

# Méthodes graphiques pour l'analyse des débits de crue

## Graphical methods of flood discharge analysis

PAR E.-J. GUMBEL

Columbia University (New-York)

Conférence faite à Paris, 12, place des Etats-Unis (16<sup>e</sup>), le 22 avril 1956  
sous les auspices de la Société Hydrotechnique de France

*Emploi de la méthode graphique, avec l'aide d'un papier à probabilité, et permettant de déterminer les durées de retour des crues. Confrontation des résultats théoriques et des observations. Théorie des valeurs extrêmes. Application aux débits de crue. Estimation des paramètres à utiliser. Comparaison des débits de crue de divers cours d'eau.*

*DISCUSSION. Hypothèse d'un champ de variation illimité pour la distribution des débits. Remarque théorique sur la rapidité de convergence vers la loi asymptotique. Erreurs possibles dans l'application des méthodes graphiques aux débits de crue.*

*INTERVENTION DE M. BERNIER. — Formes limites générales de la loi de la plus grande valeur d'un échantillon de  $n$  variables aléatoires indépendantes. Remarques sur la convergence vers les lois limites des valeurs extrêmes; cas des lois : a) normale; b) de Laplace; c) de Galton-Gibrat. Propriétés et ajustements de la loi de Fréchet. Applications numériques. Conclusion.*

*The graphical method using probability paper and enabling the periods after which floods recur to be determined. Comparison of theoretical results with observations. The theory of extreme values. Application to flood discharges. Estimation of the parameters to be used. Comparison of flood discharges in various water-courses.*

*DISCUSSION.—Hypothesis of an unbounded field of variation for the distribution of flood discharges. Theoretical remark on the speed at which it converges to the asymptotic law. Possible errors due to the application of graphical methods to flood discharges.*

*Mr. BERNIER'S COMMENT.—General limiting forms of the law of the greatest value of a sample of  $n$  independent random variables. Remarks concerning the convergence of the extreme values towards limiting laws; types of laws : a) normal; b) Laplace; c) Galton-Gibrat. Properties of and adjustments for Fréchet's law. Numerical applications. Conclusion.*

### I. — Méthodes graphiques

Soit  $x$  une variable aléatoire continue, soit  $\alpha$  et  $u$  des paramètres, soit  $F(x)$  une fonction de probabilité totale telle qu'on puisse écrire :

$$(1.1) \quad F(\alpha, u, x) = \Phi(y),$$

où :

$$(1.2) \quad y = \alpha(x - u)$$

est une *variable réduite* et  $\Phi(y)$  ne contient plus de paramètres.

On construit un papier à probabilité en choisissant  $x$  comme ordonnée et  $y$  comme abscisse, tous les deux en échelle linéaire. En traçant une échelle  $\Phi(y)$  parallèle à  $y$ , on obtient les valeurs :

$$(1.3) \quad x = u + y/\alpha$$

en fonction des probabilités  $\Phi$ . Si les observations se passent dans le temps et si les intervalles entre deux observations sont constants, la fonction :

$$(1.4) \quad T(y) = [1 - \Phi(y)]^{-1} = T(x)$$

peut être appelée la *durée du retour*. C'est le nombre d'observations qu'on doit faire en moyenne pour obtenir une valeur supérieure ou égale à  $x$ . On trace  $T(x)$  sur une échelle parallèle à l'abscisse. De cette manière on obtient  $x$  en fonction du temps, pourvu que la distribution initiale soit donnée.

Pour l'emploi de ce papier, il nous faut une règle (10) qui décide les positions de chaque observation. Soit  $x_i$  la  $i$ ème ( $i = 1, 2, \dots, n$ ) parmi  $n$

observations. Une méthode générale valable pour n'importe quelle distribution continue est de choisir comme position la moyenne  $\bar{F}_i$  des fréquences des  $i$ èmes valeurs qui est :

$$(1.5) \quad \bar{F}_i = i/(n + 1).$$

Il en résulte que les durées de retour empiriques deviennent :

$$(1.6) \quad T(x_i) = (n + 1)/(n + 1 - i)$$

ce qui donne  $T(x_1) = 1 + n^{-1}$  pour la plus petite valeur et  $T(x_n) = n + 1$  pour la plus grande valeur.

Si la fonction de probabilité  $F(x)$  est bien choisie, les observations  $x$  tracées de cette manière se trouvent dans l'entourage de la droite théorique (1.3). Pour tracer cette ligne, nous avons besoin d'une estimation des paramètres  $\mu$  et  $\sigma$ . La manière la plus simple est l'usage de la méthode des moindres carrés. On peut rendre minimum soit les distances horizontales, soit les distances verticales. Un compromis entre les deux méthodes conduit à l'estimation :

$$(1.7) \quad 1/\hat{\alpha} = s/\sigma_n(y) \quad ; \quad \hat{\mu} = m - \bar{y}_n/\hat{\alpha}$$

Dans ce système,  $m$  et  $s$  sont la moyenne arithmétique et l'écart-type :

$$(1.8) \quad s = \sqrt{\sum (x_i - m)^2 / (n - 1)}$$

des valeurs observées, tandis que  $\bar{y}_n$  et  $\sigma_n(y)$  sont l'espérance mathématique et l'écart-type des valeurs théoriques réduites  $y$  obtenus à l'aide de :

$$(1.9) \quad \Phi(y_i) = i/(n + 1)$$

Reste à savoir si les déviations entre la théo-

rie et les observations peuvent être tolérées ou non. Cela nous intéresse en l'espèce pour les plus grandes valeurs de la variable. Car la ligne droite (1.3) servira à l'extrapolation dans le temps faite à l'aide de l'échelle des durées de retour. Il est d'usage commun parmi les statisticiens d'utiliser le critère dit  $\chi^2$ . On obtient ainsi la probabilité  $P(\chi^2)$  pour que les déviations entre la théorie et les observations soient dues au hasard. Mais ce critère dépend de deux décisions arbitraires, la longueur des classes dont l'influence est connue et le commencement de la première classe dont l'influence est inconnue. J'ai montré par un exemple numérique (9), valable pour les mêmes valeurs des paramètres et les mêmes longueurs des classes, que des petites variations du commencement entraînent des variations de  $P(\chi^2)$  de 0,023 jusqu'à 0,705. C'est pour ces raisons que nous nous opposons à l'emploi de cette méthode pour des variables continues. Il nous faut donc une autre méthode pour contrôler l'accord entre la théorie et les observations.

