

REVUE DE STATISTIQUE APPLIQUÉE

S. FERRY

**L'utilisation des méthodes statistiques à l'Électricité
de France pour la prévision des apports dans les
réservoirs. Examen des résultats obtenus**

Revue de statistique appliquée, tome 4, n° 2 (1956), p. 63-74

http://www.numdam.org/item?id=RSA_1956__4_2_63_0

© Société française de statistique, 1956, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

L'UTILISATION DES MÉTHODES STATISTIQUES A L'ÉLECTRICITÉ DE FRANCE POUR LA PRÉVISION DES APPORTS DANS LES RÉSERVOIRS

EXAMEN DES RÉSULTATS OBTENUS

par

S. FERRY

*Chef de la Division Technique Générale
à la Production Hydraulique d'Électricité de France*

Dans le cadre des séances spéciales de la présente Session du Comité Technique de la SOCIÉTÉ HYDROTECHNIQUE DE FRANCE, j'ai été convié à exposer ici l'emploi que nous faisons, à l'ELECTRICITE DE FRANCE, des méthodes statistiques pour la prévision des apports dans nos réservoirs.

A l'origine de nos travaux, j'ai déjà eu l'occasion de traiter le sujet devant le Comité Technique. Mais, depuis lors, nous avons émis de nombreuses prévisions relatives à des bassins des Alpes et des Pyrénées. Que valent-elles ? C'est cet aspect éminemment pratique de la question que je me propose d'examiner plus particulièrement aujourd'hui.

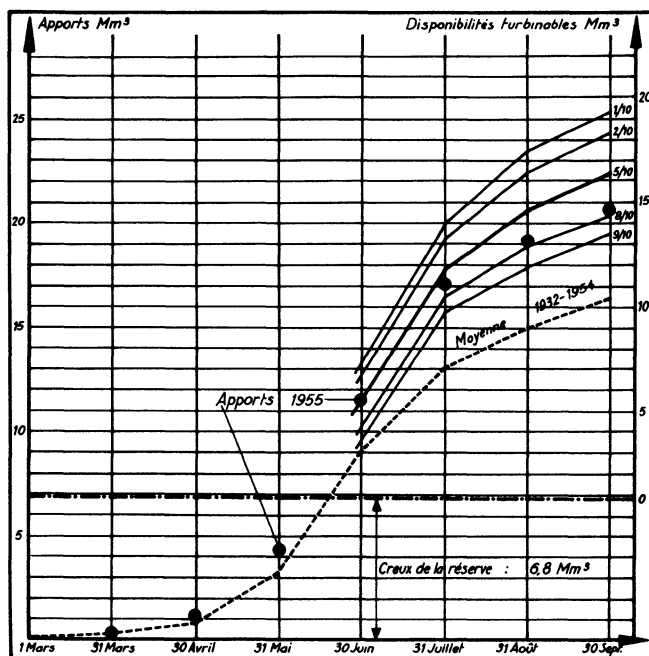
Je devrai, néanmoins, rappeler pour commencer en quoi consistent nos prévisions et quels sont les principes de base simples sur lesquels elles s'appuient. Je présente mes excuses, à ce propos, aux statisticiens ici présents, si, pour permettre à ceux des auditeurs qui ne sont pas familiers avec le langage statistique et les principes de la corrélation, je suis amené à exposer rapidement quelques notions élémentaires sous une forme volontairement schématique, et à traduire de manière plus concrète certaines grandeurs utilisées par les statisticiens.

Tout d'abord, je dois insister sur le fait que, les phénomènes à prévoir ayant un caractère aléatoire très accusé, le mot "prévision" ne doit absolument pas être assimilé à "prédiction". Prévoir les apports signifie pour nous : donner une estimation des probabilités de trouver les apports à venir d'une période déterminée entre telles ou telles limites, en tenant compte au mieux de tous les éléments d'information dont on dispose, au moment de la prévision.

Pour concrétiser ces indications, je vais vous projeter la reproduction d'un **bulletin de prévision**, émis par notre Service le 1er mars 1955, pour le bassin alimentant le réservoir des Sept-Laux dans les Alpes du Dauphiné (Fig. 1). La prévision est représentée sous forme graphique par 5 courbes correspondant aux apports cumulés depuis le 1er mars jusqu'à fin juin, juillet, août, septembre que nous estimons avoir respectivement les probabilités 9/10 ; 8/10 ; 5/10 et 1/10 d'être dépassés. A titre de comparaison, on a reporté la valeur moyenne des apports de la période 1932-54 ; on voit que, en raison de l'enneigement exceptionnel de l'hiver 1954-55, les apports prévus en 1955 se situaient très au-dessus de la moyenne. Après coup, nous avons également reporté sur le graphique les

Fig. 1 - BASSIN : SEPT-LAUX -
Prévisions d'apports à la date du 1^{er} Mars 1955

GRAPHIQUE - donnant à chaque fin de mois du 30/6 au 30/9 les valeurs des apports globaux ayant 1 - 2 - 5 - 8 et 9 chances sur 10 d'être dépassées.



