

Analyse de fiabilité en modélisation hydrologique : concepts et applications au modèle pluies-débits GR3

Reliability analysis in hydrological modelling : concepts and applications to the GR3 rainfall-runoff model

X. YANG^{1,2*} et E. PARENT²

Reçu le 21 octobre 1994, accepté le 6 juin 1995*.

SUMMARY

Reliability analysis is a method for evaluating the risk of failure of a system and this concept is widely used in structure design. The theory can be conveniently transposed to the field of water sciences so as to evaluate the quality of hydrologic models. Hydrological reliability of a model is defined as the probability that there will not be significant different performances between this model, with parameters assessed on the basis of a large sample including various hydrological conditions, and the same model running with ad hoc parameters estimated from the test sample. Consequently a tolerance threshold has to be introduced in reliability analysis for hydrological models. A number of conceptual differences between validation and reliability analysis are reviewed and categorized.

This paper shows how to adapt the concepts of reliability to hydrology and gives details of the first two moment technique that can be implemented to compute the reliability of a rainfall runoff model. This technique is based on a Taylor series developed in the neighborhood of the model working point. This simplification allows for an explicit expression of the expectation and the variance of the hydrological model performance as functions of the marginal sensitivity of the criterion and the parameter uncertainties. A Student ratio can be easily computed to evaluate the quality of a hydrological model from a reliability-based point of view.

The GR3 model is used to illustrate this technique. A reliability analysis of this three parameter rainfall-runoff model gives new insight into the performance of the model for short-term flood prediction in the Réal Collobrier river basin located in the French Côte d'Azur. Numerical results show that the reliability analysis approach is very simple and of practical interest.

-
1. Department of water resources, IWHR, 1A Fuxing Rd., P.O. Box 366, Beijing 100038, P.R. China.
 2. Ecole Nationale du Génie Rural, des Eaux et des Forêts, 19 Avenue du Maine, 75732 Paris Cédex 15, France.

* Les commentaires seront reçus jusqu'au 30 août 1996.

** Communication présentée au Colloque International sur la « Matière Organique Biodégradable ». Ecole Polytechnique de Montréal, juin 1994.

For 16 of the 18 samples used for validation, the GR3 rainfall runoff model proved to be reliable on the basis of the three criteria used in this paper to evaluate the quality of a hydrological model (general quadratic deviation loss of runoffs, flood volume error and flood peak error).

Using reliability analysis, contributions of each parameter to the variance of the performance function can be conveniently identified. For the GR3 model and the Réal Collobrier case study, the impacts of each parameter on the model efficiency are as follows:

– The contribution of a parameter varies from one criterion to another. This indicates that a parameter originally designed to play a specific role in the modelling of the rainfall runoff transformation process does contribute to the criterion that best describes this function, which confirms the conceptual « good sense » of GR3.

– For each criterion, the cross contributions are very low: the GR3 model satisfies the principle of parsimony and each parameter adds something different to the general behavior of GR3.

In this case study, the GR3 model exhibits a good balance between complexity and quality for hydrological models.

Finally, reliability analysis can provide a multicriteria point of view of the model qualities by changing the various indices of performance. Extension of reliability analysis towards engineering risk would include the estimation of joint probability distributions for these multiple criteria of interest.

Key words : *uncertainty, risk, reliability, hydrological modelling, rainfall-runoff model, model validation.*

RÉSUMÉ

Pour étudier les incertitudes d'un modèle hydrologique, on peut employer l'analyse de fiabilité, méthode couramment utilisée dans le domaine de la résistance des structures et du génie hydraulique. Cet outil peut être transposé dans le domaine hydrologique pour juger de la qualité d'un modèle. Un modèle hydrologique sera fiable si : sur une série de données test, l'ajustement obtenu avec un jeu de paramètres recommandé pour un fonctionnement passe-partout est, selon toute probabilité, d'une qualité proche de l'ajustement idéal correspondant au calage du jeu de paramètres sur l'échantillon test.

Ce papier montre comment adapter chacun des concepts de l'analyse de fiabilité en hydrologie et détaille la technique des deux premiers moments afin de calculer explicitement la fiabilité d'un modèle-pluie débit en réalisant un développement limité au voisinage du point de fonctionnement du modèle.

Cette approche par analyse de fiabilité est appliquée au modèle GR3 à titre d'illustration pour juger de la pertinence de ce modèle pluie-débit conceptuel à trois paramètres en situation de prédiction des crues à court pas de temps sur un bassin versant situé dans la région de la Côte d'Azur en France.

Les résultats numériques obtenus montrent le caractère opérationnel de cette approche très simple. D'autre part, ces calculs de fiabilité mettent en évidence la réponse du modèle à chacun des paramètres.

Enfin en mesurant la performance de représentativité d'un modèle selon plusieurs dimensions telles l'erreur quadratique d'ajustement, l'erreur sur le volume de la crue prévue ou l'importance de l'écart de pointe de crue, l'analyse de fiabilité peut être naturellement étendue vers une approche multicritère en considérant des probabilités conjointes de satisfaction du modèle sur chacun de ces critères.

Mots clés : *incertitude, risque, fiabilité, modélisation hydrologique, modèle pluie-débit, validation de modèle.*

1 – INTRODUCTION

La fiabilité est une mesure de l'adéquation entre le comportement d'un système et le rôle que l'ingénieur lui a assigné. Depuis plus de 30 ans, les méthodes numériques de l'analyse de fiabilité sont développées par les ingénieurs spécialisés dans le domaine de la sécurité des structures. Des investigations spécifiques aux sciences de l'eau ont été aussi menées autour de cette notion d'ingénierie. YEN (1986), ainsi que PLATE et DUCKSTEIN (1988) ont posé les principes de l'analyse de fiabilité pour le domaine du génie hydraulique. L'analyse de fiabilité offre aussi des perspectives intéressantes dans le domaine de la modélisation hydrologique, et plus généralement dans l'ingénierie des ressources en eau. Les cas d'application rapportés dans les Comptes-Rendus de DUCKSTEIN et PLATE (1987), GANOU LIS (1991) et DUCKSTEIN et PARENT (1994), montrent le caractère opérationnel du concept de fiabilité dans les sciences de l'eau.

Pour ce qui concerne la modélisation hydrologique, trois sortes de techniques de calcul de fiabilité paraissent potentiellement utilisables du point de vue numérique : l'intégration directe, les simulations de type Monte-Carlo, et la méthode du premier ordre et des premiers moments (POPM). WOOD (1976) a utilisé l'intégration directe pour analyser les effets des incertitudes d'un paramètre dans un modèle pluie-débit simple sur les fréquences des crues générées par des pluies mesurées. Dans le même esprit, DAVIS et NNAJI (1982) ont étudié la fiabilité de la prédétermination des crues issues d'un modèle pluie-débit sur de petits bassins versants. Pour cette même méthode, TUNG et MAYS (1980, 1981) ont proposé des innovations intéressantes : ces auteurs exploitent les relations entre les variables fondamentales du système et les fonctions de performance du système afin de trouver les fonctions de distribution.

