

1 – Introduction

La distribution généralisée de Pareto (GPD) est largement utilisée pour l'ajustement des extrêmes en l'hydrologie. Parmi ses principales utilisations se trouve l'ajustement des dépassements d'un seuil dans la modélisation des inondations ou d'autres phénomènes hydrologiques extrêmes. La matrice de variance-covariance des estimateurs du maximum de vraisemblance (MV) des paramètres d'échelle (β) et de forme (α) de la loi GPD est :

$$\Sigma = \begin{pmatrix} \text{var}(\hat{\alpha}) & \text{cov}(\hat{\alpha}, \hat{\beta}) \\ \text{cov}(\hat{\alpha}, \hat{\beta}) & \text{var}(\hat{\beta}) \end{pmatrix}$$

$$= \begin{pmatrix} (1-\alpha)^2 & \beta(1-\alpha) \\ \beta(1-\alpha) & 2\beta^2(1-\alpha) \end{pmatrix}, \alpha < 0.5$$

2 - Intervalle de confiance du quantile d'ordre p d'une loi GPD

En se basant sur la normalité asymptotique des estimateurs du MV, une distribution approximative est obtenue pour $Q_n(p; \hat{\alpha}, \hat{\beta})$, le quantile estimé d'ordre p, basé sur un échantillon de taille n. On suppose que $Q_n(p; \hat{\alpha}, \hat{\beta})$ suit asymptotiquement une distribution normale de moyenne $Q(p; \alpha_0, \beta_0)$; où α_0 et β_0 sont les vrais valeurs des paramètres. Ceci nous permettra de trouver un intervalle de confiance pour le quantile (ICQ).

3- Méthodologie

Traditionnellement, la normalité asymptotique de $Q_n(p; \hat{\alpha}, \hat{\beta})$ a servi de base de construction de cet ICQ. La probabilité suivante écrite au niveau de 95% nous permet d'écrire l'intervalle de confiance

$$\Pr[-1.96 \leq \frac{Q(p; \hat{\alpha}, \hat{\beta}) - Q(p; \alpha_0, \beta_0)}{\hat{\sigma}} \leq 1.96] \approx 95\% \quad (1)$$

Par conséquent, l'intervalle de confiance conventionnel au niveau de 95% est donnée par:

$$\text{CIQ}_{\text{conventionnel}} = [Q(p; \hat{\alpha}, \hat{\beta}) \pm 1.96\hat{\sigma}] \quad (2)$$

4 - Est-il nécessaire d'ajuster les ICQ conventionnels avec des petits échantillons ?

Il est souhaitable d'avoir les ICQ les plus précis possibles; il est donc important d'évaluer la précision des ICQ représentés par l'équation (2). On peut montrer que le ICQ approche la normalité très lentement. Ceci est vu à partir de la Fig. 1, tirée de Ashkar et El-Adlouni (2015), qui représente les histogrammes de $Q_n(p; \hat{\alpha}, \hat{\beta})$ pour $\beta_0 = 1$, $p = 0,99$ dans deux cas:

$$\alpha_0 = -0,3, n = 100 \quad \text{et} \quad \alpha_0 = 0,4, n = 80$$

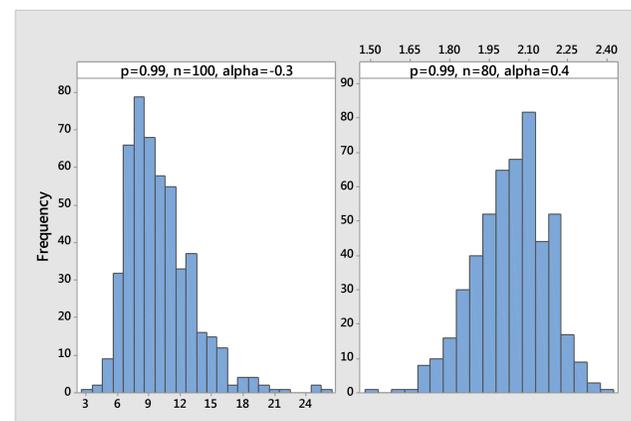


Fig. 1 (tiré de Ashkar et El-Adlouni, 2015)

