

MODÉLISATION DES RISQUES COMBINÉS

GUIDE DE L'USAGER DE RISCHYDRO

Rapport de recherche No R-847

Mars 2006

MODÉLISATION DES RISQUES COMBINÉS

GUIDE DE L'USAGER DE RISCHYDRO

Par :

Rémi Gagné

Salaheddine El Adlouni

et Taha B.M.J. Ouarda

Chaire en Hydrologie statistique Hydro-Québec / CRSNG, INRS-ETE
Chaire de recherche du Canada en estimation des variables hydrologiques, INRS-ETE
490, rue de la Couronne, QUÉBEC (Québec) G1K 9A9

Rapport de recherche N° R-847

Tous droits réservés

© Copyright 2006

Mars 2006

Version

3

MODELISATION DES RISQUES COMBINES

Préparé par Rémi Gagné, Salaheddine El Adlouni et Taha Ouarda

Chaire en hydrologie statistique Hydro-Québec / CRSNG

**Chaire de recherche du Canada en estimation des variables
hydrologiques**

Guide de l'utilisateur de
RISCHYDRO

Table des matières

Introduction	1
Modèle normal bivarié	1
Modèle gamma bivarié	3
Ajustement de distributions de probabilités	3
Copules	5
Comment utiliser RISCHYDRO	10
Commencer un nouveau projet	10
Modèle normal	11
Modèle Gamma	18
Ajustement des lois marginales	24
Modèle bivarié	26
Enregistrer un projet	37
Ouvrir un projet	37
Exportation des résultats	38
Quitter RISCHYDRO	38
Références	38

Introduction

RISCHYDRO est une interface graphique développée sous l'environnement MATLAB pour modéliser les risques combinés en hydrologie. Il permet dans un premier temps de déterminer la distribution jointe des débits de deux sites et d'obtenir la distribution d'une variable d'intérêt qui dépend des deux débits. Pour ce faire, RISCHYDRO utilise deux modèles : normal et gamma bivarié. Le développement de ces deux modèles est basé sur la notion de copule pour la génération d'observations. En effet, on utilise des algorithmes de génération des lois normale et gamma bivariées pour calculer les probabilités, ou les quantiles, correspondants à une variable Z, fonction de X et Y. Dans un deuxième temps, il permet d'ajuster plusieurs distributions marginales à la série de débits de chacun des deux sites. En comparant leur ajustement à l'aide de différents critères, RISCHYDRO sélectionne la distribution marginale qui s'adapte le mieux aux débits. Avec ces distributions marginales, il est en mesure de modéliser la structure de dépendance entre les deux séries de données grâce aux copules. Encore ici, une comparaison de l'ajustement de plusieurs copules à l'aide de critère permet à RISCHYDRO de choisir la meilleure structure de dépendance pour les deux séries de débits. Il détermine ensuite les risques appropriés à partir de cette structure.

Modèle normal bivarié

Soit X_i et Y_i ($i=1, 2, \dots, n$), deux séries des données de taille n , le but est d'obtenir la probabilité conjointe des variables aléatoires X et Y dont la distribution suit la loi normale. Les variables peuvent ne pas être distribuées suivant une loi normale, on procède donc à la normalisation avant d'ajuster le modèle. RISCHYDRO utilise la méthodologie du modèle SMPLNORM (Bruneau et al., 1994) pour vérifier la normalité bivariée. Afin d'obtenir la probabilité conjointe $P(X \leq x, Y \leq y)$, RISCHYDRO suit les étapes suivantes:

- Classer les deux séries en ordre ascendant et associer à chaque valeur son rang.
- Normaliser les deux séries X et Y en utilisant la transformation de Box-Cox (Box et Cox, 1964) :

$$U = \begin{cases} \frac{X^{\lambda_x} - 1}{\lambda_x} & \lambda_x \neq 0 \\ \log(X) & \lambda_x = 0 \end{cases}$$

$$V = \begin{cases} \frac{Y^{\lambda_y} - 1}{\lambda_y} & \lambda_y \neq 0 \\ \log(Y) & \lambda_y = 0 \end{cases}$$

Les valeurs λ_x et λ_y peuvent être estimées par la méthode du maximum de vraisemblance (Linnet, 1988).

- Utiliser le test de Filliben pour vérifier si les nouvelles séries U et V consécutives à la normalisation de X et Y suivent bien une distribution normale. La procédure de ce test est décrite aux paragraphes ci-dessous:

- Standardiser les nouvelles séries U et V

$$U^* = \frac{U - \mu_U}{\sigma_U}, \quad V^* = \frac{V - \mu_V}{\sigma_V}$$

où μ_U et σ_U sont respectivement la moyenne et l'écart type de série U et μ_V et σ_V sont la moyenne et l'écart type de série V .

- Calculer la probabilité empirique cumulée (P_e) pour les deux séries U^* et V^*

$$P_e = \frac{m - 0.4}{n + 0.2}$$

où m est le rang.

- Calculer l'inverse de la probabilité théorique cumulée pour les deux séries

$$t = F^{-1}(P_e)$$

où $F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{s^2}{2}} ds$.

- Obtenir le coefficient de détermination entre l'inverse de la probabilité théorique cumulée et la série normale standardisée. Ce coefficient est un indice pour vérifier la normalité des données.

- Identifier la loi bivariée du couple (U^*, V^*) . Il s'agira de vérifier si cette paire suit une loi normale bivariée de la forme:

$$f_{U^*, V^*}(u^*, v^*) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(u^{*2} - 2\rho U^* V^* + V^{*2})}$$

où ρ est le coefficient de corrélation. Pour vérifier la normalité bivariée, on utilise la procédure décrite ci-dessous:

- Calcul des coefficients B

$$B = \frac{U^{*2} - 2\rho U^* V^* + V^{*2}}{1 - \rho^2}$$

- Vérifier si les coefficients B suivent une loi χ^2 à 2 degrés de liberté (Hald, 1967).

