

**UNE MÉTHODE BAYÉSIENNE DE  
COMBINAISON DE L'INFORMATION  
LOCALE ET RÉGIONALE EN ANALYSE  
FRÉQUENTIELLE DES CRUES**

*Rapport de recherche No R-789*

*Janvier 2005*

**UNE MÉTHODE BAYÉSIENNE DE  
COMBINAISON DE L'INFORMATION  
LOCALE ET RÉGIONALE EN ANALYSE  
FRÉQUENTIELLE DES CRUES**

*Par :*

**Ousmane Seidou  
Taha B.M.J Ouarda  
Marc Barbet  
Pierre Bruneau  
Bernard Bobée**

**Chaire industrielle Hydro-Québec/CRSNG en hydrologie statistique**

**INRS-ETE**

**490, rue de la Couronne, Québec (Québec) G1K 9A9**

*Rapport présenté à :*

**HYDRO-QUÉBEC**

**Rapport de recherche R-789**

**Janvier 2005**

*ISBN : 2-89146-498-2*

## TABLE DES MATIÈRES

---

Table des matières .....	2
Liste des tableaux .....	3
Liste des figures .....	4
RÉSUMÉ .....	6
Introduction.....	8
1. REVUE DE LITTÉRATURE .....	13
2. L'APPROCHE BAYÉSIENNE .....	16
3. LE MODÈLE RÉGIONAL.....	19
4. MÉTHODOLOGIE DE VALIDATION.....	23
5. LES SITES D'APPLICATION.....	24
6. RÉSULTATS ET DISCUSSIONS .....	24
7. CONCLUSION .....	29
8. BIBLIOGRAPHIE .....	29

---

## LISTE DES TABLEAUX

---

Tableau 1: résultats de l'analyse régionale (rivière Harricana, station 080101).....	37
Tableau 2 : résultats de l'analyse régionale (rivière Matane, station 023401).....	38
Tableau 3 : résultats de l'analyse régionale (rivière Beaurivage, station 021601).....	39
Tableau 4 : <i>REQM</i> des estimateurs des paramètres de la distribution GEV (rivière).....	40
Tableau 5 : <i>REQM</i> des estimateurs des paramètres de la distribution GEV (rivière Matane).....	41
Tableau 6 : <i>REQM</i> des estimateurs des paramètres de la distribution GEV (rivière Beaurivage).....	42
Tableau 7 : <i>REQM</i> et biais des estimateurs des quantiles (rivière Harricana).....	43
Tableau 8 : <i>REQM</i> et biais des estimateurs des quantiles (rivière Matane).....	44
Tableau 9 : <i>REQM</i> et biais des estimateurs des quantiles (rivière Beaurivage).....	45

## LISTE DES FIGURES

---

Figure 1: débits maximaux annuels observés aux trois stations: a) Harricana; b) Matane; c) Beaurivage;.....	46
Figure 2: histogramme du quantile $q_{10}$ , comparé à la densité de probabilité (mise à l'échelle) des estimateurs local, régional et bayésien empirique (rivière Harricana): a) 1915-1934; b) 1924-1944; c) 1935-1954; d) 1945-1964; e) 1955-1974; f) 1965-1984. ....	47
Figure 3: histogramme du quantile $q_{100}$ , comparé à la densité de probabilité (mise à l'échelle) des estimateurs local, régional et bayésien empirique (rivière Harricana): a) 1915-1934; b) 1924-1944; c) 1935-1954; d) 1945-1964; e) 1955-1974; f) 1965-1984. ....	48
Figure 4: histogramme du quantile $q_{1000}$ , comparé à la densité de probabilité (mise à l'échelle) des estimateurs local, régional et bayésien empirique (rivière Harricana): a) 1915-1934; b) 1924-1944; c) 1935-1954; d) 1945-1964; e) 1955-1974; f) 1965-1984. ....	49
Figure 5: variation des paramètres de la distribution GEV en fonction de la période de 20 années consécutives choisie (rivière Harricana): a) $\mu$ ; b) $\sigma$ ; c) $\xi$ . ....	50
Figure 6: variation des paramètres de la GEV en fonction de la période de 20 années consécutives choisie (rivière Matane): a) $\mu$ ; b) $\sigma$ ; c) $\xi$ . ....	51
Figure 7: variation des paramètres de la distribution GEV en fonction de la période de 20 années consécutives choisie (Rivière Beaurivage): a) $\mu$ ; b) $\sigma$ ; c) $\xi$ .....	52
Figure 8: variation du quantile $q_{10}$ estimé en fonction de la période de 20 années consécutives choisie (rivière Harricana). ....	53
Figure 9: variation du quantile $q_{100}$ estimé en fonction de la période de 20 années consécutives choisie (Rivière Harricana). ....	54
Figure 10: variation du quantile $q_{1000}$ estimé en fonction de la période de 20 années consécutives choisie (Rivière Harricana). ....	55
Figure 11: variation du quantile $q_{10}$ estimé en fonction de la période de 20 années consécutives choisie (rivière Matane). ....	56

Figure 12: variation du quantile  $q_{100}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Matane)..... 57

Figure 13: variation du quantile  $q_{1000}$  estimé en fonction de la période de 20 années consécutives choisie (Rivière Matane). ..... 58

Figure 14: variation du quantile  $q_{10}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Beaurivage)..... 59

Figure 15: variation du quantile  $q_{100}$  estimé en fonction de la période de 20 années consécutives choisie (Rivière Beaurivage). ..... 60

Figure 16: variation du quantile  $q_{1000}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Beaurivage)..... 61

Figure 17: Paramètre de forme  $\hat{\xi}$  de la distribution GEV calculé à partir de  $\frac{q_{1000} - q_{10}}{q_{100} - q_{10}}$  ..... 62

## RÉSUMÉ

---

Du fait de leur impact sur la conception des ouvrages hydrauliques ainsi que sur l'aménagement des plaines inondables, les quantiles de crue doivent être estimés avec autant de précision que possible. Lorsqu'on dispose d'une série de données historiques relativement longue, on peut utiliser l'analyse fréquentielle locale et obtenir des quantiles avec une bonne précision. Lorsque trop peu d'observations sont disponibles, on utilise généralement une méthode d'estimation régionale pour pallier au manque de données, ignorant ainsi l'information locale. Une méthode communément utilisée pour combiner les deux types d'information consiste à utiliser un estimateur bayésien empirique linéaire : sous réserve que les deux estimateurs aient une distribution normale, l'estimateur bayésien empirique du vrai quantile est la moyenne pondérée des deux estimations. Le facteur de pondération de chaque estimateur est inversement proportionnel à sa variance.

On propose dans cet article une méthode bayésienne de combinaison de l'information locale et régionale. L'application est faite avec la loi GEV, mais elle peut facilement être généralisée à d'autres types de distributions. Contrairement à l'approche bayésienne empirique, elle fonctionne même avec une seule observation locale en plus de relaxer l'hypothèse de normalité de la distribution des quantiles locaux. Les

deux approches sont comparées sur de courtes séries (20 années consécutives) extraites des données historiques des rivières Harricana (84 ans), Beaurivage (76 ans) et Matane (74 ans) situées dans la province du Québec, Canada. Les résultats montrent l'effet stabilisateur de la méthode présentée sur les paramètres de la loi GEV lorsqu'ils sont estimés avec de courtes séries. De plus, les estimateurs construits avec la nouvelle méthode donnent des valeurs plus proches des quantiles estimés avec la totalité des données disponibles au site pour les deux sites ayant un paramètre de forme de la GEV proche du paramètre de forme moyen observé au Québec.

Mots clés : régionalisation, distribution GEV, régression linéaire, méthode bayésienne empirique, analyse fréquentielle des crues, combinaison.

## **INTRODUCTION**

---

Selon la quantité et la nature des données disponibles, les quantiles de crue à un site peuvent être estimés en utilisant l'analyse fréquentielle locale, l'analyse fréquentielle régionale ou une combinaison des deux. Les propriétés statistiques des distributions pouvant décrire les débits de pointe en rivière ont été intensivement étudiées pendant les dernières décades, mais le manque de séries de données de longueur suffisante limite toujours la précision des résultats (Bobée et Rasmussen, 1995). Les méthodes de régionalisation, introduites de Dalrymple (1960) et continuellement développées depuis par plusieurs chercheurs (e.g. Benson, 1962; Wallis, 1980; Tasker, 1980; Greis et Wood, 1981; Hosking et al. 1985; Burn, 1990a,b; GREHYS 1996a,b; Ribeiro-Corréa, 1995; Ouarda et al. 2001), permettent de pallier au manque de données en transposant l'information des sites jaugés vers les sites non-jaugés d'intérêt. Il est cependant difficile de déterminer à quel moment on a suffisamment de données pour privilégier l'information locale par rapport à l'information régionale. De ce fait, Matalas et Gilroy (1968) recommandaient de choisir l'estimateur qui avait la moindre variance. Il serait cependant plus judicieux de les combiner systématiquement dans le souci d'avoir une estimation plus précise de la quantité hydrologique à

estimer, en faisant cependant attention au fait que dans une région hétérogène, l'ajout de l'information régionale peut introduire un biais et de ce fait être contre-productif.

Deux types d'approches ont été utilisées dans la littérature hydrologique pour faire la combinaison de l'information régionale et locale: a) une approche mixte qui consiste à estimer certains paramètres avec les données locales et d'autres avec les données régionales, b) les approches bayésiennes qui peuvent consister soit en la construction d'un estimateur bayésien empirique des quantités à estimer, soit c) l'estimation bayésienne complète des distributions a posteriori de façon analytique ou numérique. Selon le type de distributions choisies pour les estimateurs locaux et régionaux, les distributions a posteriori peuvent être obtenues analytiquement (cas idéal), ou via une intégration numérique plus ou moins complexe.

La méthode régionale d'indice de crue (Dalrymple, 1960; NERC, 1975) fait partie des approches mixtes lorsqu'elle est appliquée à un site jaugeé puisque le débit moyen au site est estimé avec les données locales, tandis que les paramètres de la distribution du quantile normalisé le sont avec les données régionales. Lettenmaier et al., 1987 ont montré par simulation que si la distribution régionale sous-jacente est la loi GEV, et si les paramètres de cette loi sont estimés avec les L-moments ou les moments pondérés (Probability Weighted Moments, PWM), alors la

méthode régionale d'indice de crue est plus efficace que l'estimation locale, même en cas d'hétérogénéité régionale modérée. Un autre exemple d'approche mixte est la méthode GEV/PWM à deux paramètres où le paramètre de forme de la distribution GEV est estimé par une approche régionale, les deux autres paramètres avec les données locales. Cette méthode s'est révélée être supérieure à la méthode régionale d'indice de crue GEV/PWM à trois paramètres (pour laquelle tous les trois paramètres sont libres) pour l'estimation du débit centenaire quand le nombre de données historiques augmente, ou quand l'hétérogénéité régionale est significative (Lettenmaier *et al.*, 1987; Stedinger et Lu, 1995; Fill et Stedinger, 1998).

Dans le cas des approches bayésiennes, la connaissance a priori sur les quantités à estimer (paramètres ou quantiles de la distribution locale) est décrite par des densités de probabilité déduites à partir d'analyses régionales. Ces distributions de probabilité sont ensuite mises à jour en utilisant les observations locales avec le théorème de Bayes. Une approche usuelle consiste à attribuer aux deux estimateurs local et régional des distributions judicieusement choisies (appelées distributions conjuguées) de façon à ce que les distributions a posteriori de certaines (ou de toutes) les inconnues puissent être calculées de façon analytique. Une telle approche diminue (ou élimine) la charge de calcul nécessaire pour l'intégration dans des espaces de dimensionnalité élevées qui

résulteraient autrement de la formulation du problème. Par contre, les distributions les plus adéquates pour les phénomènes considérés sont parfois non conjuguées. Un compromis doit alors être fait entre la facilité de la manipulation mathématique et le réalisme du modèle. Des distributions conjuguées ont été utilisées dans les travaux de Shane et Gaver (1970), Vicens *et al.*(1975), Kuczera (1982), Fill et Stedinger (1998) et Fortin *et al.* (1998).

Dans le cas où les distributions des deux estimateurs sont normales, on démontre facilement (voir par ex. GREHYS, 1996b; Bernier, 1993) que le quantile combiné suit également une loi normale dont la moyenne et la variance sont respectivement donnés par:

$$E(q_p | \hat{q}_p^{(L)}, \hat{q}_p^{(R)}) = \frac{\sigma_R^2 \hat{q}_p^{(L)} + \sigma_L^2 \hat{q}_p^{(R)}}{\sigma_R^2 + \sigma_L^2} \quad [1]$$

$$Var(q_p | \hat{q}_p^{(L)}, \hat{q}_p^{(R)}) = \frac{\sigma_R^2 \sigma_L^2}{\sigma_R^2 + \sigma_L^2} \quad [2]$$

où  $q_p$  est le quantile de crue qu'on cherche à estimer,  $\hat{q}_p^{(L)}$  ( resp.  $\hat{q}_p^{(R)}$  ) l'estimation locale ( resp. régionale) de  $q_p$ ,  $\sigma_L$  ( resp.  $\sigma_R$  ) sa variance d'estimation locale ( resp. régionale). L'estimateur présenté à l'équation [1] est également appelé estimateur bayésien empirique linéaire (Linear Empirical Bayes) et son utilisation est suggérée même dans des cas non normaux (Kuczera, 1982).