Nous nous bornons ici aux plus grandes valeurs de la variable. Soit  $T$  la durée de retour théorique, alors il existe une probabilité de 2/3 pour qu'une durée de retour observée soit contenue dans l'intervalle :

$$.32 T \quad \text{et} \quad 3.13 T$$

La probabilité de 2/3 est choisie ici parce qu'elle correspond à la probabilité de l'intervalle  $\bar{x} \pm \sigma$  pour la distribution normale. On tracera cet intervalle autour des plus grandes valeurs théoriques de  $x$  et on acceptera la théorie pourvu que les observations soient contenues dans cet intervalle qu'on utilisera aussi pour l'extrapolation.

## II. — Théorie des valeurs extrêmes

Ces méthodes graphiques seront employées maintenant pour une des distributions asymptotiques de la plus grande valeur. Soit  $F(x)$  une fonction de probabilité; alors la probabilité, pour que la plus grande de  $n$  observations indépendantes soit inférieure à  $x$ , est  $F^n(x)$ . La question de savoir s'il existe une distribution asymptotique valable si  $n$  augmente a été étudiée par plusieurs auteurs parmi lesquels FRÉCHET (8), FISHER et TIPPETT (7) qui ont montré qu'il existe trois distributions asymptotiques liées à certaines conditions analytiques. Nous nous bornons ici à une de ces distributions.

Soit  $x$  une variable illimitée vers la droite et soit :

$$f(x) = F'(x)$$

la distribution initiale; alors on introduit deux valeurs,  $u_n$  et  $\alpha_n$  étant les solutions de .

$$(2.1) \quad F(u_n) = 1 - 1/n$$

$$(2.2) \quad \alpha_n = n f(u_n)$$

Admettons enfin que la probabilité initiale pour des grandes valeurs de  $x$  soit telle que :

$$(2.3) \quad F(x) = 1 - \frac{1}{n} e^{-\alpha_n(x - u_n)}$$

alors la distribution asymptotique de la plus grande valeur est :

$$(2.4) \quad \Phi(x) = \exp[-e^{-y}] \quad ; \quad y = \alpha_n(x - u_n)$$

La signification de  $u_n$  est bien simple, c'est la plus grande valeur la plus probable. La signification de  $\alpha_n$  sera donnée plus tard. Des tableaux numériques donnant  $\Phi$  en fonction de  $y$ , et  $y$  en fonction de  $\Phi$ , ont été publiés par le U.S. National Bureau of Standards (22). Ces tables nous permettent la construction d'un papier à probabilité décrit dans le premier paragraphe.

Montrons d'abord une application théorique de cette formule. Une opinion courante due à l'éminent statisticien R. A. FISHER (7) est qu'on ne peut pas utiliser la distribution asymptotique (2.4) pour les plus grandes valeurs normales. Or, tout dépend de ce qu'on entend par l'accord entre cette théorie et les valeurs numériques valables pour la distribution normale. La figure 1 compare les médianes des plus grandes valeurs normales en fonction de  $n$  calculées par B. DE FINETTI (6) aux valeurs obtenues par l'usage de (2.4) à l'aide des paramètres donnés par (2.1) et (2.2).

On voit immédiatement que cette convergence est rapide.

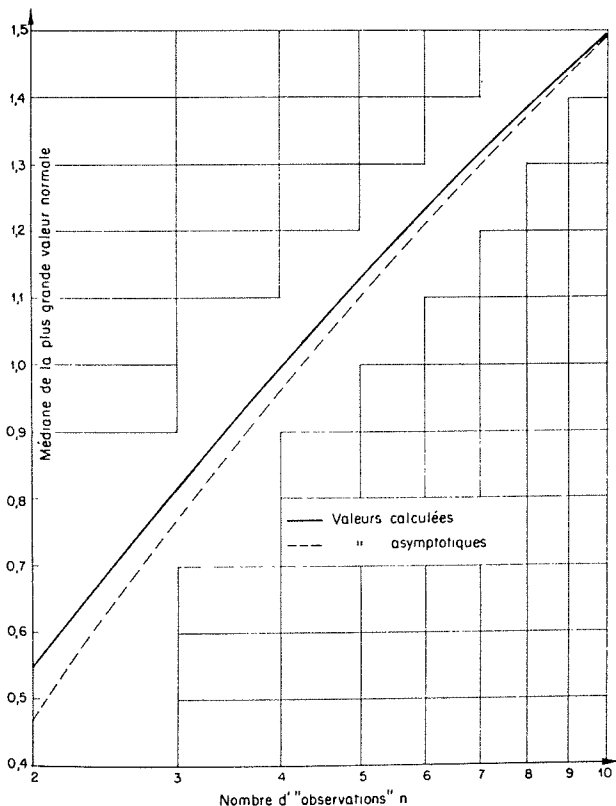


FIG. 1

La médiane des extrêmes normales.

