Comparaison des lois de Gumbel et de Fréchet sur l'estimation des débits maxima de crues

Comportement asymptotique des courbes de débits classés (1)

Comparison between Gumbel's and Fréchet's laws on the estimation of maximum flood discharges

The asymptotic behaviour of classified discharge curves

PAR M. BERNIER

INGÉNIEUR AU SERVICE DES ÉTUDES ET RECHERCHES HYDRAULIQUES D'ÉLECTRICITÉ DE FRANCE

Communication présentée lors de la réunion partielle de la Société Hydrotechnique de France (Section d'Hydrologie statistique. - Commission pour l'étude des débits de crue) organisée à Paris le 12 septembre 1957.

La théorie statistique des valeurs extrêmes a montré que deux types de lois de probabilité sont susceptibles d'être ajustées aux distributions de fréquence des débits de crue : ce sont les lois de Gumbel et de Fréchet. Il est en général très difficile de faire un choix cohérent entre ces deux lois à partir des trop courtes séries d'observations de débits maxima annuels dont on dispose actuellement.
On étudie ici certaines propriétés asymptotiques des courbes de débits classés qui pourraient être utilisées pour le choix a priori entre les lois de Gumbel et de Fréchet. Cette étude comporte :

a) *Des ajustements de lois de Galton composées aux courbes de débits classés observés de huit cours d'eau à régimes variés.*
b) *La comparaison de ces lois avec la distribution de Gumbel pour les grandes valeurs des débits.*

En conclusion, le plus ou moins bon ajustement de la loi de Gumbel aux débits de crue, semble être lié à certaines caractéristiques hydrologiques des bassins étudiés. Cette liaison demandera cependant à être précisée par des applications à des exemples plus nombreux et par un perfectionnement de la méthode statistique utilisée.

The statistical theory of extreme values has shown that the Gumbel and Fréchet probability laws can be made to fit flood discharge frequency distributions. It is usually very difficult to base a coherent decision as to which of these two laws to use on presently available maximum annual discharge data which does not cover sufficiently long periods.
Certain asymptotic properties of classified discharge curves, which can be used to make an a priori choice between the Gumbel and Fréchet laws, are studied here. This study includes:

a) *Examples of fitting compound Galton laws to classified discharge curves for eight streams having varying regimens.*
b) *A comparison of these laws with the Gumbel distribution for high discharge values.*

In conclusion how well the Gumbel law can be made to fit flood discharges, seems to be dependent upon certain hydrological characteristics of the river basins being investigated. However, this connection will have to be determined by applying the procedure to a greater number of examples and perfecting the statistical method which is used.

Dans la présente note, nous nous proposons d'étudier un aspect particulier de la courbe des débits classés : son comportement pour les grandes valeurs du débit.

---

(1) M. le Professeur GUMBEL a traité le 23 avril 1956 devant la Société Hydrotechnique de France, la théorie des valeurs extrêmes appliquée aux débits maxima de crue. Cet exposé avait été suivi d'une discussion au cours de laquelle M. BERNIER avait, dans une importante intervention, comparé la loi de Gumbel à diverses autres lois (Galton-Gibrat, Fréchet) dans leurs applications à l'évaluation des débits maxima de crues. Ces exposés ont paru dans *la Houille Blanche* de 1956, n° 5.
Au cours d'une seconde réunion de la S.H.F., le 12 septembre 1957, M. Bernier présenta un nouvel exposé sur la comparaison des lois de Gumbel et de Fréchet, tandis que M. GUMBEL consacrait aux débits d'étiage une communication qui constitue une application inverse de la théorie des valeurs extrêmes. (Cf. p. 57 du présent numéro.)

Rappelons que les applications de la théorie des valeurs extrêmes nous ont montré (2) que deux types de lois de probabilité (lois de Gumbel et Fréchet) peuvent s'ajuster aux courbes de fréquences empiriques des débits de crue. Le problème du choix *a priori* entre ces deux lois se pose donc.

En effet, les formes de ces deux lois sont très voisines pour les valeurs centrales et ne diffèrent essentiellement que par leurs comportements asymptotiques. Il est par ce fait impossible de détecter des différences sur les trop courtes séries d'observations de crue existantes, qui ne renseignent précisément que sur les parties centrales des distributions.

Théoriquement, le choix entre les lois de Fréchet et Gumbel, dont les fonctions de répartition s'expriment simplement en fonction de variables réduites :

$$G(y) = \exp\left[-e^{-y}\right]$$

$$H(z) = \exp\left[-z^{-k}\right] \qquad (k \text{ positif})$$

dépend de la forme analytique prise par la loi initiale des débits journaliers pour les grandes valeurs. On ne peut confondre cette loi, telle qu'elle est définie dans la théorie de Gumbel, avec l'expression mathématique de la courbe des débits classés, étant donné l'hétérogénéité de celle-ci résultant du mélange des débits d'années différentes et des liaisons entre débits de jours successifs. On conçoit cependant, de manière intuitive, que le comportement de cette courbe pour les grandes valeurs du débit est identique à celui de la loi initiale des débits journaliers ou, mieux, à celui de la loi des crues annuelles.

Nous ferons donc cette hypothèse, qui peut d'ailleurs être vérifiée par certaines justifications théoriques.

Le comportement asymptotique de la courbe des débits classés guidera donc le choix entre les deux lois des valeurs extrêmes.