Beaucoup moins souvent utilisée du fait de sa complexité, l'approche bayésienne complète pour des distributions non normales consiste à calculer la densité de probabilité a posteriori des paramètres et des quantiles. Elle nécessite une intégration numérique et a été utilisée par exemple pour les séries de durées partielles (PDS) avec des dépassements exponentiellement distribués (Rousselle and Hindie, 1976, Madsen *et al.*, 1994). Elle a également été utilisé dans Madsen *et al.*, 1994 et Madsen and Rojsberg (1997) pour des series de durées partielles dont les dépassements suivent la distribution de Pareto généralisée.

On présente dans cet article une méthode bayésienne de combinaison de l'information locale et régionale analogue à celle de Madsen et Rosjberg (1997), mais pour la loi GEV. On fournit aussi les indications nécessaires pour son adaptation à d'autres types de distributions. L'information a priori est spécifiée par la distribution des quantités  $q_{10}$ ,  $q_{100} - q_{10}$ ,  $q_{1000} - q_{100}$  comme proposé dans Coles et Tawn (1996) et Stephenson (2002). Ces quantités hydrologiques seront obtenues grâce à un modèle régional de régression multiple.

Le présent document est divisé en cinq parties. La première partie est une revue de littérature sur les approches bayésiennes de combinaison de l'information locale et régionale. Dans la seconde partie, le modèle

bayésien développé dans cet article est présenté. La troisième partie est consacrée au modèle régional et à la spécification de l'information a priori. Les sites d'application sont présentés dans la quatrième partie tandis que les résultats sont discutés dans la cinquième. Une conclusion est finalement présentée.

## **1. REVUE DE LITTÉRATURE**

---

Kuczera (1982) a utilisé la méthode bayésienne empirique avec des débits de crue supposés suivre la loi lognormale pour stabiliser l'estimation de la variance. L'information a priori est obtenue en ajustant une loi gamma aux variances des débits estimées localement. Il a ensuite montré en considérant des données simulées sans corrélation entre sites que la racine de l'erreur quadratique relative moyenne (*REQRM*) de l'estimation du quantile  $q_{100}$  est ainsi réduite, le gain étant cependant moins élevé à mesure que l'hétérogénéité régionale augmente. Lettenmaier et Potter (1985) ont montré que cette approche était aussi sensible au type de distribution choisi.

Le calcul du *REQRM* dans Kuczera (1982) n'est possible que parce que la vraie valeur du quantile des données simulées est connue. Dans une deuxième application il utilise des données réelles de rivières de la Nouvelle Angleterre. Comme dans ce cas les quantiles réels ne sont pas

connus, il a seulement été capable de montrer que la combinaison de l'information régionale et locale stabilise l'estimation du quantile (i.e. la distribution a posteriori a une variance moins élevée).

Fill et Stedinger (1998) ont utilisé la méthode bayésienne empirique pour combiner le résultat obtenu par régression sur les quantiles normalisés (NQR) avec l'estimateur régional GEV/PWM à deux paramètres. La méthode NQR, suggérée par Stedinger (1989), consiste à estimer le quantile normalisé (le quantile divisé par la moyenne au site) par régression linéaire sur les variables physiographiques. Fill et Stedinger (1998) ont ensuite montré par simulation que l'estimateur bayésien empirique était plus robuste et, en termes de *REQM*, performait aussi bien ou mieux que la méthode NQR ou la méthode GEV/PWM à deux paramètres.

Vicens et al. (1975) ont considéré que le débit moyen annuel des rivières de la Nouvelle Angleterre peuvent être décrits par une loi normale. La connaissance à priori sur la moyenne et la variance de cette loi est obtenue par régression linéaire multiple sur les variables physiographiques. Ils se sont ensuite contentés de discuter de la variation de la forme des distributions a posteriori des débits, de leur moyenne et de leur variance en fonction de la précision des distributions locales et régionales. Il ressort de leur analyse que la combinaison des deux sources d'information réduisait la variance d'estimation autant sur les

paramètres que sur le débit moyen, et que le résultat était dominé par l'estimation qui avait la plus petite variance.

Madsen et Rojsberg (1997) ont quant à eux appliqué deux estimateurs bayésiens des quantiles de crue pour à un modèle régional d'indice de crue dont les dépassements sont supposées suivre la loi de Pareto Généralisée. Le premier est l'estimateur bayésien empirique donné par l'équation [1] tandis que le second est la moyenne de la distribution à posteriori du quantile obtenu avec une approche bayésienne complète. Dans les deux cas, l'information à priori sur les paramètres de la distribution de Pareto généralisée est obtenue par régression linéaire sur les variables physiographiques. Leurs résultats montrent que l'estimateur bayésien empirique donnait des valeurs légèrement plus petites que celles obtenues avec l'approche bayésienne complète. Ils expliquèrent ce fait par l'asymétrie positive introduite par le choix des distributions a priori.

Shane et Gaver (1970) ont considéré que les dépassements au dessus d'un seuil donné suivent une loi exponentielle et ont déduit les équivalents des équations [1] et [2] pour cette loi en cherchant la combinaison linéaire des estimations locale et régionale qui donnaient le plus petit *REQM*. Ils ont également considéré une approche bayésienne où l'information à priori pour la loi exponentielle décrivant l'amplitude des dépassements est décrite par une loi Gamma. La moyenne et la

variance de la distribution a priori sont obtenues par régression linéaire régionale. Ils ont ensuite comparé l'impact de l'utilisation de l'un ou l'autre des deux estimateurs sur le calcul de la hauteur optimale d'une digue de protection contre les crues, et trouvé que les deux méthodes donnent essentiellement le même résultat.

D'autres travaux traitant de la combinaison bayésienne de l'information locale et régionale peuvent être trouvés dans Cunnane and Nash (1971), Rousselle and Hindie (1976), Rasmussen and Rojsberg (1991), et Madsen *et al.* (1994, 1995).

## **2. L'APPROCHE BAYÉSIENNE**

---

Dans une approche bayésienne, toutes les quantités inconnues (paramètres du modèle et variables non observées) sont traitées comme des variables stochastiques pour tenir compte de leur caractère aléatoire ou de la connaissance imparfaite qu'on a de leur véritable valeur. Pour cela, la connaissance ou la croyance qu'on a sur les paramètres est décrite par une distribution de probabilité dite a priori, qui doit être formulée indépendamment des observations. Cette densité de probabilité est ensuite mise à jour avec les observations en utilisant le théorème de Bayes :

$$p(\boldsymbol{\theta} | \mathbf{x}) = \frac{f(\mathbf{x} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int f(\mathbf{x} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})d(\boldsymbol{\theta})} \quad [3]$$

où  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  est le vecteur des observations,  $\pi(\boldsymbol{\theta})$  la densité de probabilité a priori des paramètres,  $f(\mathbf{x} | \boldsymbol{\theta})$  est la vraisemblance des observations, et  $p(\boldsymbol{\theta} | \mathbf{x})$  la densité de probabilité a posteriori des paramètres étant données les observations. Dans des cas très simples, l'expression de  $p(\boldsymbol{\theta} | \mathbf{x})$  peut être déduite de façon analytique. Dans la plupart des cas, elle est obtenue numériquement en utilisant l'intégration de Monte Carlo par Chaîne de Markov (Monte-Carlo Markov Chain, MCMC : voir par exemple Gilks *et al.* 1996). Dans l'application considérée ici, on a :  $\boldsymbol{\theta} = (\mu, \sigma, \xi)$  où  $\mu$ ,  $\sigma$  et  $\xi$  sont respectivement les paramètres de position, d'échelle et de forme de la loi GEV dont la densité de probabilité est donnée par :

$$f(x_i; \boldsymbol{\theta}) = \frac{1}{\sigma} \left( 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right)^{-\frac{\xi+1}{\xi}} \quad [4]$$

Sa fonction de distribution est donnée par

$$F(x_i; \boldsymbol{\theta}) = \begin{cases} \exp\left(-\left(1 + \xi \frac{x_i - \mu}{\sigma}\right)\right)^{\frac{1}{\xi}} & \xi \neq 0 \\ \exp\left(-\exp\left(\frac{x_i - \mu}{\sigma}\right)\right) & \xi = 0 \end{cases} \quad [5]$$

Ses quantiles sont donnés par

$$q_p = \mu + \sigma \frac{(-\log(1-p))^{-\xi} - 1}{\xi} \quad \xi \neq 0 \quad [6]$$

Puisque les observations sont indépendantes, la vraisemblance correspondant à un échantillon observé  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  est donnée par :

$$L_n(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^n f(x_i; \boldsymbol{\theta}) \quad [7]$$

La spécification de l'information a priori via  $\pi(\boldsymbol{\theta})$  peut se faire de plusieurs manières, par exemple :

- En spécifiant la distribution conjointe des paramètres  $\xi$ ,  $\mu$  et  $\sigma$  (Coles et Powell 1996)
- En utilisant deux rapports de probabilités de quantiles donnés  $q_{p1}$ ,  $q_{p2}$  et  $q_{p3}$  (par ex.  $\frac{P_3}{P_2}$  et  $\frac{P_2}{P_1}$ ) auxquels on attribue une distribution de probabilité (Crowder 1992).

- En utilisant un quantile et deux différences de quantiles (par ex.  $q_{p_3} - q_{p_2}$ ,  $q_{p_2} - q_{p_1}$  et  $q_{p_1}$  ) auxquels on attribue une distribution de probabilité (Coles et Tawn 1996).

C'est la dernière méthode qui a été choisie du fait que  $q_{p_3} - q_{p_2}$ ,  $q_{p_2} - q_{p_1}$  et  $q_{p_1}$  sont des quantités hydrologiques faciles à obtenir par régression linéaire multiple.

### **3. LE MODÈLE RÉGIONAL**

---

Un modèle d'estimation régionale comporte deux parties (GREHYS 1996a):

- Une méthode de détermination des régions homogènes (composées de stations ayant un comportement hydrologique similaire).
- Une méthode d'estimation régionale.

Les régions homogènes sont des sous-ensembles de stations ayant des comportements hydrologiques similaires, et plusieurs méthodes ont été proposées dans la littérature pour les délimiter. On peut sans être exhaustif citer la méthode ad hoc (découpage arbitraire en fonction par exemple des limites administratives) utilisée par exemple par Dalrymple (1960), la méthode des régions d'influence (Burn, 1988, 1990a, b), l'analyse des correspondances et la classification ascendante

hiérarchique (Birikundavyi et al, 1993), l'analyse canonique des corrélations (Cavadias, 1989, 1990; Ouarda et al. 2001), et la méthode des L-moments (Hosking et Wallis, 1993). Les méthodes d'estimation régionale comprennent entre autres la méthode d'indice de crue (Dalrymple, 1960), l'analyse régionale non paramétrique (Adamowski, 1985, 1989), l'estimation régionale des crues par la méthode des dépassements (Lang et al. 1997; Rasmussen, 1994; Madsen, et al. 1997) et la régression multiple directe (Matalas et Gilroy, 1968; Thomas et Benson, 1970).

L'analyse canonique des correspondances a été choisie dans cette étude pour la détermination des régions homogènes parce qu'elle donne de bons résultats pour les sites non jaugés (Ouarda et al. 2001). La méthode d'estimation régionale choisie est la régression linéaire multiple directe sur le logarithme de différences de quantiles. L'information a priori est spécifiée à partir du modèle régional comme suit : soient  $q_{p_1}$ ,  $q_{p_2}$ ,  $q_{p_3}$  trois quantiles tels que  $p_1 < p_2 < p_3$  et  $\tilde{q}_{p_1}$ ,  $\tilde{q}_{p_2}$ ,  $\tilde{q}_{p_3}$  définis par

$$\tilde{q}_{p_1} = q_{p_1} = \mu - (-\log(1 - p_1))^{-\xi} \frac{\sigma}{\xi} \quad [8]$$

$$\tilde{q}_{p_2} = q_{p_2} - q_{p_1} = \sigma((- \log(1 - p_2))^{-\xi} - (- \log(1 - p_1))^{-\xi}) / \xi \quad [9]$$

$$\tilde{q}_{p_3} = q_{p_3} - q_{p_1} = \sigma((- \log(1 - p_3))^{-\xi} - (- \log(1 - p_1))^{-\xi}) / \xi \quad [10]$$

Le modèle loglinéaire est habituellement utilisé pour décrire la relation

entre les quantités hydrologiques (quantiles de crue notamment) du fait que la transformation logarithmique permet généralement de normaliser des données qui ne le sont pas :

$$\log(\tilde{q}_p^R) = \begin{bmatrix} \log(\tilde{q}_{p1}^R) \\ \log(\tilde{q}_{p2}^R) \\ \log(\tilde{q}_{p3}^R) \end{bmatrix} \sim MVN(\tilde{\beta}X, \Sigma) \quad [11]$$

où  $MVN(\tilde{\beta}X, \Sigma)$  représente la loi normale multivariée de moyenne

$\tilde{\beta}X$  et de matrice de variance-covariance  $\Sigma$ . Le vecteur des variables

$$\text{physiographiques est } X = \begin{bmatrix} 1 \\ A_1 \\ \cdot \\ \cdot \\ A_m \end{bmatrix} \text{ et } \tilde{\beta} = \begin{pmatrix} \tilde{\beta}^{(1)} \\ \tilde{\beta}^{(2)} \\ \tilde{\beta}^{(3)} \end{pmatrix}$$

avec  $\tilde{\beta}^{(i)} = (\tilde{\beta}_0^{(i)}, \tilde{\beta}_1^{(i)}, \dots, \tilde{\beta}_m^{(i)})$ ;  $A_k$  désigne la valeur de la  $k^{\text{ème}}$  variable

hydrologique au site d'intérêt,  $\tilde{\beta}_k^{(i)}$  un coefficient et  $m$  le nombre de variables physiographiques.

