

AJUSTEMENT ET RÉGLAGE DES MODÈLES DÉTERMINISTES MÉTHODE DE CALAGE DES PARAMÈTRES

par **Y. CORMARY**

Chef du Centre Hydrométéorologique de Montpellier

et **A. GUILBOT**

Centre Hydrométéorologique de Montpellier
Université des Sciences et Techniques

Introduction

Il ne m'est pas possible de faire un exposé complet et détaillé des méthodes d'optimisation qu'il serait intéressant d'utiliser en hydrologie pour résoudre les problèmes que recouvrent ce terme.

En effet, il s'agit d'élaborer une méthode objective permettant :

a) L'identification du système de relation entre variables adéquates;

b) Le calage des paramètres des fonctions qui expriment ces liaisons sur l'échantillon entrée-sortie utilisé;

Des calages effectués indépendamment sur divers bassins ne présentent tout leur intérêt que s'ils conduisent sur des zones réputées homogènes au point de vue hydrologique à la définition d'un modèle le plus souple possible et à des jeux de paramètres homologues cohérents;

Il doit être prouvé que le calage n'est pas fondamentalement remis en question quand la série utilisée change (échantillon d'années différent ou changement de postes pluviométriques). Un test de comparaison des paramètres des lois de distribution des débits reconstitués (à divers pas de temps) à ceux obtenus à partir des débits réelles observés, est indispensable et ceci pour divers échantillons d'années.

Ce serait beaucoup trop long mais, si nous nous contentons simplement d'exposer sommairement les premiers résultats que nous avons obtenus par la méthode que nous avons choisi, le problème ne serait vu que trop partielle-

ment. Nous ferons donc auparavant un rappel des définitions et des schémas d'optimisation possibles :

Un modèle déterministe s'exprime de la manière suivante :

$$Y_e = T(X_1 \dots X_J \dots X_m, a_1 \dots a_i \dots a_n)$$

Les $X_1 \dots X_m$ étant les variables et $a_1 \dots a_n$ les paramètres liant de manière mathématique la fonction Y aux m variables X , T l'expression de cette liaison.

Si Y_0 est la valeur réelle observée pour le groupe des valeurs $X \dots X_J$, il s'agit en général de minimiser une fonction critère F . Cette fonction critère dépend des écarts successifs observés par l réalisations Y_c de Y_0 .

$$\text{Pour une réalisation : } F = (Y_0 - Y_c)^2$$

$$\text{Pour } l \text{ réalisations : } F = \sum_1^l (Y_0 - Y_c)^2$$

F est donc fonction de la nature des $mX_1 \dots X_m$ et des n paramètres $a_1 \dots a_n$ et dépend de la forme de T mais aussi de la série de l groupes de valeurs $(Y_{0l}, X_{1l} \dots X_{Jl})$ pris en compte.

Hydrologie : cas particuliers

En hydrologie ces problèmes présentent deux particularités :

a) Dans l'identification du système des relations entre variables hydrologiques (en général pluie et débit), la complexité du comportement de l'eau en milieu non saturé, la non-stationarité de ces relations au cours des sai-

sons ne permettent pas encore, pour l'instant, l'identification objective. Il s'agit en fait de choisir entre divers schémas raisonnables mais subjectifs, qui viennent d'être décrits, celui qui avec le minimum de paramètres permet, non le calcul le plus simple, mais d'effectuer des extrapolations efficaces aux années où seule la pluie est disponible, stables et cohérentes régionalement sur plusieurs bassins versants semblables. Seuls les modèles de nappes et de propagations des crues sont plus solides sur le plan théorique mais certaines simplifications que l'on fait subir au cas complet les ramènent quelquefois au cas général de l'hydrologie.

b) L'hydrologue n'a pas la ressource d'effectuer des mesures expérimentales permettant par le jeu de tel ou tel paramètre seul d'opérer des explorations partielles, d'apprécier l'erreur qui entache des mesures entrées et sorties à l'échelle du bassin et ne peut effectuer que des appréciations encore bien insuffisantes (piézométrie, humidité) sur certaines variables intermédiaires du cycle schématisé choisi.

Par contre, il se trouve confronté au cas général de la minimisation d'une fonction critère F qui est sensée représenter l'adéquation de ces modèles et l'efficacité du calage des n paramètres.

Comme ses collègues chimistes et économistes, il ne sait pas exprimer la fonction F en fonction des paramètres et il admet en première approximation que le modèle est parfaitement déterministe et adéquat. C'est-à-dire en particulier que les données ne sont pas entachées d'erreurs et que l'échantillon qu'il en possède est suffisant pour trouver les paramètres qui minimiseront cette fonction.

La fonction F réelle dans l'hyper-espace des n paramètres a_i a des chances d'être complexe et dissymétrique et sans doute non unimodale (présentant plusieurs dépressions de différentes profondeurs). De toute manière elle présente un point bas qui, si le modèle n'est pas parfaitement adéquat et les données exemptes d'erreurs ne peut pas être 0 mais où toutes les n dérivées dF/da s'annulent.

Il n'est pas plus courant en hydrologie qu'en matière de Génie chimique ou d'optimisation économique de trouver raisonnablement ou théoriquement des fonctions T liant variables dépendantes aux variables indépendantes qui s'expriment sous une forme linéaire et additive en a_i (et non forcément en X_j) et l'on a à résoudre le problème de la régression non linéaire.

Dans le cas rare où l'on sait d'abord écrire F en fonction des na_i et où le système de n équations à n inconnues est linéaire, la solution est immédiate. Elle est encore possible quand r contraintes entre variables s'expriment de manière linéaire en a (résolution de n équations à $n-r$ inconnues). Ce cas particulier de la régression linéaire, le seul praticable en l'absence d'ordinateur, a été très utilisé et a donc l'avantage d'avoir conduit à des études de variabilité statistique du calage qui ne sont d'ailleurs pas parfaitement résolues, dans le cas de la colinéarité des variables X_j . Le cas de la régression linéaire laisse lui aussi entier le problème de l'adéquation du modèle et des erreurs sur les données.

Notons que l'aplatissement artificiel par transformations des variables X de l'hyper surface représentant la fonction $T [T(X)]$ et non $F(T, a_n \dots)$ paraît à nos collègues chimistes (M. Vignes et M. Kitrel) être très dangereux car il déforme la fonction F et la répartition réelle des erreurs et conduit à des conclusions erronées tant dans le domaine de l'exploitation que dans celui de la compréhension du phénomène.