L'étude de ce comportement est effectuée ici en plusieurs étapes :

a) Ajustement d'une loi de probabilité assez simple pour représenter convenablement les grandes valeurs, à la courbe observée des débits classés de certains cours d'eau. Critique de l'ajustement.

b) Etude des distributions des valeurs extrêmes d'échantillons tirés au hasard de la loi ainsi ajustée. Ajustement des lois de Gumbel et Fréchet à ces distributions dans la zone du champ de variation des débits maxima annuels.

I. — REPRÉSENTATION MATHÉMATIQUE DE LA COURBE DES DÉBITS CLASSÉS

1. — Choix d'un type de loi.

Il y a intérêt à remplacer la courbe des débits classés observée, qui n'est pas d'un maniement facile, par une forme mathématique convenable. Cette loi de probabilité théorique doit être assez souple, de façon à représenter fidèlement les grands débits. L'utilisation d'une simple loi de Galton, comme l'a préconisé M. Gibrat, ne s'est pas avérée satisfaisante dans le cas présent. Il est préférable d'employer une loi mixte (composition de deux lois de Galton), tenant compte ainsi de l'hétérogénéité des éléments constitutifs de la courbe des débits classés. Cette loi mixte a d'ailleurs le mérite de posséder un plus grand nombre de paramètres, d'où sa plus grande souplesse.

(2) *La Houille Blanche*, n° 5, 1956, page 718.

2. — Ajustement de la somme de deux lois de Galton.

Les applications de cette loi, déjà étudiée par M. Bloch, n'ont pu être généralisées par suite des complications de la méthode de calcul des paramètres mise au point par cet auteur. Mais les calculs peuvent être de beaucoup simplifiés si on utilise comme donnée initiale le débit logarithmique (logarithme du débit suivant la dénomination d'Halphen) à la place du débit réel.

Soit donc x le débit logarithmique.

On considère la fonction de répartition :

$$F(x) = 1/2\, F_1(x) + 1/2\, F_2(x) \qquad (1)$$

(on suppose, pour simplifier, même répartition de deux lois composantes, d'où le facteur 1/2).

Dans cette formule :

F_1 est une loi normale (moyenne μ_1, écart-type σ_1).

F_2 est une loi normale (moyenne μ_2, écart-type σ_2).

Il s'agit donc d'estimer μ_1, μ_2, σ_1, σ_2.

De la façon classique, on les détermine en fonction de la moyenne m, de la variance σ^2 et des coefficients de Pearson β_1 et β_2 de la distribution empirique des débits observés :

Il est commode de poser :

$$\lambda = \operatorname{tg} \theta = \sigma_1/\sigma_2 \qquad (\sigma_1 \geqslant \sigma_2)$$

$$\delta = \sqrt{2} \operatorname{tg} \varphi = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Avec ces conventions :

$$\left.\begin{aligned} m &= (\mu_1+\mu_2)/2 \\ \sigma &= \sigma_2^2/4\,(1+\lambda^2)\,(2+\delta^2) \\ \sqrt{\beta_1} &= -3 \sin\varphi \cos^2\varphi \cos 2\theta \\ 1/2\,(\beta_2-3) &= 3 \cos^2 2\theta \cos^4\varphi - 2 \sin^4\varphi \end{aligned}\right\} \qquad (2)$$

Après élimination de θ dans les deux dernières équations et après avoir posé $u = \sin^2\varphi$, on a l'équation résolvante du système (2) :

$$6\,u^3 + 3\,u\,(\beta_2 - 3) - \beta_1 = 0 \qquad (3)$$

Il est aisé de voir que, quelles que soient les distributions, cette équation possède toujours une racine et une seule comprise entre 0 et 1. Les paramètres sont donc entièrement déterminés par les équations (2) et (3).

REMARQUE. — Dans la troisième équation (2), apparaît la quantité $\sqrt{\beta_1}$. On désigne en général avec cette expression le rapport μ_3/σ^3 (μ_3 : moment centré du troisième ordre), dont le carré est égal à β_1. Avec cette convention, on voit que $\sqrt{\beta_1}$ peut être négatif.

3. — Critique de l'ajustement.

Les débits journaliers ne constituant pas un échantillon de variables indépendantes issues d'une loi homogène, il n'est pas possible d'utiliser le test du χ^2 pour juger de l'ajustement effectué sur l'ensemble des observations. Cette difficulté peut cependant être tournée en effectuant un tirage au hasard, avec remise, d'un échantillon de taille donnée dans l'ensemble de tous les débits observés de la période étudiée.

Puisque la valeur de l'ajustement nous intéresse surtout pour les grands débits, il semble tout indiqué d'effectuer un tirage au hasard des m plus grandes valeurs de m échantillons de n débits observés. La comparaison de cet échantillon avec la loi théorique de la plus grande de n variables issues de la loi F dont la fonction de répartition est $G(x) = [F(x)]^n$, fournit alors un test d'ajustement pour les grands débits.

Le test du χ^2 est bien applicable ici. Il reste cependant une difficulté dans le choix du nombre de degrés de liberté intervenant dans la loi du χ^2. On sait, d'après la règle classique, que ce nombre ν est égal à :

$$\nu = k - p - 1$$

où :

k est le nombre de classes construites avec l'échantillon;

p et le nombre des paramètres de la loi théorique estimés sur l'échantillon.

Or, les paramètres de la loi théorique sont ici estimés sur tous les débits et le test du χ^2 ne porte que sur une sous-population tirée au hasard de la population totale. En toute rigueur : ν est donc compris entre $k-5$ et $k-1$. Quoi qu'il en soit, on peut prendre k assez grand compatible avec un échantillon assez petit par rapport au nombre total de débits de façon à ne commettre qu'une erreur négligeable en utilisant un χ^2 avec $\nu = k - 1$ degrés de liberté.