$\tilde{\beta}^{(i)}$  est obtenu en résolvant par la méthode des moindres carrés ordinaires l'équation

$$\log(\tilde{q}_{pi}^R) = \tilde{\beta}^{(i)}X + \varepsilon^{(i)} \quad [12]$$

Les éléments de  $\Sigma$  sont calculés directement à partir des données par :

$$\Sigma_{ij} = \text{cov}(\varepsilon^{(i)}, \varepsilon^{(j)}) \quad [13]$$

On tire de [11] et [12] que

$$\pi(\mu, \sigma, \xi) \propto J \frac{1}{\tilde{q}_{p1} \tilde{q}_{p2} \tilde{q}_{p3}} \exp\left(\frac{(\log(\tilde{q}_p) - \tilde{\beta}X)' \Sigma^{-1} (\log(\tilde{q}_p) - \tilde{\beta}X)}{2}\right) \quad [14]$$

où  $J$  est le jacobien de la transformation de  $(\tilde{q}_{p1}, \tilde{q}_{p2}, \tilde{q}_{p3})$  vers  $(\mu, \sigma, \xi)$ . L'expression de  $J$  est démontrée dans Stephenson (2002) :

$$J = \begin{cases} \frac{\sigma}{\xi^2} \left| \sum_{i,j \in \{1,2,3\}; i < j} (-1)^{i+j} (x_i x_j)^{-\xi} \log\left(\frac{x_j}{x_i}\right) \right| & \text{si } \xi \neq 0 \\ \frac{\sigma}{2} \left| \sum_{i,j \in \{1,2,3\}; i < j} (-1)^{i+j} \log(x_i) \log(x_j) \log\left(\frac{x_j}{x_i}\right) \right| & \text{si } \xi = 0 \end{cases} \quad [15]$$

Pour une comparaison avec la méthode bayésienne empirique, on calcule également  $E(q_p^R)$  et  $Var(q_p^R)$  en résolvant par la méthode des moindres carrés:

$$\log(q_{pi}^R) = \beta^{(i)} X + \varepsilon_i^{(i)} \quad [16]$$

La transformation logarithmique dans [16] introduit un biais dans l'estimation de  $q_{pi}^R$  qui sera corrigé :

$$q_{pi}^R \sim N(b_i^r \exp(\beta^{(i)} X) + b_i^a, \sigma_{pi}^R) \quad [17]$$

où  $b_i^r$  et  $b_i^a$  sont les biais relatifs et absolus introduits par la transformation logarithmique, et  $\sigma_{pi}$  la variance d'estimation du quantile.  $b_i^r$ ,  $b_i^a$  et  $\sigma_{pi}$  sont estimés par la méthode des moindres carrés ordinaires en utilisant la solution de l'équation [16].

#### **4. MÉTHODOLOGIE DE VALIDATION**

---

Pour démontrer que la combinaison de l'information régionale et locale donne des résultats plus précis, il est nécessaire de connaître la vraie valeur du quantile au site. Ce n'est jamais le cas en pratique, sauf si les données ont été obtenues par simulation. De tous les travaux cités dans la revue de littérature, seuls Kuczera(1982) et Fill et Stedinger (1998) ont démontré que le quantile estimé par la méthode bayésienne empirique s'approchait plus du quantile réel que le quantile estimé avec les données locales. La méthode de validation adoptée dans cet article est la suivante : on utilisera des séries de données historiques suffisamment longues (plus de 70 ans de données) pour qu'on puisse considérer les quantiles locaux comme étant une bonne estimation des vrais quantiles. Les données locales sont ensuite combinées par tranches de vingt années consécutives avec l'information régionale en utilisant autant la méthode bayésienne empirique que la méthode bayésienne complète. Le *REQM* servira à comparer trois estimateurs ponctuels déduits de la méthode bayésienne complète (le mode, la médiane et la moyenne de la distribution a priori) avec l'estimateur bayésien empirique et l'estimateur local. On vérifiera également si les paramètres  $\mu$ ,  $\sigma$  et  $\xi$  obtenus avec la méthode bayésienne complète sont plus proches

des paramètres 'réels' que ceux obtenus avec les séries courtes de vingt années de données.

## **5. LES SITES D'APPLICATION**

---

Les trois procédures (estimation locale, estimation bayésienne empirique et estimation bayésienne complète) ont été appliquées en considérant les données de trois rivières situées dans la province du Québec, Canada: la rivière Harricana à la station 080101, la rivière Beaurivage à la station 023401 et la rivière Matane à la station 021601 qui comportent respectivement 84, 76 et 74 années de données. Les débits aux rivières Harricana et Beaurivage sont naturels, tandis que ceux à la rivière Matane sont influencés journalièrement. Les débits maximaux annuels observés aux trois stations sont présentés à la figure 1. Ils ont été extraites d'une base de données de 168 stations hydrologiques maintenue par le Ministère de l'Environnement du Québec et pour lesquelles on dispose de l'estimation des quantiles et de certaines variables physiographiques : l'aire du bassin versant, le pourcentage de la superficie couvert par les lacs, la pente moyenne du bassin versant, la précipitation totale moyenne annuelle et la moyenne des cumuls annuels de degrés-jours de gel.

## 6. RÉSULTATS ET DISCUSSIONS

---

Le modèle régional a été appliqué en utilisant le logiciel RÉGIONS (Gingras et al., 2002). Les voisinages ont été déterminés par analyse canonique des corrélations (ACC) et on a obtenu 71 (resp. 74, 69) stations dans le voisinage de 080101 (resp. 023401, 021601). Les résultats de cette analyse sont présentés au Tableaux 1, 2 et 3 pour les stations 080101, 023401 et 021601 respectivement. Les quantiles locaux et leur variance d'estimation ont été calculés selon la méthode des moments pondérés (GEV/PWM). Les distributions a posteriori des quantiles ont été obtenues par MCMC en utilisant l'algorithme de Métropolis-Hasting tiré de Stephenson (2002). Au total 20000 itérations sont effectuées à chaque fois, et les 10000 dernières sont utilisées pour faire l'inférence sur les paramètres  $\mu$ ,  $\sigma$  et  $\xi$  ainsi que les quantiles  $q_{10}$  (crue de 10 ans),  $q_{100}$  (crue de 100 ans) et  $q_{1000}$  (crue de 1000 ans). Les diagnostics de convergence menés en utilisant la méthode de Geweke (1992) sur les 10000 dernières itérations ont donné des résultats satisfaisants. A titre d'exemple, on présente aux figures 2 à 4 les histogrammes (calculés à partir des chaînes MCMC) des quantiles  $q_{10}$ ,  $q_{100}$  et  $q_{1000}$  pour la rivière Harricana, superposée avec les densités de probabilité des estimateurs locaux, régionaux et bayésiens empiriques.

Chacun des paramètres et chacun des quantiles est estimé avec le mode, la médiane et la moyenne des distributions a posteriori. L'écart entre la valeur de référence (calculée avec toutes les données disponibles) et la valeur estimée du quantile ou du paramètre est mesuré par la *REQM* (racine de l'erreur quadratique moyenne) défini par :

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{\theta}_i - \theta_i)^2} \quad [18]$$

où  $\theta_i$  représente la variable (quantile ou paramètre),  $\hat{\theta}_i$  son estimation.

Les *REQM* des estimateurs des paramètres de la loi GEV sont données aux Tableaux 4 (rivière Harricana), 5 (rivière Matane) et 6 (rivière Beaurivage). Celui des estimateurs des quantiles  $q_{10}$ ,  $q_{100}$  et  $q_{1000}$  sont données aux Tableaux 7 (rivière Harricana) et 8 (rivière Matane) et 9 (rivière Beaurivage). La supériorité de la méthode bayésienne complète par rapport à l'estimation locale est très claire pour les rivières Harricana et Matane, autant pour le paramètre de forme  $\zeta$  de la GEV (tableaux 4 et 5; figures 5c et 6c) que pour les quantiles (tableaux 7 et 8; figures 8,9,10,11,12,13). Par contre, l'estimateur local est meilleur sur les données des rivières Matane et Beaurivage pour l'estimation des paramètres de position et d'échelle  $\mu$  et  $\sigma$  (tableaux 4 et 5; figures 5a, 5b, 6a, 6b). L'estimateur local performe moins bien que les estimateurs bayésiens pour le calcul de  $\mu$  et  $\sigma$  sur les données de la rivière Harricana.

Le bon comportement de la méthode bayésienne complète au niveau des rivières Harricana et Beaurivage est en concordance avec la bien connue bonne performance de la méthode GEV/PWM à deux paramètres (Lettenmaier et al., 1987; Stedinger and Lu, 1995; Fill and Stedinger, 1998) pour laquelle  $\mu$  et  $\sigma$  sont estimés avec les données locales, tandis que  $\xi$  l'est par une approche régionale. On peut également voir sur les figures 5, 6 et 7 que le paramètre de forme  $\xi$  est stabilisé (à moins de variance) par la combinaison de l'information locale et régionale. Par contre,  $\xi$  est surestimé dans le cas de la rivière Beaurivage (figure 7.c). Cela ne mène pas à une surestimation de la crue décennale  $q_{10}$ , mais les quantiles  $q_{100}$  et  $q_{1000}$  sont clairement biaisés vers le haut (tableau 9; figures 15 et 16). De ce fait, la REQM des estimateurs du paramètre de forme est nettement plus basse pour l'estimateur local que pour les estimateurs bayésiens (table 6). Le 'vrai' paramètre de forme de la distribution locale au niveau de la rivière Beaurivage est 0.2529. Elle est donc une GEV type III avec une borne supérieure. Afin d'expliquer le phénomène, on a calculé pour chacun des sites de la base de donnée un paramètre de forme  $\hat{\xi}$  estimé numériquement à partir de  $\frac{q_{1000} - q_{10}}{q_{100} - q_{10}}$  (cette expression est une fonction monotone de  $\xi$  si on suppose que les débits de pointe annuels peuvent être décrits par une distribution GEV :

voir équations [9] et [10]). L'histogramme des valeurs de  $\hat{\xi}$  obtenues pour toutes les stations de la base de données est présenté à la figure 17. On remarque que les rivières Harricana et Matane ont des valeurs de  $\hat{\xi}$  proches de la moyenne, tandis que Beaurivage s'en écarte de façon significative. La raison de la mauvaise performance de la méthode proposée sur la rivière Beaurivage est donc qu'elle diffère des autres sites, et que de ce fait l'utilisation de l'information régionale y est contre-productive.

En termes de REQM, l'estimateur bayésien empirique est quasi équivalent ou meilleur que l'estimateur locale sur les données des rivières Harricana (tableau 7) et Matane (tableau 8). On peut cependant observer de grosses différences au niveau du biais sur les données de la rivière Matane (tableau 8). A Beaurivage (tableau 9), l'estimateur local et l'estimateur bayésien empirique ont des performances quasi-équivalentes, loin devant les estimateurs issus de la méthode bayésienne complète.

La méthodologie présentée dans ce rapport peut être adaptée à d'autres distributions. Pour ce faire, on doit juste réécrire l'expression du jacobien  $J$  (équation [15]) de la transformation de  $\tilde{q}_{pi}$ ,  $i=1, \dots, n$  vers les paramètres de la distribution, où  $n$  est le nombre de paramètres. On peut également utiliser d'autres expressions pour  $\tilde{q}_{pi}(q_{p1}, \dots, q_{pn})$  en

faisant attention à ce que  $J$  ne prenne pas de valeurs nulles dans l'espace des paramètres.

## **7. CONCLUSION**

---

On a présenté dans cet article une méthode bayésienne pour faire la combinaison de l'information locale et régionale. L'application a été faite avec la loi GEV, mais elle peut être adaptée à d'autres types de distributions. Elle améliore significativement l'estimation des paramètres et des quantiles lorsqu'on utilise des séries courtes (20 ans de données) si l'asymétrie réelle au site ne diffère pas beaucoup de celle du voisinage. Elle s'est révélée être supérieure en termes de *REQM* à l'estimation locale, régionale, et à celle utilisant la méthode bayésienne empirique utilisée par Kuczera (1982) pourvu que l'hypothèse d'homogénéité régionale est respectée (i.e le paramètre de forme est sensiblement le même dans la région considérée). C'est donc une voie intéressante pour l'estimation des quantiles aux sites peu jaugés.

## **8. BIBLIOGRAPHIE**

---

Adamowski, K., 1985. Nonparametric kernel estimation of flood frequencies. *Water Resources Research* **21**(11):1585-1590.

- Adamowski, k. 1989. A Monte-Carlo comparison of parametric and non-parametric estimation of flood frequencies. *Journal of hydrology* **108**: 295-308.
- Benson, M.A. 1962. Evaluation methods for evaluating the occurrence of floods, U. S. Geological survey, Water supply paper 1543-A.
- Benzecri, J.P. 1973. L'analyse des données, T.2: L'analyse des correspondances. Dunod, Paris, 619 p.
- Bernier, J. 1993. Combinaison des informations locales et spatiales dans les modèles de régionalisation des crues. Personal communication (13-08-1993), INRS-ETE, Quebec. 6 p.
- Burn, D.H. 1988. Delineation of groups for regional flood frequency analysis. *Journal of Hydrology* **104**: 343-361.
- Burn, D.H., 1990a. An appraisal of the 'region of influence' approach to flood frequency analysis, *Hydrol. Sci. J.*, **35**(2), 149-165,
- Birikundavyi, S., Rousselle, J. et Nguyen, V.T.V. 1993. Détermination des régions homogènes pour le Québec et l'Ontario : une approche par l'analyse des correspondances et la classification ascendante hiérarchique. École Polytechnique de Montréal, 50 p.
- Bobée, B. et Rasmussen, P. 1995. Recent advances in flood frequency analysis. *Reviews of Geophysics* **33** (Supplement) [<http://www.agu.org/revgeophys/bobee01/bobee01.html>].