Notons aussi que les méthodes graphiques telles que les « Coaxial - correlations » ou l'assouplissement des corrélations multiples linéaires cherchent en fait à identifier par tâtonnement la fonction T et les paramètres a .

Simplification possible en hydrologie

On peut introduire des découpages matriciels géographiques liés à des différences matricielles de la pluviométrie (intensité et distance à l'exutoire) ou à un rendement des pluies (divers complexes sols-végétation) ou de la topographie (hydrogramme synthétique de Larrieu).

On peut aussi décomposer le modèle lui-même.

Citons trois exemples concrets :

Nous utilisons un modèle Hamon pour suivre l'évolution d'un stock d'eau dans un profil soumis à une évapotranspiration fournie par le déficit au-dessous de la capacité de rétention et au contraire drainé quand la teneur en eau dépasse ce seuil.

On a pu apprécier le rapport ETR/ETP par les expérimentations agricoles en été; des mesures de percolation identifiées par remontée de la nappe (piézométrie) en hiver, nous ont permis de caler indépendamment la deuxième partie du modèle par ailleurs.

Dans un deuxième cas il est possible d'identifier une partie du système sous forme de fonction de transfert linéaire (analyse des courbes de décrues après arrêt de la pluie). Une fois ce système déterminé pour plusieurs crues il est possible pour chaque crue de calculer inversement les entrées E_i qui, transformées par un système moyen linéaire, ont produit la crue observée. En fait, la série discrète des mE_i qui sont les parties efficaces des m valeur P_i de la pluie qui a causé la crue peut être identifiée en se donnant des contraintes de cohésion entre les valeurs E_i et de non-dépassement du P_i homologue grâce à l'emploi de la programmation dynamique. Celle-ci s'adapte bien car la décision se fait par récurrence de E_{m-1} à E_1 (les m paramètres sont ici les mE_i), se fait séquentiellement et l'état du système ne dépend que de la décision précédente.

La fonction critère se décompose en une somme de carrés des écarts. Ainsi, en faisant le minimum d'hypothèses sur la fonction de transfert il reste à identifier le système non linéaire et non invariant à mémoire de passage des P_i aux E_i en phase. Pour cela des essais d'identification ont utilisé le calcul analogique. Nous y reviendrons.

Troisième exemple : pour l'annonce des crues la même politique peut être suivie. Sur un certain nombre de crues, on peut choisir la fonction d'amortissement ou l'hydrogramme unitaire (nous utilisons par exemple la méthode du simplex sur l'ensemble des crues) ou les paramètres de la fonction (type Holtan) de production d'une crue autres que ceux de l'humidité du sol u avant la crue.

Une inversion du système de convolution dans le premier cas où l'optimisation crue par crue de la valeur de u dans le second permet d'obtenir de nouvelles données qui servent d'entrée à un modèle de relation pluie-débit dans un cas, à un modèle d'évolution du stock de l'eau dans le second (type Hamon).

Cet effort d'éclatement du modèle en modèles déterministes plus simples permet de limiter le nombre d'hypothèses, de laisser le maximum de chances de possibilités de linéarisation, et de mieux se rendre compte de ce que l'on fait.

Bien entendu, il faut dans cette pratique séparer les données en une série qui servira à ajuster le modèle et l'autre à le tester.

Optimisation dans le cas général

Une fois ces tentatives de division du problème, on revient toujours tôt ou tard au problème général évoqué qui met en cause une stratégie développant quatre tactiques différentes.

Il faut :

1. Définir la zone d'attaque.

Définir la place que l'on attribue à la variation de chaque paramètre. Les bornes peuvent être fixées par des considérations physiques mais le plus souvent par des impossibilités de calcul (par exemple : log de chiffres négatifs ou des absurdités : capacité maximale négative d'un réservoir). On peut d'ailleurs se demander s'il faut effectivement masquer les incohérences du modèle ou les éliminer.

Ce bornage peut se faire en ajoutant à la fonction F une fonction G qui, dans le domaine permis, garde des valeurs très faibles et devient rapidement important en dehors.

Il se peut que les contraintes entre paramètres doivent être respectées. Par exemple entre deux paramètres A_1 et A_2 :

$$A_1 = 2(A_2)^2$$

on ajoute à F une fonction G (v) qui, lorsque :

$$v = 2(A_2)^2 - A_1$$

est positif, est très grande et qui est nulle quand il est négatif. On définit ainsi des segments $L_{(n)}$ pour chaque paramètre a_n .

2. Sonder le front pour trouver le point propice.

Choisir dans le domaine un point de départ (n paramètres + valeur des variables d'état du système).

Toutes les techniques que nous allons passer en revue postulent que la fonction est unimodale dans le domaine où l'on explore sa variation. Plus le point de départ sera situé bas sur l'hypersurface, plus on a de chance que les faux minimums d'ordre plus élevés soient éliminés, ce qui converge vers un vrai minimum dans la zone où la fonction F est unimodale.

D'où l'intérêt de déterminer d'abord dans la zone de creux principal si l'on n'a pas de raisons physiques qui permettent de le choisir (expérimentation antérieure locale, analyse d'un état particulier) *a priori*. Il y a la solution simple, qui consiste à diviser en s tronçons la longueur permise à chaque paramètre, mais l'hypervolume défini par les domaines précédents est divisé en S^n (volumes partiels et égaux) et il devient vite prohibitif en temps machine d'explorer sériellement tous ces cas.

On peut gagner en se donnant une politique d'expérience séquentielle destinée à économiser les expériences (voir paragraphe suivant), mais l'on admet alors l'unimodalité de la surface. Une politique d'exploration au hasard (méthode de Brooks), permet d'éliminer avec une certaine probabilité la zone qui ne contient pas le minimum. Supposons que l'on délimite le meilleur des cent cubes possibles d'une fonction de trois paramètres. Ce volume :

$$0,4_{L,1} \times 0,5_{L,2} = 0,5_{L,3}$$

représente une réduction d'incertitude de seulement 4/10 sur un des paramètres et 5/10 sur chacun des deux autres. Il a fallu pourtant d'après la méthode de Brooks, K essais pour que avec une probabilité P_f (ici $f = 1/100$) la valeur F minimum soit située dans le cube retenu :

$$P_f = 1 - (1 - f)^k$$

Ici $P = 0,866$ pour $K = 200$.