4. — Applications.

a) PRÉSENTATION DES EXEMPLES CHOISIS.

Ces applications ont porté sur 8 stations de jaugeages de caractères assez variés pour fournir un bon échantillonnage de divers régimes existant dans le pays.

Le tableau ci-après donne les valeurs des divers paramètres estimés à partir des observations. Ces valeurs ont été calculées en effectuant un découpage en classes du champ de variation des débits logarithmiques, ce qui a nécessité certaines corrections (corrections de Sheppard).

De ce tableau, on peut faire ressortir les valeurs de β_1 et β_2 du Rhin pour lequel les deux lois de Galton composantes sont pratiquement confondues.

Il faut aussi souligner les chiffres de la dernière colonne (coefficient de variation). Ce coefficient met en évidence la forte dispersion des débits des stations à régime purement pluvial (Blavet, Creuse) et de celles soumises aux influences méditerranéennes, exception faite pour le Verdon, dont la faible valeur de la dispersion est pour le moins étonnante.

b) CALCUL DU χ^2.

Il nous reste à tester ces divers ajustements par la méthode décrite au paragraphe précédent.

Nous avons pris ici $m=100$. $n=10$.

STATIONS	B. V. km²	Période	$m^{(1)}$	$\sigma^{2(1)}$	$\sqrt{\beta_1}^{(1)}$	$\beta_2^{(1)}$	μ_1	μ_2	σ_1	σ_2	Coefficient de variation CV(2)
RHIN (Rheinfelden).	34.550	1939-54	2,9578	0,0342	0	3	2,9578	2,9578	0,1849	0,1849	0,45
BLAVET (Guerledan).	620	1939-54	0,7499	0,2126	— 0,022	2,273	1,1078	0,3920	0,2906	0,2906	1,44
CREUSE (Eguzon)...	2.400	1939-54	1,0317	0.4262	— 0,528	3,208	0,6948	1,3685	0,6768	0,4090	2,92
TARN (Pinet)......	2.677	1939-54	1,4903	0,1484	— 0,484	2,917	1,7272	1,2535	0,3621	0,2309	1,08
TRUYÈRE (Sarrans).	2.462	1939-54	1,2197	0,3422	— 0,415	2,820	0,8489	1,5906	0,5285	0,3606	2,26
GAVE DU BROUSSET (Allias).	61	1939-54	0,3049	0.1586	0	2,108	0,6303	0,0205	0,2295	0,2295	1,15
DORON DE BOZEL.. (La Perrière).	301	1939-54	0,9424	0,0592	0,201	1,760	1,1591	0,7257	0,1294	0,0884	0,61
VERDON (Quinson).	1.661	1939-48	1,3385	0,1081	0,361	2,337	1,5946	1,0825	0,2455	0,1609	0,88

(1) Ces paramètres ont été calculés sur les débits logarithmiques (logarithmes décimaux).
(2) Ces coefficients de variation sont ceux des débits et non des débits logarithmiques.

Les figures I_1 à I_8 montrent les histogrammes empiriques des 100 plus grandes valeurs de 100 échantillons de 10 débits tirés au hasard parmi les débits observés de la période étudiée. Les courbes continues figurent les lois théoriques déduites de l'équation : $G(x) = [F(x)]^{10}$.

Le test du χ^2 a donné les résultats suivants :

COURS D'EAU	Degrés de liberté	χ_0^2	Prob. $[\chi^2 \geqslant \chi_0^2]$
RHIN	10	5,22	0,88
BLAVET	9	4,92	0,84
CREUSE	10	4,63	0,91
TARN	9	12,95	0,17
TRUYÈRE	9	9,40	0,40
GAVE DU BROUSSET	9	15,01	0,09
DORON DE BOZEL..	8	5,04	0,76
VERDON	9	13,76	0,13

Pour la majorité des exemples traités, la probabilité de dépasser par hasard la valeur du χ^2 observée est assez forte. C'est donc l'indice d'un très bon ajustement. Seuls les χ^2 du Gave du Brousset et du Verdon soulèvent quelques difficultés du fait qu'ils sont proches du seuil de signification à 5 %. Après examen des graphiques I_5 et I_7, on peut cependant conclure que, si la distribution observée s'écarte significativement de la loi théorique, c'est dans le sens d'une plus petite dispersion des valeurs maxima.

En conclusion, la somme de deux lois de Galton s'adapte parfaitement bien aux courbes observées des débits classés. L'ajustement est particulièrement bon pour les grandes valeurs. Il est donc parfaitement légitime de substituer l'étude asymptotique de cette loi à l'étude de la courbe empirique, qui reste d'un maniement difficile pour les forts débits.

Dans la suite, cette loi ajustée sera désignée par la notation : loi F.

Les distributions des valeurs extrêmes des échantillons de taille n issus de cette loi seront désignées par : F_n.