- Calcul des probabilités normales bivariées puis des quantiles de X et Y correspondant à une période de retour T par simulation à partir d'une copule Gaussienne.

- Calcul des probabilités, ou des quantiles, correspondants à une variable Z , fonction de X et Y .

Modèle gamma bivarié

Les étapes de la méthode d'analyse des risques combinés à partir de la distribution gamma bivariée sont les suivantes:

- Ajustement et test d'adéquation de la loi gamma à chacune des séries
 - Calculer la probabilité empirique cumulée pour les deux séries X et Y .
 - Calculer l'inverse de la probabilité théorique cumulée pour les deux séries

$$t = F^{-1}(P_e | a, b)$$

où

$$F(t | a, b) = \frac{1}{b^a \Gamma(a)} \int_0^t s^{a-1} e^{-\frac{s}{b}} ds$$

et a et b sont les paramètres de la distribution Gamma (Bobée et Ashkar, 1991):

$$a = \frac{\mu^2}{\sigma^2}$$

$$b = \frac{\sigma^2}{\mu}$$

Obtenir le coefficient de détermination entre l'inverse de la probabilité théorique cumulée et la série.

- Calcul des probabilités gamma bivariées puis des quantiles de X et Y correspondant à une période de retour T par simulation à partir d'une copule Clayton.
- Calcul des probabilités, ou des quantiles, correspondants à une variable Z , fonction de X et Y .

Ajustement de distributions de probabilités

Pour ajuster les distributions de probabilités, RISCHYDRO suit les étapes et les méthodes suivantes :

- Estimation des paramètres de chacune des distributions par la méthode du maximum de vraisemblance.

- L'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ pour une distribution est donné par :

$$\hat{\theta} = \arg \max l(\theta)$$

où θ est le vecteur des paramètres de la distribution et $l(\theta)$ est la fonction de log-vraisemblance donnée par

$$l(\theta) = \sum_{i=1}^n \ln[f(x_i, \theta)]$$

où n est le nombre d'observations de la série de données, $f(x_i; \theta)$ est la fonction de densité de la distribution dont on estime les paramètres et x_i sont les observations.

- Comparaison de l'ajustement des distributions à la série de données à l'aide des critères.

- Critère d'information d'Akaike :

- Le critère d'Akaike est d'abord calculé pour chaque distribution de la manière suivante :

$$AIC = -2 \left[l(\hat{\theta}) \right] + 2k$$

où $l(\hat{\theta})$ est la fonction de log-vraisemblance évaluée avec les paramètres estimés et k est le nombre de paramètres de la distribution.

- La distribution ayant la plus faible valeur pour le critère d'Akaike est celle qui s'ajuste le mieux aux observations.

- Critère d'information bayésien :

- Le critère d'information bayésien est calculé pour chaque distribution de la manière suivante :

$$BIC = -2 \left[l(\hat{\theta}) \right] + 2k \ln(n)$$

où $l(\hat{\theta})$ est la fonction de log-vraisemblance évaluée avec les paramètres estimés, k est le nombre de paramètres de la distribution et n est le nombre d'observations.

- La distribution ayant la plus faible valeur pour le critère d'information bayésien est celle qui s'ajuste le mieux aux observations.

- Critère d'Anderson-Darling :

- La statistique d'Anderson-Darling est d'abord calculée pour chaque distribution comme suit :

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \left\{ (2i-1) \ln[F(x_i; \hat{\theta})] + (2n+1-2i) \ln[1-F(x_i; \hat{\theta})] \right\}$$

où n est le nombre d'observations, $F(x_i; \hat{\theta})$ est la fonction de distribution cumulative, x_i sont les observations et $\hat{\theta}$ est le vecteur des paramètres de la distribution.

- La méthodologie donnée par Laio (2004) est ensuite appliquée pour déterminer la distribution de la statistique A^2 et son seuil observé. Le test d'Anderson-Darling est effectué avec un seuil de confiance de 5%.
- La distribution avec le plus grand seuil observé est la distribution qui s'ajuste le mieux aux observations.

○ Critère de Cramer-von Mises :

- La statistique de Cramer-von Mises est d'abord calculée pour chaque distribution comme suit :

$$W^2 = \sum_{i=1}^n \left[F(x_i; \hat{\theta}) - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n}$$

- La méthodologie donnée par Laio (2004) est ensuite appliquée pour déterminer la distribution de la statistique W^2 et son seuil observé. Le test de Cramer-von Mises est effectué avec un seuil de confiance de 5%.
- La distribution avec le plus grand seuil observé est la distribution qui s'ajuste le mieux aux observations.

Copules

- La théorie

- Les copules permettent de modéliser la structure de dépendance entre plusieurs variables ayant des distributions quelconques. Le logiciel RISCHYDRO emploie la théorie des copules dans le cas particulier de deux variables. La pierre angulaire de cette théorie est le théorème de Sklar.

Théorème : Soit F une loi de probabilité en dimension n , avec des distributions marginales continues, F_1, \dots, F_n . Alors F a la représentation unique, sous forme de copule, suivante :

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (1)$$

Pour le cas de dimension 2, on a donc que $F(x,y) = C(F_x(x), F_y(y))$. Ainsi, la copule C est une fonction de distribution bivariee verifiant les conditions suivantes :

- $C : [0,1]^2 \longrightarrow [0,1]$;
 - C est bornee et 2-croissante;
 - C a des distributions marginales C_i telles que : $C_i(u) = C(1,u) = C(u,1) = u$ quelque soit $u \in [0,1]$.
- Un corollaire du theoreme de Sklar permet de separer les distributions marginales des variables et la structure de dependance.