Pour les techniques de type Monte-Carlo, GAREN et BURGESS (1981) ont présenté une application de la simulation pour analyser les incertitudes existant dans un système de prévision des crues fondé sur le modèle de Stanford.

Quant à la méthode POPM, il faut citer les travaux de YEN et TANG (1977), relatifs à l'analyse des incertitudes dans un modèle de propagation de crue en vue d'évaluer la fiabilité de ce modèle pour la prévision des crues. MELCHING et YEN (1986) ont utilisé la méthode POPM pour analyser la fiabilité d'un réseau de collecte des eaux pluviales. MELCHING *et al.* (1990, 1991) ont exposé des applications de cette méthode à la fiabilité de la prédétermination de crue et à la sélection des modèles hydrologiques.

Le choix d'une technique d'évaluation de la fiabilité est un problème délicat.

Pour utiliser l'intégration directe il faut trouver la fonction de densité de probabilité décrivant l'incertitude associée aux paramètres d'un modèle hydrologique sous une forme explicite. Si ces fonctions de distribution sont bien connues, le risque de défaillance du système évalué par l'intégration directe est théoriquement exact, ce qui représente le grand avantage de cette méthode. Malheureusement, ces fonctions de distribution sont en réalité rarement connues.

La simulation par techniques de type Monte-Carlo est une méthode beaucoup plus souple. Utilisée pour de nombreux problèmes, elle est souvent prise comme une méthode passe-partout. En dépit de sa souplesse, cette méthode n'est pas très recommandée pour l'analyse du risque d'un système : en effet la valeur

estimée du risque fournie par cette technique est elle-même une grandeur aléatoire dépendant de la taille de l'échantillon et du nombre de tirages aléatoires.

Notre contribution développe une application possible de l'analyse de fiabilité à la modélisation hydrologique.

Nous avons retenu la méthode POPM dans le cadre d'une application à la modélisation hydrologique, pour les raisons suivantes :

- 1) elle est relativement simple dans sa formulation ;
- 2) elle permet d'utiliser autant de paramètres que nécessaire ;
- 3) le besoin en données est assez modéré.

Elle nécessite seulement des estimations de l'espérance et de l'écart-type des variables, mais pas de leur distribution, bien qu'elle puisse utiliser les informations sur la distribution si celles-ci sont disponibles.

Nous ne présenterons les méthodes d'intégration directe et de Monte-Carlo qu'à titre de référence.

Nous définirons d'abord la fiabilité pour un modèle hydrologique. Nous indiquerons ensuite les traits principaux de la technique POPM pour le calcul pratique de cette fiabilité. Dans la section 4, nous détaillerons la correspondance entre les concepts de fiabilité et les grandeurs manipulées lors du calage et l'utilisation en mode prédictif d'un modèle hydrologique. Nous avons utilisé le modèle GR3 (EDITJATNO et MICHEL, 1989) comme exemple d'application pour cette analyse de fiabilité. Les résultats obtenus permettent de porter un jugement de la qualité globale de ce modèle quand il est utilisé pour la prédétermination des crues et pour leur prévision en temps réel.

2 – FIABILITÉ D'UN MODÈLE HYDROLOGIQUE

Nous soulignerons ici les différences entre le problème de la résistance des structures (par exemple celles d'un système hydraulique décrit par YEN en 1986) et celui de la modélisation hydrologique. Il nous faut donc définir la fiabilité d'un modèle hydrologique et ses liens avec les techniques classiques de calage et de validation de modèles.

2.1 Rappel de la définition usuelle de la fiabilité

Rappelons la définition usuelle de fiabilité pour l'analyse d'un système hydraulique. Selon YEN *et al.* (1986), la défaillance du système est l'événement pour lequel la charge L , agissant sur le système (par exemple une demande pour l'irrigation), dépasse la capacité (ou la résistance) R du système (telle la ressource prélevable dans la rivière) avant que le système n'atteigne l'objectif voulu. Le risque peut être défini comme la probabilité de défaillance du système P_r .

Dans ce cas, il s'agit d'un risque de non-satisfaction (de l'irrigation). On peut écrire :

$$P_r = \text{Prob}(L > R) \quad (1)$$

La fiabilité est la probabilité de l'événement complémentaire :

$$P_f = 1 - P_r = \text{Prob}(L < R) \quad (2)$$

En introduisant la fonction de sécurité, Z , par exemple $Z = R - L$ ou $Z = \ln(R/L)$, on peut réécrire l'équation 2 sous la forme :

$$P_f = \text{Prob}(Z \geq 0) \quad (3)$$

2.2 Qu'est-ce que la fiabilité d'un modèle hydrologique ?

La fiabilité d'un modèle hydrologique dépend de ce que l'on attend du modèle. Nous nous intéresserons à la qualité du calage d'un modèle pluie-débit que l'on veut utiliser en crue, à court pas de temps. Il est d'usage de diviser les données d'épisodes de crues en deux parties : échantillon de calage et échantillon de validation. On juge de la qualité du modèle suivant les différences de calage correspondant à ces deux échantillons. La qualité de ce calage est mesurée par une fonction de performance traduisant l'écart entre les données réelles de débit et celles générées par le modèle à partir des pluies, par exemple en utilisant la moyenne des écarts quadratiques. Plus ces différences sont faibles, plus le modèle est fiable.

On prend en compte ici deux calages particuliers, à savoir le calage sur l'ensemble des épisodes de crues de l'échantillon de calage (appelé calage global) et le calage sur un épisode de l'échantillon de validation (appelé calage par épisode). La différence entre ces deux calages caractérise la capacité d'extrapolation de ce modèle.

La procédure adoptée ici est la suivante :

1) On effectue un calage global en estimant les coefficients x_g du modèle s'ajustant au mieux à l'ensemble des épisodes de crues de l'échantillon de calage, au sens d'une fonction de performance $G(x_g)$.

2) On prend un sous-ensemble des épisodes de l'échantillon de validation (par exemple un épisode unique et complet), et l'on réalise le calage local du même modèle. Cela fournit un jeu de paramètres x_c pour chaque épisode de l'échantillon de validation.

La comparaison des écarts entre données de l'échantillon test et valeurs prévues par le modèle utilisant le jeu de paramètres x_g ou x_c peut être utilisée comme mesure de la distance entre x_g et x_c .

Pour définir la fiabilité d'un modèle définissons $L = G(x_g)$, $R = G(x_c)$ et $Z = L - R$, où $G(\cdot)$ est une fonction mesurant la performance du modèle pour un jeu de paramètres donné (les détails des calculs seront développés dans la section 4).