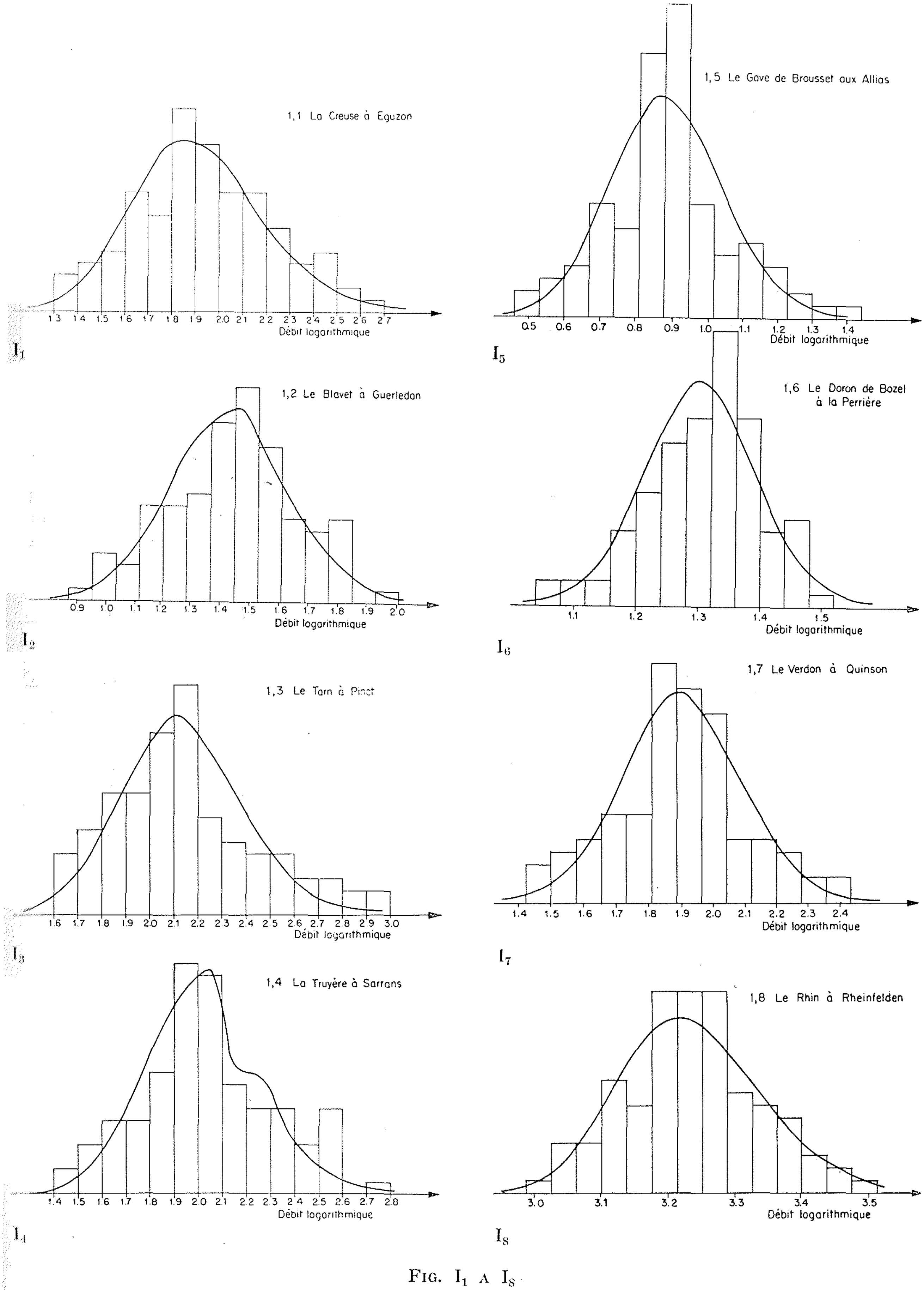
1,1 La Creuse à Eguzon
Débit logarithmique
I1
1,2 Le Blavet à Guerledan
Débit logarithmique
I2
1,3 Le Tarn à Pinet
Débit logarithmique
I3
1,4 La Truyère à Sarrans
Débit logarithmique
I4
1,5 Le Gave de Brousset aux Allias
Débit logarithmique
I5
1,6 Le Doron de Bozel à la Perrière
Débit logarithmique
I6
1,7 Le Verdon à Quinson
Débit logarithmique
I7
1,8 Le Rhin à Rheinfelden
Débit logarithmique
I8

FIG. I_1 A I_8

II. — COMPORTEMENTS ASYMPTOTIQUES COMPARÉS DES LOIS AJUSTÉES ET DE LA LOI DE GUMBEL

La dernière étape de l'étude est la comparaison des comportements asymptotiques de la loi de Gumbel et de la loi des débits classés. De façon plus précise, il s'agit de savoir si, dans les limites les plus probables du champ de variation des débits de crue, la courbe des débits classés et la loi de Gumbel ont des comportements semblables.

Il est commode pour cela de prendre la loi de la plus grande de n observations issues de la *loi ajustée* sur les débits classés et de la comparer à la loi de Gumbel.

Entendons-nous bien : dans le paragraphe précédent, nous avons pris des échantillons de grandes valeurs *tirés de la distribution observée* pour tester l'ajustement. Nous faisons maintenant des *tirages à partir de la loi ajustée* pour entrer dans le cadre de la théorie de Gumbel.

Nous prenons ici $n=100$. Ce chiffre n'a pas de valeur absolue. Il ne signifie pas qu'on fait l'hypothèse que la crue annuelle est la plus grande de 100 observations issues de la loi ajustée. On peut seulement dire que le champ de variation de cette loi F_{100} recouvre à peu près, dans les cas les plus généraux, le champ de variation des débits de crue. Cela étant, les conclusions relatives à F_{100} pourront alors être étendues à la loi des débits maxima annuels.

