UTILISATION DE L´ANALYSE FRÉQUENTIELLE POUR SOLUTIONNER QUELQUES PROBLÈMES EN HYDROLOGIE

UTILIZAREA ANALIZEI FRECVENȚIALE PENTRU REZOLVAREA UNOR PROBLEME ÎN HIDROLOGIE

DINU GABRIELA
‘‘Valahia’’ University of Târgovişte

Abstract. *The frequential analysis, as a statical method, can be applied to any type of data; in hydrology it is used especially to solve problems related to extreme events.*

This paper presents the most important techniques to adjust a frequential model such as: the graphical method; the method of the moments; the Gumbel method.

Rezumat. *Analiza frecvențială, ca metodă statistică, se poate aplica oricărui tip de date; în hidrologie se foloseşte, în special, pentru rezolvarea problemelor legate de evenimentele extreme.*

În această lucrare sunt prezentate cele mai importante tehnici de ajustarea a unui model frecvențial, printre care: metoda grafică; metoda momentelor; metoda Gumbel.

Se face o comparație între aceste metode, folosind diferiți parametri, şi se stabilesc principalele criterii pentru alegerea uneia dintre ele.

L´ANALYSE FRÉQUENTIELLE

Prévision et prédiction

Deux approches fort différentes des événements futurs sont utilisées en hydrologie. D´une part les *prévisions* à relativement court terme et, d´autre part, les *prèdictions*, généralement à plus long terme. Cette distinction correspond à des approches différents, mais aussi à des problèmes différents.

Dans le premier cas il s´agit d´un problème de *gestion* d´ouvrage ou de système d´ouvrages, alors que dans le second il s´agit d´un problème de *planification* d´aménagements et de dimensionnement d´ouvrages.

L´analyse fréquentielle est une méthode statistique de prédiction consistant à étudier les événements passés, caractéristiques d´un processus donné (hydrologique ou autre), afin d´en définir les probabilités d´apparition future.

Cette prédiction repose sur la définition et la mise en œuvre d´un *modèle fréquentiel,* qui est une équation décrivant (modélisant) le comportement statistique d´un processus.

Ces modèles décrivent la probabilité d´apparition d´un événement de valeur donnée.

La loi GUMBEL est l´exemple le plus commun de modèle fréquentiel utilisé en hydrolgie:

$$F(x) = e^{-e^{\frac{x-a}{b}}} \tag{1}$$

Dans ce modèle $F(x)$ est la fonction de répartition, ou *fréquence cumulée*, alors que *a* et *b* sont les deux paramètres du modèle de GUMBEL.

La probabilité annuelle d´observer un événement supérieur ou égal à la valeur *x* vaut $p(x) = 1 - F(x)$. On parle de *probabilité de dépassement* ou encore de *probabilité au dépassement.*

En hydrologie on utilise volontiers la notion de *temps de retour* défini par:

$$T(x) = \frac{1}{p(x)} = \frac{1}{1 - F(x)} \tag{2}$$

Cela signifie que, considérant une très longue période, l´événement d´amplitude *x* ou supérieure se produit, *en moyenne*, une fois toutes les $T(x)$ années.

Il convient de noter:

- que la notation de *temps de retour* n´implique aucune régularité dans la survenance des événements: ils n´apparaissent pas régulièrement toutes les $T(x)$ années, ce n´est qu´une valeur moyenne sur une longue période.
- que la probabilité de survenance d´un événement supérieur ou égal à la valeur *x* au cours d´une période donnée doit se calculer en utilisant la loi binomiale.

L´ajustement du modèle, ou encore son *calage* ou sa *spécification* permet de définir les valeurs prises par ses paramètres (a,b,..). Il sera dès lors possible d´exploiter le modèle, par exemple pour déterminer la valeur x correspondant au temps de retour choisi.

L´analyse fréquentielle puisse s´appliquer à n´importe quel type de données. L´hydrologie applique cette méthode statistique pour traiter un problème très important: les événements extrêmes.