Dans le cas de l'analyse de fiabilité pour un modèle hydrologique, on introduit alors un seuil ϵ de tolérance sur l'écart de performance et l'on définit les grandeurs suivantes :

$$\text{le risque } P_r(\epsilon) = \text{Prob}(|L - R| > \epsilon) \quad (4)$$

$$\text{la fiabilité} = P_f(\epsilon) = 1 - P_r = \text{Prob}(|L - R| \leq \epsilon) \quad (5)$$

$$P_f(\epsilon) = \text{Prob}(|Z| \leq \epsilon) \quad (6)$$

L'équation 5 définissant la fiabilité s'interprète comme la probabilité que l'écart de performance entre les deux calages définis par x_g et x_c ne soit pas sensible (car inférieur au seuil de tolérance ϵ). Les équations 4, 5 et 6 sont spécifiques à la

fiabilité d'un modèle en hydrologie et correspondent respectivement aux équations 1, 2 et 3 de la définition générale.

La définition précédente de la fiabilité d'un modèle hydrologique concerne le test de deux capacités importantes d'un modèle de simulation :

1) Dans l'absolu, si x_g et x_c ne jouent pas un rôle particulier de référence, on compare simplement la capacité à prendre en compte, à la fois le caractère universonnel (jeu x_g), et le caractère spécifique des événements de crues (jeu x_c).

2) En termes relatifs, puisque x_c est le jeu de paramètres a priori le plus acceptable pour l'épisode de validation considéré pour le calcul de performance, on analyse la capacité d'extrapolation de x_g au-delà de l'échantillon qui a été utilisé pour le caler.

2.3 Validation et fiabilité

On peut présenter l'analyse de fiabilité comme un prolongement de la validation d'un modèle de simulation. Néanmoins l'analyse de fiabilité est, stricto sensu, différente du travail de validation. En décrivant qualitativement par « bons » ou « mauvais » les résultats possibles issus de simulations d'un épisode de crue de l'échantillon de validation avec les jeux de paramètres x_g et de x_c , on peut juger de la fiabilité du modèle comme le montre le tableau 1.

Tableau 1 Différence entre la fiabilité et la validation d'un modèle.

Table 1 Difference between reliability and validation for a model.

Simulation avec X_c	Simulation avec X_g	
	Bons	Mauvais
Bons	(1) fiable et valide	(2) non fiable et non valide
Mauvais	(3) non-fiable et valide	(4) fiable et non valide

On remarque que pour un tel épisode c'est le jeu de paramètres x_c qui devrait être a priori le mieux adapté. Pour une analyse de validation, on retient les résultats des cas (1) et (3), mais pour une analyse de fiabilité le modèle est considéré fiable dans les cas (1) et (4).

Bien sûr, le cas (3) est rarissime parce que le calage par épisode est toujours meilleur qu'un calage global sauf optimum secondaire piégeant la procédure de recherche d'optimum. Quant au cas (4), considéré non valide, il relève probablement d'un problème d'extrapolation, mais il est néanmoins reconnu acceptable par l'analyse de fiabilité du fait que les résultats issus de x_g possèdent la même précision que ceux issus de x_c , c'est-à-dire que les jeux x_g et x_c sont comparables en terme de capacité du modèle pour cet épisode de crue.

En général, le cas (3) est provoqué par la non-efficacité de la méthode d'optimisation utilisée pour le calage, mais pas par un problème d'extrapolation du modèle. Le cas (4) est dû probablement à des raisons hors modèle : il nous semble donc peu raisonnable de rejeter le modèle dans le cas (4) et de le confirmer dans le cas (3). Dans ce cas 4, la conception même du modèle peut être à rejeter, car son comportement hydrologique n'est pas valable par exemple vis à vis de propriétés de conservation des masses de la pluie transformées en débit, néanmoins il peut être fiable au sens de la définition proposée (c'est-à-dire stricto sensu une distance faible entre les résultats fournis par x_c et x_g).

En dernier lieu, le fait que le modèle soit fiable implique simplement que x_g n'est pas très loin de x_c . Du point de vue de l'hydrologue, cette proximité nous renseigne sur la possibilité d'ajuster globalement, et une fois pour toute, les paramètres d'un modèle de prévision des crues en temps réel.

3 - LA MÉTHODE DU PREMIER ORDRE ET DES PREMIERS MOMENTS (POPM)

BENJAMIN et CORNELL (1970) ont proposé la méthode du premier ordre pour évaluer la performance des modèles hydrologiques. Considérons une relation simple entre la fonction de performance Z et les paramètres du modèle notés ici x :

$$Z = G(x) \quad (7)$$

Développons $G(x)$ avec la série de Taylor autour de x_0 , point de fonctionnement de référence :

$$Z = G(x_0) + \frac{dG}{dx_0}(x - x_0) + \frac{1}{2!} \frac{d^2G}{dx_0^2}(x - x_0)^2 + \dots \quad (8)$$

où dG/dx_0 et d^2G/dx_0^2 sont respectivement les valeurs de dG/dx et d^2G/dx^2 évaluées en x_0 . Si l'on se contente d'une approximation linéaire au premier ordre, l'équation 8 devient simplement :

$$Z \approx G(x_0) + \frac{dG}{dx_0}(x - x_0) \quad (9)$$

Cette méthode n'est donc acceptable que pour les fonctions faiblement non-linéaires, si bien que l'erreur d'extrapolation n'affecte pas trop les analyses fondées sur l'équation 8.

De plus, à partir de l'équation 8, on peut obtenir les estimations de l'espérance et de la variance de Z :

$$\bar{Z} \approx G(x_0) + \frac{dG}{dx_0}(E(x) - x_0) \quad (10)$$

$$sz^2 \approx \left(\frac{dG}{dx_0}\right)^2 \text{VAR}(x) \quad (11)$$

où $E(\cdot)$ et $\text{VAR}(\cdot)$ sont respectivement les opérateurs d'espérance et de variance.

Dans un cas multivariable, on a une extension directe de l'équation 8 sous la forme :

$$Z \approx G(\underline{x}_0) + \sum_{i=1}^n D_i(x_i - x_{i0}) \quad (12)$$

où D_i représente la dérivée partielle $\partial G/\partial x_i$ évaluée à x_{i0} . On peut obtenir les estimations de l'espérance mathématique et de la variance de Z :

$$\bar{Z} \approx G(\underline{x}_0) + \sum_{i=1}^n D_i(E(x)_i - x_{i0}) \quad (13)$$

$$s_z^2 = \sum_{i=1}^n \sum_{j=1}^n D_i D_j \text{COV}(x_i, x_j) \quad (14)$$

où COV est l'opérateur de variance-covariance.