1. — Loi de Gumbel ajustée et loi de Gumbel asymptotique.

La loi de Gumbel s'exprime le plus généralement en fonction d'une variable réduite :

$$\text{Prob}\,[Y \leqslant y] = \exp\,[-e^{-y}]$$

variable réduite liée au débit D par la relation :

$$y = \alpha\,(D - u)$$

Les paramètres α et u sont asymptotiquement fonction de la loi initiale F :

$$F\,(u) = I - 1/n \qquad \alpha = n f\,(u) \qquad (4)$$

Dans la deuxième formule :

$$f\,(x) = dF/dx$$

Ces relations fournissent une loi de Gumbel asymptotique qui est très éloignée dans le cas présent de la loi exacte F_{100}, comme le montrent les valeurs de α calculées par la formule précédente, comparées à celles obtenues en ajustant une loi de Gumbel aux débits de probabilité 0,5 et 0,9 de la loi exacte F_{100}.

Soit :

$$0{,}36651 = \alpha_{100}\,(D_5 - u)$$
$$2{,}25037 = \alpha_{100}\,(D_9 - u)$$

COURS D'EAU	α (formule 4)	α (formule 5)
RHIN	0,0059	0,0024
BLAVET	0,1644	0,0607
TRUYÈRE	0,0308	0,0101

Il y a donc intérêt à utiliser la formule (5) de préférence à (4) pour déterminer la loi de Gumbel la plus apte à représenter correctement F_{100}. Le tableau ci-dessous donne les valeurs de α et u calculés par cette méthode :

COURS D'EAU	α_{100}	u_{100}
RHIN	0,0024	2.438 m³/s
BLAVET	0,0607	49,91 m³/s
CREUSE	0,0071	183,3 m³/s
TARN	0,0077	287,0 m³/s
TRUYÈRE	0,0101	215,7 m³/s
GAVE DU BROUSSET	0,3305	12,56 m³/s
DORON DE BOZEL	0,3119	26,62 m³/s
VERDON	0,0311	123,3 m³/s

2. — Comparaison graphique.

Pour comparer la loi F_{100} et la loi de Gumbel ajustée, nous utilisons une méthode graphique, qui a l'avantage sur les méthodes analytiques de présenter de manière plus simple et plus concrète les écarts entre les deux lois. Nous devons aussi nous attacher à présenter des courbes indépendantes des champs de variations différents des divers cours d'eau pour pouvoir effectuer des comparaisons valables. Le graphique II répond à cette exigence.

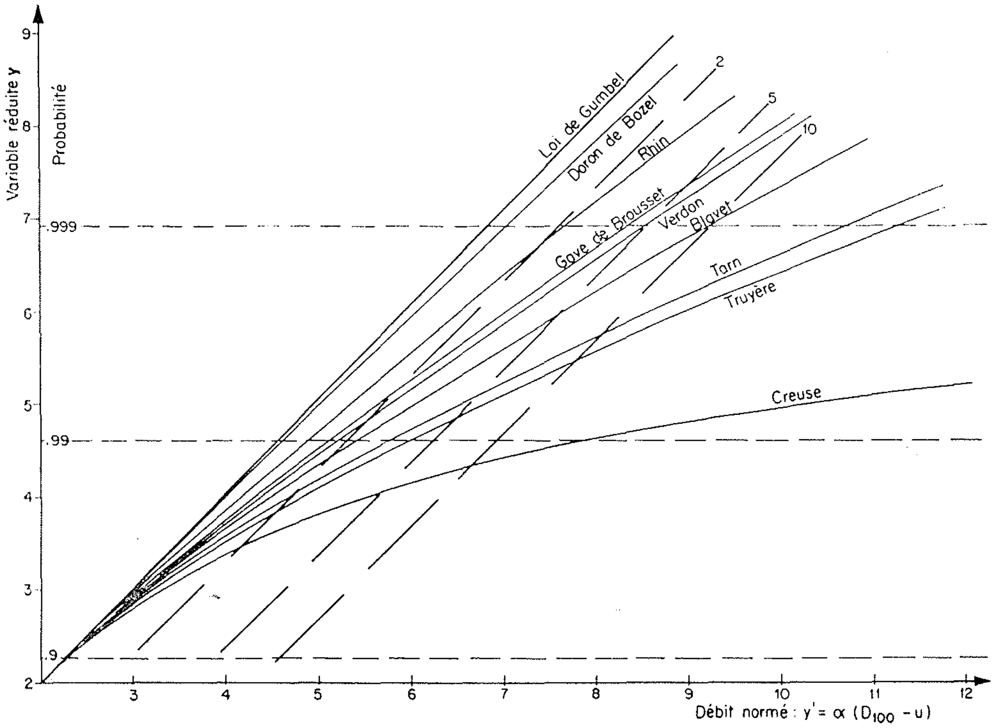

FIG. II

La figure présente :

a) En ordonnées, la variable réduite de Gumbel y (l'échelle est aussi graduée en probabilités);

b) En abscisses, la variable $y' = \alpha_{100} (D_{100} - u_{100})$ où u_{100} et α_{100} sont déterminés de la façon décrite au paragraphe ci-dessus et D_{100} est le débit de la loi exacte F_{100} correspondant à une probabilité donnée.