Corollaire : Soit F une loi de probabilite en dimension n , avec des distributions marginales continues, F_1, \dots, F_n et de copule C (verifiant (1)). Alors, pour tout $u=(u_1, \dots, u_n)$ dans $[0,1]^n$:

$$C(u_1, \dots, u_n) = F^{-1}(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad (2)$$

où F_i^{-1} est l'inverse generalise de F_i .

- Ce corollaire donne donc une expression mathematique pour les copules en fonction des inverses des distributions marginales.

• Copules Archimediennes

- À partir de ces deux resultats, il est possible de definir plusieurs classes de copules. Le logiciel RISCHYDRO utilise trois copules de la famille Archimedienne soient les copules de Gumbel, de Clayton et de Frank.

Une copule Archimedienne s'exprime sous la forme suivante dans un cas bivarie :

$$C(u_1, u_2) = \psi^{-1}(\psi(u_1) + \psi(u_2))$$

pour tout $0 \leq u_1, u_2 \leq 1$, et où ψ est une fonction souvent appelee « Générateur », qui verifie :

- $\psi(1) = 0$;

- pour tout $t \in]0,1[$, $\psi'(t) < 0$, (ψ est décroissante);
- pour tout $t \in]0,1[$, $\psi''(t) \geq 0$, (ψ est convexe);

Le tableau ci-dessous donne les expressions des paramètres, des « Générateurs » et des copules pour celles de Gumbel, Clayton et de Frank.

Copule	Paramètre α (En fonction du tau de Kendall)	Générateur $\psi(t)$	$C(u,v)$
Gumbel	$(1-\tau)^{-1}$	$(-\ln t)^\theta$	$\exp\left(-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{1/\theta}\right)$
Clayton	$\frac{2\tau}{1-\tau}$	$t^{-\theta}-1$	$\max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{1/\theta}, 0\right)$
Frank	$\tau(\theta) = 1 - \frac{4}{\theta} \left(1 - \frac{1}{\theta} \int_0^1 \frac{t}{e^t - 1} dt\right)$	$-\ln\left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right)$	$\frac{-1}{\theta} \ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right)$

- Estimation du paramètre

- Les paramètres de ces copules peuvent être estimés de différentes manières.

RISCHYDRO utilise le fait que le paramètre α peut s'exprimer en fonction du tau de Kendall, un coefficient de corrélation défini par :

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

où les indices 1 et 2 réfèrent à deux observations différentes des variables X et Y. Ainsi, en mesurant le tau de Kendall pour des variables X et Y, on peut obtenir une estimation du paramètre α .

- Comparaison de l'ajustement

Pour déterminer quelle copule décrit le mieux la structure de dépendance entre deux variables, RISCHYDRO utilise deux critères : K et S.

- Le critère K

- Développé par Genest et Rivest (1993) dans le cas bivarié, ce critère calcule la corrélation entre la fonction K univariée (3) et son estimateur non paramétrique (4). La fonction K théorique est déterminée à l'aide des « Générateurs » des copules Archimédiennes.

$$K_{\psi}(z) = z - \frac{\varphi(z)}{\varphi'(z)} \quad (3)$$

$$\hat{K}(z) = \frac{1}{T} \sum_{i=1}^T 1_{[w_i \leq z]} \quad \text{avec } w_i = \frac{1}{T-1} \sum_{t=1}^T 1_{[x_1^t < x_1^i, x_2^t < x_2^i]}, \quad i=1, \dots, T$$

(4)

- Pour plus de détails sur ce critère, consulter Genest et Rivest (1993).

- Le critère S

- Ce critère permet de comparer directement les copules à l'aide de la statistique S (5). La copule qui minimise l'estimateur de cette statistique donne le meilleur ajustement aux observations (6).

$$S_n = \int_0^1 |k_n(z)|^2 \hat{K}(z) dz \quad \text{où } k_n(z) = \sqrt{n} \{K_{\psi}(z) - \hat{K}(z)\} \quad (5)$$

$$S_n = \frac{n}{3} + n \sum_{j=1}^{n-1} K_{\psi}^2\left(\frac{j}{n}\right) \left\{ \hat{K}\left(\frac{j+1}{n}\right) - \hat{K}\left(\frac{j}{n}\right) \right\} - n \sum_{j=1}^{n-1} K_{\psi}^2\left(\frac{j}{n}\right) \left\{ \hat{K}^2\left(\frac{j+1}{n}\right) - \hat{K}^2\left(\frac{j}{n}\right) \right\}$$

(6)

- Fonction de répartition et de densité des copules

- La fonction de répartition des copules est simplement donnée par $C(u_1, u_2)$.
- La fonction de densité des copules est quant à elle donnée par

$$c(u_1, u_2) = \frac{\partial C(u_1, u_2)}{\partial u_1 \partial u_2}.$$

- Probabilité conjointe, probabilité conditionnelle et période de retour conditionnelle

À l'aide des copules, il est possible de calculer la probabilité conjointe de plusieurs événements.

- Probabilité conjointe d'un événement « ET » :

$$P(X > x, Y > y) = (1 - u) + (1 - v) - 1 + C(u, v) \text{ où } u = P(X \leq x) = F_X(x) \text{ et}$$

$$v = P(Y \leq y) = F_Y(y)$$

- Probabilité conjointe d'un événement « OU » :

$$P(X > x, Y > y) = 1 - C(u, v) \text{ où } u = P(X \leq x) = F_X(x) \text{ et}$$

$$v = P(Y \leq y) = F_Y(y)$$

- Probabilité conditionnelle de $X | Y > y$

$$P(X > x | Y > y) = \frac{P(X > x, Y > y)}{P(Y > y)}$$

- Probabilité conditionnelle de $Y | X > x$

$$P(Y > y | X > x) = \frac{P(X > x, Y > y)}{P(X > x)}$$

- Période de retour conditionnelle de $X | Y > y$: $T_{X|Y>y} =$

$$\frac{1}{P(X > x | Y > y)}$$

- Période de retour conditionnelle de $Y \mid X > x$: $T_{Y|X>x} =$

$$\frac{1}{P(Y > y \mid X > x)}$$

Comment utiliser RISCHYDRO

Ce manuel contient des exemples qui vous permettent de découvrir pas à pas les différentes fonctionnalités du logiciel RISCHYDRO.