Caractéristiques du B. V. et du Réservoir

Superficie du B. V.	9,6 Km ²
Capacité utile du réservoir	13,7 Mm ³
Situation du	{ 6,9 Mm ³
Réservoir au 1 ^{er} Mars 1955	{ 50 %
Débit Maximum turbinable)	{ 4,7 m ³ /sec
à l'usine de tête (D. M. T.))	{ 16,9 × 10 ³ m ³ /heure
Equivalence	1 Mm ³ = 59 h. D. M. T.

TABIEAU - donnant pour la période 1/3 - 31/8 et 1/3 - 30/9 les probabilités d'apports et les disponibilités globales résultant des conditions hydrologiques et météorologiques antérieures au 1^{er} Mars 1955.

Probabilités de dépasser les valeurs indiquées	Période : 1 Mars - 31 Août			Période : 1 Mars - 30 Sept.		
	Apports globaux Mm ³	Disponibilités turbinables globales correspond.		Apports globaux Mm ³	Disponibilités turbinables globales correspond.	
		Mm ³	Heures de marche au D. M. T.		Mm ³	Heures de marche au D. M. T.
1/10	24	17	980	25	18	1 100
2/10	23	16	930	24	17	1 040
5/10	21	14	820	22	15	920
8/10	19	12	710	20	13	800
9/10	18	11	660	19	12	740

points représentatifs des apports effectivement constatés depuis le 1^{er} mars jusqu'à fin septembre 1955. On voit qu'ils sont en parfait accord avec les prévisions

Le tableau figurant à la partie inférieure du bulletin ne fait qu'expliquer les principales valeurs numériques que l'on peut tirer des courbes.

Des prévisions analogues ont été émises pour ce bassin, chaque mois, du 1^{er} mars au 1^{er} juillet, soit 5 bulletins de prévisions comportant chacun 4 ou 5 échéances, et au total, 22 prévisions distinctes. Il en a été de même pour 23 bassins des Alpes et des Pyrénées. Nous avons ainsi publié, en 1955, une centaine de bulletins comportant 450 prévisions.

Vous ayant ainsi donné une idée assez concrète du produit de notre travail, je vais résumer succinctement les méthodes statistiques, d'ailleurs très simples qui sont à la base de nos études :

Nos prévisions sont fondées sur la connaissance des apports constatés chaque année dans le passé pendant les périodes étudiées (c'est-à-dire en fait, les périodes influencées par l'écoulement des réserves neigeuses) en corrélation avec une donnée représentative de ces réserves naturelles, également pour chaque année.

Un exemple concret va me permettre de préciser le schéma par lequel on transpose numériquement le passé pour en tirer une prévision de l'avenir.

Le graphique de la figure 2 représente le nuage de corrélation entre l'écoulement du 1^{er} mars au 31 août dans le bassin des SEPT-LAUX et un élément hydrométéorologique représentatif des réserves naturelles (solides ou liquides) au 1^{er} mars de chaque année.

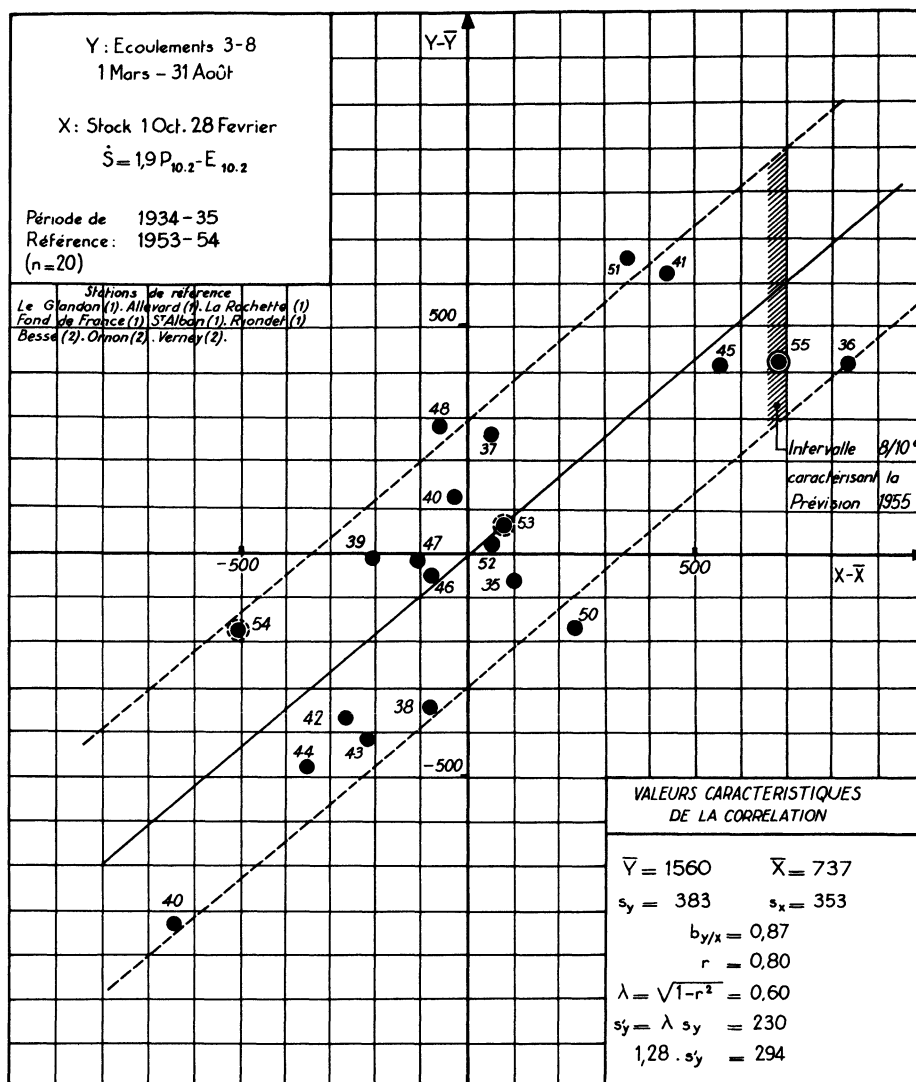
Chaque point du graphique concerne une année déterminée. Son ordonnée est égale à l'écart entre l'écoulement de cette année et la moyenne de la série de vingt années du nuage, pour la période 1^{er} mars - 31 août considérée ; l'unité employée est le millimètre d'eau sur la surface du bassin versant. Son abscisse est de même égale à l'écart à la moyenne, exprimée en mm d'eau, d'une quantité S représentative des réserves naturelles au 1^{er} mars.