Burn, D.H., 1990b. Evaluation of regional flood frequency analysis with a region of influence approach, *Water Resour. Res.*, **26**(10), 2257-2265.

Cavadias, G.S. 1989. The Canonical correlation approach to regional flood estimation. Regionalisation in hydrology: In: Proceedings of the Ljubljana Symposium. IAHS, Publ. No. 191 IAHS, Wallingford, 171-178.

Cavadias, G.S. 1990. Regional flood estimation by canonical correlation. Conférence annuelle de la société Canadienne de Génie Civil, St-John's, Newfoundland.

Coles, S. G. and Powell, E. A. (1996). Bayesian methods in extreme value modelling. A review and new developments. *Int. Statist. Review* **43**: 1-48.

Coles, S.G and Tawn, J. A. 1996. A bayesian analysis of extreme rainfall data. *Appl. Statist.* **45**, 463-478.

Crowder, M. (1992). Bayesian priors based on a parameters transformation using the distribution function. *Ann. Inst. Statsit. Math.* **44**: 405-416.

Cunane, C. and Nash, J.E. 1971. Bayesian estimation of frequency of hydrological events. *Mathematical models in hydrology*, IAHS Publ. **100**: 47-55.

Dalrymple, T. 1960. Flood frequency analysis. U.S. Geological Survey, Water Supply Paper 1543-A.

Fill, H.D. and Stedinger, J.R. 1998. Using regional regression within index flood procedures and an empirical bayes estimator. *Journal of hydrology* **210**: 128:145.

Fortin, V., Bernier, J. et Bobée, B. (1998). Projet C3. Détermination des crues de conception. Rapport de recherche No R-535 (confidentiel). Chaire industrielle Hydro-Québec/CRSNG en hydrologie statistique, INRS-Eau.

Geweke, J. 1992. *Bayesian Statistics 4*, chapter Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, pages 69–193. Oxford University Press, Oxford.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. (eds.) (1996). Markov Chain MonteCarlo in Practice. Chapman and Hall: London.

Gingras, H., Ouarda, T.B.M.J., Ristic-Rudolf, Z., Kouider, A., Haché, M. et Bobée, B. 2002. Guide de l'utilisateur du logiciel REGIONS. Rapport scientifique INRS-Eau 638-g1. ISBN 2891464974.

Greis, N.P. et Wood, E.F. 1981. Regional flood frequency analysis and network design, *Water Resources Research* **17** (4): 1167-1177.

GREHYS. 1996a. Presentation and review of some methods for regional flood frequency analysis. *Journal of Hydrology* **186**: 63-84.

GREHYS. 1996b. Inter-comparison of regional flood frequency procedures for Canadian rivers. *Journal of Hydrology* **186**: 85-103.

Hosking, J. R. M. Wallis, J.R. et Wood, E. F. 1985. An appraisal of regional flood frequency procedure in the U.K. Flood Studies Report, *Hydrological Sciences Journal* **30** (1). 85-109.

Hosking, J. R. M. et Wallis, J.R. 1993. Some statistics useful in regional frequency analyses. *Water Resources Research*. **29**(2): 271-281.

Kuczera, G. 1982. Combining site-specific and regional information: an empirical bayes approach. *Water resources research* **18**(2): 306-314.

Lang M., Rasmussen P., Oberlin G. & Bobée, B. 1997. Échantillonnage par valeurs supérieures à un seuil : modélisation des occurrences par la méthode du renouvellement. *Revue des Sciences de l'Eau* **10** (3) : 279-32

Lettenmaier, D. P. and Potter, K. W. 1985. Testing flood frequency estimation methods using a regional flood generation model. *Water Resources Research* **21**(12), 1903-1914.

Lettenmaier, D.P., Wallis, J.R. and Wood, E.F. 1987. Effect of regional heterogeneity on flood frequency estimation. *Water Resour. Res.*, **23**(2): 313-323.

Madsen, H., Rojsberg, H.D. and Harremøe, P. 1994. PDS modeling and regional bayesian estimation of extreme rainfall. *Nordic hydrology* **25**(4), 279-300.

Madsen, H., Rojsberg, H.D. and Harremøe, P. 1995. Applications of the Bayesian approach in regional analysis of extreme rainfall. *Stochastic hydrology and hydraulics* **9**(1): 77-88.

Madsen, H. et Rojsberg, D. 1997. Generalized least square and empirical bayes estimation in regional partial duration series index flood modeling. *Water resources research* **33** (4): 771-781.

Matalas, N.C. et Gilroy, E.J. 1968. Some comments on regionalization in hydrologic studies. *Water Resources Research* **4**: 1361-1369.

National Environment Research Council (NERC). 1975. *Hydrological studies, Volume I*. London.

Ouarda, T.B.M.J., Girard, C., Cavadias, G. S. et Bobée, B. 2001. Regional flood frequency estimation with canonical correlation analysis. *Journal of hydrology* **254**: 157-173.

Rasmussen, P.F. 1994. A regional index-flood method based on partial duration series. Rapport de recherche INRS-EAU, NSERC Strategic Grant no STR0118482, 17 p.

Rasmussen, P. F. and Rojsberg, D. 1991. Application of Bayesian principles in regional flood frequency estimation. In *Advances in water resources Technology*, edited by G. Tsakiris, pp 65-75, A.A. Balkema, Rotherdam, Netherland.

Ribeiro-Corréa, J., Cavadias, J. S., Clément, B. et Rousselle, J. 1995.

Identification of hydrological neighbourhood using canonical correlation analysis. *Journal of hydrology* **173** : 71-89.

Rousselle, J. et Hindie, F. 1976. Incertitude dans les débits de crue : approche bayésienne. *Journal of hydrology* **30** : 341-349.

Shane et Gaver (1970). Statistical decision theory techniques for the revision of mean flood flow revision estimates, *Water Resources Research* **6(6)**: 1649-1654.

Stedinger, J.R. 1989. Using historical and regional information in flood frequency analysis. Dans: *Proceedings of the Pacific International Seminar on Water Resources Systems*, Tomanu, Japan.

Stedinger, J.R. et Lu, L.H. 1995. Appraisal of regional and index flood quantile estimators, *Stochastic hydrology and hydraulics*, **9 (1)** : 49-75.

Stephenson, A. G. (2002). A user guide to the evdbayes package (version 1.0). [<http://cran.r-project.org/>].

Tasker, G. D. 1980. Hydrologic regression and weighted least squares, *Water Resources Research* **16 (6)**: 1107-1113.

Thomas, D. M. et Benson, M. A. 1970. Generalization of streamflow characteristic from drainage-basin characteristics, U. S. Geological survey, Water supply paper 1975.

Vicens, G. J., Rodriguez-Iturbe, I et Shaake, J.C. 1975. A bayesian framework for the use of regional information in hydrology.

Wallis, J. R. (1980). Risk and uncertainty in the evaluation of flood events for the design of hydraulic structures, in *Piene e Siccita*, édité par E. Giggino, E. Rossi, et E. Todini: 3-36.

**Tableau 1: résultats de l'analyse régionale (rivière Harricana, station 080101)**

Station	$\Sigma$	Variable	Moyenne	Variance
080101	$\begin{pmatrix} 0.055 & 0.069 & 0.074 \\ 0.069 & 0.185 & 0.235 \\ 0.074 & 0.235 & 0.305 \end{pmatrix}$	$\log(\tilde{q}_{p1}^R)$	5.731	0.055
		$\log(\tilde{q}_{p2}^R)$	4.456	0.185
		$\log(\tilde{q}_{p3}^R)$	4.281	0.305
	$\log(q_{p1}^R)$	5.731	0.055	
	$\log(q_{p2}^R)$	5.980	0.070	
	$\log(q_{p3}^R)$	6.150	0.093	
	$q_{p1}^R$	317.167	6432.361	
	$q_{p2}^R$	410.137	15217.344	
	$q_{p3}^R$	493.166	32908.899	
	$b_1^r$	1.011	-	
	$b_1^a$	5.433	-	
	$b_2^r$	1.015	-	
	$b_2^a$	8.836	-	
	$b_3^r$	1.015	-	
	$b_3^a$	17.333	-	

**Tableau 2 : résultats de l'analyse régionale (rivière Matane, station 021601)**

Station	$\Sigma$	Variable	Moyenne	Variance
021601	$\begin{pmatrix} 0.055 & 0.072 & 0.077 \\ 0.072 & 0.210 & 0.263 \\ 0.077 & 0.263 & 0.341 \end{pmatrix}$	$\log(\tilde{q}_{p1}^R)$	5.907	0.055
		$\log(\tilde{q}_{p2}^R)$	5.101	0.210
		$\log(\tilde{q}_{p3}^R)$	5.079	0.341
	$\log(q_{p1}^R)$	5.907	0.2342	
	$\log(q_{p2}^R)$	6.287	0.2713	
	$\log(q_{p3}^R)$	6.617	0.3168	
	$q_{p1}^R$	367.589	8292.11	
	$q_{p2}^R$	537.329	22412.72	
	$q_{p3}^R$	747.684	50630.31	
	$b_1^r$	1.018		
	$b_1^a$	3.739		
	$b_2^r$	1.053		
	$b_2^a$	-6.993		
	$b_3^r$	1.084		
	$b_3^a$	-18.985		

**Tableau 3 : résultats de l'analyse régionale (rivière Beaurivage, station 023401)**

Station	$\Sigma$	Variable	Moyenne	Variance
023401	(0.047 0.063 0.067)	$\log(\tilde{q}_{p1}^R)$	6.218	0.047
		$\log(\tilde{q}_{p2}^R)$	5.262	0.205
		$\log(\tilde{q}_{p3}^R)$	5.201	0.339
	$\log(q_{p1}^R)$	6.218	0.2455	
	$\log(q_{p2}^R)$	6.559	0.286	
	$\log(q_{p3}^R)$	6.803	0.3369	
	$q_{p1}^R$	200.606	3772.88	
	$q_{p2}^R$	277.205	10935.32	
	$q_{p3}^R$	353.139	26384.25	
	$b_1^r$	1.028	-	
	$b_1^a$	-1.571	-	
	$b_2^r$	1.064	-	
	$b_2^a$	-19.109	-	
	$b_3^r$	1.100	-	
	$b_3^a$	-42.271	-	

**Tableau 4 : REQM des estimateurs des paramètres de la distribution GEV (rivière Harricana)**

	<b>Estimateur</b>	$\mu$	$\sigma$	$\xi$
<i>REQ</i>	Estimateur local (20 ans)	8.130	8.694	0.203
<i>M</i>	Estimateur bayésien complet (mode)	7.496	4.728	0.045
	Estimateur bayésien complet (médiane)	6.928	4.512	0.045
	Estimateur bayésien complet (moyenne)	6.951	4.572	0.045

**Tableau 5 : REQM des estimateurs des paramètres de la distribution GEV (rivière Matane)**

Estimateur		$\mu$	$\sigma$	$\xi$
<i>REQM</i>	Estimateur local (20 ans)	25.309	10.773	0.209
	Estimateur bayésien complet (mode)	31.379	13.832	0.042
	Estimateur bayésien complet (médiane)	31.060	11.596	0.042
	Estimateur bayésien complet (moyenne)	31.038	10.758	0.043

**Tableau 6 : REQM des estimateurs des paramètres de la distribution GEV (rivière Beaurivage)**

Estimateur		$\mu$	$\sigma$	$\xi$
<i>REQM</i>	Estimateur local (20 ans)	16.514	6.163	0.070
	Estimateur bayésien complet (mode)	19.818	11.293	0.228
	Estimateur bayésien complet (médiane)	19.386	10.669	0.228
	Estimateur bayésien complet (moyenne)	19.300	10.041	0.228

**Tableau 7 : *REQM* et biais des estimateurs des quantiles (rivière Harricana)**

Estimateur		Crue de 10 ans	Crue de 100 ans	Crue de 1000 ans
<i>REQM</i> (m <sup>3</sup> /s)	Estimateur local	16.423	40.847	92.273
	Estimateur régional	68.167	90.137	117.166
	Estimateur bayésien complet (mode)	12.798	25.237	41.538
	Estimateur bayésien complet (médiane)	14.000	30.737	52.990
	Estimateur bayésien complet (moyenne)	15.187	34.370	61.101
	Estimateur bayésien empirique	16.652	46.199	85.860
<i>biais</i> (m <sup>3</sup> /s)	Estimateur local	2.764	10.345	35.291
	Estimateur régional	68.167	90.137	117.166
	Estimateur bayésien complet (mode)	5.397	10.747	21.143
	Estimateur bayésien complet (médiane)	9.234	21.643	40.657
	Estimateur bayésien complet (moyenne)	11.335	27.370	51.979
	Estimateur bayésien empirique	4.717	21.357	43.097

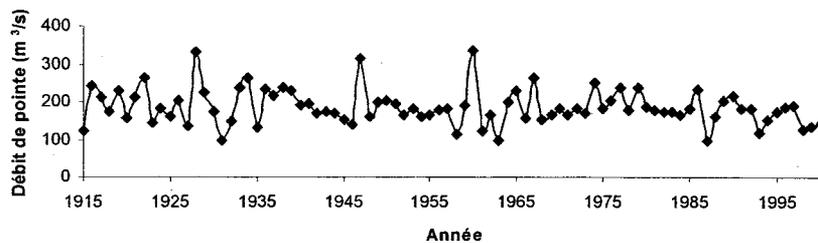
**Tableau 8 : REQ<sub>M</sub> et biais des estimateurs des quantiles (rivière Matane)**