L.H.C. TIPPETT (21) a calculé les probabilités  $F^n(x)$  des plus grandes valeurs normales en fonction du nombre d'observations  $n$ . Prenons  $n = 100$ . Si l'on calcule les différences des probabilités on obtient une approximation de la

distribution exacte des plus grandes valeurs normales. D'autre part, on obtient une approximation de la distribution théorique en prenant les différences des probabilités (2.4). Les valeurs numériques et théoriques ainsi obtenues sont tra-

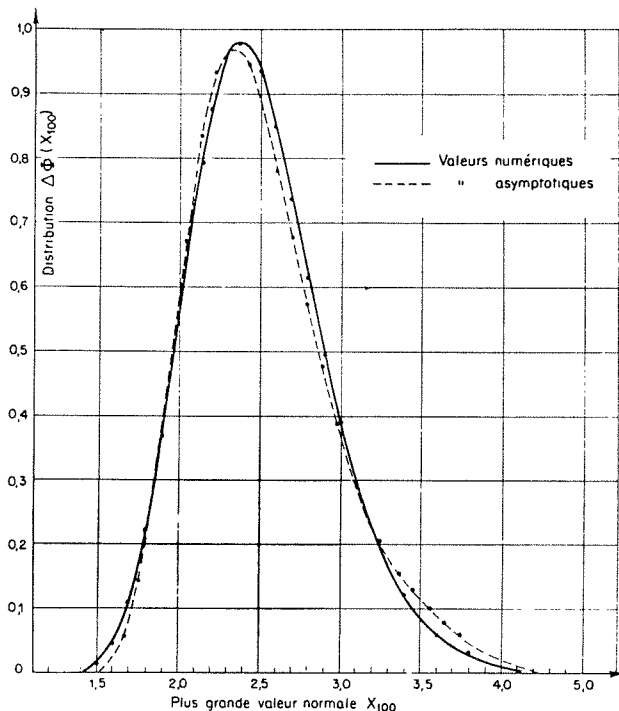


FIG. 2

Distribution des plus grandes valeurs normales pour 100 observations.

cées dans la figure 2 qui prouve que les deux distributions sont assez proches.

Cette application de la théorie des valeurs extrêmes est basée sur la connaissance de la distribution initiale. Dans les applications pratiques, on l'ignore. Mais l'hypothèse du comportement asymptotique (2.3) est assez naturelle puisqu'elle vaut pour la majorité des distributions usuelles en statistique. Dans la première application, les paramètres  $\alpha_n$  et  $u_n$  étaient connus, maintenant ils sont inconnus. Ecrivons dorénavant  $\alpha$  et  $u$ . L'estimation de ces deux paramètres doit être basée sur les plus grandes valeurs observées, seule source de renseignements qui soit à notre disposition.

Les deux premiers moments qu'on obtient à l'aide de (2.4) sont :

$$\bar{x} = u + \gamma/\alpha \quad ; \quad \sigma^2 = \pi^2/(6 \alpha^2)$$

Une modification légère de ces deux équations, d'après la méthode introduite dans le premier paragraphe, mène aux estimations :

$$(2.5) \quad \hat{u} = m - \bar{y}_n/\hat{\alpha} \quad ; \quad 1/\hat{\alpha} = s/\sigma_n$$

où  $m$  et  $s$  sont la moyenne et l'écart-type des plus grandes valeurs observées. Un tableau de  $\bar{y}_n$  et de  $\sigma_n$  se trouve dans la brochure (10). La durée de retour  $T$  converge pour des grandes valeurs de  $x$  vers :

$$(2.6) \quad \log T = y$$

Si l'on trace les durées de retour sur une parallèle à l'abscisse, cette échelle devient logarithmique pour des grandes valeurs de  $T$ . Il en résulte que les maxima des plus grandes valeurs :

$$(2.7) \quad x = u + (\log T)/\alpha$$

augmentent en fonction linéaire du logarithme du temps et que  $1/\alpha$  est la pente de cette ligne. Les courbes de contrôle établies dans le premier paragraphe deviennent des lignes droites parallèles à la ligne théorique (2.7). L'écart entre les courbes de contrôle augmente considérablement avec  $T$ .

Une propriété importante de la distribution (2.4) est la stabilité : le maximum de la plus grande valeur la plus probable augmente avec le nombre d'observations, mais l'écart-type reste le même.

### III. — Application aux débits des crues

La probabilité asymptotique (2.4) a été utilisée souvent pour l'analyse des débits de crue. (Voir les articles 1 à 5 et 10 à 20 de la bibliographie.) Admettons d'abord que les conditions nécessaires pour l'application de la théorie ne sont remplies que d'une façon imparfaite. D'abord, il nous faut des observations précises; beaucoup ne le sont pas. Puis nous avons besoin de longues séries d'observations comparables. C'est rarement le cas puisque tout aménagement du cours d'eau change la rivière. Enfin, nous supposons que les observations initiales (les débits journaliers) dont on tire les plus grandes valeurs annuelles (les crues) sont indépendantes et que le nombre d'observations est très élevé. Ce n'est pas le cas. Le nombre d'observations, fixé par le rythme annuel, est 365. En plus, le débit d'aujourd'hui dépend fortement du débit d'hier. Le nombre d'observations indépendantes est donc de beaucoup inférieur à 365.

D'autre part, la distribution de la plus grande valeur ne dépend que des propriétés analytiques de la distribution initiale valable pour des grandes valeurs de la variable. Il en résulte qu'elle peut être appliquée même à des observations dépendantes pourvu que la distribution résultante des dépendances ait les propriétés asymptotiques nécessaires. Enfin, pour maintes rivières, la distribution initiale des débits ressemble déjà à la distribution asymptotique des crues. Dans ce cas, la stabilité de la distribution asymptotique permet son usage.

Evidemment, cette théorie ne peut être appliquée qu'à condition qu'il n'existe qu'un seul maximum annuel. S'il y a deux influences saisonnières, disons la pluie en automne et la fonte des neiges au printemps, cette méthode doit être appliquée séparément aux deux maxima annuels.

Examinons enfin quelques résultats. Les cour-

bes en zigzag (fig. 3 à 7) tracées sur le papier à probabilité représentent les observations, les lignes droites la théorie. Pour justifier la prévision, nous traçons encore les courbes de contrôle des plus grandes crues.

Le tableau suivant donne les détails de l'estimation des paramètres pour les maxima annuels du Verdon à Quinson, de 1874 à 1954, et pour le Rhin à Rheinfelden, 1869 à 1954.