3. Développement de l'attaque en profondeur.

Choix de la méthode d'attaque jusqu'à un point proche du minimum. Dans cette plage on admet que la surface n'a pas une forme suffisamment régulière pour que la manière de converger puisse être prise en compte dès le départ, mais que l'hypersurface F est suffisamment abrupte pour que les gains que l'on détecte par de nouveaux essais sur les paramètres a soient d'un ordre de grandeur supérieur à ceux provoqués sur F par les erreurs de mesures sur les valeurs $X_{j,l}$ ou Y_l .

Il n'est d'ailleurs pas évident que la fonction critère à optimiser soit la même pour les points 2 et 3 (calage grossier d'abord, raffiné ensuite, suivant le but recherché), ni qu'il ne soit pas possible de négliger certains paramètres secondaires au stade 2.

3.1 *Le cas simple* : 1 seul paramètre, 3 types de méthodes. Méthodes de recherche des racines :

Certaines supposent que F(a) prend la valeur 0 entre les bornes a_0 et a_1 .

Certaines appliquent soit des déductions géométriques soit des ajustements de plus en plus complexes et les résultats précédents progressent vers la valeur de a où F s'annule, mais elles paraissent mal s'appliquer puisqu'il faut qu'une valeur de l'intervalle donne une valeur négative.

D'autres sont déduites des méthodes de Newton Raphson qui met en jeu la dérivée de la fonction.

Les plus courantes sont sérielles et systématiques, et la zone d'incertitude au bout de n coups équidistants est réduite à :

$$2(a_1 - a_0)/(n - 1)$$

de l'intervalle initial a_1 et a_0 .

D'autres méthodes de recherche simultanée ne partent pas d'intervalles *a priori* égaux, mais s'appuient sur la théorie des jeux, minimisant les écarts maximaux d'incertitude (zone résiduelle d'incertitude) dans le domaine primitivement défini.

Il est nécessaire de se donner la plus petite différence significative qui traduit au niveau de la précision sur le paramètre a , celle sur la fonction F due aux erreurs de mesures sur les variables Y et X.

Ceci pour trouver le nombre d'expériences nécessaires pour optimiser un schéma adéquatement choisi.

Deux types de recherches séquentielles sont courantes :

— la recherche dichotomique permet de situer au bout de n essais l'optimum dans l'intervalle de longueur :

$$[2^{-n/2} + (1 - 2^{-n/2}) \epsilon]$$

— la recherche de Fibonacci, basée sur la série des chiffres de Fibonacci 1, 2, 3, 5, 8, etc. :

$$[F(k) = F(k - 1) + F(k - 2)]$$

place la première expérience à une distance :

$$[F(n - 1) + (-1)^n \epsilon]/F(n)$$

d'un bout de l'intervalle.

Le deuxième essai se place symétriquement par rapport à l'autre bout, le troisième est placé de la même manière que le premier mais $n = 3$ et deux intervalles sur trois sont ainsi éliminés et le processus se réitère sur l'intervalle restant.

Au bout de n essais l'intervalle est réduit à :

$$[1 + F(n-2)/F(n)]\epsilon$$

de l'intervalle original. En pratique le nombre maximal pour que l'expérience soit rentable est :

$$n \leq 4,785 \log(1/\epsilon) - 0,328$$

La méthode de la section dorée ne se réfère pas à la valeur *a priori*, mais est très semblable à celle de Fibonacci.

Le premier intervalle est placé à 0,618 d'une extrémité, ce qui conduit l'intervalle d'incertitude à $(0,618)^n$ au bout de n coups. Elle est de 17 % moins efficace que celle de Fibonacci. A titre d'illustration, une réduction à 1/100 de l'intervalle initial est obtenue au bout de onze essais de Fibonacci, quatorze expériences dichotomiques et cent quatre vingt dix-huit simultanées.

Une transposition de la méthode de Fibonacci permet de résoudre le problème quand le paramètre ne peut prendre que des valeurs discrètes, méthode plus simple que d'élaborer par la théorie des jeux un plan d'expérience spécial.

3.2 La recherche d'un optimum multidimensionnel.

Premier cas : pas d'interaction entre paramètres, forme symétrique ou régulière, il est possible d'optimiser tour à tour les n paramètres.

Cas général : l'utilisation de deux variables, tour à tour, risque de bloquer sur une arête de résolution. Le changement de plusieurs variables à la fois définit un hyperplan qui n'a aucune raison de passer par le sommet de l'hyper-surface C . Donc deux solutions :

- détecter les arêtes par une méthode directe;
- mieux utiliser la structure de la surface (utiliser les plans tangents).

Recherche d'une arête par la méthode Hooke et Jeeves.

On cherche par accroissement préalable d'un pas déterminé la direction unidimensionnelle la plus efficace, et on accélère cette direction jusqu'à ce que l'on se heurte, malgré une réduction éventuelle du pas d'incrément, à l'impossibilité de faire décroître la fonction F .

Le dernier point obtenu sert de départ à une nouvelle exploration.

Méthode de Rosenbrock.

Au lieu de choisir l'axe d'exploration par rapport aux axes des coordonnées, on fait tourner ceux-ci de manière à ce que l'un d'eux soit celui de l'arête de résolution. Le pas d'incrément et d'exploration n'est pas le même sur chaque axe, mais l'exploration se fait successivement sur tous les axes.

On ne poursuit la recherche suivant l'un des axes qu'après avoir exploré tous les autres tour à tour.

Un succès dans une direction entraînant une multiplication par trois du pas dans cette direction et un échec entraînant une multiplication par $1/2$ de ce même pas.

On poursuit la recherche jusqu'à ce que dans toutes les directions on ait enregistré au moins un succès suivi d'un échec.

On détermine alors de nouveaux axes tels que le nouveau système soit automatiquement orthogonal et unitaire par l'orthogonalisation de Schmidt : la direction des nouveaux axes est pondérée par le chemin parcouru sur chacun d'eux et la première direction par le cheminement global du jeu d'essais précédents.

On recommence l'exploration jusqu'à l'obtention d'un minimum acceptable.

Méthode des gradients et des tangentes.