<b>Estimateur</b>		<b>Crue de 10 ans</b>	<b>Crue de 100 ans</b>	<b>Crue de 1000 ans</b>
<i>REQ<sub>M</sub></i> (m <sup>3</sup> /s)	Estimateur local	62.447	196.752	455.689
	Estimateur régional	159.099	168.370	138.316
	Estimateur bayésien complet (mode)	52.762	77.579	103.158
	Estimateur bayésien complet (médiane)	50.283	68.660	95.290
	Estimateur bayésien complet (moyenne)	48.789	68.317	102.827
	Estimateur bayésien empirique	50.716	151.059	166.595
	<i>biais</i> (m <sup>3</sup> /s)	Estimateur local	0.036	45.545
Estimateur régional		-159.09	-168.370	-138.316
Estimateur bayésien complet (mode)		-25.365	-30.244	-13.703
Estimateur bayésien complet (médiane)		-19.245	-6.958	21.648
Estimateur bayésien complet (moyenne)		-15.175	4.212	43.860
Estimateur bayésien empirique		-28.494	-148.983	-157.340

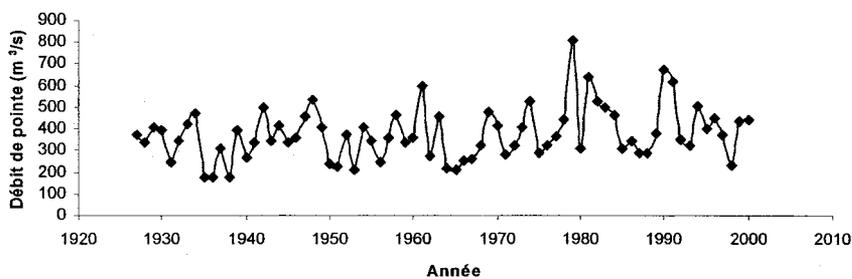
**Tableau 9 : *REQM* et biais des estimateurs des quantiles (rivière Beauvillage)**

<b>Estimateur</b>		Crue de 10 ans	Crue de 100 ans	Crue de 1000 ans
<i>REQM</i> (m <sup>3</sup> /s)	Estimateur local	16.692	28.635	52.894
	Estimateur régional	48.394	42.795	22.861
	Estimateur bayésien complet (mode)	17.179	25.881	50.406
	Estimateur bayésien complet (médiane)	15.841	29.479	62.046
	Estimateur bayésien complet (moyenne)	15.726	33.382	71.418
	Estimateur bayésien empirique	15.611	28.891	51.336
<i>biais</i> (m <sup>3</sup> /s)	Estimateur local	-3.800	-5.873	-5.200
	Estimateur régional	-48.394	-42.795	22.861
	Estimateur bayésien complet (mode)	-9.975	20.298	69.066
	Estimateur bayésien complet (médiane)	-6.966	31.328	88.922
	Estimateur bayésien complet (moyenne)	-5.056	36.711	99.885
	Estimateur bayésien empirique	-6.528	-8.937	-4.270

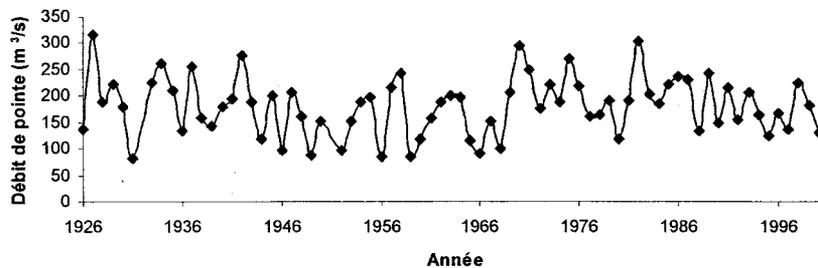
a)



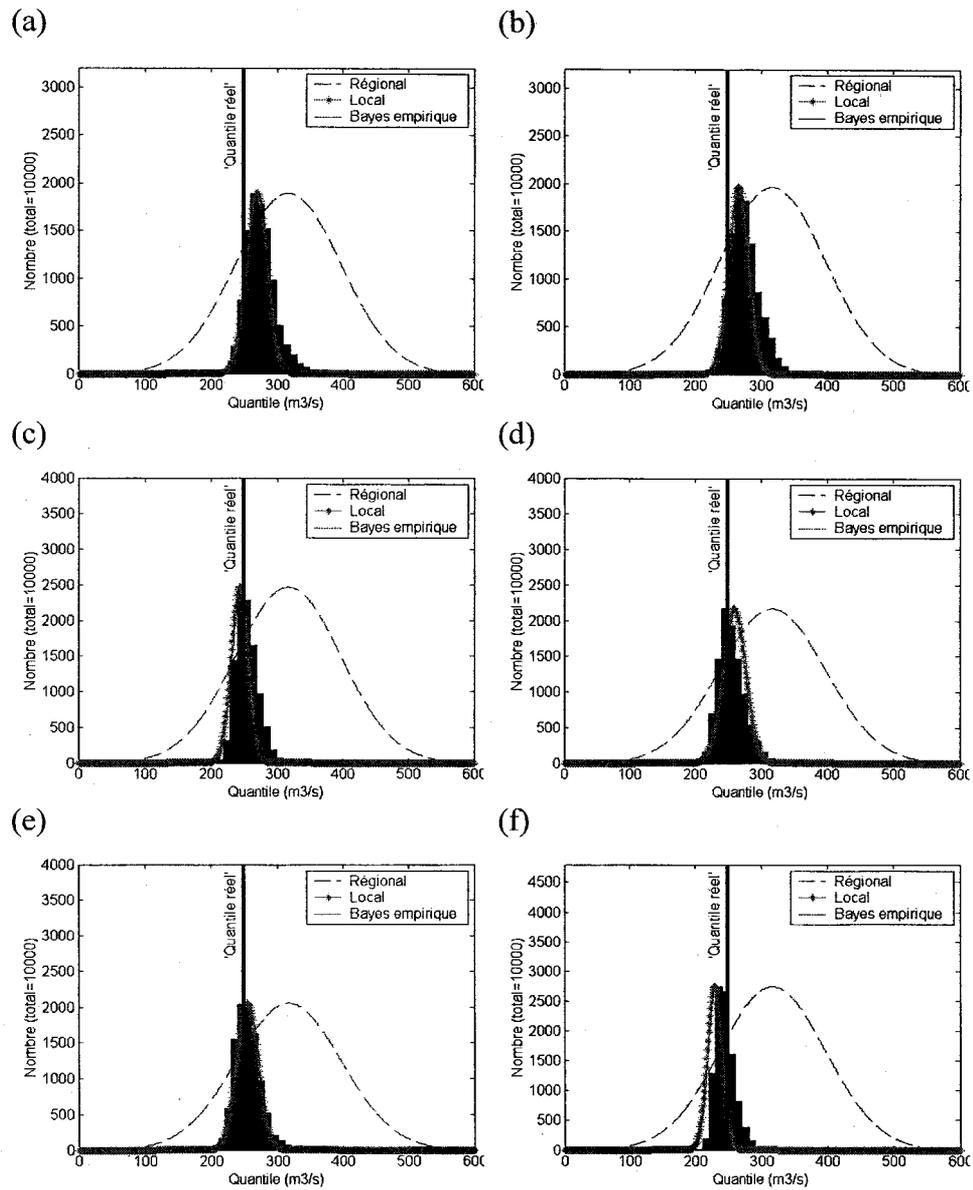
b)



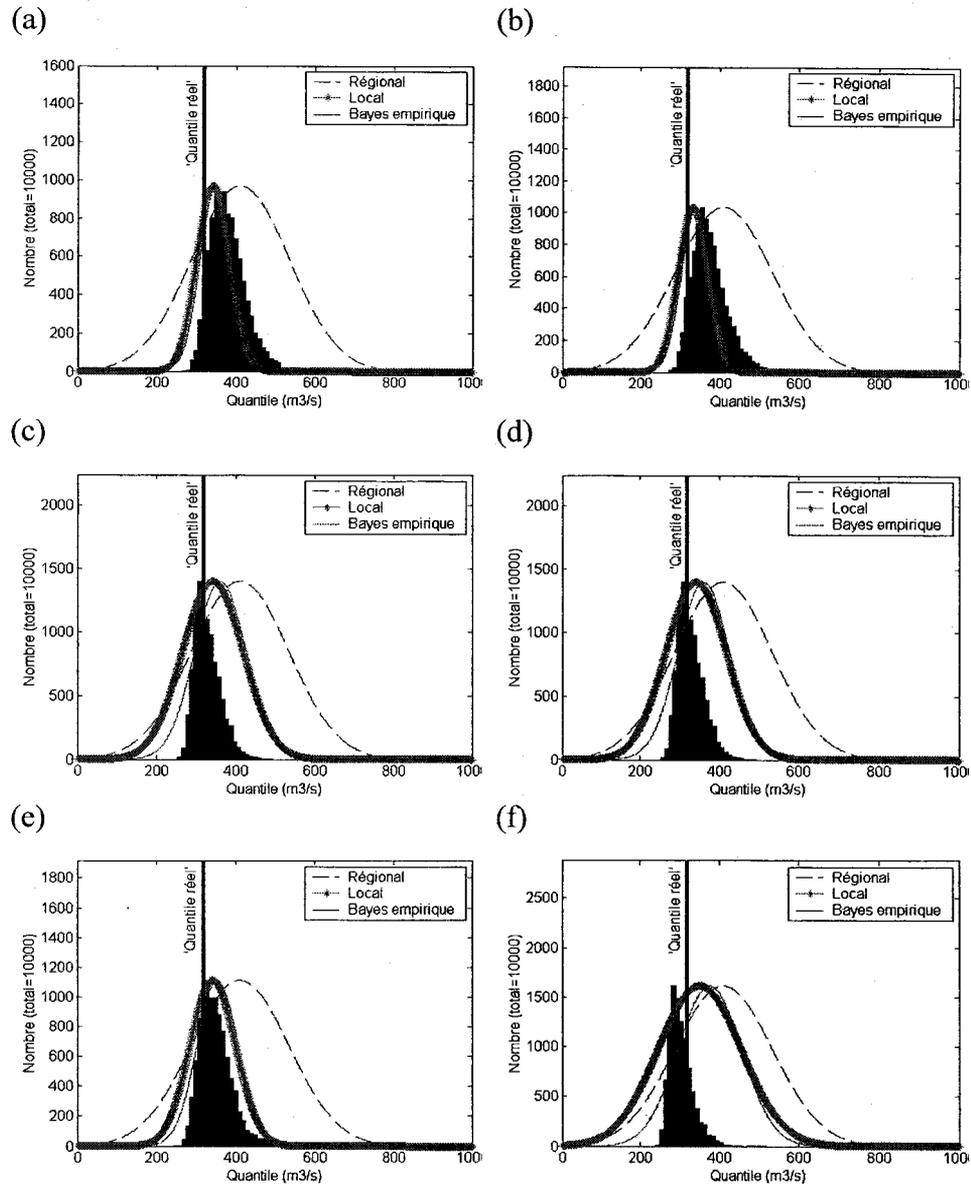
c)



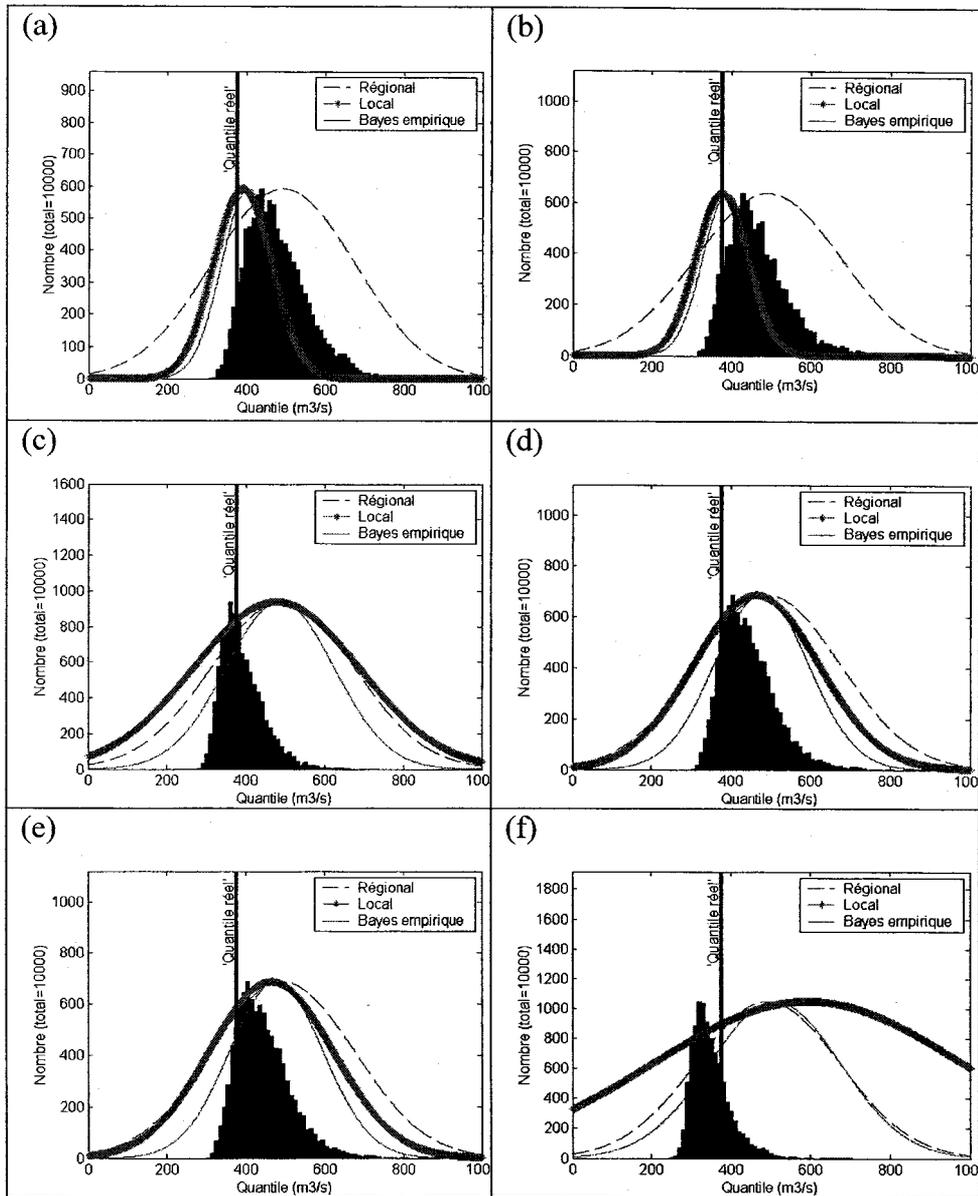
**Figure 1: débits maximaux annuels observés aux trois stations: a) Harricana; b) Matane; c) Beaurivage;**