#### Estimation des paramètres

SYMBOLE	LE VERDON	LE RHIN
Nombre d'années : $n$ ....	81	86
Moyenne : $m$ .....	303,815	2579,54
Ecart-type : $s$ .....	193,148	684,80
Moyenne réduite : $\bar{y}_n$ ....	0,55705	0,5580
Ecart-type réduit : $\sigma_n$ ....	1,1945	1,1980
Paramètre : $1/\alpha$ .....	= 165,16	571,62
Paramètre : $u$ .....	= 211,81	2260,57

Les observations et les droites théoriques sont tracées dans les figures (3) et (4). L'accord général entre la théorie et les observations est tout à fait satisfaisant. Pourtant les deux graphiques montrent quelques particularités. Pour le Verdon, les minima de débits de crue sont assez faibles. Or, la distribution des plus grandes valeurs est illimitée vers la droite. Mais la probabilité pour la valeur réduite  $y = -2.5$  est déjà  $5 \cdot 10^{-6}$ . Si le paramètre  $u$  est petit et si  $1/\alpha$  est du même ordre de grandeur, l'extrapolation vers les petites valeurs devient illégale puisqu'elle conduirait à des valeurs négatives.

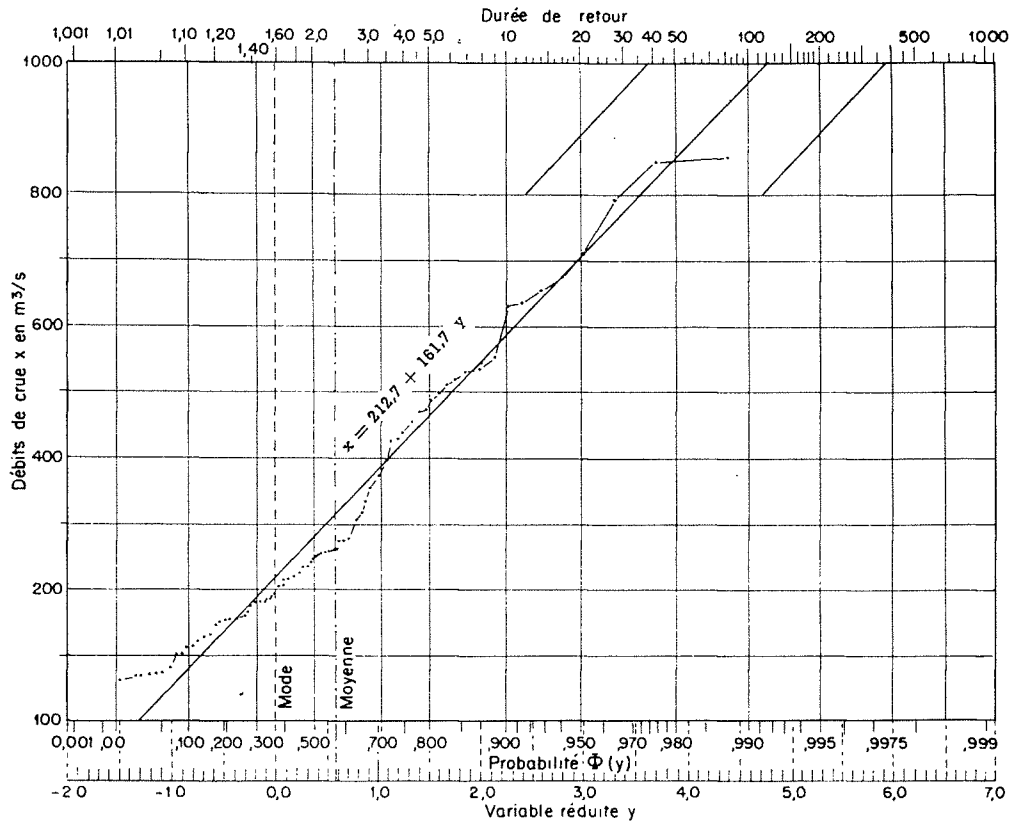


FIG. 3  
Le Verdon à Quinson.

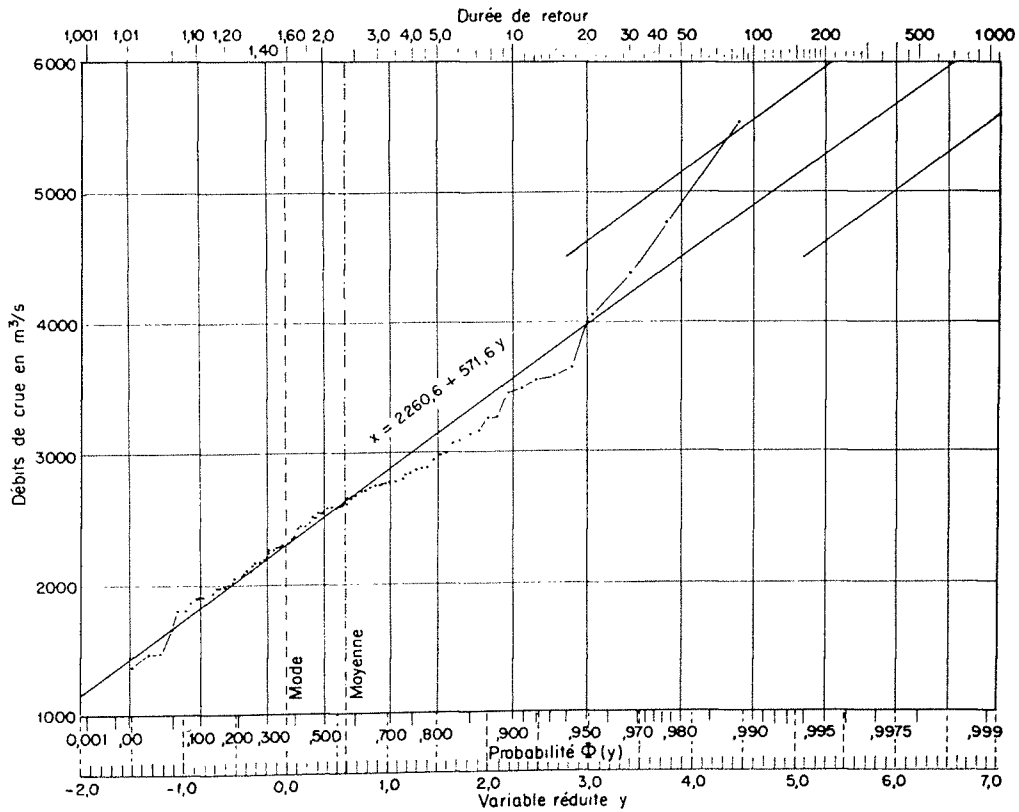


FIG. 4  
Le Rhin à Rheinfelden.