Commencer un nouveau projet

Dans RISCHYDRO, cliquez sur l'option Nouveau du menu Fichier et ensuite sur les options Première Série et Deuxième Série pour charger deux séries de débits (Figure 1).

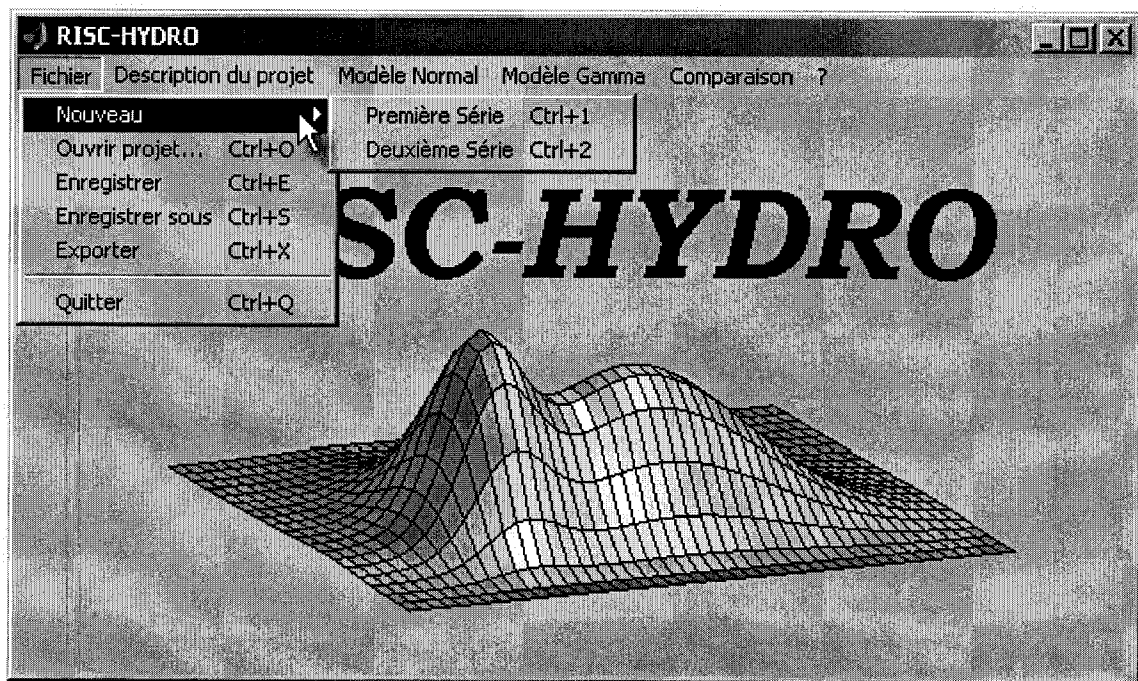


Figure 1- Menu Fichier du logiciel RISCHYDRO.

Pour l'exemple que nous avons choisi, le point d'intérêt se situe sur le fleuve Saint-Laurent en aval de son confluent avec la rivière Saint-Maurice. Les débits pour ces deux bassins sont disponibles sur une période de 65 années de 1922 à 1986.

Après le chargement des données, vous pouvez aller au menu 'Description du projet' pour entrer les commentaires sur le projet.

Modèle normal

Dans RISCHYDRO, allez au menu Modèle Normal. Ce menu vous offre les commandes suivantes : Normalisation, Test de normalité et Simulation (Figure 2). Dans le menu Modèle Normal, cliquez sur l'option Normalisation et ensuite sur l'option Première Série pour ouvrir la fenêtre 'Normalisation d'une série'. Cette fenêtre vous offre les options suivantes: Série Initiale, Ln(X), Normalisation, Quantile Quantile, Imprimer et Copier.