**Figure 2: histogramme du quantile  $q_{10}$ , comparé à la densité de probabilité (mise à l'échelle) des estimateurs local, régional et bayésien empirique (rivière Harricana): a) 1915-1934; b) 1924-1944; c) 1935-1954; d) 1945-1964; e) 1955-1974; f) 1965-1984.**

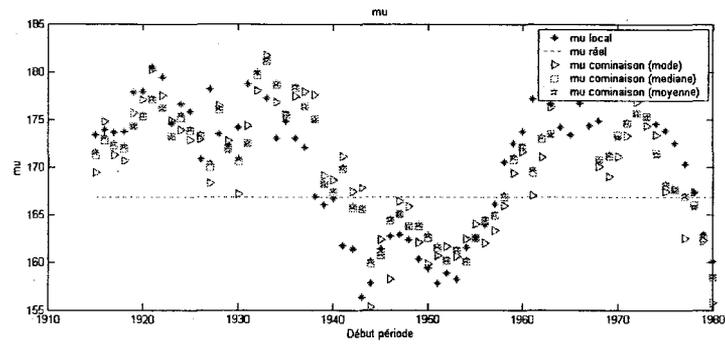


**Figure 3: histogramme du quantile  $q_{100}$ , comparé à la densité de probabilité (mise à l'échelle) des estimateurs local, régional et bayésien empirique (rivière Harricana): a) 1915-1934; b) 1924-1944; c) 1935-1954; d) 1945-1964; e) 1955-1974; f) 1965-1984.**

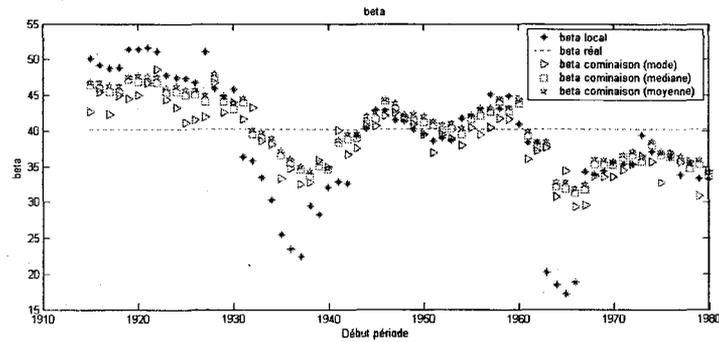


**Figure 4: histogramme du quantile  $q_{1000}$ , comparé à la densité de probabilité (mise à l'échelle) des estimateurs local, régional et bayésien empirique (rivière Harricana): a) 1915-1934; b) 1924-1944; c) 1935-1954; d) 1945-1964; e) 1955-1974; f) 1965-1984.**

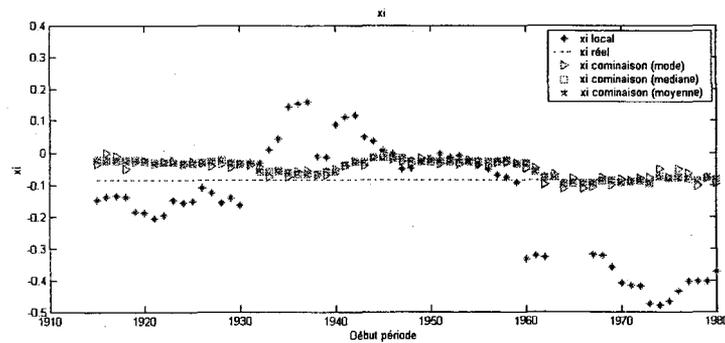
(a)



(b)

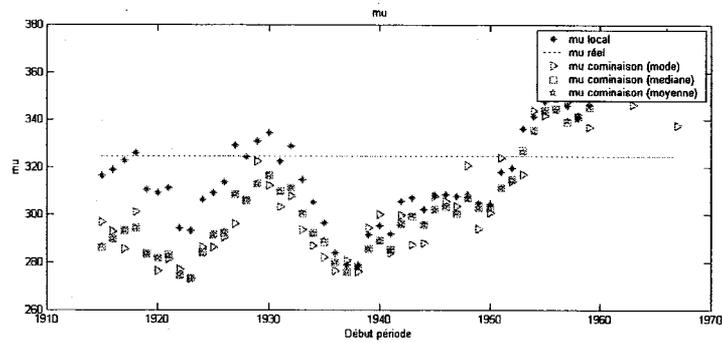


(c)

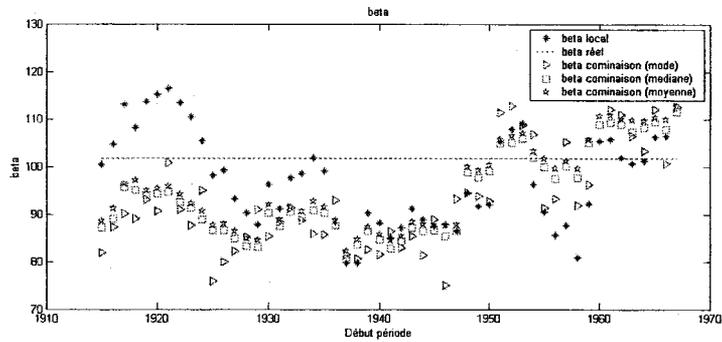


**Figure 5: variation des paramètres de la distribution GEV en fonction de la période de 20 années consécutives choisie (rivière Harricana): a)  $\mu$ ; b)  $\sigma$ ; c)  $\xi$ .**

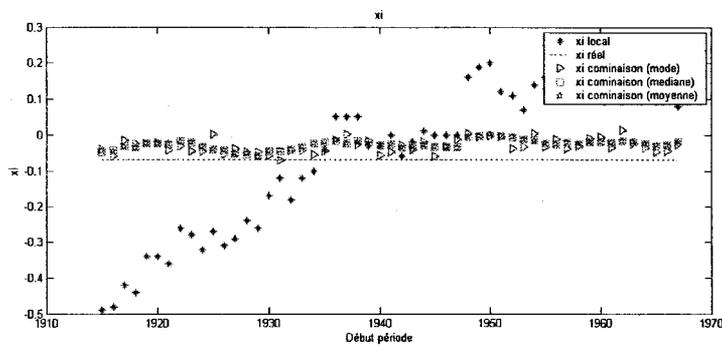
(a)



(b)

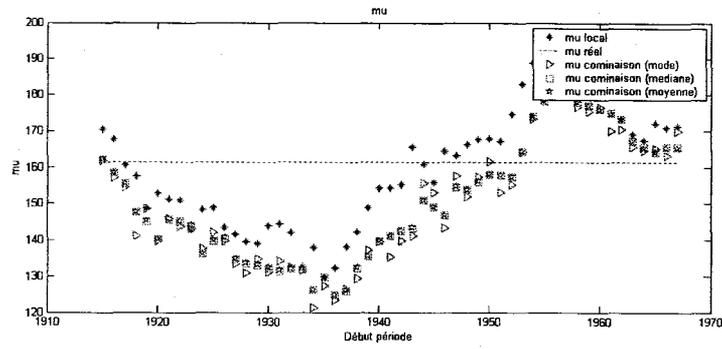


(c)

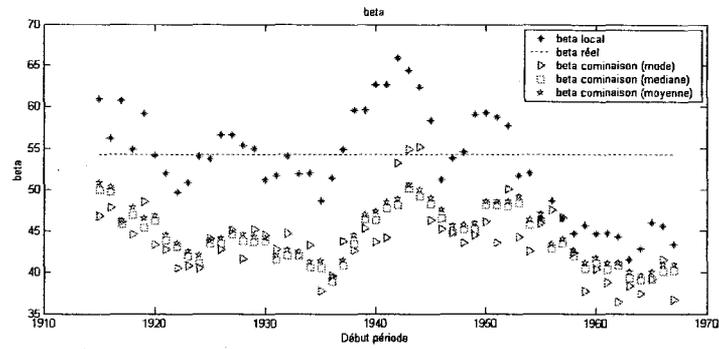


**Figure 6: variation des paramètres de la GEV en fonction de la période de 20 années consécutives choisie (rivière Matane): a)  $\mu$ ; b)  $\sigma$ ; c)  $\xi$ .**

(a)



(b)



(c)

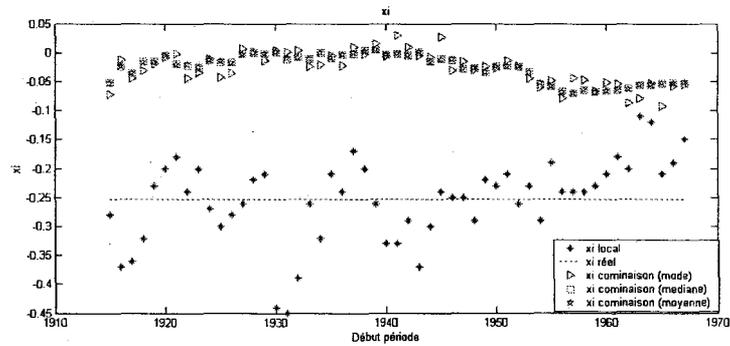


Figure 7: variation des paramètres de la distribution GEV en fonction de la période de 20 années consécutives choisie (Rivière Beauvillage): a)  $\mu$ ; b)  $\sigma$ ; c)  $\xi$ .

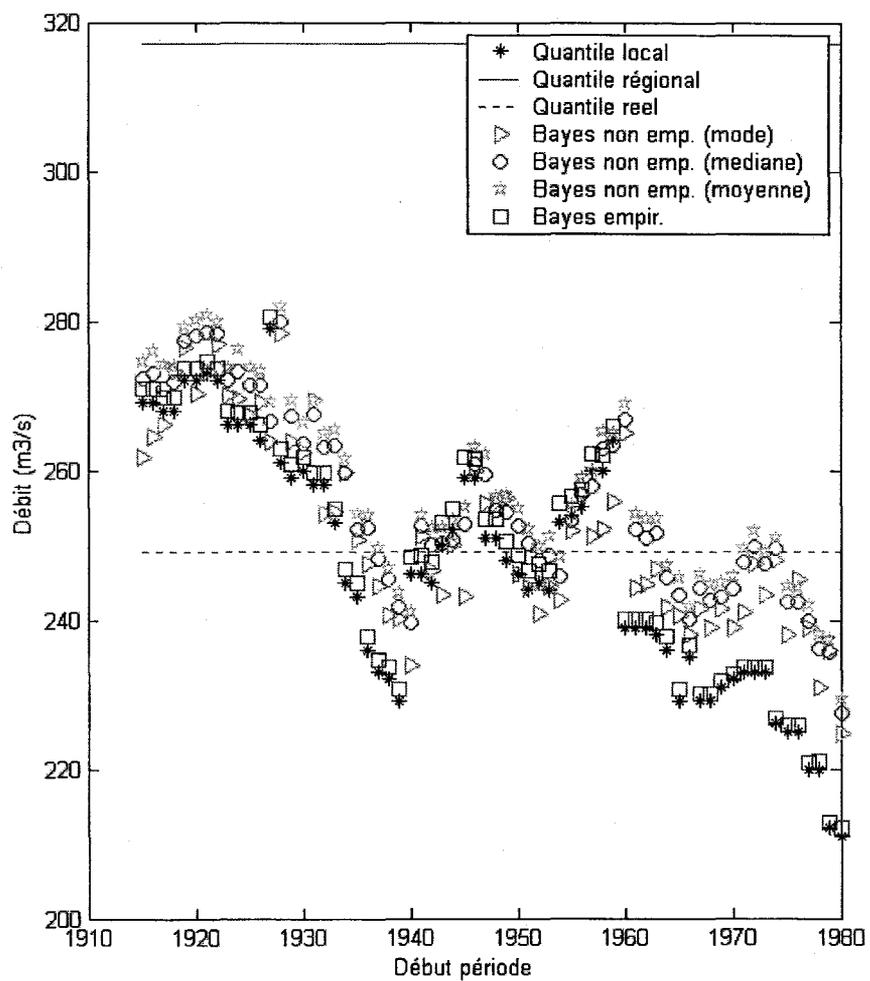


Figure 8: variation du quantile  $q_{10}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Harricana).