S est une estimation des précipitations reçues par le bassin pendant la période d'hiver du 1^{er} octobre au 1^{er} mars de l'année considérée, diminuée de l'écoulement du bassin pendant la même période.

L'estimation des précipitations reçues par le bassin est elle-même le résultat d'une moyenne pondérée des précipitations reçues dans 9 stations sélectionnées aux environs du bassin, moyenne multipliée par le coefficient 1,9 qui tient compte de la différence importante entre l'altitude moyenne du bassin et celle des stations d'observations. La sélection des stations, leur pondération, ont été effectuées en utilisant concurremment des méthodes statistiques et des considérations physiques.

Si le nombre de point du nuage était très grand, on pourrait décomposer celui-ci en tranches verticales de faible largeur Δx et calculer, à l'intérieur de chacune de ces tranches, l'ordonnée moyenne des points correspondants. On obtiendrait ainsi une courbe moyenne $Y(X)$ appelée ligne de régression de Y en X , donnant l'"espérance mathématique" de Y pour une valeur fixée de X . Mais, le nombre de points dont nous disposons est toujours beaucoup trop petit pour permettre un tel mode de calcul. Nous sommes donc obligés de faire une hypothèse sur la forme de la ligne de régression. Nous admettons en général que c'est une droite, ce qui est l'hypothèse la plus simple et d'ailleurs conforme à la nature physique du problème.

Fig. 2 - Graphique de Corrélation - SEPT-LAUX -
Prévisions d'apports au 1^{er} Mars 1955



Moyennant cette hypothèse, la droite de régression qui s'accorde le mieux avec les données de l'échantillon se calcule facilement. Elle passe par le centre de gravité des points de nuage et sa pente résulte d'une condition de moindres carrés.

Mais la connaissance de la droite de régression ne nous suffit pas. Nous ne voulons pas seulement annoncer l'"espérance mathématique" des apports connaissant X (écart de S à sa valeur moyenne), ce qui aurait l'apparence d'une "prédiction", mais donner la loi de probabilité des apports, comme nous l'avons vu sur le bulletin précédemment projeté. Nous sommes obligés d'introduire une hypothèse nouvelle qui est le plus souvent la suivante : la loi de distribution des apports autour de la droite de régression quand X est donné est gaussienne et sa dispersion est indépendante de X.

Le problème de la prévision, tel que nous l'avons posé, se trouve ainsi complètement résolu. En particulier, nous pouvons tracer sur le graphique de

Fig. 3 - PREVISIONS D'APPORTS
1^{er} MARS - 31 AOÛT

- ALPES -

Bassins	Apports moyens mm.	Intervalle de Probabilité 8/10 en % des apports moyens		Gain dû à la corrélacion en %
		sans corrélacion	avec corrélacion	
Girotte (*)	1 203	± 20	± 12	40
Tignes	875	± 17	± 9	49
Bissorte	899	± 31	± 17,5	44
Sept-Laux	1 560	± 31	± 19	40
Romanche + Ferrand	774	± 38	± 17,5	47
Drac	672	± 42	± 33	22
Durance	422	± 43	± 28	34
Verdon	393	± 51	± 36	28
Rhin à Rheinfelden	539	± 21	± 13	40
Rhône à Chancy (*)	662	± 18	± 9	50
Rhône à Genève	647	± 17	± 12	29