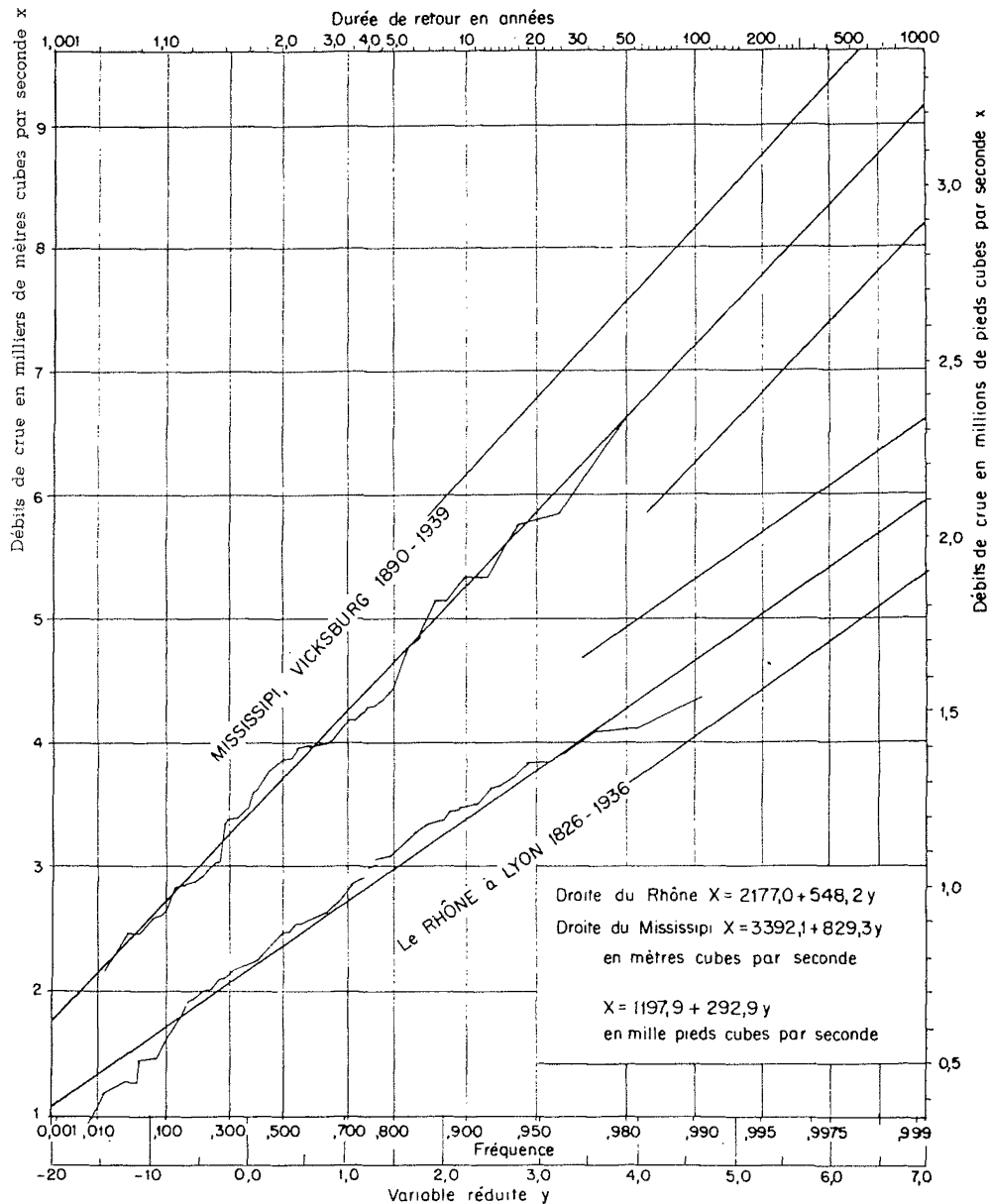


FIG. 5

C'est un manque de beauté de la part de la formule, mais il n'entraîne pas de conséquences graves. Ce qui nous intéresse, ce sont les maxima et non les minima des crues.

Dans la figure 4, on réalise que la durée de retour du maximum surpasse un petit peu les limites données par les courbes de contrôle. En effet, la plus grande crue observée en 86 ans correspond à une durée de retour d'à peu près 300 ans, un phénomène qui est tout à fait possible. La droite théorique s'écarte sensiblement de la courbe observée pour les valeurs supérieures à 4,500 m<sup>3</sup>/s. Pourtant, on peut accepter la théorie puisque la plus grande crue ne dépasse presque pas la courbe de contrôle.

La figure 5 compare les débits de crue du Mis-

sissippi à Vicksburg et du Rhône à Lyon. A l'encontre de la figure 4, le maximum observé pour le Rhône est inférieur à sa valeur théorique, tandis que les deux valeurs coïncident pour le Mississippi. La figure 6 montre les crues du Mississippi à une station du nord située en Minnesota. L'accord entre la théorie et les observations est remarquable puisqu'il s'agit seulement de 20 observations. La dernière figure (7) montre enfin deux rivières situées dans le Tennessee. Ces observations ont servi à la construction des grands barrages et des usines hydro-électriques qui ont tellement contribué au bien-être de cette vallée.