Avec ces échelles, la loi de Gumbel exacte est représentée par la première bissectrice. Les segments de droite, parallèles à cette bissectrice, marqués 2, 5, 10, sont tels que les lois F_{100} qui les traversent présentent alors des périodes de récurrence respectivement 2, 5, 10 fois plus faibles au moins que la loi de Gumbel ajustée. Ceci a pour intérêt d'indiquer de façon concrète les écarts entre les lois :

En se rapportant aux débits maxima annuels, on peut, pour fixer les idées, admettre des écarts allant jusqu'à 5 pour les valeurs millénaires (probabilité : 0,999), ceci, compte tenu des erreurs de mesure sur les grands débits. Si les écarts dépassent 5 et atteignent 10 et plus, ils deviennent inacceptables. (La crue centenaire deviendrait millénaire avec la loi de Gumbel.)

3. — Analyse du graphique II.

En se basant sur les critères établis ci-dessus, on voit que la loi de Gumbel donne un ajustement :

— excellent pour le Doron de Bozel,

— assez bon pour le Rhin;

— acceptable pour le Gave de Brousset et le Verdon;

— inacceptable pour le Blavet, le Tarn, la Truyère et la Creuse.

Pour ces quatre derniers cours d'eau, la loi de Fréchet, dont la décroissance asymptotique est plus lente, devra donner des ajustements meilleurs. Ces conclusions restent valables si le Verdon et le Gave du Brousset ont des lois moins dispersées.

Il est intéressant de confronter ces résultats avec les valeurs des coefficients de variation calculés plus haut. Il semble que la loi de Gumbel est d'autant moins valable que ces coefficients sont plus forts. On sait que cette forte dispersion est liée le plus souvent à l'influence pluviale que subit le cours d'eau. Là encore, le Verdon fait exception, qui présente un ajustement acceptable par la loi de Gumbel, bien que soumis fortement aux pluies méditerranéennes.

CONCLUSION

Malgré les imperfections de la méthode, qui demande à être approfondie, les résultats de cette étude sont assez encourageants : ainsi les écarts des lois F_{100} à la loi de Gumbel semblent obéir à des règles précises en rapport avec le régime du cours d'eau. Cependant, le nombre des exemples traités est trop petit pour pouvoir lier de façon sûre le choix d'un type de loi (Fréchet ou Gumbel) aux caractères hydrologiques du bassin étudié. Pourtant la conclusion la plus claire qui se dégage de cette étude est que les influences pluviales doivent jouer un rôle important surtout parce qu'elles sont la cause de la plus ou moins grande variabilité des débits.

RÉFÉRENCES

[1] BLOCK (M.). — La distribution des débits d'un cours d'eau considérée comme somme de deux distributions de Galton. *La Météorologie,* juillet 1939.

[2] GIBRAT (R.). — Aménagement hydroélectrique des cours d'eau, statistique mathématique et calcul des probabilités. *Revue Générale de l'Electricité,* octobre 1932.

[3] GUMBEL (E.-J.). — Méthodes graphiques pour l'analyse des débits de crue. *La Houille Blanche,* n° 5, novembre 1956.

A la suite de cette communication à la Société Hydrotechnique de France, une discussion s'est engagée, dont nous donnons ci-après le compte rendu.

DISCUSSION

Président : M. HUPNER

Le Président remercie M. BERNIER et pose quelques questions :

1° Pourquoi ne prend-on, pour comparer les deux lois, que la valeur médiane et celle de probabilité 0,9, et non tous les éléments de l'échantillonnage?

M. MORLAT répond que ce choix, est en effet largement conventionnel. Il se justifie cependant par ceci : lorsqu'on étudie une série de crues observées, on dispose au maximum de quelques dizaines d'observations, et les valeurs déciles et médianes ont un poids important dans l'ajustement. Il y a donc intérêt à voir, ces valeurs étant fixées, comment les diverses lois divergent.

2° Peut-on conclure de cette comparaison qu'il n'y a pas de formule universelle et que les projeteurs d'ouvrages doivent appliquer l'une ou l'autre loi suivant le régime de la rivière et le mécanisme de ses crues, et après avoir fait un certain nombre d'essais tendant à rechercher si une loi est plus convenable qu'une autre?

M. BERNIER pense qu'il en est bien ainsi et que l'on doit rechercher pour choisir entre les lois un critère qui ne soit pas soumis aux erreurs d'échantillonnage auxquelles sont sujettes les crues observées.

M. MORLAT ajoute que, dans le cadre de la comparaison des lois de Gumbel et de Fréchet faite par M. BERNIER, on peut conclure que la plupart des courbes de débits classés peuvent être ajustées avec ces deux seules lois, qui sont d'ailleurs de la même famille, la loi de Fréchet pouvant être considérée comme une généralisation de la loi de Gumbel. Cette conclusion limite donc, en principe, le choix de l'utilisateur à ces deux seules lois.

3° Pourquoi les lois du type Galton sont-elles à rejeter, a priori pour l'étude asymptotique des débits classés.

M. BERNIER répond que ces lois paraissent insuffisantes du fait qu'elles ne mettent en jeu que deux paramètres; on pourrait songer à utiliser ces lois en introduisant une valeur inférieure des débits, mais l'estimation physique de cette valeur ne peut être faite que graphiquement, c'est-à-dire d'après une répartition empirique des débits observés, toujours soumise à un certain arbitraire qui devient, précisément, très important pour les grandes valeurs.

M. CHAPOUTHIER est frappé de la grande variété des comportements asymptotiques des courbes suivant les lois théoriques dont on est parti. La méthode probabiliste apporte-t-elle plus que les méthodes grossières appliquées jusqu'ici par les ingénieurs? Peut-on espérer que le calcul statistique serrera un jour d'assez près la réalité pour qu'on puisse éviter les marges de sécurité souvent considérables que l'on est obligé de garder, faute de mieux?