- PYRENEES -

Bassins	Apports moyens mm.	Intervalle de Probabilité 8/10 en % des apports moyens		Gain dû à la corrélacion en %
		sans corrélacion	avec corrélacion	
Fabrèges	1 086	± 31	± 16	46
Artouste	1 345	± 34	± 20	39
Oule	841	± 35	± 16,5	44
Caillaouas (*)	1 374	± 31	± 26	16
Portillon + Glace	1 805	± 18	± 15	17
Oo	1 217	± 36	± 20	26
Arraing	869	± 31	± 23	25
Izourt	1 109	± 33	± 11	66
Gnioure	1 059	± 31	± 16	49
Naguilhes en Beys	959	± 24,5	± 16	36
Puyvalador	506	± 42	± 32	22
Bouillouses	723	± 37	± 24	35

(*) Caractéristiques 1^{er} Avril - 31 Août.

corrélation les deux parallèles à la droite de régression, située de part et d'autre de celle-ci et entre lesquelles le point représentatif d'un écoulement futur aura 8 chances sur 10 de tomber pour toute valeur de X donnée.

*

* *

Nous pouvons maintenant définir un certain nombre de notions permettant de caractériser la qualité des prévisions. Tout d'abord, en supposant que les probabilités annoncées soient parfaitement exactes, la précision de la prévision peut être caractérisée, par exemple, par la $1/2$ largeur de l'intervalle dans lequel nous estimons que les apports ont 8 chances sur 10 de tomber.

Dans l'exemple donné par les Sept-Laux, ce $1/2$ intervalle est égal à 294 mm soit 19 % de la valeur moyenne des apports de la période faisant l'objet de la prévision.

En l'absence de corrélation, ce $1/2$ intervalle aurait été de 490 mm soit 31 % de la moyenne des apports. On peut donc porter au crédit de la corrélation le gain qu'elle a permis, soit 40 %.

Pour l'ensemble des corrélations utilisées, le gain ainsi obtenu varie suivant le bassin et la période de 20 % à 65 %, la moyenne étant de 40 % environ. Le $1/2$ intervalle de probabilité 8/10 varie de son côté de 35 % à 10 % avec une moyenne de 18 % (voir tableau - Fig. 3).

Mais ces indications de gain et de précision de la prévision supposent l'exactitude des probabilités annoncées. Or, il se peut que ces probabilités soient erronées, soit parce que les hypothèses de base précédemment rappelées ne sont pas entièrement correctes, soit parce que, même avec des hypothèses valables, l'échantillon limité dont on dispose conduit à calculer des paramètres de corrélation s'écartant accidentellement par trop des paramètres "vrais" qui seraient obtenus sur un échantillon très nombreux. Il convient donc de vérifier la concordance de la distribution des apports réels avec les probabilités annoncées.

Une première vérification qualitative consiste à contrôler que, pour chaque bassin, l'ensemble des prévisions émises ne sort que rarement et modérément de l'intervalle auquel nous avons attribué une probabilité de contenir les apports effectifs égale à 8/10. Les graphiques 4A, 4B, 4C font ressortir ce contrôle pour le bassin des Sept-Laux et les principales prévisions émises au cours des 3 dernières années. Les graphiques 5A, 5B, 5C donnent les mêmes éléments pour le bassin du lac d'Oo dans les Pyrénées.

Pour une échéance déterminée (31 août par exemple), les diverses prévisions énoncées le 1^{er} mars, 1^{er} avril, 1^{er} mai et 1^{er} juin, se présentent sous la forme d'une série d'échancrures (ou fenêtres) accolées, limitées par les hâchures et correspondant chacune à un intervalle de probabilité 8/10 considéré précédemment. Le graphique permet ainsi de se rendre compte de l'évolution de la prévision au fur et à mesure que la saison s'avance, du gain généralement obtenu dans la précision (l'intervalle devenant plus resserré pour les prévisions émises en Avril et Mai) et de la position respective des prévisions et des apports effectifs.

A titre accessoire, le graphique donne également les apports de référence : valeurs cumulées à partir du 1^{er} Mars, d'une part moyenne et d'autre part correspondant aux fréquences 1/10 et 9/10.

Le type de graphique que vous venez de voir permet d'apprécier, dans une certaine mesure, la concordance de l'ensemble des prévisions émises pour une année et un bassin déterminé, tant entre elles, qu'avec la réalisation effective. Mais il est plus instructif d'examiner synthétiquement, si, pour l'ensemble des bassins étudiés et les différentes années, la distribution des apports effectifs peut être considérée comme conforme à nos prévisions.