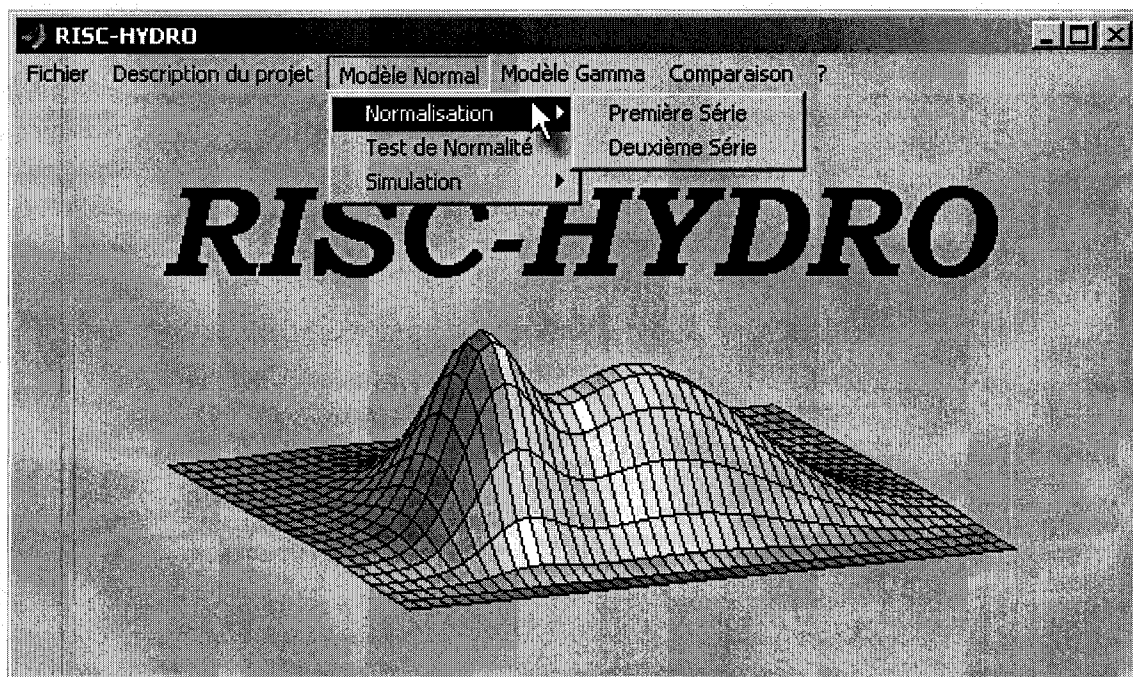


Figure 2- Menu Modèle Normal du logiciel RISCHYDRO

Série Initial: Cette option affiche la série initiale sur le papier de probabilité normal. Le but est d'évaluer graphiquement si les données de la série initiale pourraient venir d'une distribution normale. Si les données sont normales, le graphe sera linéaire et le coefficient de détermination (R^2) sera près de 1.

Ln(X): Cette option affiche le logarithme naturel de la série initiale sur le papier de probabilité normal.

Normalisation: Cette option utilise la transformation de Box-Cox pour normaliser la série et affiche les résultats sur le papier de probabilité normal (Figure 3).

Quantile Quantile: Cette option crée un graphique quantile-quantile (test graphique de normalité) (Figure 4).

Imprimer: Cette option vous permet d'imprimer la figure en cours.

Copier: Cette option vous permet de copier la figure en cours dans le presse-papier.

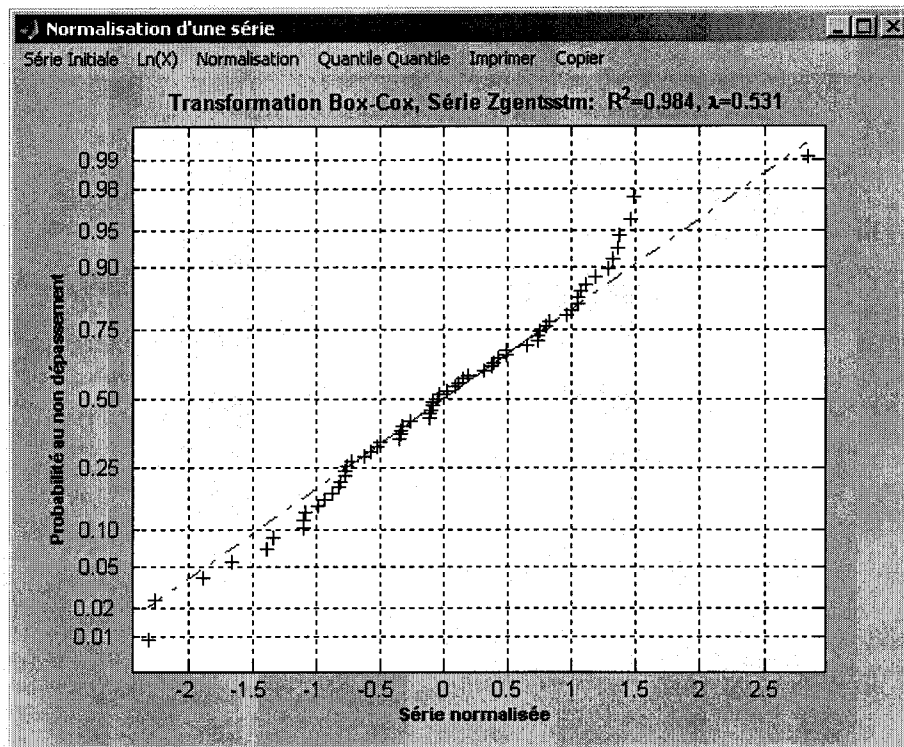


Figure 3- Valeurs normalisées pour la série du fleuve Saint-Laurent

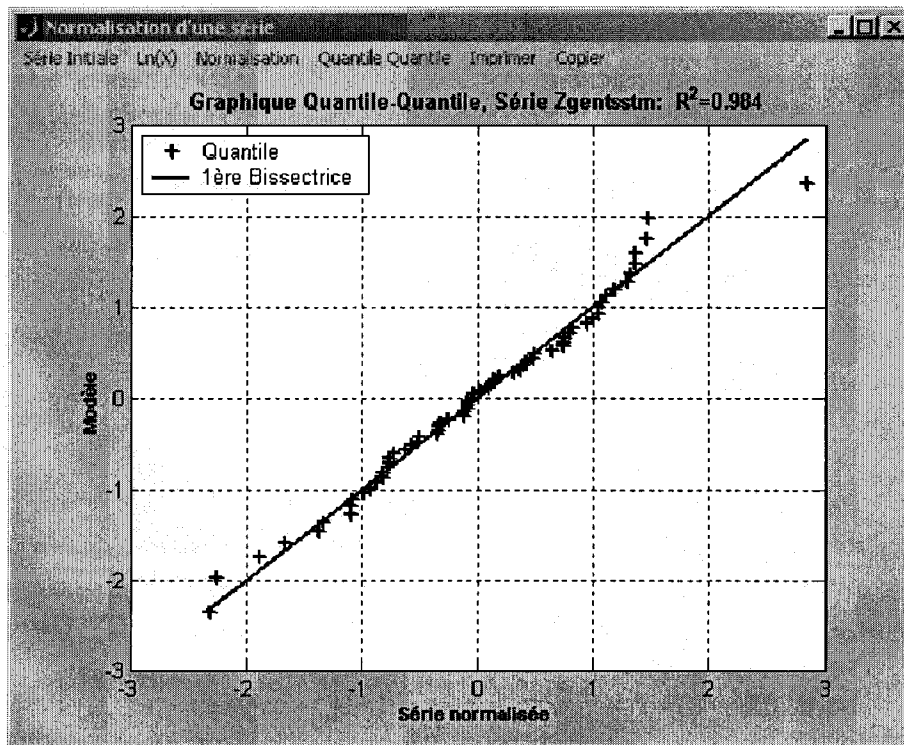


Figure 4- Graphique QQ pour les débits du Saint-Laurent.