Si l'on suppose (dans un plan à deux dimensions pour les facilités d'expression) que la surface de F est régulière (ellipsoïde ou sphère), on peut utiliser les plans tangents à la surface ou les gradients perpendiculaires aux contours (intersection de la surface avec un plan horizontal).

On explore la ligne de plus grande pente de la colline (ou du vallon) tant que l'on peut monter ou descendre. Ensuite on redéfinit une nouvelle direction d'exploration suivant la nouvelle plus grande pente.

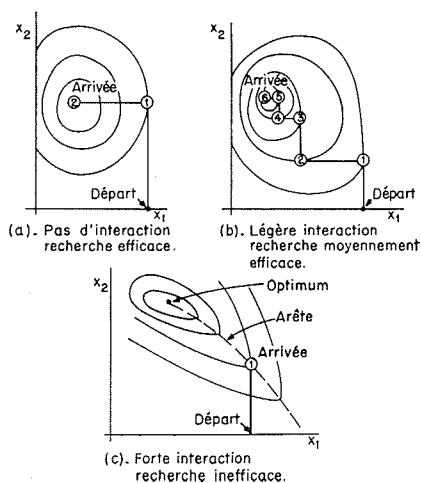
Des variantes comme les méthodes de Both et de Marquart, perfectionnent la méthode en fonction de l'angle de deux directions de montée successives.

La convergence est d'autant plus rapide que l'échelle est choisie de manière à rendre la figure la plus régulière possible.

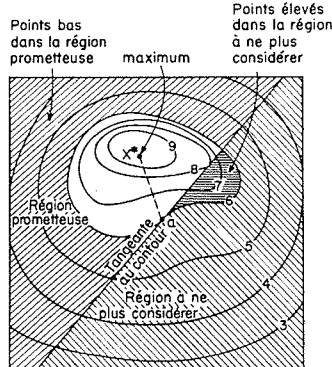
La méthode Partan (tangentes parallèles).

Elle repose sur la propriété des ellipsoïdes près d'un sommet.

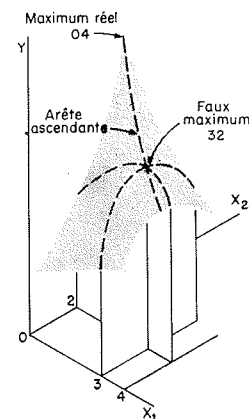
Recherche pas-à-pas



Fonction fortement unimodale.



Surface de réponse possédant une arête ascendante. Vue isométrique.



D'après D. J. Wilde

**Problèmes posés
par les erreurs expérimentales**

Il est possible de déterminer l'axe de symétrie et d'accélérer dans cette direction.

Au lieu d'établir les dérivées en plusieurs points de contours, il est plus facile d'établir la tangente contour en un seul point P_0 et ensuite d'explorer suivant une ligne ou un plan parallèle à cette tangente. Celle-ci est tracée à partir d'un point P_2 quelconque. La variation de F suivant cette ligne donne un maximum au point P_3 . Le sommet se trouve sur la ligne $P_0 P_3$, que l'on explore de manière accélérée ou unidimensionnelle.

L'intérêt est d'aboutir directement au sommet, avec :

- $2_n - 1$: recherche unidimensionnelle;
- n : mesures de tangentes au contour.

La méthode Fletcher et Powell est particulièrement intéressante lorsque la fonction critère prend une forme quadratique. Dans l'hyper-espace à n dimensions, défini par les n paramètres, la recherche dans une direction permet de définir, en rendant minimale la forme quadratique à chaque itération, un nouveau système de coordonnées orthonormé. Comme dans la méthode de Rosenbrock, on a effectué une rotation des axes, mais cette fois-ci au sens de la convergence quadratique, ce qui rapproche cette méthode de l'approximation stochastique de M. Blum.

4 Phase finale.

Il est inutile de poursuivre indéfiniment l'optimisation, alors que l'on n'est pas sûr que le modèle est adéquat et que l'effet des variations sur les paramètres n'est pas du même ordre que celui des erreurs de mesures sur les variables X et Y .

Par contre, près du sommet, courbures asymétriques et autres phénomènes non linéaires, revêtent une importance plus grande.

L'analyse de la courbure en utilisant des termes du deuxième degré de la formule de Taylor, et en faisant intervenir ou non les termes d'intersection entre paramètres, permet de déterminer par la mise sous forme de sommes de carrés si l'on est à un sommet ou à un col, non d'après une exploration unique, mais d'après un plus grand nombre d'expériences.

Ces méthodes permettent d'établir la stabilité statistique des mouvements du sommet en repartant de nouvelles données (Box).

Nous avons étudié jusqu'à présent la question de l'erreur expérimentale; or dans la plupart des situations pratiques, les mesures et les prévisions ne sont que des estimations plus ou moins précises des grandeurs considérées.

Les erreurs aléatoires d'estimation compliquent sérieusement toute recherche d'un optimum, conduisant souvent à l'échec de méthodes précédemment citées pourtant réputées comme efficaces. C'est pourquoi il importe en présence de bruit (erreur de mesure ou incertitude de prévision) de considérer la notion de convergence avant celle d'efficacité. C'est à Robbins et Monro pour la détermination de la racine d'une fonction et à Kiefer - Wolfowitz pour la recherche du maximum d'une fonction unimodale en présence de bruit, que l'on doit ces procédés d'approximation stochastique dont les idées de base sont toujours le choix d'une direction favorable et la détermination de la distance à parcourir dans cette direction.

Dans le cas de la détermination de la racine d'une fonction, et après avoir fait quelques hypothèses sur le bruit (Robbins et Monro) :

1° Valeurs des paramètres sans biais, c'est-à-dire :

$$E \{ F(a) \} = Z(a)$$

$F(a)$ étant la valeur estimée et $Z(a)$ la vraie valeur.