Fig. 4 A - SEPT-LAUX -
Comparaison entre Prévisions et Réalisation

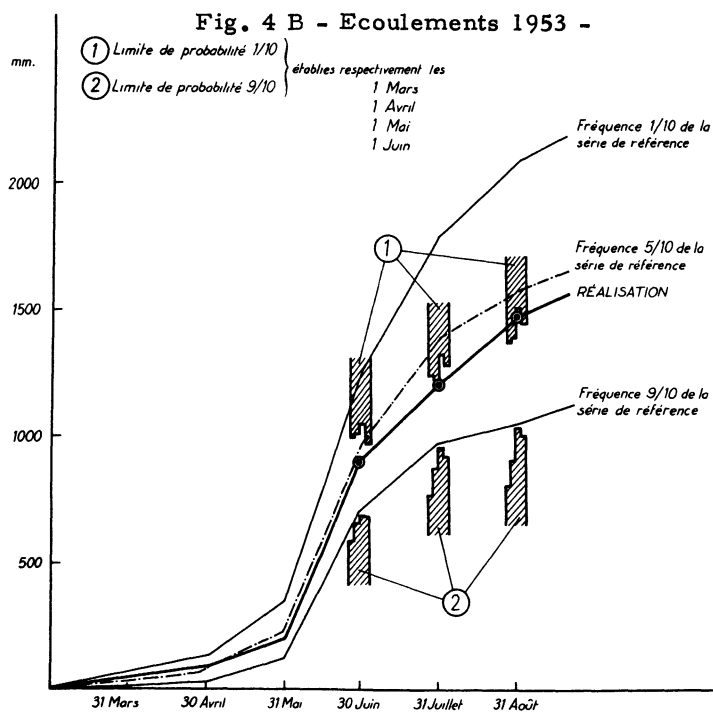
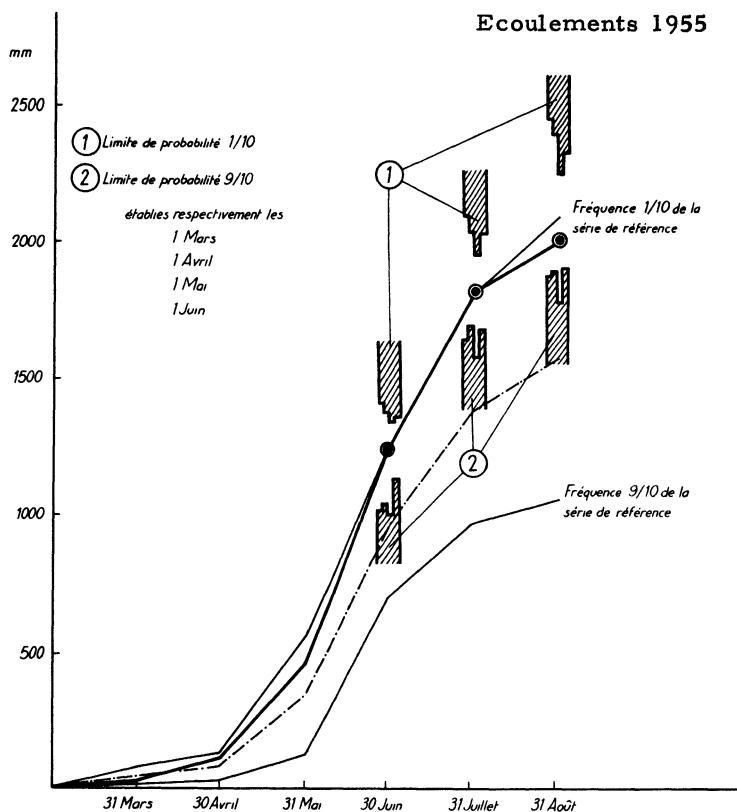