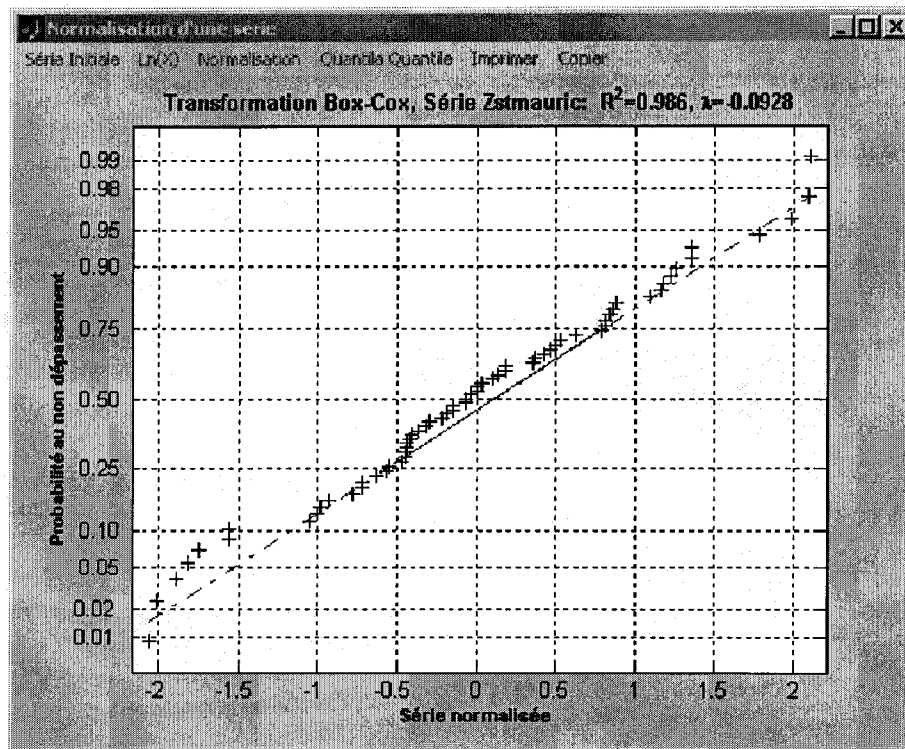


Figure 5- Valeurs normalisées pour la série de la rivière Saint-Maurice

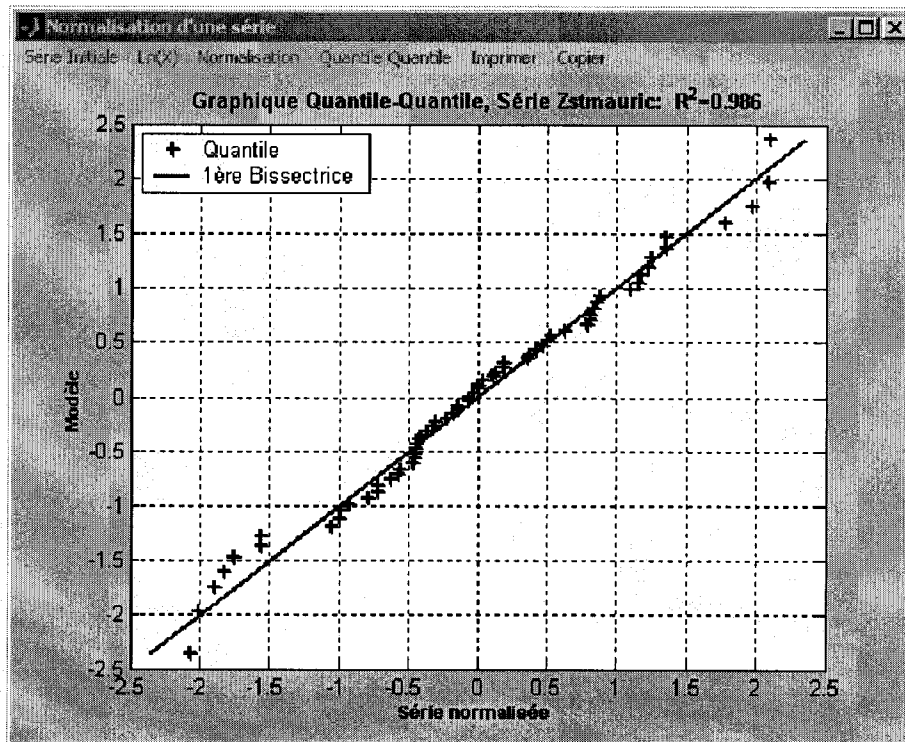


Figure 6- Graphique QQ pour les débits du Saint-Maurice

Dans le menu Modèle Normal, cliquez sur l'option Normalisation et ensuite sur l'option Deuxième Série pour ouvrir la fenêtre 'Normalisation d'une série' et normalisez les données de débit de la deuxième série (Figures 5 et 6).