AJUSTEMENT DU MODÈLE FRÉQUENTIEL

La loi GUMBEL se prête particulièrement bien à la modélisation des événements extrêmes, les pluies notamment.

1. Présentation de la loi de GUMBEL

La loi de GUMBEL est un cas particulier de la loi de JENKINSON.

- Fonction de dénsité

$$f(x) = \frac{1}{b} \cdot e^{\frac{x-a}{b}} \cdot e^{-e^{\frac{x-a}{b}}} \tag{3}$$

- Fonction de répartition

$$F(x) = e^{-e^{\frac{x-a}{b}}} \tag{4}$$

- Paramètres

  a – paramètre de position;
  b – paramètre d´échelle ou de dispersion

- Variable réduite u

$$u = \frac{x-a}{b} \tag{5}$$

La fréquence cumulée est donnée par les relations pratiques suivantes:

$$\begin{aligned} F(x) &= e^{-e^{-u}} \\ u &= -\ln[-\ln F(x)] \end{aligned} \tag{6}$$

- Expression d´un quantile

Pour trouver la valeur x_q (quantile), correspondant à la fréquence cumulée $F(x_q) = q$, en fonction des deux paramètres a et b il suffit d´inverser la relation (5):

$$x_q = a + b \cdot u_q \tag{7}$$

puis de remplacer u_q par son expresion (6-b).

2. Techniques d´ajustement

Méthode graphique

La méthode graphique repose sur le fait que l´expresion d´un quantile (7) correspond à l´équation d´une droite. En conséquence, dès lors que les points de la série à ajuster peuvent être reportés dans un système d´axes x- u, il est possible de tracer la droite qui passe le mieux par ces points et d´en déduire les deux paramètres a et b de la loi (fig.1).

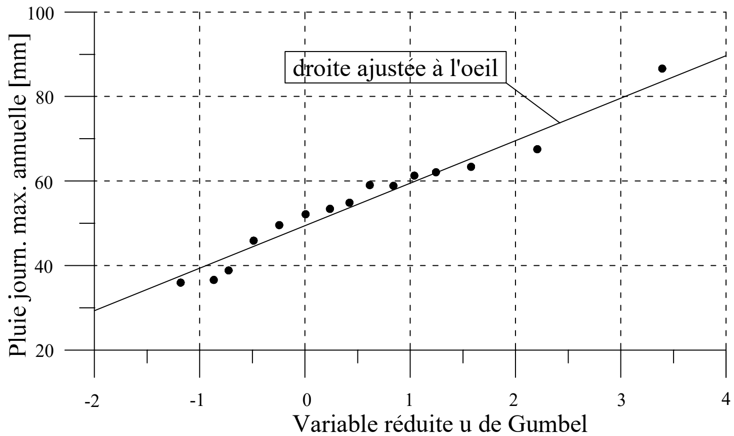

Figure 1 Principe de la méthode d´ajustement graphique

Dans la mesure où les points x_i sont connus (ils font partie de la donnée du problème), il suffit de définir les coordonnées u_i correspondant à chaque point pour pouvoir le positionner dans le graphique.

Ces coordonnées se déterminent à partir de la relation inverse de la fonction de répartition qui donne u en fonction de la fréquence cumulée $F(x)$. Il s´agit donc essentiellement d´estimer la probabilité de non dépassement $F(x_i)$ qu´il convient d´attribuer à chaque valeur x_i.

Il existe de nombreuses formules d´estimation de la fréquence cumulée $\hat{F}(x)$: Formules de *positionnement du point figuratif* ou formule d´évaluation de la *fréquence empirique.*

Elles reposent toutes sur un tri de la série par valeurs croissantes (ou décroissantes), permettant d'associer à chaque valeur son rang r. Ces formules peuvent pratiquement toutes être résumées par une relation générale qui garantit la symétrie autour de la médiane:

$$\hat{F}\left(x_{[r]}\right) = \frac{r-\alpha}{n+1-2\alpha} \tag{8}$$

où: n est la taille de l´échantillon ;

$x_{[r]}$ - la valeur de rang r ;

α - un coefficient compris entre 0 et 0,5.