En somme, l'application de la distribution asymptotique (2.4) des plus grandes valeurs aux débits de crue a donné de bons résultats.

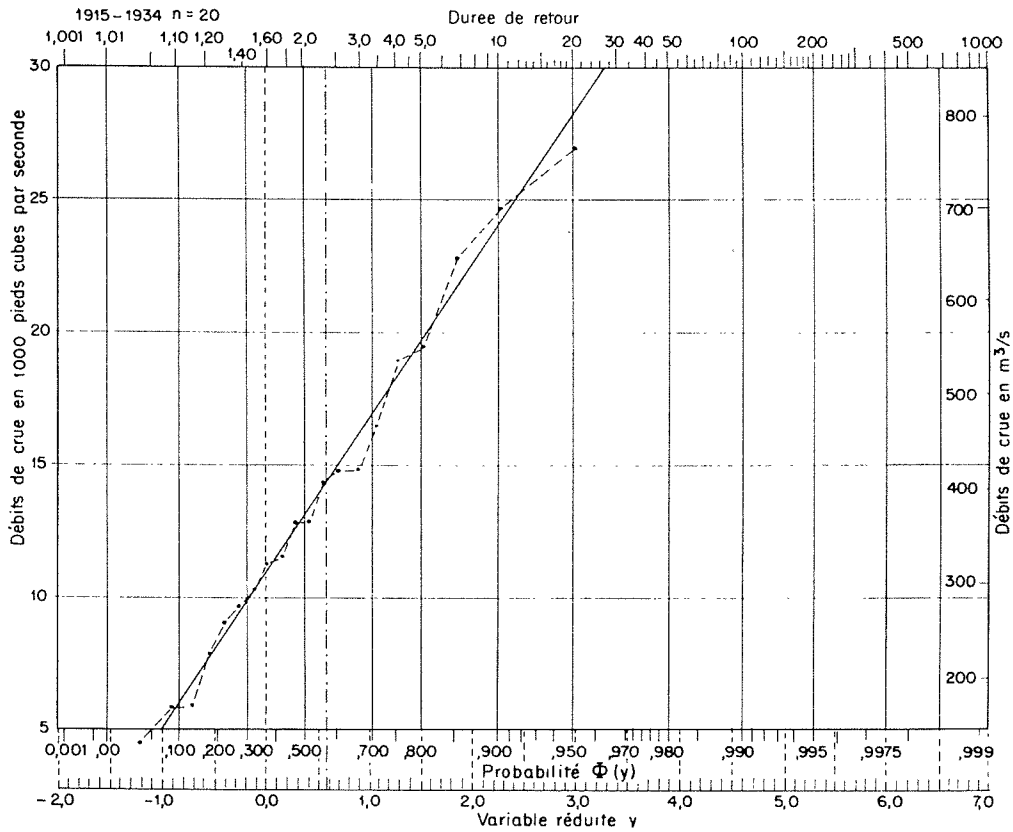


Fig. 6 Mississippi à Elk River (Minnesota).