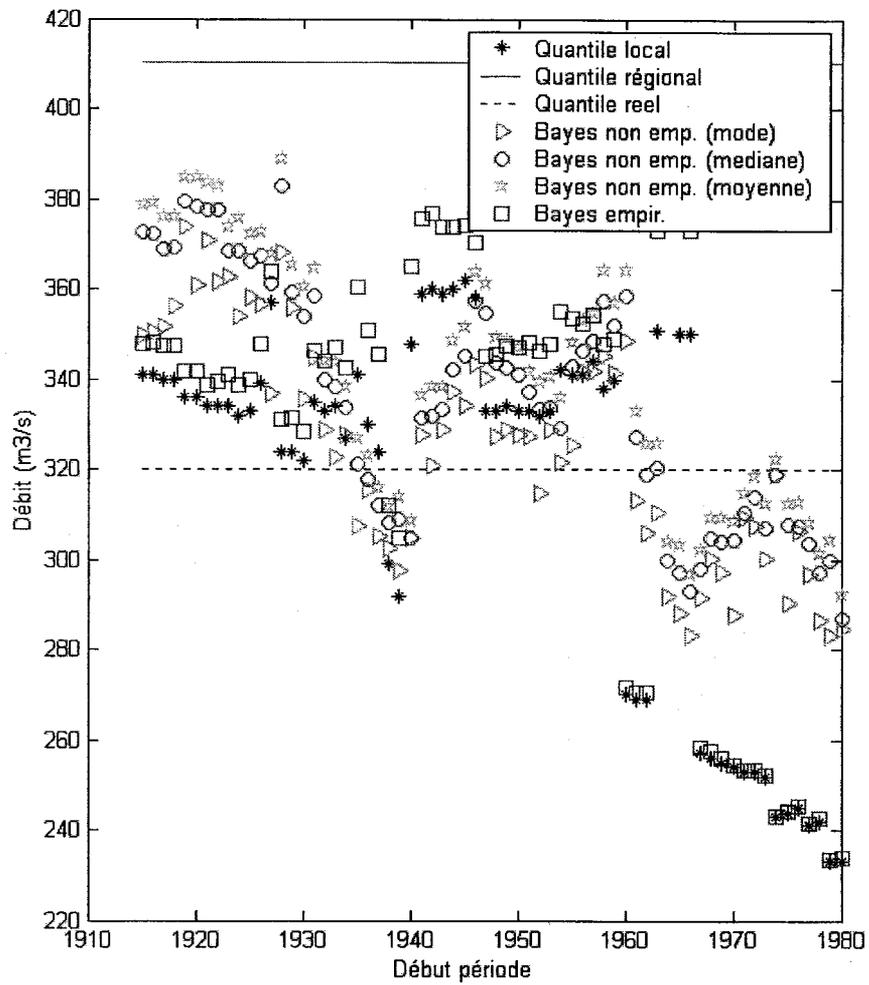


Figure 9: variation du quantile  $q_{100}$  estimé en fonction de la période de 20 années consécutives choisie (Rivière Harricana).

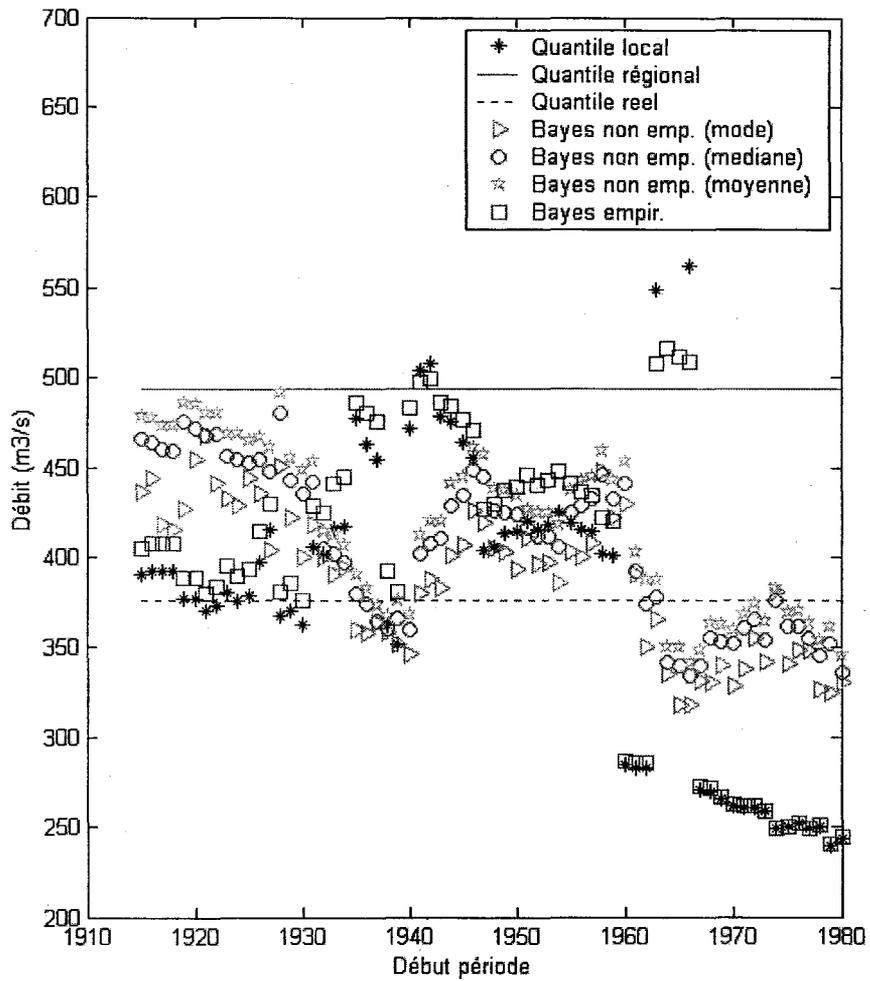


Figure 10: variation du quantile  $q_{1000}$  estimé en fonction de la période de 20 années consécutives choisie (Rivière Harricana).

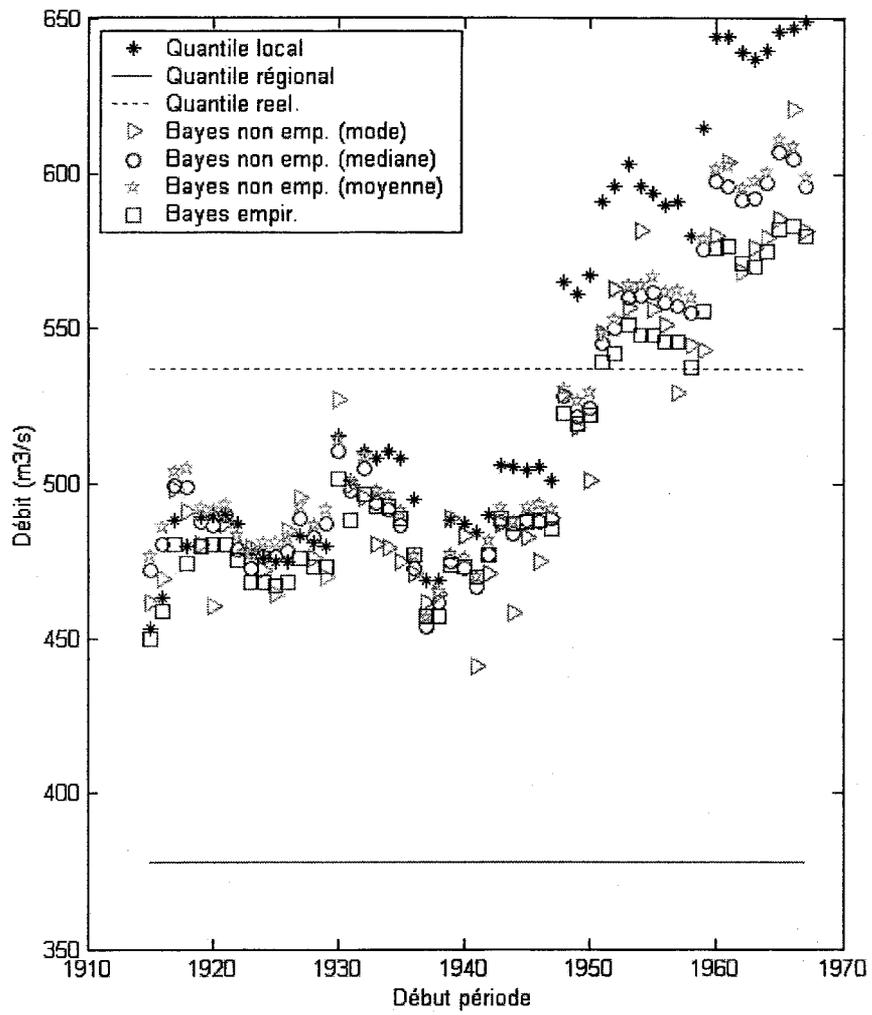
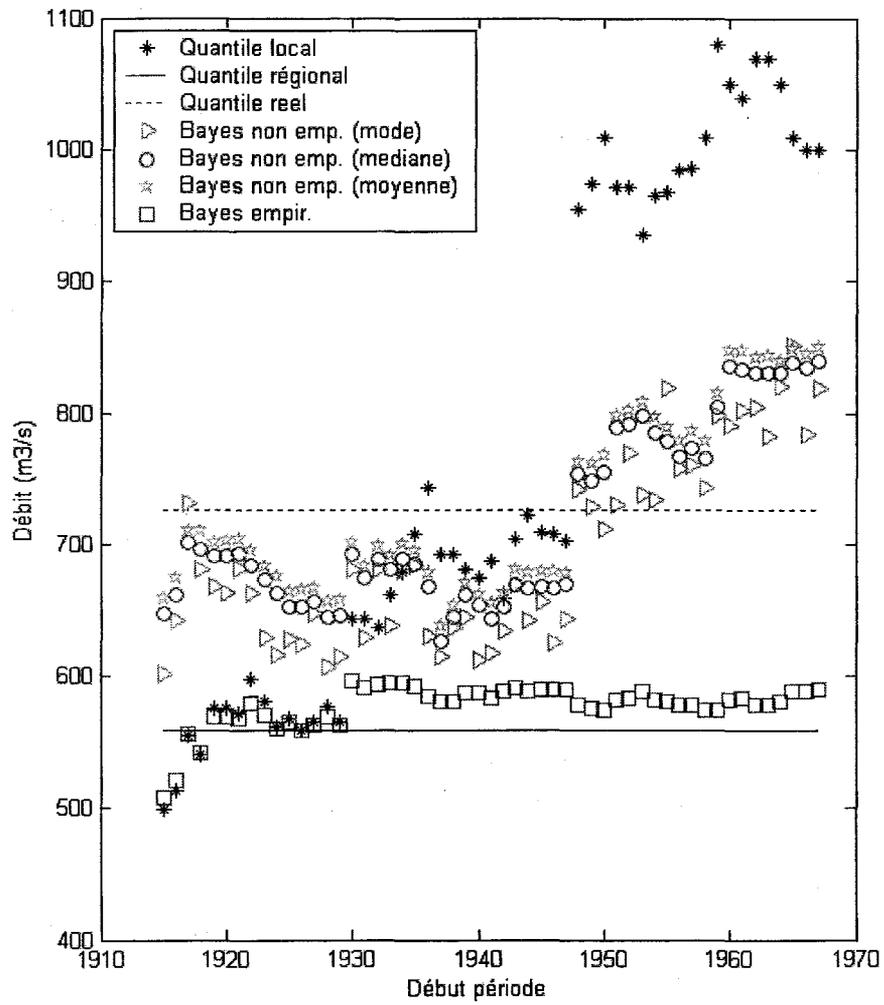
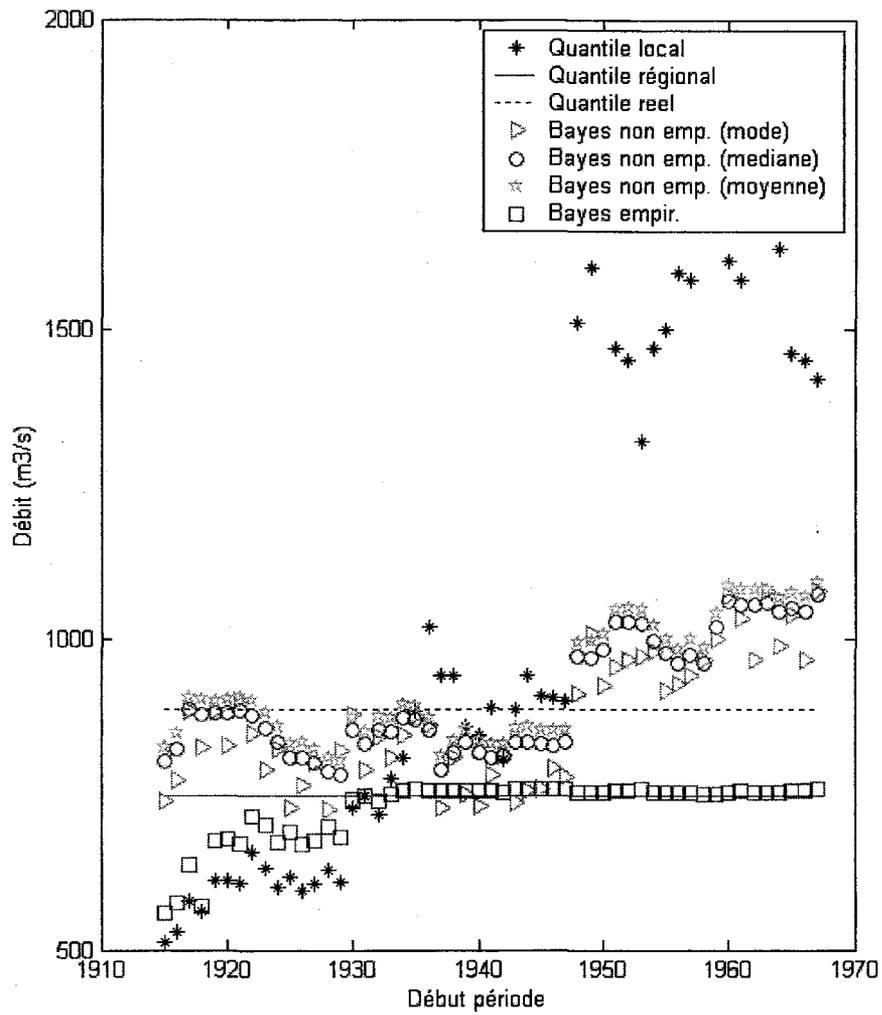


Figure 11: variation du quantile  $q_{10}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Matane).