Le choix de la formule ne fait pas l´unanimité.

Méthode des moments

La méthode des moments consiste à utiliser lestimation « classique » $\hat{\mu}$ et $\hat{\sigma}^2$ des deux premiers moments de la population, basées sur les caractéristiques de l´échantillon $\bar{x}$ et s :

$$\hat{\mu} = \bar{x}; \quad \hat{\sigma} = s \cdot \sqrt{\frac{n}{n-1}} \tag{9}$$

Compte tenu des caractéristiques de la loi de GUMBEL :

$$\begin{aligned} \mu &= a + b \cdot 0{,}5772 \\ \sigma^2 &= 1{,}645 \cdot b^2 \end{aligned} \tag{10}$$

les formules relatives à l´estimation par la méthode des moments s´obtiennent facilement:

$$\begin{aligned} \hat{b} &= 0{,}7797 \cdot \hat{\sigma} \\ \hat{a} &= \hat{\mu} - \hat{b} \cdot 0{,}5773 \end{aligned} \tag{11}$$

Méthodes des moindres rectangles ou de GUMBEL

La méthode des moindres rectangles consiste à «numériser» la technique de l´ajustement graphique. Elle a été introduise par GUMBEL.

La solution des *moindres rectangles* conduit à trouver la droite bissectrice des solutions classiques de la régression par moindres carrés de y en x d´une part et de x en y d´autre part. Elle a l´avantage de ne pas faire intervenir les produits croisés (soit la covariance entre les deux variables). Il s´agit donc d´une solution particulièrement simple pour l´écuation de la droite $y = a + b \cdot x$:

$$b = \frac{s_y}{s_x} \qquad a = \bar{y} - b \cdot \bar{x} \tag{12}$$

Application à l´espace u-x pour la loi de GUMBEL

Dans le cas présent l´axe x est remplacé par l´axe u de la variable standardisée de GUMBEL et l´axe y est remplacé par celui de la variable hydrologique étudiée que nous notons ici x .

Nous aurons donc:

$$\hat{b} = \frac{s_x}{s_y}; \quad \hat{a} = \overline{x} - \hat{b} \cdot \overline{u} \tag{13}$$

GUMBEL a remarqué que, si la taille n de l'échantillon est connue, toutes les positions u_i sont connues. Il est donc possible de calculer une fois pour toutes les valeurs $\overline{u}$ et s_u en fonction de r.

En effet nous avons :

$$u_i = -\ln\left[-\ln\frac{r_i - \alpha}{n + 1 - 2\alpha}\right], \quad \forall i = 1,...n \tag{14}$$

Il est donc facile de tabuler les grandeurs que nous noterons $\overline{u}(n)$ et $s_u(n)$.

Finalement les estimateurs sont donnés par les deux relations :

$$\hat{b} = \frac{s_x}{s_u(n)} \quad \text{et} \quad \hat{a} = \overline{x} - \hat{b} \cdot \overline{u}(n) \tag{15}$$

COMPARAISON DES MÉTHODES, CRITÈRES DE CHOIX

Il est toujours délicat de recommander l'utilisation de telle ou telle méthode.

- Pour un calcul manuel on recommande la méthode des moindres rectangles (avec un coefficient $\alpha = 0,5$).
- Pour un calcul informatisé on recommande la méthode du maximum de vraisemblance.
- La méthode de moments est plus sensible à une éventuelle autocorrélation de la série de données que les autres.

BIBLIOGRAPHIE

1. Gumbel, E.-J., 1958 **–** *Statistics of extremes*. Columbia University Press, New-York

2. Lubès, H., Masson, J-M., 1991 – *Méthode des moments de probabilité pondérés-Application à la loi de Jenkinson,* Hydrol. Continent. Vol.6, No.1

3. **xxx,** 1983 – OMM. Guide des pratiques hydrologiques, Volume II: Analyse, prévision et autres applications, OMM No.168, Genève, 1983