2° Variance

$$\sigma^2(a) = E \{ [F(a) - Z(a)]^2 \}$$

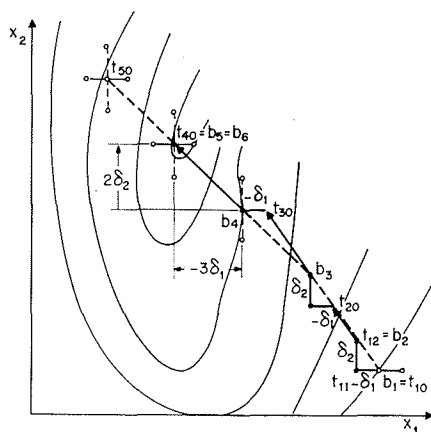
bornée, soit $\sigma^2 < M < \infty$, afin que les valeurs $F(a)$ ne s'écartent que rarement de façon sensible de $Z(a)$, la technique de Robbins et Monro consiste à placer l'expérience suivante selon :

$$a_{n+1} = a_n - \lambda_n F(a_n)$$

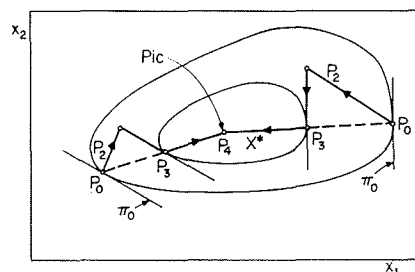
(λ_n étant le n^e élément d'une suite harmonique).

Dans le cas de la recherche d'un maximum (Kiefer-Wolfowitz) unimodale en présence de bruit, plutôt que de

Recherche d'une arête



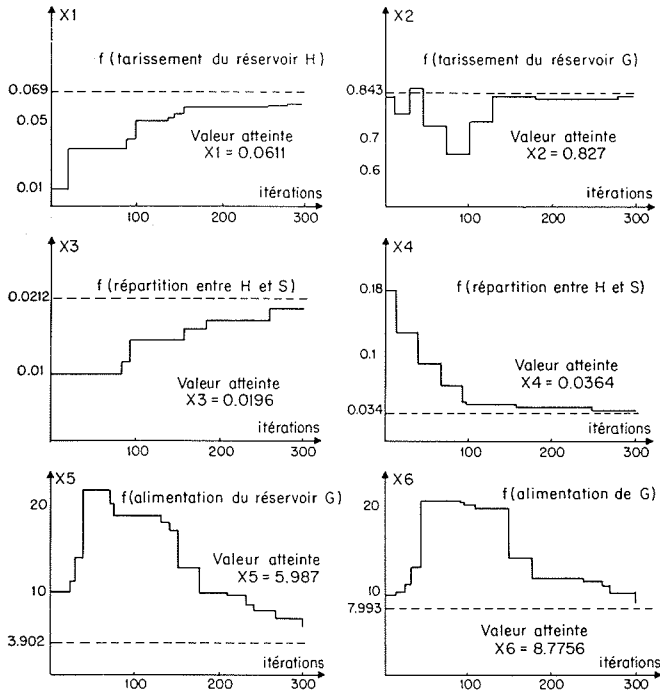
Partant dans le cas d'un critère à 2 variables indépendantes présentant des contours homothétiques.



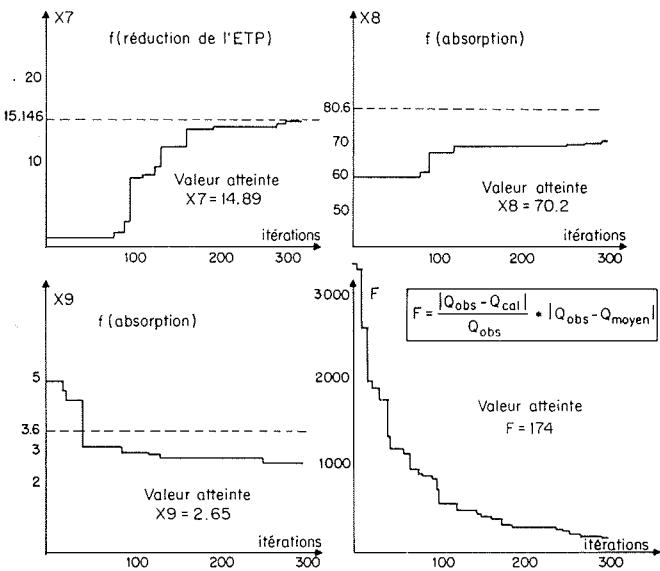
D'après D. J. Wilde

Modèle CREC

Étude de la convergence des paramètres par optimisation Rosenbrock sur un jeu simulé fictif (5 ans) à partir des pluies 1955-1959.



3/

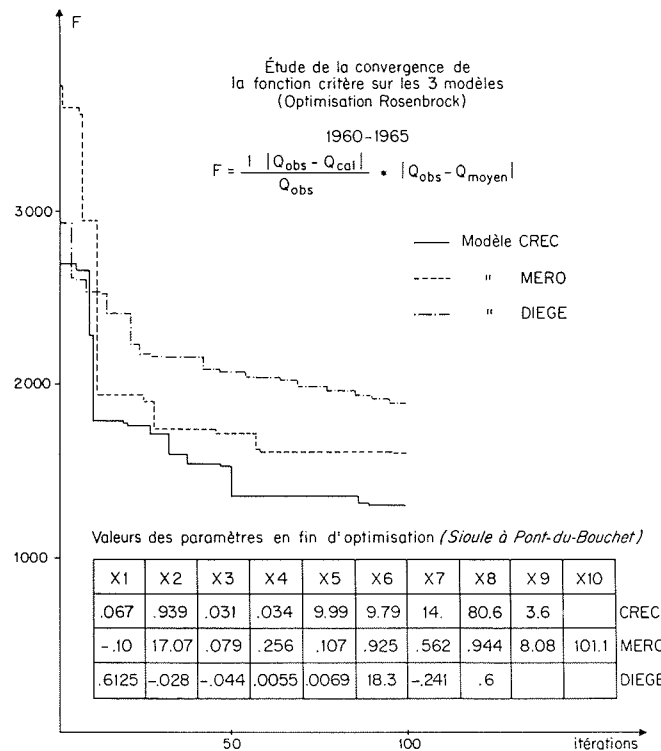


3 bis

Bilan de l'étude de la convergence des paramètres vers leur valeur vraie sur une série de 5 ans de débits simulés fictifs (à partir des pluies de 1955-1959)

| Para-mètre | Valeurs vraies | Valeurs de départ | 1° Optim. 100 itér. | 2° Optim. 100 itér. | 3° Optim. 100 itér. | Erreur relative (%) |
|------------|----------------|-------------------|---------------------|---------------------|---------------------|---------------------|
| X7 | 15.146 | 1. | 7.92 | 14.18 | 14.89 | 1.6 |
| X2 | 0.843 | 0.826 | 0.654 | 0.8233 | 0.827 | 1.8 |
| X4 | 0.0344 | 0.18 | 0.0433 | 0.0402 | 0.0364 | 5.5 |
| X3 | 0.0212 | 0.01 | 0.0151 | 0.018 | 0.0196 | 7.5 |
| X6 | 7.992 | 10. | 20.74 | 11.59 | 8.77 | 8.8 |
| X1 | 0.069 | 0.01 | 0.0502 | 0.0583 | 0.061 | 11. |
| X8 | 80.6 | 60. | 66.57 | 67.99 | 70.21 | 12.4 |
| X9 | 3.6 | 5. | 2.89 | 2.69 | 2.65 | 26.4 |
| X5 | 3.902 | 9.9 | 18.96 | 9.83 | 5.98 | 34.8 |
| F | 0. | 3339. | 613. | 303. | 174. | |