M. MORLAT répond que les efforts de M. BERNIER tendent précisément à fixer, d'après les éléments connus, comme la courbe des débits classés, la forme de la loi qu'il faut ajuster; dans ces conditions, si l'on peut analyser davantage les caractéristiques physiques de la rivière et la classer suivant son régime, on voit, par exemple, que les régimes glaciaires correspondent bien à la loi de Gumbel, tandis que les régimes à prédominance pluviale s'accommodent mieux de la loi de Fréchet. Ce résultat représente donc un certain progrès dans la réduction de la marge d'incertitude que le calcul des probabilités laisse encore aux ingénieurs et qui leur impose sans doute une certaine prudence.

M. Chapouthier conclut que la physique doit s'associer aux calculs de probabilités pour les compléter.

M. le Président remercie M. Morlat et le félicite d'avoir bien tiré la philosophie de la discussion sur la valeur du calcul des probabilités.

Pour M. Gumbel, l'incertitude des résultats du calcul des probabilités en matière d'évaluation des débits maxima de crues n'est pas imputable à une faille quelconque dans la structure mathématique des lois, mais à une carence des données physiques : distributions initiales pratiquement inconnues, interdépendance d'un grand nombre d'observations supposées indépendantes au cours d'une année et prises pour base, extrapolation de données portant sur un nombre assez faible d'années (20 à 30 ans) à des prévisions millénaires. Les mathématiques ne sauraient suppléer ni à un manque de pensée, ni à un manque d'observations. Cependant, il est légitime de supposer connues certaines propriétés analytiques de ces distributions initiales, elles-même inconnues, par exemple, la diminution de la densité de probabilité des valeurs extrêmes suivant une loi exponentielle qui suffira à établir la distribution asymptotique de la plus grande valeur et définit la classe d'observations nécessaires. Le nombre de ces observations dépend du caractère de la distribution initiale; d'autre part, l'aménagement hydro-électrique progressif des rivières détruit ou réduit les possibilités de comparaison.

L'idéal de l'ingénieur réside dans une seule observation qui donne la réalité; mais statistiquement, une observation unique ne donne rien; il faut donc un grand nombre d'observations, des études supplémentaires sur le caractère de la distribution initiale, un départ du nombre des observations indépendantes, toutes choses qui sont du domaine de la météorologie, non de la statistique.

Revenant à l'exposé de M. Bernier, M. Gumbel classe les distributions des plus grandes valeurs en trois types : exponentielle (étudiée par M. Gumbel), distribution de Cauchy (étudiée par M. Fréchet), distribution à limite supérieure.

Il ne lui semble pas nécessaire d'introduire une limite supérieure pour les crues; par ailleurs, l'introduction d'une limite inférieure est sans intérêt pratique (« Mathusalem ne meurt pas d'une maladie infantile »). Il reste donc à comparer les deux premiers types de distribution. Comme l'a dit M. Bernier, s'il y a un petit nombre d'observations, et si le quotient de la plus grande valeur à la plus petite n'est que de l'ordre de 2 ou 3, la distinction entre ces deux types est très difficile, parce que l'allure des fonctions qui les représentent est pratiquement confondue dans un intervalle aussi réduit Dans ces conditions, il semble plus prudent d'utiliser la loi de Fréchet qui aboutit à des extrapolations de débits plus grandes, quoique, du point de vue logique, elle présente la difficulté que le débit moyen et l'écart-type n'y existent pas. Pourtant on peut l'utiliser, pourvu qu'elle corresponde à un meilleur ajustement, ce qui, *par hypothèse* seulement, doit donner une meilleure extrapolation.

En deuxième lieu, M. Gumbel remarque que le travail de M. Bloch, introduit dans l'exposé de M. Bernier, a été fait à l'Université de Lyon, du temps où ce dernier y était son élève. M. Gumbel estime que l'on pourrait substituer aux facteurs 1/2 et 1/2 de M. Bloch, les poids B et 1-B en vue d'un meilleur ajustement.

Enfin, la troisième partie du travail de M. Bernier consiste à traiter la distribution initiale d'une façon non asymptotique. M. Gumbel demande quelle est la distribution *exacte* de la plus grande valeur à partir de la distribution initiale connue.

Il demande d'autre part à M. Bernier :

a) Pourquoi il emploie tantôt $n = 10$ et tantôt $n = 100$;

b) Pourquoi, contrairement à toutes les autres, la courbe de la Truyère présente un point d'inflexion;

c) S'il a essayé de faire la comparaison avec une échelle linéaire, qui, selon M. Gumbel, donne une meilleure représentation qu'en employant l'échelle logarithmique.

M. le Président remercie M. Gumbel de son exposé montrant les difficultés du problème; toutefois, parmi celles-ci, la détérioration des données initiales des rivières par suite des aménagements ne lui paraît pas insurmontable, car les ingénieurs doivent pouvoir rétablir ces données et les reconstituer comme si la rivière était dans son état naturel.

En ce qui concerne les poids B et 1-B proposés par M. Gumbel, M. Bernier estime que le fait d'utiliser des facteurs multiplicatifs différents pour les deux lois de Galton composantes introduit un cinquième paramètre qu'il faut estimer à partir des observations. On court le risque d'avoir une loi ajustée trop souple. Mais il serait intéressant, à défaut de détermination statistique, de chercher une détermination physique du rapport des deux lois composantes.