**Figure 12: variation du quantile  $q_{100}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Matane).**



**Figure 13: variation du quantile  $q_{1000}$  estimé en fonction de la période de 20 années consécutives choisie (Rivière Matane).**

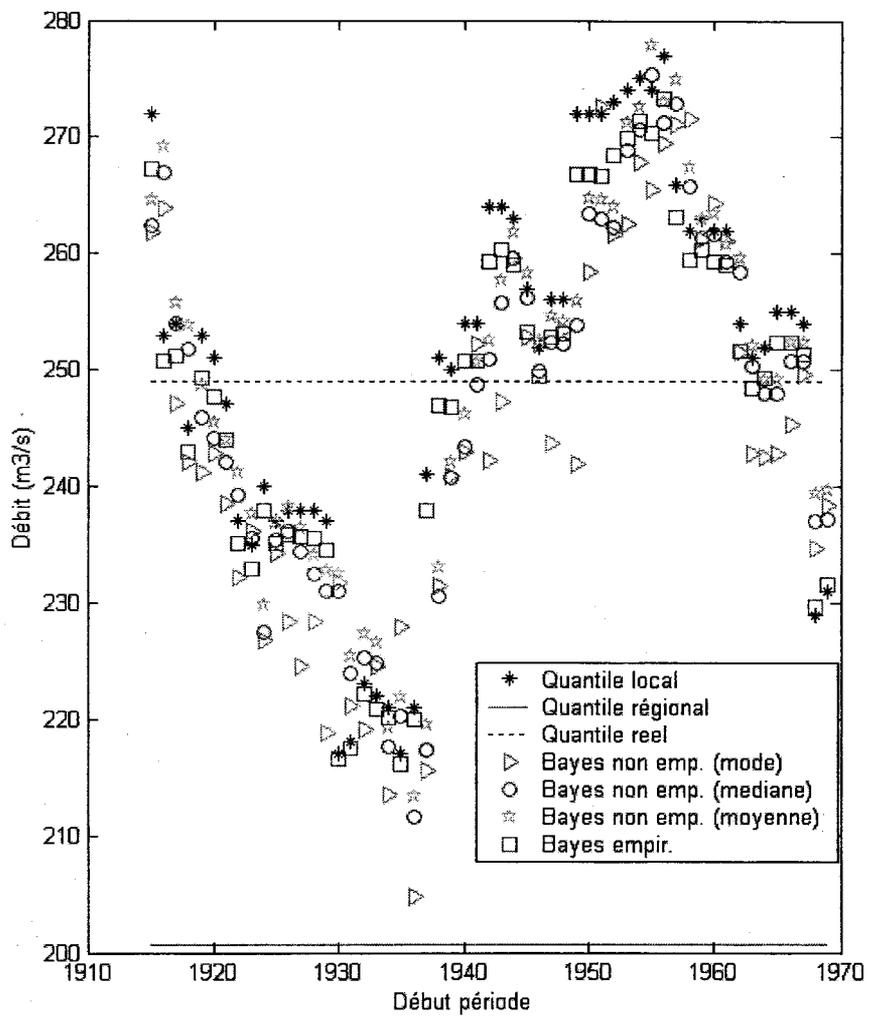


Figure 14: variation du quantile  $q_{10}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Beaurivage).

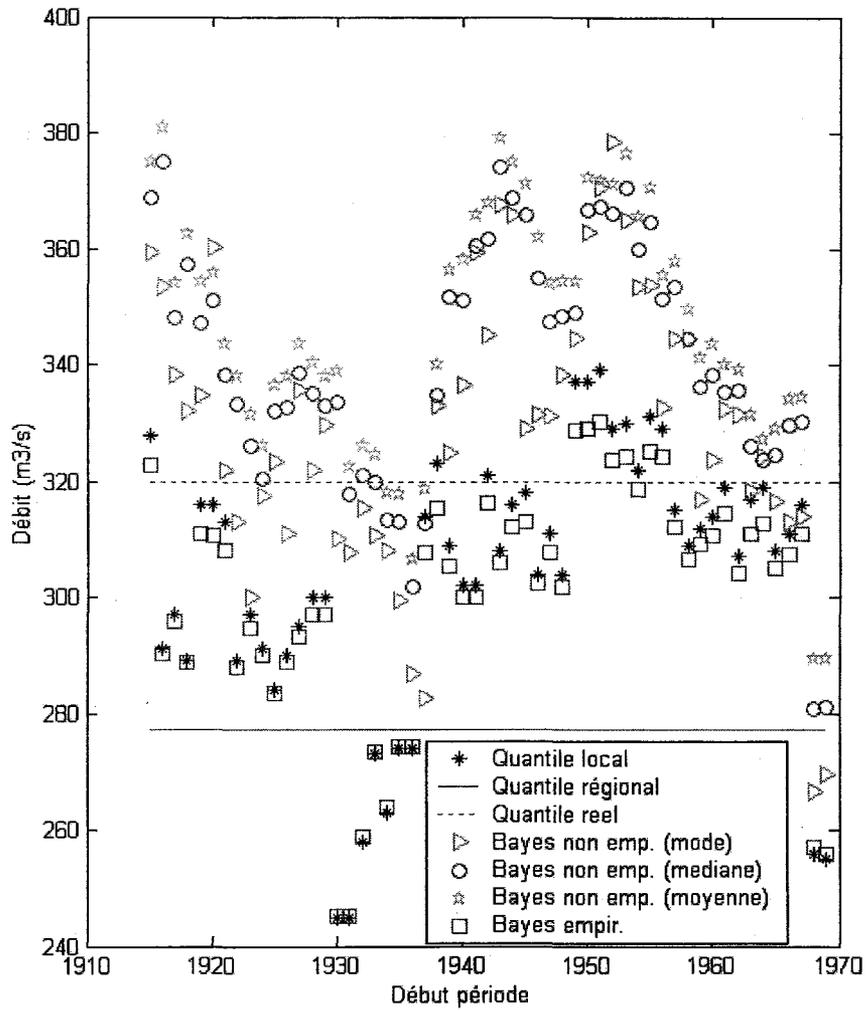


Figure 15: variation du quantile  $q_{100}$  estimé en fonction de la période de 20 années consécutives choisie (Rivière Beaurivage).

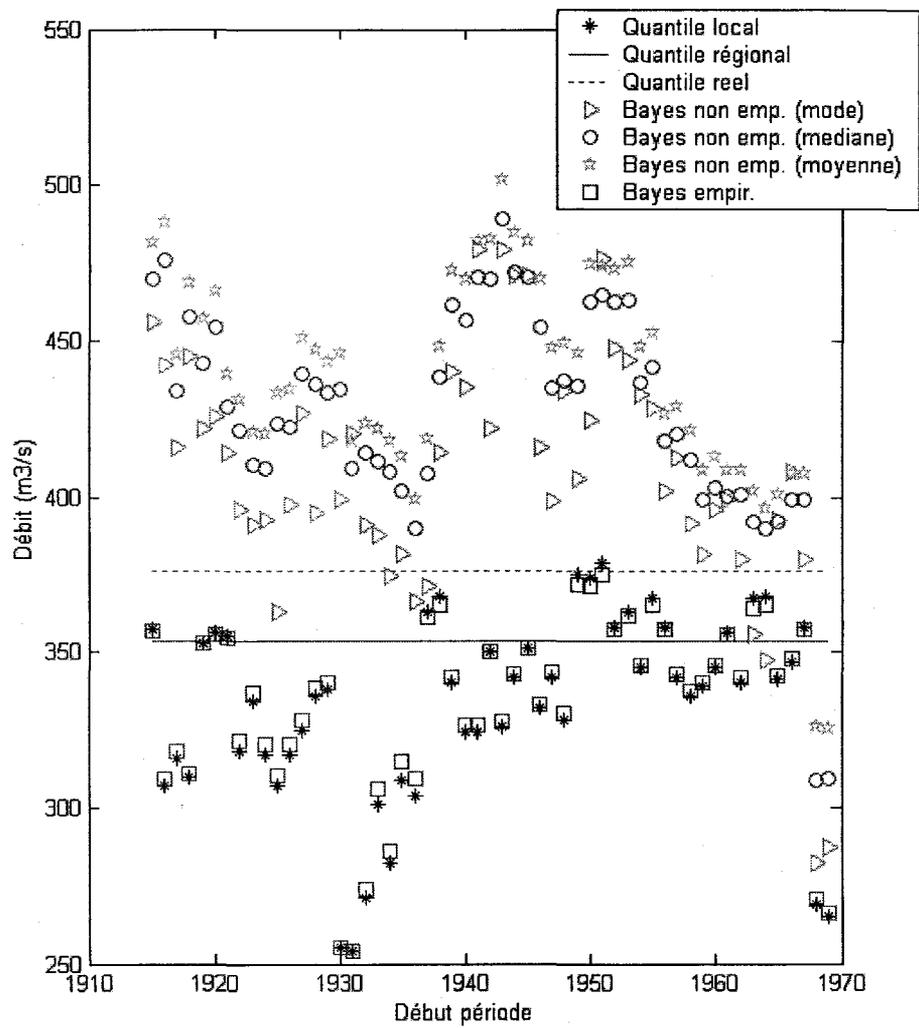


Figure 16: variation du quantile  $q_{1000}$  estimé en fonction de la période de 20 années consécutives choisie (rivière Beaurivage).

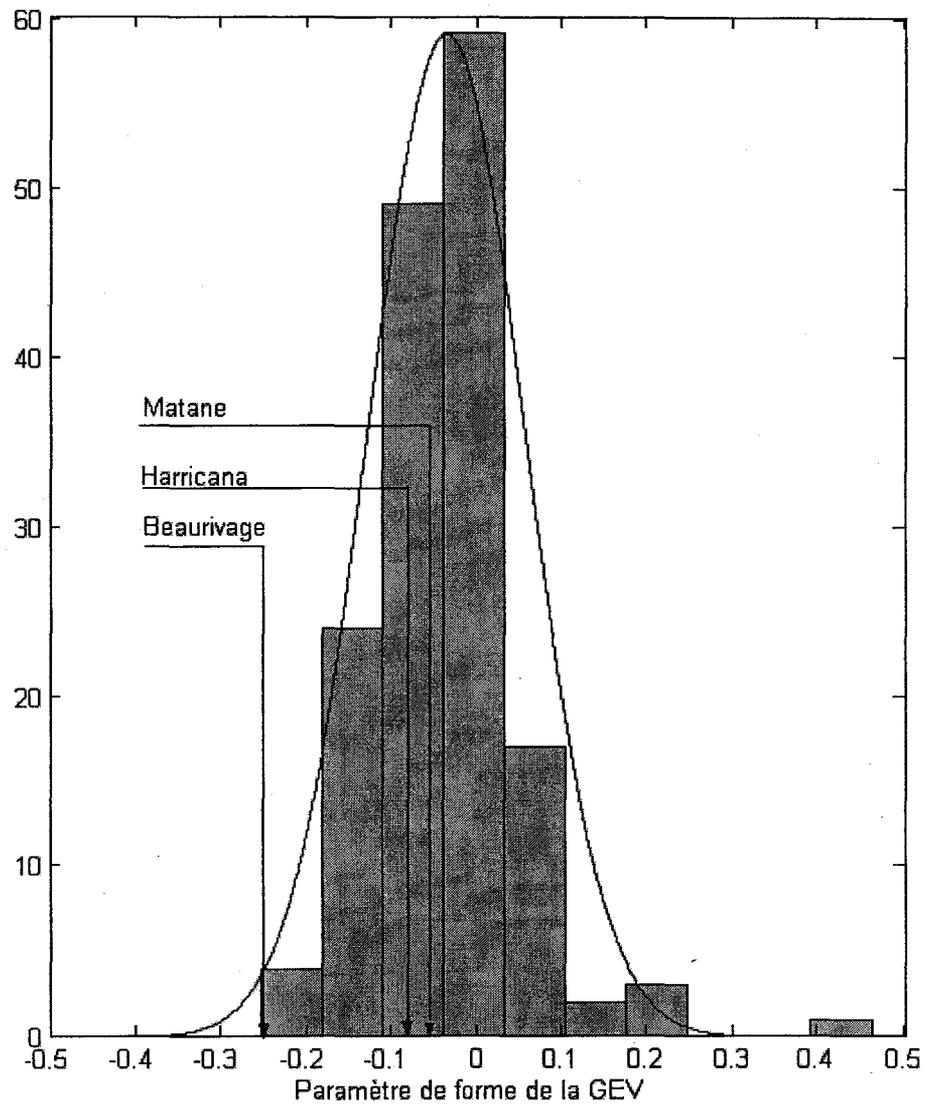


Figure 17: Paramètre de forme  $\hat{\xi}$  de la distribution GEV calculé à partir de  $\frac{q_{1000} - q_{10}}{q_{100} - q_{10}}$