Fig. 4 C - Ecoulements 1955 -

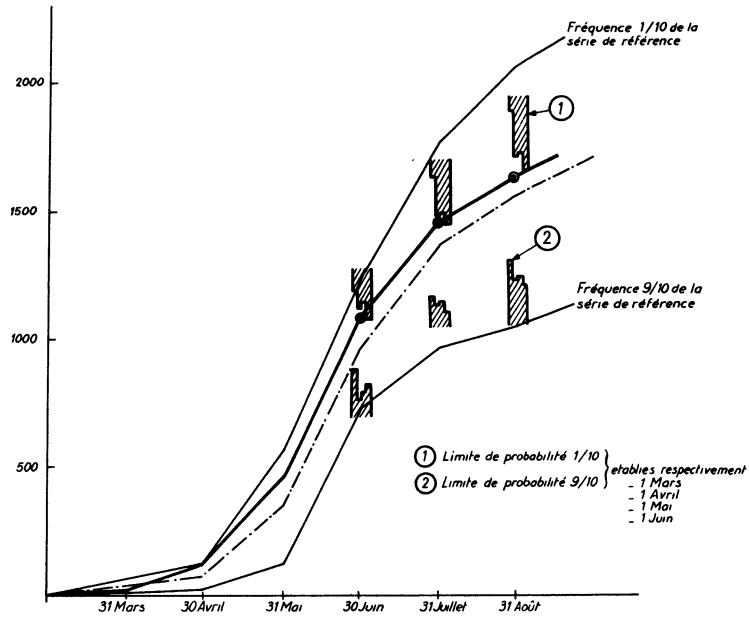


Fig. 5 A - OÔ -

Comparaison entre Prévisions et Réalisation

Ecoulements 1955

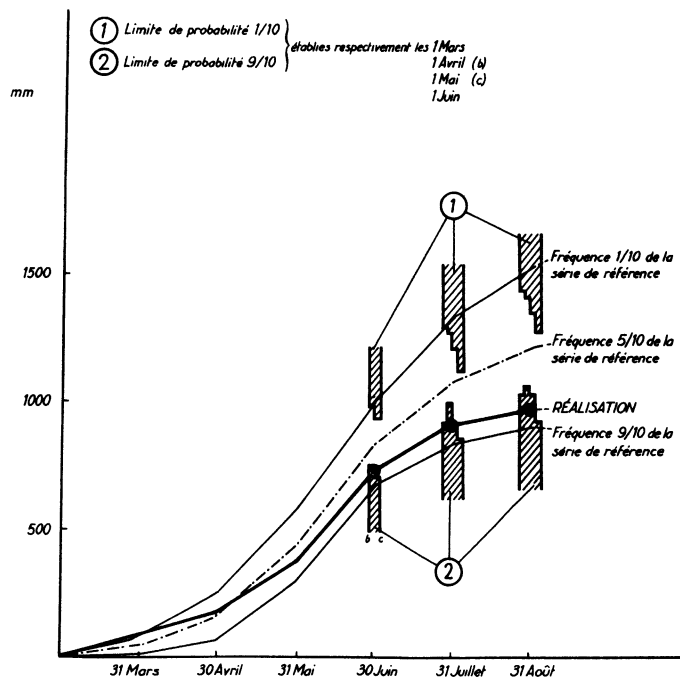


Fig. 5 B
Ecoulements 1954

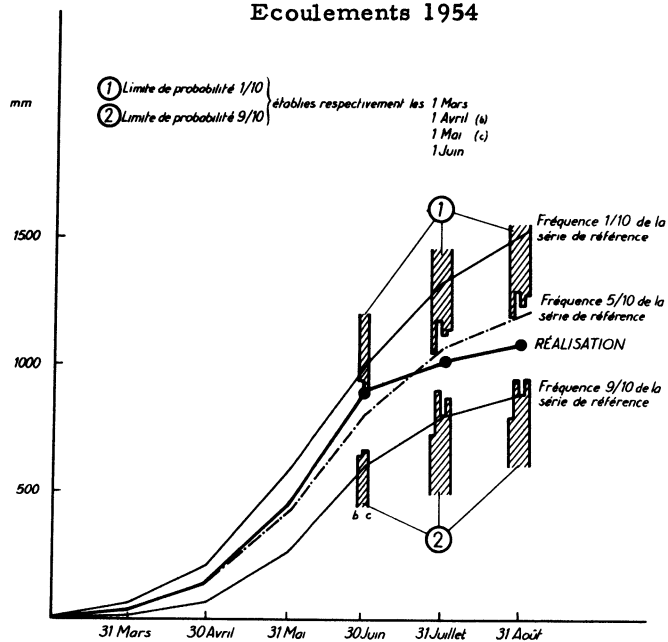
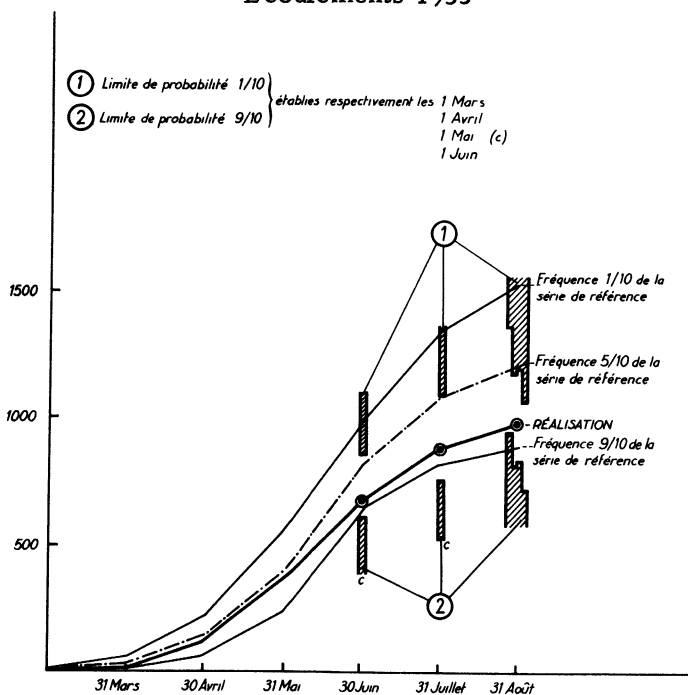


Fig. 5 C
Ecoulements 1953



Les graphiques 6A et 6B répondent à cet objet, ils intéressent l'ensemble des prévisions énoncées depuis 1951 pour les divers bassins, d'une part dans les Alpes, d'autre part dans les Pyrénées, pour une période déterminée.