M. Bernier répond, d'autre part aux questions *a*, *b*, et *c* de M. Gumbel :

a) Les chiffres $n = 10$ et $n = 100$ de son exposé n'ont pas la même signification.

Dans le premier cas, il s'agit de tester l'ajustement de la plus grande valeur de n débits observés à la loi théorique : il a donc pris un n assez petit (en l'occurrence 10) pour avoir un échantillon de grandes valeurs observées assez important.

Dans le second cas, il s'agit d'avoir une représentation de la loi ajustée aux débits classés dans la zone du champ de variation des débits de crue. C'est ce qui impose un n différent du premier chiffre utilisé et l'emploi de $n = 100$.

b) Le point d'inflexion de la distribution relative à la Truyère tient à ce que la loi initiale de distribution est hétérogène (elle correspond à deux lois de Galton superposées) la construction de ces courbes repose sur l'arbitraire du découpage en classes, de telle sorte que si les classes étaient plus petites pour les autres cours d'eau, on pourrait aussi distinguer des points d'inflexion dans les courbes qui représentent leur répartition.

c) Il n'a pas fait la comparaison avec l'échelle linéaire, mais il pense qu'il serait intéressant de voir l'allure des courbes à cette échelle.

M. Gumbel fait alors, sur le sujet étudié, l'intervention suivante :

M. Bernier a employé la deuxième distribution asymptotique des plus grandes valeurs pour l'analyse des débits de crue. Mais il n'a pas utilisé la forme la plus générale de cette distribution.

Puisqu'une transformation linéaire ne change pas la forme d'une distribution statistique, on peut écrire la probabilité de Fréchet comme :

$$F(x) = e^{-(u - \varepsilon / x - \varepsilon)^{\alpha}}$$

Cette distribution est de nouveau illimitée vers la droite. La signification de u reste la même, mais, cette fois, il existe un troisième paramètre, à savoir la limite inférieure ε des crues. Dans certains cas, l'introduction de ε permettra un meilleur ajustement.

Malheureusement, l'estimation de ε semble être difficile; ε doit être positif; d'un autre côté, ε doit être supérieur au plus petit débit de crue observé.

Pour le moment, je ne vois pas d'autres méthodes d'estimation de ε que des approximations successives. On pose d'abord $\varepsilon = 0$ et l'on obtient ainsi des estimations préliminaires $\hat{\alpha}_1$ et $\hat{u}_1$. Puis on doit choisir une nouvelle équation qui relie des valeurs observées ou une observation x proprement choisie à $\hat{\alpha}_1$, $\hat{u}_1$ et $\hat{\varepsilon}$. Evidemment, le plus petit débit de crue sera voisin de la limite inférieure.

On obtiendra donc une estimation de ε en calculant la médiane $\breve{x}_1$ (N) du plus petit débit de crue pour N années. Cette valeur est la solution : $x = \breve{x}_1(N)$ de :

$$1 - e^{-(u-\varepsilon/x-\varepsilon)^{\alpha}} = e^{-\log 2/N} \approx 1 - (\log 2)\, N$$

dont il résulte :

$$(u-\varepsilon)/(x-\varepsilon)) = [\log(1{,}442\ 69\ N)]^{1/\alpha}$$

On remplacera $x = \breve{x}_1$ (N) par le plus petit débit de crue observé x_1, u par $\hat{u}_1$, α par $\hat{\alpha}_1$; alors l'estimation de ε est :

$$\frac{x_1\,[\log(1{,}442\ 69\ N)]^{1/\hat{\alpha}_1} - u_1}{[\log(1{,}444\ 629\ N)]^{1/\hat{\alpha}_1} - 1}$$

A la limite, à savoir pour $N \to \infty$, on obtient $\hat{\varepsilon}_1 = x_1$, quoique cette convergence soit assez lente. L'estimation de $\hat{\varepsilon}$ est asymptotiquement « unbiased ». Elle peut être employée pourvu que $\hat{\varepsilon} > 0$, à savoir :

$$(u/x_1)^{\alpha} < \log(1{,}442\ 69\ N)$$

Cette condition peut être vérifiée immédiatement d'après les observations.

Après l'estimation de $\hat{\varepsilon}_1$, on obtiendra une seconde estimation $\hat{\alpha}_2$ et $\hat{u}_2$ en traitant les « observations » $\log(x - \hat{\varepsilon}_1)$ d'après la théorie linéaire employée par M. Bernier (1).

L'introduction du troisième paramètre réduira la valeur de $g(x_1)$ à $\log(x_1 - \varepsilon)$. L'influence diminue pour les grandes crues. Il en résulte que l'introduction de ε n'améliorera pas l'accord entre la théorie et les observations si le plus petit débit de crue observé est inférieur à la valeur théorique obtenue par l'ajustement linéaire. C'est le cas dans les exemples traités par M. Bernier. Par contre, l'introduction de la limite inférieure des crues pourra améliorer l'accord entre la théorie et les observations, si le plus petit débit de crue est sensiblement supérieur à la valeur prévue par la théorie linéaire.

M. le Président remercie M. BERNIER ainsi que M. GUMBEL, et donne la parole à M. GUMBEL pour son deuxième exposé sur les débits d'étiage.

---

(1) (E.-J.) GUMBEL. — Détermination commune des constantes dans les distributions des plus grandes valeurs, *C.R.Ac.Sc.*, vol. 222, n° 1, pp. 34-36, Paris, 1946.