3 ter



4/

calculer la dérivée au point a_n , on calcule la pente moyenne en considérant l'intervalle :

$$a_n + \epsilon_n, a_n - \epsilon_n$$

pente qui a pour expression

$$P = (F(a_n + \epsilon_n) - F(a_n - \epsilon_n)) / 2\epsilon$$

Le signe de la pente détermine la direction favorable de recherche et le couple d'expérience suivant est placé en :

$$a_{n+1} = a_n + \lambda_n P$$

ou λ_n , longueur du pas de recherche, est le n^e élément d'une suite harmonique.

Le procédé converge vers l'optimum d'une fonction unimodale avec bruit à la fois en moyenne quadratique et avec probabilité 1, sous réserve de restrictions raisonnables sur le bruit et la fonction de régression.

Une généralisation multidimensionnelle a été faite par Blum qui effectue une recherche de sommet stochastique.

Les conditions assurant la convergence d'un procédé d'approximation stochastique sont assez souples, mais il faut, en présence d'erreurs, se contenter de probabilité à défaut de certitude. Choissant une probabilité élevée p

jugée satisfaisante (0,95 ou plus), on dit qu'il y a une probabilité p pour que le sommet appartienne à l'intervalle d'incertitude probabiliste.

Conclusion sur les méthodes

Avant de conclure que dans ce domaine comme dans bien d'autres la méthode de travail et le passage prudent et pas à pas du simple au complexe n'est pas facile à respecter spontanément mais qu'il est absolument impératif, notons que nous ne sommes pas sûrs que cette solution de régression non linéaire par algorithmes de calcul pour ordinateur digital soit la meilleure méthode adaptée à nos problèmes.

En effet, dans la mesure où le système hydrologique résulte à chaque instant d'un équilibre qui ne dépend pas seulement d'équations algébriques entre variables mais où les divers flux qui s'équilibrent sont à chaque instant réglés par des potentiels, le système est un système dynamique : les formes mathématiques qui le décrivent contiennent des équations différentielles dont le temps est une variable indépendante.

Un calage analogique du système s'impose, ne serait-ce que pour éliminer l'erreur quadratique, et les calculateurs analogiques sont bien plus performants dans ce domaine. L'appoint d'extension hybride à conversion, les éléments de décisions destinés à explorer le comportement du modèle et à en déduire une gestion optimale sont relativement simples à appliquer et à exploiter.

Nous pensons quant à nous, poursuivre essentiellement dans cette dernière voie qui a bien d'autres applications d'ailleurs que l'hydrologie ou l'économie.

Conclusion de l'application

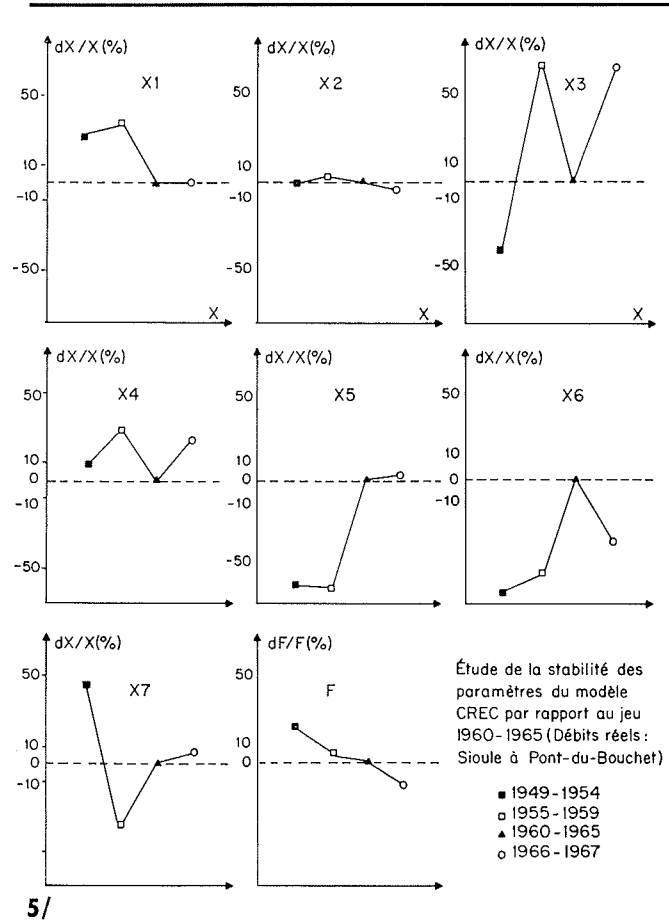
La méthode de Rosenbrock a d'abord été pratiquée depuis deux ans avec plus ou moins d'incohérence sur le bassin de la Diège (225 km²) et bien que les résultats obtenus aient été encourageants, il a semblé nécessaire de tester l'adéquation de la méthode d'optimisation et la validité du choix de la fonction critère avant d'étudier un nouveau bassin (La Sioule à Pont-du-Bouchet).

Une série chronologique de débits simulés fictifs obtenus en appliquant un jeu déterminé de paramètres à une série chronologique de pluie (1955-1959) a permis de constater la réelle convergence des paramètres vers leur vraie valeur (graphique n° 3, 3 bis, 3 ter.).