APPORTS 1^{er} AVRIL - 31 AOÛT

Ecartés réduits liés "Prévision-Réalisation"

Fig. 6 A ALPES

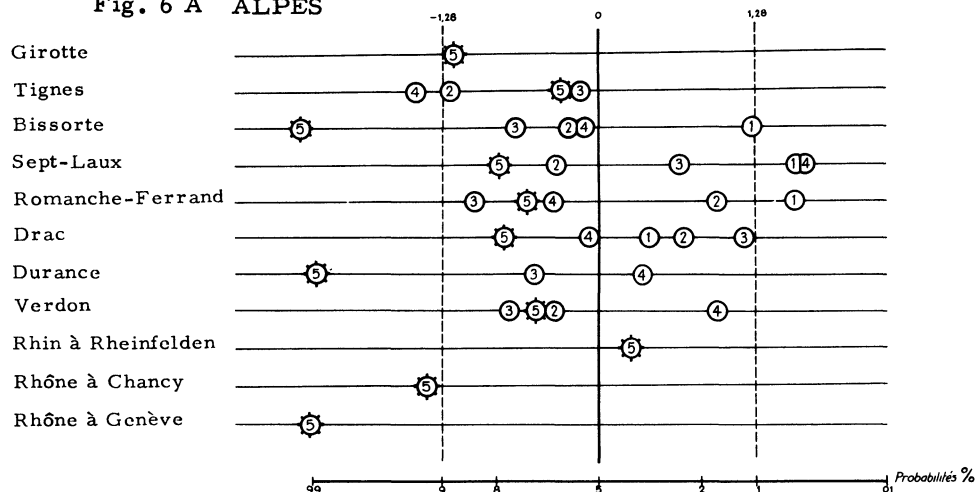
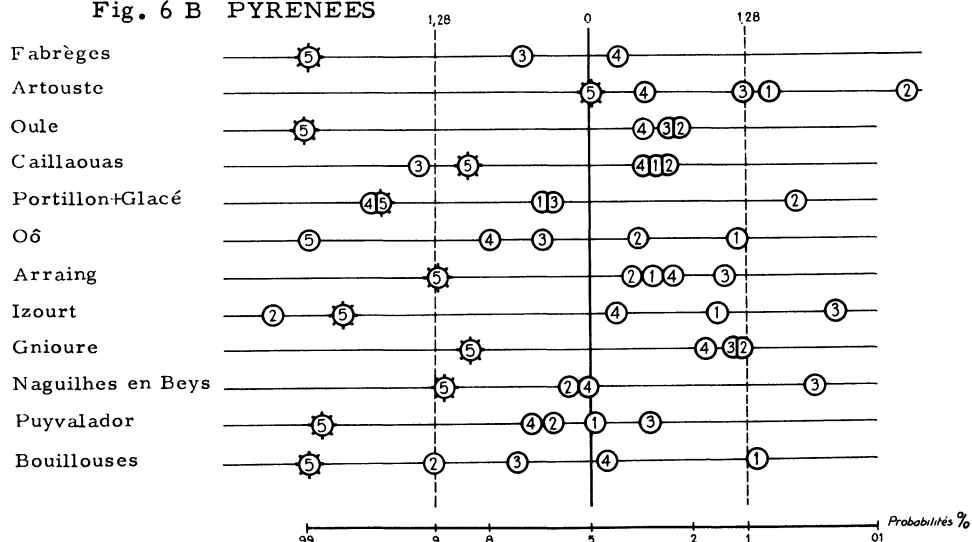


Fig. 6 B PYRENEES



Chaque ligne concerne un bassin et, sur ces lignes, un point représente, pour chaque année, la position des apports constatés par rapport à une échelle commune graduée en probabilité.

Si la loi de distribution que nous avons admise dans chaque cas est correcte, nous devons constater en particulier que 80 % environ des points se situent bien dans l'intervalle des deux droites tiretées.

Sur le graphique des Alpes, on constate que 27 points (au lieu de 28 théoriques) sur 35 sont contenus dans l'intervalle dans lequel on doit s'attendre à trouver 8 points sur 10. Mais, en réalité, 3 points relatifs à l'année 1955 s'écartent anorma-

lement de la limite inférieure de cet intervalle. En y regardant de plus près, on s'aperçoit que l'ensemble des points de 1955 accuse un déficit par rapport aux valeurs centrales de la prévision, leur dispersion restant comparable à celle des autres années.

Sur le graphique des Pyrénées, ce phénomène est encore plus accentué. Pour les années 1951-1954, 34 points sur 43 (c'est à dire la proportion théorique) sont compris dans l'intervalle de probabilité 8/10. Mais, en 1955, la prévision se trouve manifestement en défaut.

Naturellement, on peut trouver diverses explications à cet accident. Mais, à vrai dire, nous manquons d'éléments pour retenir telle hypothèse, plutôt que telle autre. Ce qui caractérise tout particulièrement l'hiver 1955, c'est la concomitance d'une température exceptionnellement élevée avec les plus importantes précipitations de la période. Cette situation paraît avoir coïncidé avec un gradient pluviométrique particulièrement faible, d'où une surestimation des précipitations en altitude.