Les coefficients de la transformation de Box et Cox sont respectivement 0.531 et -0.093 pour Saint-Laurent et Saint-Maurice. Les figures 3 et 4 montrent un bon ajustement de la loi normale à chacune des deux séries après transformation.

Remarque

Pour mieux visualiser les résultats, vous pouvez faire le zoom avant avec le bouton gauche de la souris et faire le zoom arrière avec le bouton droit sur les graphes.

Après la normalisation, vérifiez la corrélation entre les deux séries transformées en cliquant sur l'option Bivariée du menu Modèle Normal. Le coefficient de corrélation entre les deux séries transformées est de 0.542. Le test t sur le coefficient de corrélation indique le rejet de l'hypothèse nulle au niveau de signification 5%. On ne peut donc considérer les deux séries comme indépendantes (Figure 7). Cliquez ensuite sur Ok pour vérifier la normalité de la distribution jointe des deux séries transformées selon la procédure graphique de Hald (1967) (Figure 8).

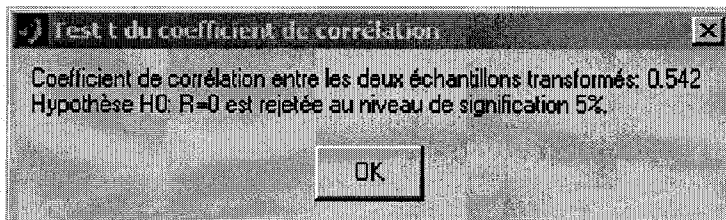


Figure 7- Test t sur le coefficient de corrélation entre les deux séries transformées

Ensuite dans le menu Modèle Normal, choisissez l'onglet Simulation. Cette onglet vous offre les options suivantes: Probabilités Jointes, Variable étudiée, Graphiques et Statistiques :

Probabilités Jointes: L'option 'Probabilités Jointes' vous permet de calculer les probabilités jointes au non dépassement pour divers couples de débits en provenance des deux bassins et les quantiles des deux séries correspondant à une période de retour. Dans le menu 'Probabilités Jointes' entrez les périodes de retour désirées et le nombre de points à simuler et puis cliquez sur le bouton Actualiser. Pour les périodes de retour 2, 10, 50, 100, 1000 et 10000 ans, nous avons calculé les débits en provenance des stations de Gentilly et St-Maurice ainsi que les probabilités conjointes au non- dépassement en générant un vecteur aléatoire suivant une distribution Gaussienne bivariée de taille 100000 (Figure 9). Il est aussi possible de calculer la probabilité au non dépassement à partir des valeurs marginales. Pour ce faire, il vous suffit d'entrer les valeurs marginales et puis cliquez sur le bouton ↗ .

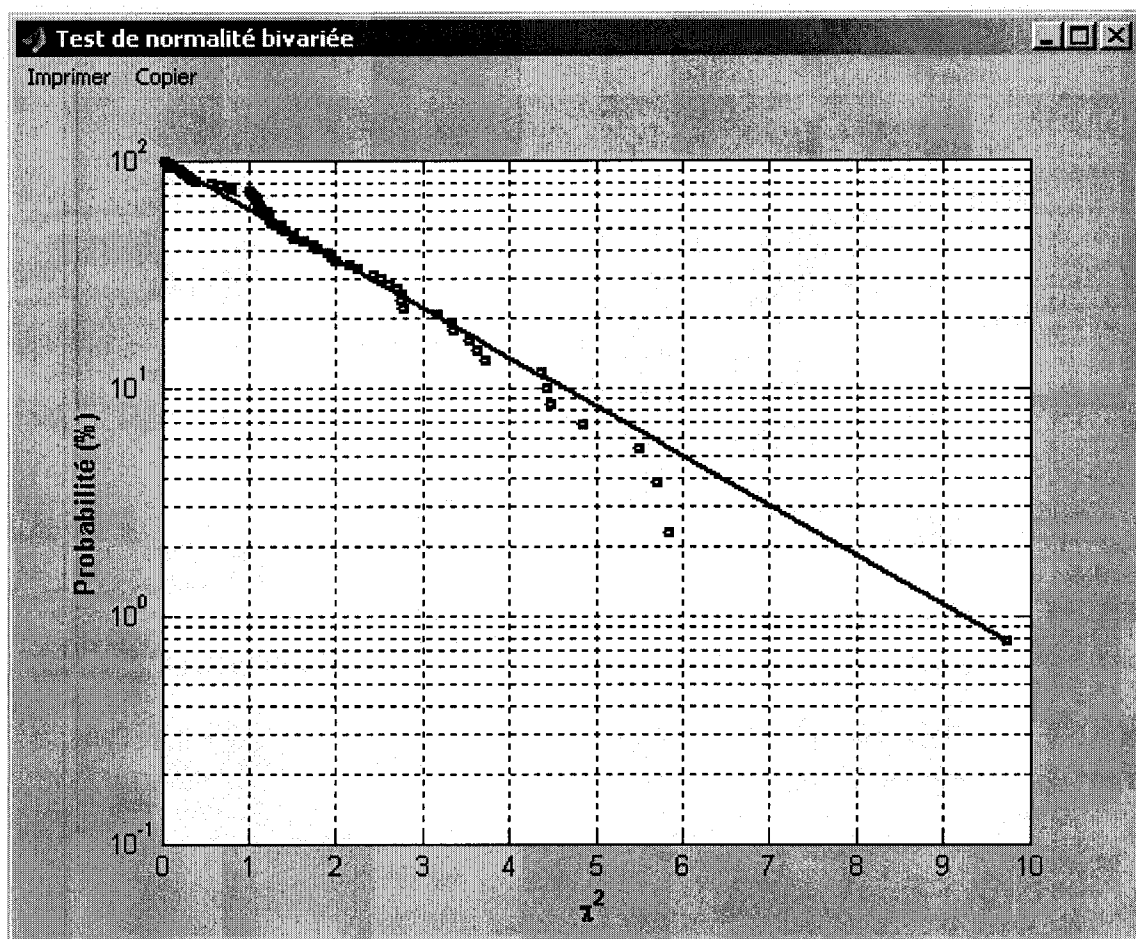


Figure 8- Vérification de la normalité de la distribution jointe des deux séries de débits transformés

