indiquent que, pour la détermination des régions homogènes, la méthode des régions d'influence ( $DRH_{11}$  et  $DRH_{12}$ ) et la méthode d'analyse des corrélations canoniques ( $DRH_2$ ) qui correspondent à des plus faibles valeurs d' $AEPL$  se distinguent des autres. En général, les résultats de l'intercomparaison ont montré que la méthode des régions

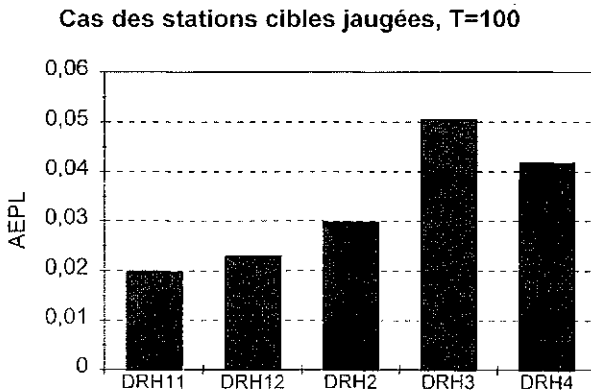


Fig. 2 Performance des différentes méthodes dans le cas des sites cibles jaugés.

d'influence et la méthode d'analyse des corrélations canoniques utilisées pour déterminer les régions homogènes produisent de bonnes estimations des quantiles de crue à des sites jaugés ou non-jaugés, quelle que soit la méthode d'estimation utilisée ensuite dans la région. Ces résultats démontrent la nette supériorité des régions du type voisinage. L'adaptation de ces méthodes peut mener au développement de méthodologies qui sont particulièrement adaptées au type de variabilité hydrologique rencontrée en Afrique. C'est pourquoi il est recommandé de valider l'utilité de ces méthodes aux données du continent africain.

**Remerciements** Les auteurs souhaitent remercier le Conseil de Recherches en Sciences Naturelles et Génie du Canada pour avoir financé cette recherche. Les auteurs souhaitent aussi remercier leurs collègues qui ont participé dans ce projet.

## REFERENCES

- Adamowski, K., Gingras, D. & Pilon, P. J. (1994) *Regional Flood Frequency Analysis by Nonparametric and L-Moment Methods for Ontario and Quebec*. Report to NSERC Strategic Grant. Faculty of Engineering, Univ. of Ottawa, Ottawa, Canada.
- Burn, D. H. (1988) Delineation of groups for regional flood frequency analysis. *J. Hydrol.* **104**, 345–361.
- Burn, D. H. (1990) Evaluation of regional flood frequency analysis with a region of influence approach. *Wat. Resour. Res.* **26**(10), 2257–2265.
- Cavadias, G. S. (1989) Regional flood estimation by canonical correlation. Paper presented at: *1989 Annual Conference of the Canadian Society for Civil Engineering* (St-John's, Newfoundland, Canada).
- Dalrymple, T. (1960) Flood frequency methods. *US Geol. Survey Wat. Supply Pap.* **1543A**, 11–51.
- GREHYS (Groupe de recherche en hydrologie statistique) (1996a) Presentation and review of some methods for regional flood frequency analysis. *J. Hydrol.* **186**, 63–84.
- GREHYS (Groupe de recherche en hydrologie statistique) (1996b) Inter-comparison of regional flood frequency procedures for Canadian rivers. *J. Hydrol.* **186**, 85–103.
- Hosking, J. R. M. & Wallis, J. R. (1993) Some statistics useful in regional frequency analysis. *Wat. Resour. Res.* **29**(2), 271–281.
- Lebart, L. & Fenelon, J. P. (1975) *Statistiques et informatique appliquées*. Dunod, Paris, France.
- Ouarda, T. B. M. J., Boucher, G., Rasmussen, P. F. & Bobee, B. (1997) Regionalization of floods by canonical correlation analysis. In: *Operational Water Management* (ed. by J. C. Refsgaard & E. A. Karalis), 297–302. A. A. Balkema, Rotterdam, The Netherlands.
- Ribeiro-Correa, J., Cavadias, G. S., Clement, B. & Rousselle, J. (1995) Identification of hydrological neighborhoods using canonical correlation analysis. *J. Hydrol.* **173**, 71–89.
- Servat, E. & Sakho, M. (1995) Instability of water resources and management of a planned water system in non-Saharan West Africa. *Hydrol. Sci. J.* **40**(2), 217–230.
- Zrinji, Z. & Burn, D. H. (1994). Flood frequency analysis for ungauged sites using a region of influence approach. *J. Hydrol.* **153**, 1–21.