Quoi qu'il en soit, ce que l'on peut conclure, c'est que, sans que le schéma statistique simple adopté soit parfaitement adapté à la réalité, il constitue en pratique un guide de prévision valable qui peut se trouver exceptionnellement en défaut. Ce n'est que lorsqu'on possèdera des échantillons beaucoup plus fournis qu'il sera justifié de compliquer le schéma.

*

* *

Il me reste pour terminer à donner quelques indications sur l'utilisation des prévisions pour l'exploitation E.D.F. Je reprendrai à ce propos l'exemple des prévisions des Sept-Laux au 1er mars 1955.

La première question que l'on peut se poser est : quel degré de vidange peut-on atteindre en Avril sans compromettre la probabilité du remplissage ?

Le bulletin de prévision donné tout à l'heure (Sept-Laux à la date du 1er mars 1955) indiquait, d'une part, qu'au 1er mars le réservoir contenait encore en volume 50 % de sa capacité utile, soit 6,9 M.m³ et que, d'autre part, on disposait jusqu'au 31 août de 11 M.m³ à turbiner, en ayant 9 chances sur 10 de remplir le réservoir au 31 août.

Dans ces conditions, on pouvait envisager dès le 1er mars de vider complètement la réserve avant l'arrivée des apports de fusion. En fait, on a vidé 30 % en mars et le reste, soit 20 %, en Avril.

Dans une situation déficitaire, les prévisions permettent, au contraire, de déterminer le creux minimum qu'il convient de ne pas dépasser pour conserver une probabilité donnée de remplir à fin août.

D'une façon plus générale, on peut apprécier l'intérêt de l'information apportée par les prévisions sur ce point, en calculant - pour une probabilité de remplissage donnée, par exemple 9/10 - le culot moyen qu'il conviendrait de laisser dans chaque réservoir en fin d'hiver et le creux moyen qu'on obtiendra au 31 août, d'une part, lorsqu'on ne connaît que la seule répartition des apports de référence d'autre part, lorsqu'on dispose des informations résultant des corrélations.

Nous avons fait le calcul dans le cas des réservoirs dont la capacité est au moins de l'ordre des apports avril-août et ne disposant pas de pompage.

Pour des réservoirs comme Bissorte, Sept-Laux, Gnioure, par exemple, dont la capacité est voisine des apports moyens d'avril à août, la valeur du culot passe de 30 % environ de la capacité à 17 % en moyenne.

Par ailleurs, en fin de remplissage, les prévisions permettent d'atteindre en moyenne un creux pratiquement nul, tandis que la seule connaissance des données de référence conduirait à un creux moyen compris entre 1 et 2 %.

On peut donc dire que les prévisions permettent, à probabilité égale, une meilleure utilisation de la capacité des réservoirs.

En outre, les prévisions permettent une meilleure utilisation des disponibilités au cours de la période de remplissage de mai à août. Sans elles, en effet, on serait conduit à une prudence en général excessive, et parfois cependant insuffisante dans les premiers mois de remplissage.

Disons encore que, dans quelques cas particuliers, concernant des installations dont on prévoyait l'indisponibilité pendant une partie de la période de remplissage ou lorsqu'une cote maximum était provisoirement imposée par des travaux, nous avons pu fournir des indications sur les programmes de turbinage qu'il convenait alors de se fixer (exemples : Portillon en 1951, Verdon en 1954).

Les prévisions peuvent encore donner des informations utiles pour :

- l'élaboration de certains programmes de fourniture d'énergie complémentaire à des consommateurs industriels ;

- la fixation de la durée des périodes d'indisponibilité pour l'exécution de travaux sur les ouvrages ;

- l'exploitation de la tranche "agriculture" de certains réservoirs.

Pour conclure, nous pouvons dire que, sans éliminer toutes les incertitudes nécessairement inhérentes à la prévision des apports printaniers et estivaux dans les hauts bassins des Alpes et des Pyrénées, nos prévisions, basées sur des méthodes statistiques simples mais efficaces, contribuent incontestablement à réduire de façon très appréciable cette marge d'incertitude. Nous continuons d'ailleurs à travailler dans diverses directions, pour permettre de la réduire encore dans la mesure du possible, dans un avenir plus ou moins lointain. Nous espérons avoir ainsi apporté une contribution utile à l'effort de gestion rationnelle de nos ressources en énergie électrique poursuivi par les Services de Production et Transport d'ELECTRICITE DE FRANCE.

Commentaires de M. le Président BATICLE

M. le Président remercie M. FERRY d'avoir apporté une contribution très intéressante à la prévision du fonctionnement d'usines et croit qu'il ne faut pas s'étonner des points qui se sont trouvés en dehors de la zone de probabilité de 80 % ; c'est une chose absolument courante. Il est très facile de voir, par exemple, par la loi de Poisson, que la probabilité qu'on échappe à cette fourchette est d'un ordre de grandeur assez fort.

Cela est tout à fait normal et se produit chaque fois que l'on veut confronter les probabilités et les fréquences observées et, comme M. FERRY l'a d'ailleurs dit au début, estimer ou donner une prévision n'est pas du tout prédire, car cette prévision est elle-même soumise à une loi de probabilité.