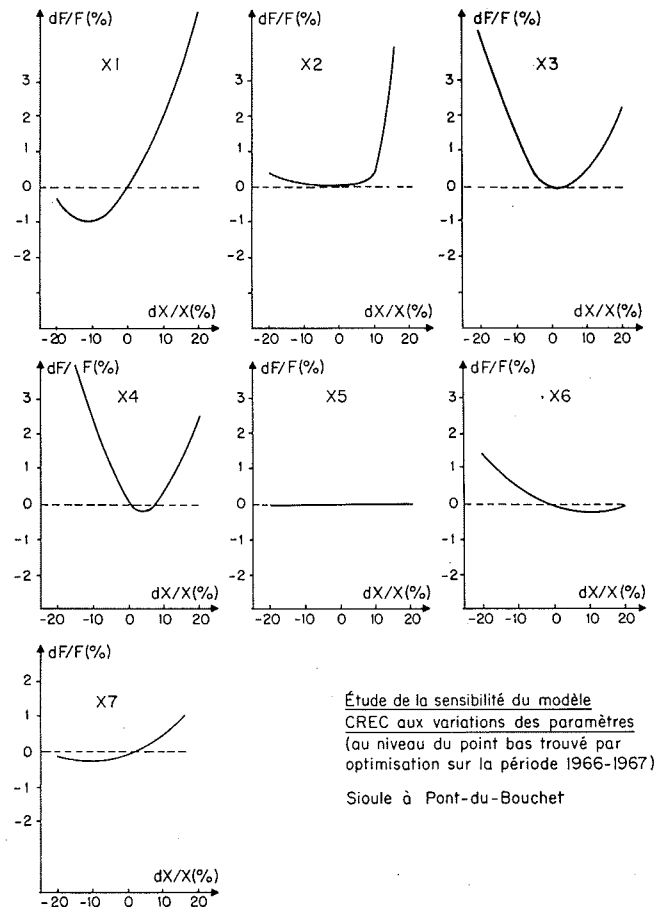
Des essais concernant le pas de variation des paramètres (pas identique ou proportionnel, changement d'unité des paramètres) ont permis de vérifier qu'au bout d'un certain nombre d'essais, les résultats obtenus sont identiques, ce qui est caractéristique de la méthode.

Il restait à tester l'adéquation et à départager le comportement des trois modèles (Crec, Mero et un modèle plus simple élaboré sur la Diège). Les deux bassins, peu éloignés, étaient *a priori* assez semblables : le classement relatif quant à la vitesse de décroissance de la fonction critère F et au niveau de sa stabilisation est le même. Cela peut tenir essentiellement au nombre plus ou moins grand de paramètres dans chaque modèle (graphique n° 4).

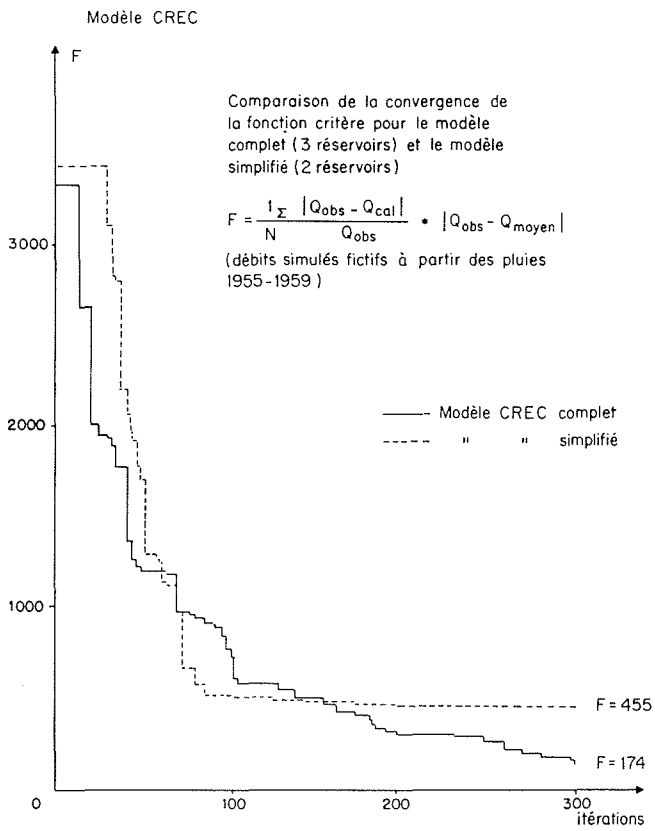
Sur le modèle Crec (exposé par M. Roche), l'étude de la stabilité des paramètres selon le choix de la période de calage (graphique n° 5) et de la sensibilité du modèle aux variations de ces paramètres (graphique n° 6) a permis



5/



6/



7/

une simplification du modèle par élimination du réservoir G dont l'alimentation est déterminée par la valeur des paramètres X 5 et X 6, peu stables et de peu d'influence sur le comportement du modèle (graphique n° 7).

La stabilité des caractéristiques des séries reconstituées à partir de la même série de pluie et d'évapotranspiration

— au niveau des moyennes et des écarts-type mensuels (graphique n° 8 et 9);

— au niveau des ajustements des lois de distribution du débit d'un mois déterminé, des débits journaliers maximaux annuels et comparativement à un ajustement graphique sur les cinq années disponibles (période de calage du modèle) (graphique n° 10 et 11), montre que l'on tire parti de manière plus performante de la série pluies-débits de cinq ans en calant le modèle sur cette période et en l'extrapolant plutôt que d'utiliser sur ces mêmes cinq années une méthode classique de corrélation du type Amélie (ceci en réponse au texte de M. Bernier).