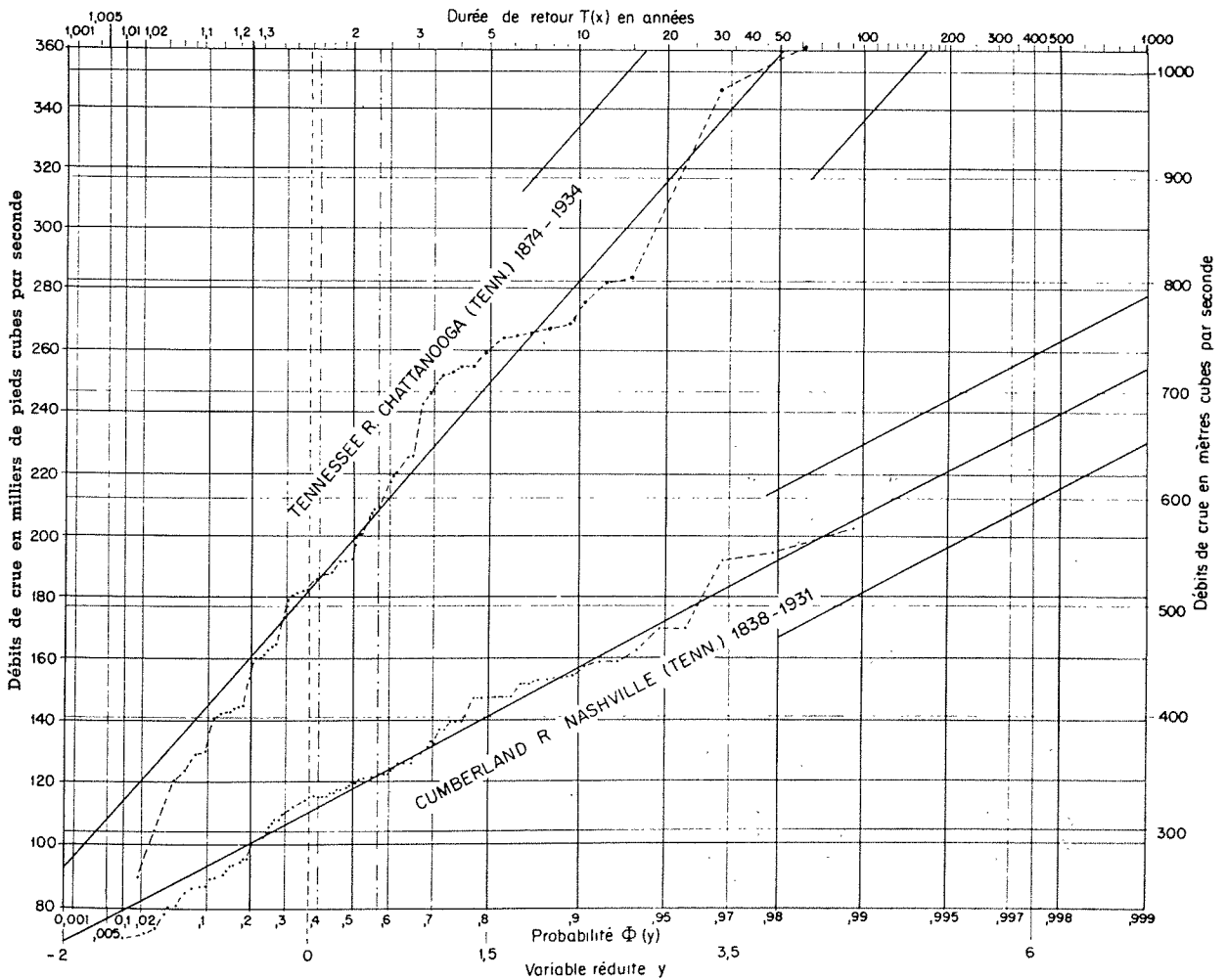


Fig. 7 Crues de la Tennessee River et de la Columbia River.

## BIBLIOGRAPHIE

- (1) BENHAM (A. D.). — The estimation of extreme flood discharges by statistical methods. *Proceedings, the New Zealand Institution of Engineers*, Wellington, 1950. Vol. 36, pp. 119-187.
- (2) CARTER (R. W.). — Floods in Georgia, frequency and magnitude. *Geological Survey Circular*, No. 100. Washington, 1951.
- (3) CRAGWELL (J. S.). — Floods in Louisiana, magnitude and frequency. *State of Louisiana, Dept. of Highway and Geological Survey*, Washington, 1952.
- (4) CROSS (William P.). — Floods in Ohio, magnitude and frequency. *Bull. No. 7, Ohio Water Resources Board*, Columbus, Ohio, 1946.
- (5) DALRYMPLE, TATE. — Use of streamflow records in design of bridge waterways. *Proc. Highway Research Board*, vol. 26, Washington D.C., 1946.
- (5a) DALRYMPLE, TATE. — Regional Flood Frequency. *Highway Research Board, Research Report*, No. II-B, *Surface Drainage*. Washington D.C., 1950.
- (6) FINETTI (B. de). — Sulla legge di probabilità degli estremi. *Metron*, vol. 9, p. 127, Roma 1932.
- (7) FISHER (R. A.) and TIPPETT (L. H. C.). — Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Cambridge Phil. Soc.*, vol. 28, page 180, Cambridge, 1928.
- (8) FRECHET (M.). — Sur la loi de probabilité de l'écart maximum. *Ann. de la Société Pol. de Math.*, vol. 6, p. 93, Cracovie, 1927.
- (9) GUMBEL (E. J.). — On the reliability of the classical Chi square test. *Ann. Math. Stat.*, vol. 14, 143, p. 253.
- (10) GUMBEL (E. J.). — Statistical theory of extreme values and some practical applications. *Nat. Bur. of Standards, Appl. Math. Series* 33, Washington, 1954.
- (11) IZZARD (Carl). — Peak discharge for highway drainage design. *Proc. Am. Soc. Civ. Engrs.*, vol. 79 Separate 320, New-York, 1953.
- (12) MITCHELL (William D.). — Floods in Illinois; magnitude and frequency, *U.S. Geological Survey and Div. of Waterways*. State of Illinois, 1954.
- (13) PIERCE (L. B.). — Floods in Alabama, magnitude and frequency. *Geological Survey Circular* 342, Washington, 1954.
- (14) POTTER (W. D.). — Simplification of the Gumbel method for computing probability curves. *Department of Agriculture, Soil Conservation Service*, Technical Paper 78, Washington, D.C., 1949.
- (15) POTTER (W. D.). — Normalcy tests of precipitation and frequency studies of runoff on small watersheds. *Department of Agriculture, Soil Conservation Service*, Tech. Bull. No. 985, Washington, D.C., 1949.
- (16) POTTER (W. D.). — Surface Runoff from Agricultural Watersheds. *Highway Research Board, Research Report* No. II-B, *Surface Drainage*. Washington, D.C., 1950.
- (17) RANTZ (S. E.) and RIGGS (H. C.). — Magnitude and Frequency of Floods in the Columbia River Basin, *Department of the Interior, Geological Survey Water Supply*, paper 1080, Washington, D.C., 1949.
- (18) ROBINSON (W. H.) and BODHAINE (G. L.). — Floods in Western Washington, frequency and magnitude in relation to drainage basin characteristics. *Geol. Survey Circular* 192, Washington, 1952.
- (19) SHUH-CHAI LEE. — The return period and variability of floods in China. *Essays and papers in memory of late President Fu Ssu-nien*, Taipan, 1952.
- (20) SCHWOB (H. H.). — Iowa floods, magnitude and frequency. *Iowa Highway Research Board*, bulletin No. 1, 1953.
- (21) TIPPETT (L. H. C.). — On the extreme individual and the range of samples taken from a normal population. *Biometrika*, vol. 17, p. 364, 1925.
- (22) Probability tables for the analysis of extreme value data. *U.S. National Bureau of Standards, Appl. Math*, series 22, Washington, 1953.

## COMPTE RENDU DE LA CONFÉRENCE DE M. GUMBEL ET DE LA DISCUSSION QUI A SUIVI

L'essentiel de la Conférence que fit M. GUMBEL le 23 avril 1956 à la Société Hydrotechnique de France se trouve dans l'article publié ci-dessus. Cependant le brillant conférencier qu'est M. GUMBEL ne s'est pas borné à cela et il a développé longuement, dans son exposé oral, d'autres idées particulièrement intéressantes que nous voudrions rapporter et résumer ici, avec quelques remarques soulevées dans la discussion qui a suivi la séance.