On constate que même avec une taille d'échantillon relativement grande de $n=100$, l'histogramme peut être assez loin de ressembler à une courbe normale. Cela rend souhaitable d'essayer de trouver une fonction de $Q_n(p; \hat{\alpha}, \hat{\beta})$ qui converge vers la normalité plus rapidement. La transformation suivante de Box-Cox (B-C) peut donc être proposée.

$$Q_n^{BC}(p; \hat{\alpha}, \hat{\beta}) = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln(Y) & \lambda = 0 \end{cases}$$

$$Y = \frac{Q(p; \hat{\alpha}, \hat{\beta})}{\hat{\beta}}$$

Le paramètre B-C, λ , doit être choisi pour normaliser approximativement $Q_n(p; \hat{\alpha}, \hat{\beta})$, ou au moins rendre sa répartition symétrique. En effectuant la transformation BOX-COX inverse, on obtient l'ICQ approximatif souhaité pour le quantile non transformé $Q(p; \alpha_0, \beta_0)$:

$$\hat{\beta}[\lambda(Q_n^{BC}(p; \hat{\alpha}, \hat{\beta}) \pm 1.96\hat{\sigma}_{BC}) + 1]^{1/\lambda} \quad (3)$$

On s'attendrait à ce ICQ ajusté pour constituer une alternative plus précise au ICQ classique de l'équation (2) dans les cas où $Q_n(p; \hat{\alpha}, \hat{\beta})$ n'est pas symétriquement distribué.

5- Comparaison des ICQ conventionnel et ajusté par la transformation Box-Cox

On compare ici les ICQ de 95% obtenues par l'approche conventionnelle et ajustée. Dans l'approche ajustée, tous les estimations $Q_n(p; \hat{\alpha}, \hat{\beta})$ ont été ajustés en utilisant une transformation B-C avec λ , où λ est donné dans (Ashkar et El-Adlouni, 2015). Les résultats de la Fig. 2 sont basés sur $k = 5000$ échantillons GPD générés pour chaque combinaison (α, β) et taille de l'échantillon n. La Fig. 2 donne le pourcentage de couverture des ICQ à 95% en fonction de α et n, $p = 0,95$.

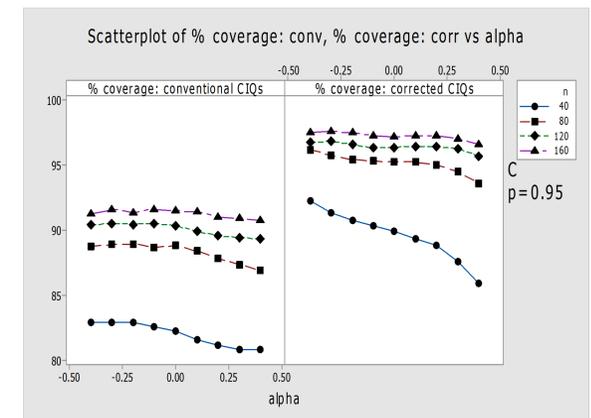


Fig. 2 (tiré de Ashkar et El-Adlouni, 2015)

6 – Conclusion et autre travail en cours

En résumé, pour les cas étudiés (combinaisons de paramètres de forme et de quantile), la méthode ajustée améliore significativement l'estimation de l'intervalle de confiance, sauf dans le cas de $p = 0,90$ où un ajustement CIQ ne semble pas nécessaire. Pour un GPD significativement à queue courte avec α proche de la valeur critique de 0.5, les performances des méthodes conventionnelles et des méthodes ajustées sont affectées négativement. La raison en est que le GPD ajusté peut dans ce cas prendre des formes qui sont rarement trouvées dans la pratique hydrologique (voir Ashkar et El-Adlouni, 2015).

Autres travaux semblables sont en cours pour ajuster les intervalles de confiance pour les quantiles de la loi Gumbel

Référence

Ashkar, F., El-Adlouni, S. (2015). Correcting for small sample non-normality of design flood estimators under a generalized Pareto distribution. Journal of Hydrology 530, 384-391