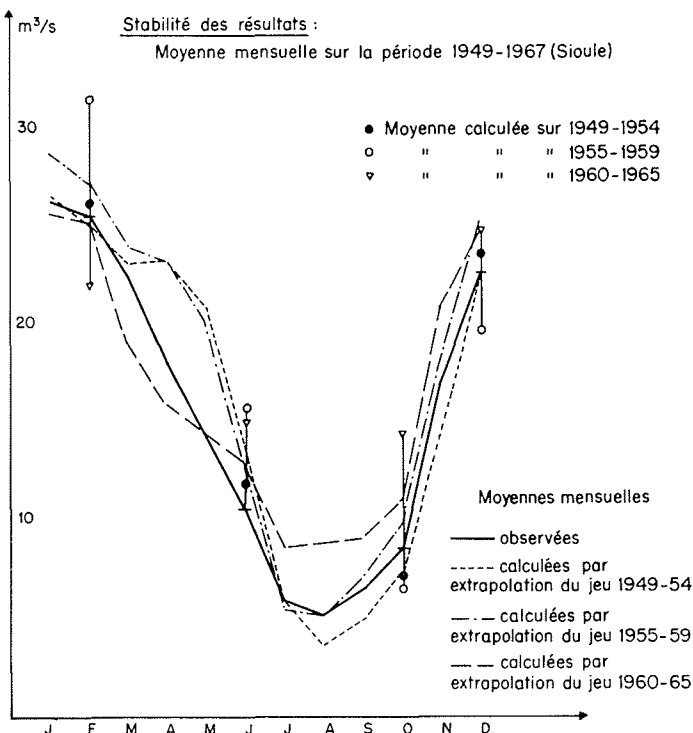
Nous nous rendons bien compte maintenant :

1° Que nous n'avons pas traité de manière assez indépendante les problèmes d'adéquation et les problèmes des méthodes de convergence.

2° Qu'une tactique d'optimisation différente doit être appliquée à chaque phase de l'optimisation et que la mise au point des premières phases doit se faire sur un modèle parfaitement adéquat (en générant des débits fictifs à partir des pluies et d'un jeu donné de paramètres).

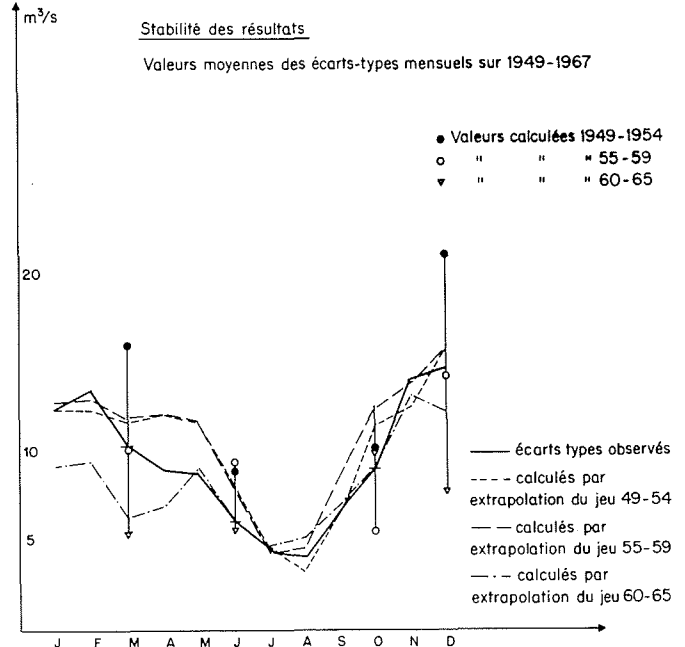
3° Que la dernière phase ne peut se définir que compte tenu du bruit aléatoire et de la détermination expérimentale à partir d'erreur sur les variables $X_1 \dots X_j$, de la déformation de la fonction F pour le jeu de paramètres

Modèle CREC : Sioule à Pont-du-Bouchet



8/

Modèle CREC : Sioule à Pont-du-Bouchet



9/

- FLETCHER (R.) and POWELL (M.J.D.). — "A rapidly convergent descent method for minimization", *Comp. J.*, 6, 2 (1963).
- ROSENBROCK (H.H.). — "An automatic method for finding the greatest or least value of a function", *Comp. J.*, 3, 3 (October 1960).
- HAMON. — *A.I.H.S.*, Assemblée de Berkeley. Publication n° 63 (1964), p. 52.

- HOLTAN. — International Hydrology Symposium, Fort-Collins (1968), p. 220.
- KIEFER-WOLFOWITZ. — Stochastic estimation of the maximum of a regression function, *Annals of Math. Stat.*, 23 (1952).
- BLUM (J.R.). — Approximation Methods which converge with probability one, *Annals of Math. Stat.*, 25 (1954).

Discussion

Président : M. J. RODIER

M. le Président remercie M. CORMARY pour son intéressant exposé où la technique du spécialiste des modèles s'allie si heureusement à sa remarquable expérience « d'hydrologue de terrain ».

M. ROCHE intervient pour souligner l'importance du choix du critère d'ajustement lequel peut influencer sur l'efficacité du réglage recherché. Si l'on choisit comme critère une fonction de la différence $Q_o - Q_c$ entre débits observés et calculés, il n'est pas sûr qu'une fonction quadratique — telle que celle employée dans la méthode des moindres carrés — soit la mieux adaptée au contrôle de la qualité de l'ajustement. Ainsi, s'il existe un simple décalage dans le temps entre un hydrogramme observé et un hydrogramme calculé — par ailleurs exactement superposables — un critère basé sur une fonction de $Q - Q_o$ conduira généralement à une conclusion très pessimiste sur l'adéquation du modèle alors que le « décalage dans le temps » qui conditionne celle-ci peut ne pas être gênant, en pratique, eu égard au but poursuivi.

M. CORMARY exprime son plein accord sur ce point qu'il n'a pas pu développer faute de temps. On peut être amené à changer la fonction-critère suivant la phase d'optimisation et suivant l'objectif final recherché. La « fonction-critère » la plus adéquate est

celle qui pénalise le plus fortement les écarts jugés les plus indésirables.

Combien de paramètres avez-vous réglé simultanément dans les applications que vous avez réalisées? demande M. DUJARDIN à M. CORMARY.

Dans les trois modèles étudiées, répond ce dernier, il y avait respectivement cinq, dix et quatre paramètres à régler; c'est le modèle à cinq paramètres qui s'est montré le plus maniable.

A la S.O.G.R.E.A.H., dit M. CAYLA, sur des modèles à six paramètres, nous tentons de régler ces derniers en adoptant *successivement* plusieurs critères; notre but est de faire réaliser ce réglage automatiquement par l'ordinateur.

M. CORMARY a procédé à quelques expériences du même genre; on arrive généralement au résultat plus rapidement mais dans certains cas il apparaît une instabilité plus forte.

Plus personne ne demandant la parole, M. RODIER clôt la discussion en remerciant tous ceux qui y ont participé.

M. RODIER passe la présidence à M. CASEAU pour le reste de la séance. Ce dernier donne immédiatement la parole à M. BERNIER pour l'exposé de sa communication.