Dans la technique dite des « premiers moments », on applique les estimations de l'espérance et de l'écart-type de Z pour analyser la fiabilité du système, définie selon le rapport :

$$\beta = \frac{\bar{Z}}{s_z} \quad (15)$$

où \bar{Z} et s_z sont respectivement les estimations de l'espérance et de l'écart-type de Z. D'après les équations 13 et 14, on a donc :

$$\beta = \frac{Gx_0 + \sum_{i=1}^n D_i (\bar{x}_i - x_{i0})}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n D_i D_j \text{COV}(x_i, x_j)}} \quad (16)$$

où β est défini comme l'indice de fiabilité de système. Il est en fait l'inverse du coefficient de variation de Z c'est-à-dire un rapport de Student pour la variable Z. Cet indice sert à comparer l'écart de \bar{Z} à l'origine des Z par rapport à s_z (YEN, 1987), ($Z = 0$ est le point critique de défaillance de système). Comme l'exprime l'équation 16, la fiabilité d'un système peut être représentée comme une fonction des variations des variables fondamentales $\text{VAR}(x_i)$ et des facteurs de sensibilité (D_i). Pour un modèle hydrologique, l'indice de fiabilité fournit un aperçu global sur les effets des incertitudes affectant l'ensemble des paramètres du modèle.

4 – DÉFINITIONS DES FONCTIONS DE CHARGE, DE RÉSISTANCE ET DE SÉCURITÉ D'UN MODÈLE HYDROLOGIQUE

Pour un modèle avec un paramètre x , la figure 1 montre deux courbes correspondant respectivement à la surface de réponse du calage global sur l'échantillon de calage (courbe E), et à la surface de réponse du calage sur un épisode quelconque de l'échantillon de validation (courbe F). Ce sont les surfaces étudiées par l'analyse de fiabilité. x_g et x_c ont les mêmes définitions que précédemment. x_g est l'abscisse de l'optimum pour la courbe E, représenté par le point H ; x_c est l'abscisse de l'optimum sur la courbe F (point N) ; M est le point d'abscisse x_g sur la courbe F. Notre analyse portera essentiellement sur les points M et N.

Considérons les points M et N, la fonction de charge L est définie comme le développement au premier ordre de $G(x)$ en M tandis que la fonction de résistance sera le développement de $G(x)$ en N. L et R sont considérées comme des variables aléatoires possédant des distributions propres. Si on appelle $f_{R,L}(r,l)$ la fonction de distribution de probabilité conjointe de L et R (cf. les équations 4, 5 et 6), et si le domaine de fiabilité ϵ , dans le domaine des variables L et R, est le pavé défini par $[l_1, l_2] \times [r_1, r_2]$, ce qui correspond pour la variable $Z = L - R$ à l'intervalle

$[\varepsilon_1 = l_1 - r_2, \varepsilon_2 = l_2 - r_1]$, la fiabilité du modèle peut être calculée théoriquement par :

$$P_f = \int_{l_1 r_1}^{l_2 r_2} f_{R,L}(r, l) dr dl = \int_{\varepsilon_1}^{\varepsilon_2} f_z(z) dz \quad (17)$$

avec : $\int_{l_1-z}^{l_2-z} f_{R,L}(t+z, t) dt = f_z(z)$

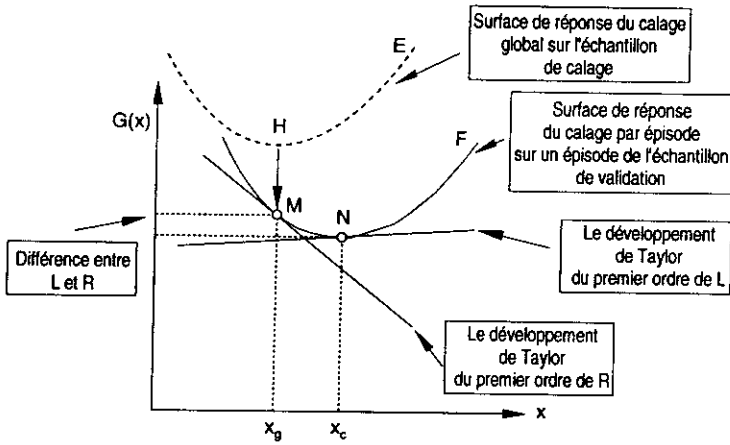


Figure 1 Description de l'analyse de fiabilité pour un modèle de simulation.
Description of reliability analysis for a simulation model.

Il faut noter les différences entre les définitions de fiabilité utilisées dans cette section et celles rencontrées en génie hydraulique. Elles sont résumées dans le tableau 2. Le risque associé à un modèle hydrologique mesure ici la probabilité que la performance de ce modèle avec un paramètre estimé sur un seul épisode soit très différente de la performance du modèle calé à partir de tous les épisodes de crues.

Tableau 2 Comparaison des différentes significations de l'analyse de fiabilité en génie hydraulique et en modélisation hydrologique.

Table 2 Comparison of terms used in reliability analysis for hydraulic engineering and hydrological modelling.

Pour l'analyse d'une structure	Pour l'analyse d'un modèle
$Z = L - R, Z = 0$ est le point de défaillance	$Z = L - R, Z = 0$ est le point de réussite
$P_f(L < R) = P(Z > 0)$: la fiabilité de système	$P_f(IL - Ri \leq \varepsilon) = P(Z \leq \varepsilon)$: la fiabilité du modèle
$P_f(L > R) = P(Z < 0)$: la probabilité de défaillance du système ou risque	$P_f(IL - Ri > \varepsilon) = P(Z > \varepsilon)$: le risque associé au modèle

5 – APPLICATION AU MODÈLE GR3

Le modèle GR3 (EDIJATNO et MICHEL, 1989) est couramment utilisé en France pour représenter la transformation pluie-débit. Une brève description de ce modèle et des données utilisées sera d'abord donnée dans cette partie. Les fonctions de charge et de résistance sont ensuite définies telles que dans le §4 et nous choisissons une fonction de sécurité de la forme $Z = L - R$ après avoir donné trois hypothèses sur ces fonctions. Enfin on cherche à savoir, sur un exemple réel, si Z peut avoir une valeur moyenne nulle par un test d'hypothèse, c'est-à-dire si $E(L) = E(R)$ est acceptable, selon la définition de Z . \bar{Z} et s_z ont été évaluées par les équations 13 et 14 d'après les estimations des moyennes et des variances de R et L fournies par la méthode POPM.