\*\*

En premier lieu M. GUMBEL a justifié l'hypothèse d'un champ de variation illimité pour la distribution des débits; nous citons :

« Il s'agit donc de savoir si nous pouvons accepter l'idée qu'il y a une limite pour la distribution des débits. Evidemment, nous l'accepterons, parce qu'enfin, même si le déluge revient, il cessera un jour, ou le monde périra : dans ce dernier cas, il n'y aurait plus de statisticiens, et nous pouvons donc nous passer de cette question! »

Il n'est donc pas question de nier l'existence des limites naturelles, mais celles-ci sont impossibles à déterminer et la prudence commande alors de dire : « S'il y a une certaine inondation, il arrivera un jour où cette inondation sera surpassée; voilà toute l'hypothèse d'une

variable illimitée... Donc si nous acceptons cette idée qu'à tout déluge, il en succédera un qui sera pire, ce qui est une idée assez pessimiste, la distribution peut être considérée comme illimitée... »

Revenons maintenant aux hypothèses nécessaires à l'application du calcul des probabilités aux débits de crue. Il faut tout d'abord supposer des conditions constantes pour que rien ne change dans le débit.

Cependant les climats changent d'une façon systématique, mais :

« Nos prévisions sont liées à l'hypothèse que les climats ne changent pas dans les temps assez courts pour lesquels nous bâtissons. »

Mais il n'y a pas que le climat qui puisse modifier la nature d'un cours d'eau :

« Tout aménagement de la rivière, tout barrage, détruit la rivière dans le sens de la continuité statistique; c'est une nouvelle rivière qui est établie par l'œuvre humaine. »

Du même point de vue de l'homogénéité statistique des observations, il faut aussi prendre quelques précautions :

« S'il y a une saison de pluie, par exemple en automne, et s'il y a une saison où la neige fond, c'est-à-dire au printemps, il y aura deux saisons, et en réalité, il ne



s'agit plus d'une rivière, il s'agit de deux rivières, c'est-à-dire la rivière « automne » et la rivière « printemps ». Et il faut séparer les deux rivières. »

\*\*\*

Signalons une remarque théorique qui porte sur la rapidité de convergence vers la loi asymptotique que M. GUMBEL utilise. Cette question a toujours soulevé de nombreuses discussions. Comme le dit l'orateur :

« Ce qu'il faut savoir du point de vue purement théorique, c'est le nombre d'observations nécessaires pour juger si l'approximation d'une distribution de la plus grande valeur vers sa forme asymptotique dépend de la distribution initiale même. »

Or, le point important est que la rapidité de convergence dépend des propriétés asymptotiques des lois initiales et l'hypothèse où se place M. GUMBEL est que la distribution initiale des débits est assez proche, dans la majorité des cas, de la loi limite pour assurer une convergence rapide.

Signalons encore quelques points d'intérêt pratique sur lesquels M. GUMBEL a insisté. Il a mis en évidence le danger qu'il y a pour des statisticiens non avertis à comparer des polygones de fréquence observés à des distributions théoriques. En réalité la notion de distribution empirique (polygone de fréquence) n'existe pas ou plutôt il y a une infinité de telles distributions, dépendant chacune du choix des intervalles de classe et de l'origine de la première classe. Cet arbitraire dans le choix des classes qui se retrouve dans le calcul du  $\chi^2$  conduit M. GUMBEL à rejeter ce test pour des variables aléatoires continues. Par contre la méthode peut être utilisée si la variable est soumise à la transformation basée sur l'intégrale de probabilité; l'origine de la première classe est alors toujours zéro et l'influence qu'exerce le choix des intervalles peut être calculée.

Mais sur une question de M. le Président, M. MORLAT a souligné, au cours de la discussion qui suivit la conférence, que le test du  $\chi^2$  a rendu et continue à rendre de très grands services dans les cas les plus généraux; de plus aucune technique de remplacement valable n'a pu être mise au point, qui rende les mêmes services. Il faut surtout retenir que l'emploi du test de

Pearson nécessite, dans le cas de petits échantillons, beaucoup de discernement et de prudence.

Un autre problème pratique se pose à propos de la probabilité empirique à affecter à la même valeur d'un échantillon. M. GUMBEL rejette les formules

$$F = \frac{i}{n}, \quad F = \frac{i-1}{n}, \quad F = \frac{i-1/2}{n}$$

les deux premières parce qu'elles affectent soit la probabilité 1 à la plus grande valeur, soit la probabilité 0 à la plus petite, la troisième parce qu'elle affecte la durée de retour  $2n$  à la plus grande valeur.

M. GUMBEL préconise alors la formule  $F = i/(n+1)$  qui n'a pas ces inconvénients et qui est susceptible en outre d'une justification théorique, puisque la quantité  $i/(n+1)$  est l'espérance mathématique de la probabilité affectée à la  $i^{\text{ème}}$  valeur.

Pour terminer, signalons un dernier point intéressant, à propos de l'application des méthodes graphiques aux débits de crue : (intervention de M. POMMIER).

Dans la majorité des cas, les débits de crue ne sont connus qu'avec une marge d'erreur considérable. La méthode graphique permet alors de se rendre compte plus aisément de l'influence des forts débits. Si un point ne tombe pas sur la courbe ajustée, c'est peut-être parce que la courbe est fautive, mais la plupart du temps, ce sont les débits qui sont faux. Ce phénomène risque de passer inaperçu lorsque l'on prend des méthodes d'ajustement analytique.

\*\*\*

Après la conférence de M. GUMBEL, M. Jacques BERNIER, Ingénieur au Service des Etudes et Recherches Hydrauliques d'Electricité de France, a présenté un exposé apportant des compléments variés et, sur certains points, assez nouveaux, aux méthodes préconisées par M. GUMBEL. M. BERNIER a montré en particulier que pour un certain nombre de séries de crues observées sur nos rivières, une autre forme limite de la loi de probabilité des valeurs extrêmes, due à M. FRECHET, donne lieu à des ajustements meilleurs que la loi de M. GUMBEL.

On trouvera ci-après le texte de l'intervention de M. BERNIER, illustré de quelques-unes des applications sur lesquelles ont porté ses travaux.