5.1 Modèle GR3

On peut citer la description du modèle GR3 faite par EDIJATNO et MICHEL (1989) :

« L'architecture du modèle GR3 repose sur deux réservoirs et un hydrogramme unitaire :

– le réservoir-sol ayant comme seule sortie le prélèvement occasionné par l'évaporation potentielle. Ce réservoir commande la répartition de la pluie nette entre lui-même et le sous-modèle de routage ;

– un hydrogramme unitaire, décrivant la propagation des débits depuis leur formation à partir des pluies nettes jusqu'au deuxième réservoir qui est :

– le réservoir-eau-gravitaire. Ce réservoir reçoit les débits routés selon l'hydrogramme unitaire. Il a comme seule sortie le débit de la rivière et sa loi vidange est du type quadratique.

Cet ensemble de trois opérateurs dont chacun dépend d'un seul paramètre semble être le schéma le plus simple pour rendre compte de façon acceptable de la transformation pluie-débit. »

Le fonctionnement du modèle GR3 est schématisé sur la figure 2, où A, B, C sont les trois paramètres du modèle. Les notations suivantes sont employées : S = niveau du réservoir-sol, R = niveau du réservoir-profond, P = pluie, E = évaporation, EN = évaporation efficace, PN = pluie nette, ES = évaporation du sol, PS = pluie entrant dans le sol, le coefficient K vaut le rapport S/A, PA = pluie d'attente, QP = écoulement pseudo-direct, QR = débit sortant du réservoir-R et $Q = QP + QR$.

5.2 Description des échantillons de calage et de validation

L'étude porte sur le bassin versant du Réal Collobrier d'une superficie de 70,6 km². Celui-ci se situe sur la façade littorale de la Méditerranée, à l'extrémité Ouest du massif montagneux des Maures, à une dizaine de kilomètres au Nord de la rade d'Hyères et à vingt kilomètres au Nord-Est de Toulon, en France. On dispose d'épisodes de crues mesurés toutes les 3 heures dont 26 épisodes constituent l'échantillon de calage (de durées comprises entre 118 et 651 heures pour un débit maximum de 2.4 à 96.4 m³/s). 18 crues forment l'échantillon de validation dont les caractéristiques sont détaillées dans les premières colonnes du tableau 3.

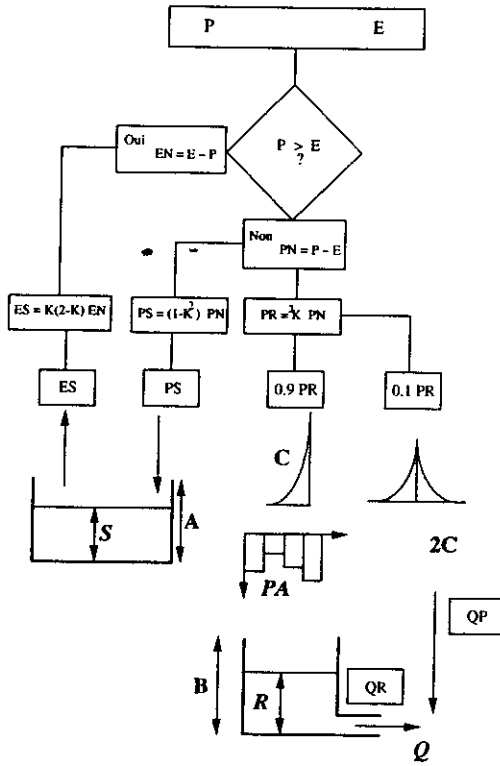


Figure 2 Schéma du modèle pluies débit GR3.
Schematic view of GR3 rainfall-runoff model.

Tableau 3 Valeurs de t de Student pour chaque épisode de validation.

Table 3 *T values (student test) for each validation input series.*

N° d'épisode	Durée (pas de temps)	Débit maximal	δ selon G_1	δ selon G_2	δ selon G_3
6 804	36	127	0,91	-1,76	-1,41
7 010	185	25,5	0,16	0,75	0,80
7 215	222	62,3	-0,20	0,41	0,45
7 422	40	2,0	2,47	0,93	2,14
7 632	212	68,7	-0,38	0,22	0,16
7 835	216	77,7	0,37	-0,27	0,02
7 113	238	14,4	0,94	0,62	0,89
7 423	77	37,1	-1,55	2,88	1,03
7 424	129	11,3	-0,14	0,14	0,41
7 631	101	15,0	-0,52	0,66	0,70
7 449	64	31,9	-0,43	0,26	1,16
7 450	98	96,74	0,51	-1,35	-1,72
7 151	65	41,3	-0,76	0,78	0,06
7 352	89	24,8	0,01	0,37	0,88
7 253	112	28,3	-0,37	0,29	1,59
6 954	50	20,8	-1,34	7,42	2,02
7 655	42	17,2	-0,30	6,05	3,48
7 456	113	9,1	-0,22	0,20	0,44

5.3 Hypothèses sur les fonctions et les distributions

Nous développons les hypothèses de comportement de la fonction de performance G , et nous donnons leur interprétation en terme hydrologique.

Hypothèse 1 : Les fonctions de charge L , et de résistance R , peuvent être approchées par leurs développements du premier ordre de la série de Taylor. Alors :

$$L \approx G(\underline{x}_g) + \sum_{i=1}^n D_i^g(x_i - x_{ig}) \quad (18)$$

$$R \approx G(\underline{x}_c) + \sum_{i=1}^n D_i^c(x_i - x_{ic}) \quad (19)$$

$$E(L) \approx G(\underline{x}_g) + \sum_{i=1}^n D_i^g(\bar{x}_i - x_{ig}) \quad (20)$$

$$\text{VAR}(L) \approx \sum_{i=1}^n \sum_{j=1}^n D_i^g D_j^g \text{COV}(x_i, x_j) \quad (21)$$

$$E(R) \approx G(\underline{x}_c) + \sum_{i=1}^n D_i^c(\bar{x}_i - x_{ic}) \quad (22)$$

$$\text{VAR}(R) \approx \sum_{i=1}^n \sum_{j=1}^n D_i^c D_j^c \text{COV}(x_i, x_j) \quad (23)$$

où \underline{x}_g et \underline{x}_c sont respectivement les vecteurs des paramètres ajustés globalement sur l'échantillon de calage et sur un épisode de l'échantillon de validation et \bar{x} est la moyenne de x . En principe, les dérivées dans l'équation 19 sont nulles car x_c est la valeur réalisant la meilleure performance.

Hypothèse 2 : L et R sont indépendants car \underline{x}_g et \underline{x}_c n'ont pas la même origine : \underline{x}_g est issu de l'échantillon de calage et \underline{x}_c est issu d'un épisode de crues de l'échantillon de validation. Cette hypothèse est liée à la manière de collecter les données. Les principes d'expérimentation stipulent que des mesures d'un même phénomène effectuées dans des conditions physiquement indépendantes fournissent des résultats statistiquement indépendants, c'est-à-dire constituent un échantillon relatif au phénomène aléatoire étudié.

Pour estimer les valeurs caractéristiques, on a alors :

$$E(Z) = E(L) - E(R) \quad (24)$$

$$\text{Var}(Z) = \text{VAR}(L) + \text{VAR}(R) \quad (25)$$

Hypothèse 3 : Les paramètres suivent la loi normale et par suite, L et R suivent une loi normale, ainsi que Z .

5.4 Test de Student pour la fiabilité du modèle hydrologique

On teste $E(Z) = 0$, c'est-à-dire que les performances moyennes du modèle avec x_c (calage par épisode) et x_g (calage global) sont équivalentes. On prend

l'hypothèse nulle $H_0 : E(Z) = 0$. A condition que Z suive une loi normale, la variable statistique β , (HAAN, 1977) suit une loi de Student à $(n - 1)$ degrés de liberté.

$$\beta = \frac{\sqrt{n}(\bar{Z} - E(Z))}{s_z} \quad (26)$$

où $E(Z)$ = l'espérance de Z , \bar{Z} = l'estimation de $E(z)$, s_z = l'estimation de l'écart-type de Z et n = la dimension de l'espace des paramètres. On prend $E(Z) = 0$ selon l'hypothèse nulle ($H_0 : E(Z) = 0$) et $\alpha = 0,05$ comme probabilité associée à l'intervalle de confiance et $n = 3$, en notant que, dans le cas du modèle GR3, Z sera une fonction de trois paramètres du modèle notés A, B et C selon les notations de EDIJATNO et MICHEL (1989).

On obtient donc $|t_{1-\alpha/2, n-1}| = |t_{0,975, 2}| = 4.30$: on doit rejeter l'hypothèse nulle si $|\beta| > 4,3$ et considérer alors que le modèle avec calage global ne donne pas un comportement fiable en crue par épisode.

5.5 Calcul

Pour les équations 20 et 21, les estimations de $E(x_i)$ et de $COV(x_i, x_j)$ dépendent des calages par épisode sur l'échantillon de calage. Pour les équations 22 et 23 les estimations de $E(x_i)$ et $COV(x_i, x_j)$ se basent sur les valeurs des paramètres calés par épisode sur l'échantillon de validation.

Les dérivées partielles sont calculées numériquement. On perturbe faiblement chaque paramètre l'un après l'autre autour de \bar{x}_0 et on mesure l'effet de cette perturbation sur G , toute chose étant égale par ailleurs : $\frac{\Delta G}{\Delta x_i}$ est utilisé comme approximation de $\frac{\partial G}{\partial x_i}$. Pour plus de précision, on a utilisé la dérivation numérique centrée en prenant $Dx_i = 0,05x_{i0}$:

$$\left. \frac{\partial G}{\partial x_i} \right|_{x = \bar{x}_0} = \frac{G(x_{i0} + \Delta x_i) - G(x_{i0} - \Delta x_i)}{2\Delta x_i} \quad (27)$$

Les difficultés du calcul numérique des dérivées partielles n'ont pas été considérées ici, mais les auteurs ont bien conscience que cette façon rustique d'approcher les dérivées doit être grandement améliorée car, en l'état actuel, elle risque d'être une source non contrôlée d'effets pervers sur les résultats de calculs de fiabilité.

5.6 Trois fonctions de performance

On a voulu étudier la fiabilité du modèle sous trois aspects : la racine de la moyenne des carrés des erreurs, le volume de crue et la pointe de crue. On peut alors considérer trois sortes de fonctions de performance : $G_1(x)$, $G_2(x)$ et $G_3(x)$:

$$G_1(x) = \sqrt{\frac{1}{n} \sum_{i=1}^n [Q_m(i) - Q_c(i)]^2} \quad (28)$$

$$G_2(x) = \sum_{i=1}^n [Q_m(i) - Q_c(i)]^2 \quad (29)$$

$$G_3(x) = Q_{mx} - Q_{cx} \quad (30)$$

Dans ces formules on note : Q_m = débit mesuré, Q_c = débit calculé, i = temps, Q_{mx} = pointe de crue mesurée et Q_{cx} = pointe de crue calculée.

5.7 Résultats

Le tableau 3 nous donne les valeurs de β pour les 18 crues de l'échantillon de validation avec $G_1(x)$, $G_2(x)$ et $G_3(x)$. Il y a 2 épisodes pour lesquels $|\beta|$, selon le critère G_2 « volume de crue », dépasse 4,30. Les enregistrements n° 6954 et 7655 correspondent à une mauvaise simulation avec les paramètres globaux issus de l'échantillon de calage. Sur ces deux crues, $E(L)$ pourrait ne pas être égale à $E(R)$, c'est-à-dire que les paramètres calés globalement sur l'ensemble de l'échantillon de calage ne permettent pas une simulation acceptable du volume de crue pour ces deux épisodes. La crue 7655 ainsi que les valeurs simulées avec x_G et x_C sont présentées à titre d'exemple d'un coefficient β fort sur la figure 3. Néanmoins, on peut dire que le modèle est fiable pour l'application au site du Réal Collobrier parce que presque tous les épisodes de l'échantillon de validation ont passé le test pour les trois critères. La figure 4 illustre la qualité obtenue pour les trois critères quand le coefficient β est faible sur la crue 7424.

6 – ANALYSE DES CONTRIBUTIONS DE CHAQUE PARAMÈTRE

Il est intéressant d'apprécier l'importance de chaque paramètre. Les contributions de chaque paramètre permettent d'estimer leur rôle dans le modèle. Ce sont des sous-produits des calculs de la méthode précédente qui fournissent néanmoins des informations importantes à la fois pour la simulation et pour la prévision.

Les équations 21, 23 et 25 indiquent que la variance de la fonction de performance est égale à la somme des covariances des paramètres pondérées par un facteur de sensibilité de la fonction critère à l'influence locale de chaque paramètre. On peut aussi détailler la variance totale en fonction des contributions de la résistance R et de la charge L , qui correspondent à deux états intéressants, respectivement :

- 1) le jeu des paramètres optimisés globalement,
- 2) les jeux des paramètres optimisés sur un épisode.

Pour mesurer la contribution de chaque paramètre aux deux états et pour trois fonctions : G_1 , G_2 et G_3 (cf. équations 28, 29 et 30), on définit les coefficients de contribution des paramètres, CB, qui se déduisent des équations 23 et 25 comme suit :

$$CB_{i,j} = \frac{|D_i D_j \text{COV}(x_i, x_j)|}{\sum_{i=1}^n \sum_{j=1}^n |D_i D_j \text{COV}(x_i, x_j)|} \quad (31)$$

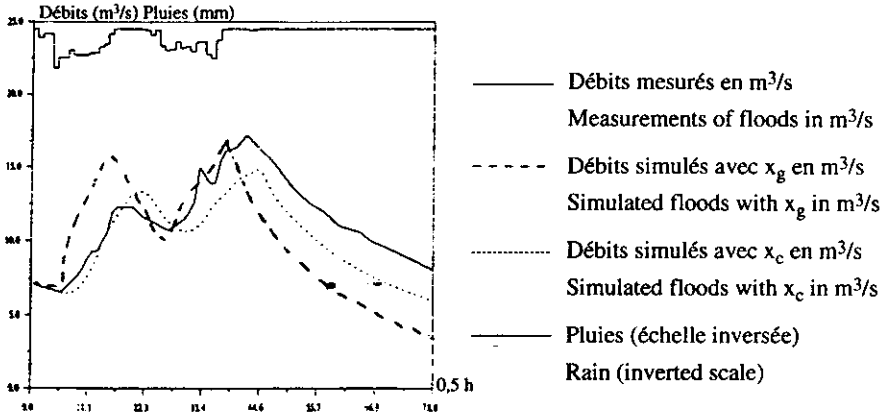


Figure 3 Crue 7655 et valeurs simulées avec x_g et x_c ($\beta = 6,05$ selon G_2).
Flood # 7655 and simulations with x_g and x_c ($\beta = 6.05$ with criterion G_2).

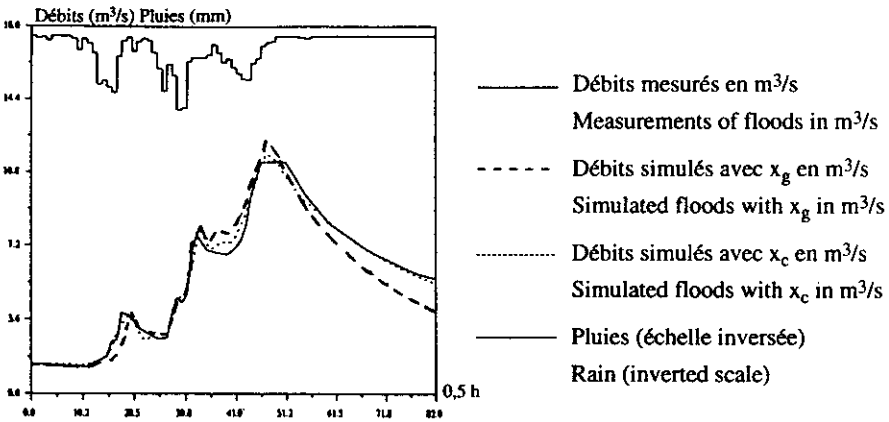


Figure 4 Crue 7424 et valeurs simulées avec x_g et x_c ($|\beta| < .41$ selon tous les critères).
Flood # 7424 and simulations with x_g and x_c ($|\beta| < .41$ for all criteria).

Les tableaux 4 et 5 montrent la valeur de CB de chaque paramètre et de leurs combinaisons relativement à x_c et x_g .

Bien évidemment, sur l'échantillon de calage, le jeu des paramètres globaux fournit de meilleurs résultats que les moyennes des jeux des paramètres calés sur les épisodes et ces derniers sont encore meilleurs que les paramètres globaux. Ces trois groupes de paramètres donnent trois simulations différentes.

Sur le tableau 4, on peut observer les faits suivants :

- 1) les contributions croisées sont relativement petites ;

2) les contributions des paramètres selon G_2 sont différentes de celles selon G_1 et G_3 , c'est-à-dire que les paramètres jouent des rôles différents pour la simulation ;

3) selon G_2 , la simulation du volume de crue dépend beaucoup du paramètre A et ne dépend presque pas du paramètre C ;

4) les paramètres A et B jouent le même rôle pour G_1 (racine de la moyenne des carrés des erreurs) et G_3 (pointe de crue) et dans ces deux cas, ils sont un peu plus importants que C.

Tableau 4 Contributions des paramètres autour de x_c .

Table 4 Parameter contribution to VAR(Z) in the neighbourhood of x_c .

	G_1			G_2			G_3		
	A	B	C	A	B	C	A	B	C
A	0,34			0,77			0,33		
B	0,09	0,34		0,06	0,15		0,06	0,32	
C	0,01	0,01	0,21	0,01	0,00	0,01	0,01	0,01	0,27

Tableau 5 Contributions des paramètres autour de x_g .

Table 5 Parameter contribution to VAR(Z) in the neighbourhood of x_g .

	G_1			G_2			G_3		
	A	B	C	A	B	C	A	B	C
A	0,12			0,80			0,47		
B	0,05	0,36		0,04	0,09		0,07	0,12	
C	0,04	0,07	0,36	0,02	0,01	0,04	0,06	0,02	0,26

Selon le tableau 5, présentant le cas où les paramètres n'arrivent pas à leur valeurs optimales, on obtient la même conclusion sur G_2 . Si on considère les critères G_1 et G_3 , on peut dire dans ce cas que B et C sont plus importants que A pour la somme des carrés des erreurs et que A est le plus important pour la pointe de crue.

En conclusion de cette analyse des contributions partielles à la variabilité, on peut donc constater que les paramètres du modèle GR3 jouent chacun un rôle bien différencié.

7 - CONCLUSION

Nous avons donné la définition de la fiabilité pour un modèle hydrologique, en effectuant une transposition à partir de la définition classique de l'analyse de fiabilité employée dans le domaine des structures. Notre démarche est une nouvelle méthode pour apprécier la réaction d'un modèle de simulation à l'ensemble des paramètres du point de vue statistique (outil statistique d'estimation) et de transfert du global au local (possibilité de passage de l'échantillon de calage à

l'échantillon de validation). Cette analyse peut être vue comme un jugement relatif à un état du modèle différent de l'état « optimal ». La possibilité de considérer une valeur non nulle pour e permettrait d'introduire de façon explicite une tolérance plus ou moins forte lors de l'établissement de ce jugement.

Dans cet article, la nature aléatoire des paramètres doit être précisée. Il est en effet important de noter que la variation des paramètres n'est qu'en très faible partie due à l'aléa naturel : pour un événement donné il n'y a que très peu d'incertitude dans l'estimation des paramètres obtenus par calage local compte tenu du nombre important de données. Il est donc peu raisonnable de traiter R et L comme des variables aléatoires au sens fréquentiste car augmenter le nombre de données n'influencera que très peu la précision des estimations obtenues. Par contre, les erreurs proviennent de la méconnaissance intrinsèque du modèle utilisé (la transformation pluie-débit pourrait être mieux représentée dans une classe beaucoup plus vaste que les modèles du type GR3, ou par un modèle GR3 à coefficients aléatoires variables) et de la non-prise en considération de divers facteurs (depuis des effets non maîtrisés d'expérimentation jusqu'à une non stationnarité éventuelle du phénomène observé). En conséquence, l'utilisation de probabilités pour traduire la mesure de l'incertitude portant sur les performances d'un modèle de transformation pluie-débit doit plutôt ici être envisagée sous l'angle de son acception bayésienne.

On peut aussi tirer des conclusions très pratiques sur deux points particuliers : le modèle GR3 et la méthode POPM.

Pour le modèle GR3, il est très instructif de voir que les paramètres obtenus à partir de l'échantillon de calage ne sont pas très éloignés de leurs valeurs optimales calées sur les épisodes de l'échantillon de validation. Cela montre que l'on peut se fier à la capacité d'extrapolation du modèle GR3 pour la prédétermination de crue. Cependant, pour la prévision, on pourra toujours partir des paramètres globaux mais il serait déraisonnable de les corriger trop fortement. L'analyse de la contribution individuelle de chaque paramètre, comme sous-produit de l'analyse de fiabilité, à différents critères de qualité apporte une connaissance intéressante : chaque paramètre du modèle GR3 possède son propre rôle.

Nous pouvons confirmer l'utilité de l'analyse de fiabilité par la méthode du premier ordre et des premiers moments dans la modélisation hydrologique. Elle présente les caractéristiques suivantes :

- 1) c'est une extension de l'analyse de validation. Elle juge la qualité d'un modèle selon non seulement la valeur du critère utilisé mais aussi la différence de ce critère entre le calage « idéal » et celui global ;
- 2) cette analyse lie la qualité d'un modèle à la complexité du modèle (nombre de paramètres) et à la sensibilité de chaque paramètre ;
- 3) son talon d'Achille est de faire une hypothèse de loi normale pour représenter les distributions d'incertitudes des paramètres et aussi que la fonction de performance ne varie pas trop brutalement au voisinage de l'optimum ;
- 4) cette analyse peut considérer différents critères de validation de façon séparée (démarche choisie dans cette étude) ou les regrouper sous un même terme de probabilité (en utilisant des lois conjointes et des combinaisons de seuils d'acceptabilité, démarche qui reste à valider). On peut discuter du choix des fonctions de performance, comme tous les choix de fonctions fournissant des « critères ». On trouvera par exemple une analyse critique des différents critères pour juger des mérites de modèles pluies-débit dans KUARK-LEITE (1990).

L'analyse de fiabilité est donc un complément permettant une synthèse des analyses existantes pour un modèle hydrologique comprenant en particulier validation et analyse de sensibilité. Nous souhaitons que cette notion puisse contribuer à promouvoir utilement la recherche dans le domaine de la modélisation hydrologique.

REMERCIEMENTS

Les auteurs ont bénéficié des critiques et suggestions de C. MICHEL. Cet article a aussi été remanié suite aux corrections et propositions de réorganisation de trois réviseurs anonymes de la Revue des Sciences de l'Eau que nous tenons à remercier pour leur appui méthodologique.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BENJAMIN J.R., CORNELL, C.A., 1970. Probability, Statistics and Decision for Civil Engineers, McGraw-Hill, New York, N.Y.
- BOBEE B., ASHKAR F., 1991. The Gamma Family and Derived Distributions Applied in Hydrology, Water Resources Publications, Colorado.
- DAVIS D.R., NNAJI S., 1982. The information needed to evaluate the worth of uncertain information, predictions and forecasts. *J. Appl. Meteorology*, 21, 461-470.
- DUCKSTEIN L., PARENT E., 1994. Engineering Risk in Natural Resources Management, with Special References to Hydro-systems under Changes of Physical or Climatic Environment, NATO ASI Series, Vol. 217, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- DUCKSTEIN L., PLATE E., 1987. Engineering Reliability and Risk in Water Resources, in Engineering Reliability and Risk in Water Resources, NATO ASI Series, E.M. Nijhoff, Dordrecht, The Netherlands.
- EDIJATNO, MICHEL C., 1989. Un modèle pluie-débit journalier à trois paramètres. *La Houille Blanche*, 2, 113-121.
- GANOULIS J., 1991. Water Resources Engineering Risk Assessment, NATO ASI series, Ecological Science, Vol.29, Springer Verlag.
- GAREN D.C., BURGESS S.J., 1981. Approximate error bounds for simulated hydrographs. *J. Hydraulics Division, ASCE*, 107 (HY11), 1519-1534.
- HAAN C.T., 1977. Statistical Methods in Hydrology. Ames, Iowa, Iowa State University Press.
- KUARK LEITE L., 1990. Réflexions sur l'utilité des modèles mathématiques dans la gestion de la pollution diffuse d'origine agricole. Thèse ENPC, Paris.
- MELCHING C.S., YEN B.C., 1986. Slope Influence on storm sewer risk. Stochastic and Risk Analysis in Hydraulic Engineering, B.C. Yen, ed., Water Resources Publications, Littleton, Colorado, 79-89.
- MELCHING C.S., YEN B.C., WENZEL H.G., 1990. A reliability estimation in modeling watershed runoff with uncertainties. *Water Resour. Res.*, 26(10), 2275-2286.
- MELCHING C.S., YEN B.C., WENZEL H.G., 1991. Output reliability as guide for selection of rainfall-runoff models. *J. Water Resour. Planning and Management*, 117(3), 383-398.

- PLATE E.J., DUCKSTEIN L., 1988. Reliability based design concepts in hydraulic engineering. *Water Resour. Bull.*, AWRA, 24, 234-245.
- TUNG Y.-K., MAYS L.W., 1981. Risk models for flood levee design. *Water Resour. Res.*, 17(4), 833-841.
- TUNG Y.-K., MAYS L.W., 1980. Risk analysis for hydraulic design. *J. Hydraulics Division*, ASCE, 106(HY5), 893-913.
- WOOD E.F., 1976. An analysis of the effects of parameter uncertainty in deterministic hydrologic models. *Water Resour. Res.*, 12(5), 925-932.
- YEN B.C., 1986. Reliability analysis in hydraulic design. *Stochastic and Risk Analysis in Hydraulic Engineering*, Water Resources Publications, Littleton, Colorado : 37-47.
- YEN B.C., 1987. Reliability of hydraulic structures possessing random loading and resistance. In : *Engineering Reliability and Risk in Water Resources*, L. Duckstein and E. Plate, eds., NATO ASI Series, E. M. Nijhoff, Dordrecht, The Netherlands, 95-113.
- YEN B.C., TANG W.H., 1977. Reliability of flood warning. *Stochastic Processes in Water Resources Engineering. Proceedings of the 2nd International Symposium on Stochastic Hydraulics*, Lund, Sweden, Water Resources Publications Littleton, Colorado, 333-337.
- YEN B.C., CHANG S.-T., MELCHING C.S., 1986. First-order reliability analysis. *Stochastic and Risk Analysis in Hydraulic Engineering*, B.C. Yen, ed., Water Resources Publications, Littleton, Colorado, 1-34.