Variable étudiée: L'option 'Variable étudiée' vous permet de calculer les probabilités, ou les quantiles, correspondants à une variable z , fonction de x et y . En utilisant la relation (Bruneau et al., 1994)

$$\log(z) = 0.5094 \times \log(x) + 0.1053 \times \log(y) - 1.8103$$

entre les débits (x et y) et le niveau (z), nous avons calculé par simulation les niveaux correspondants aux périodes de retour 2, 10, 50, 100, 1000 et 10000 ans (Figure 10).

Graphiques: L'option 'Graphiques' vous permet d'afficher les valeurs observées et simulées des séries x et y ainsi que l'histogramme de la variable étudiée z des séries (Figures 11 et 12).

Statistiques: L'option 'Statistiques' vous permet d'afficher les statistiques pour les données observées et simulées (Figure 13).

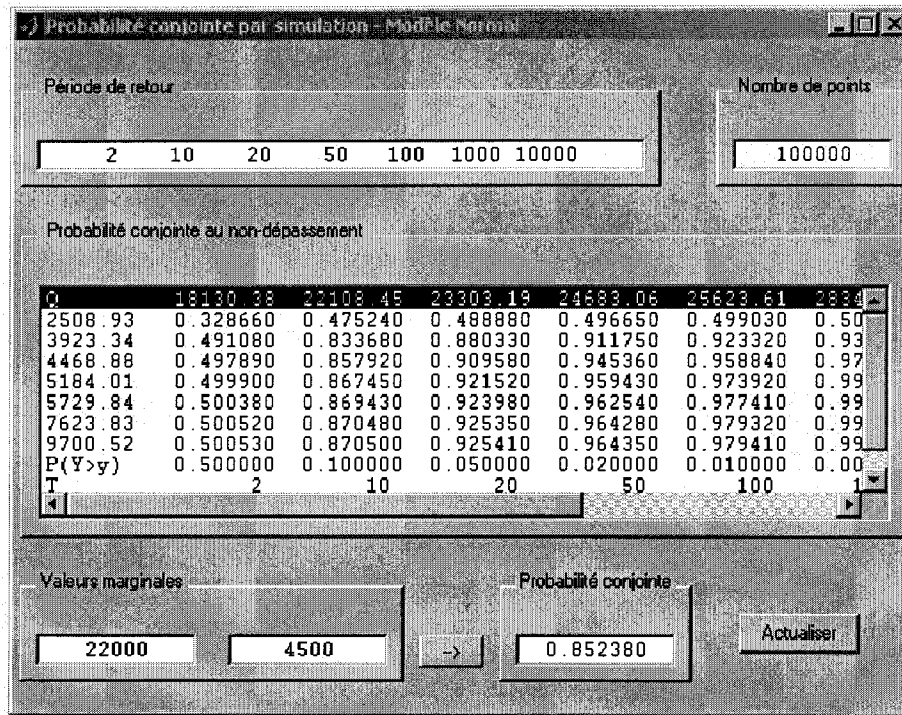


Figure 9- Fenêtre 'Probabilité conjointe' du logiciel RISCHYDRO pour le modèle normal.

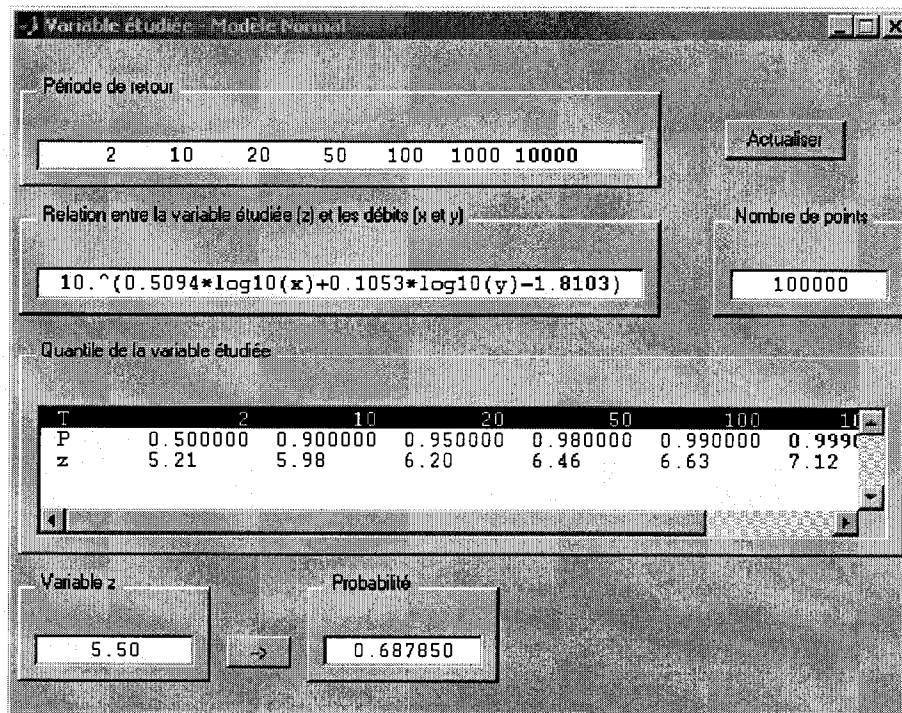


Figure 10- Fenêtre 'Variable étudiée' du logiciel RISCHYDRO pour le modèle normal.

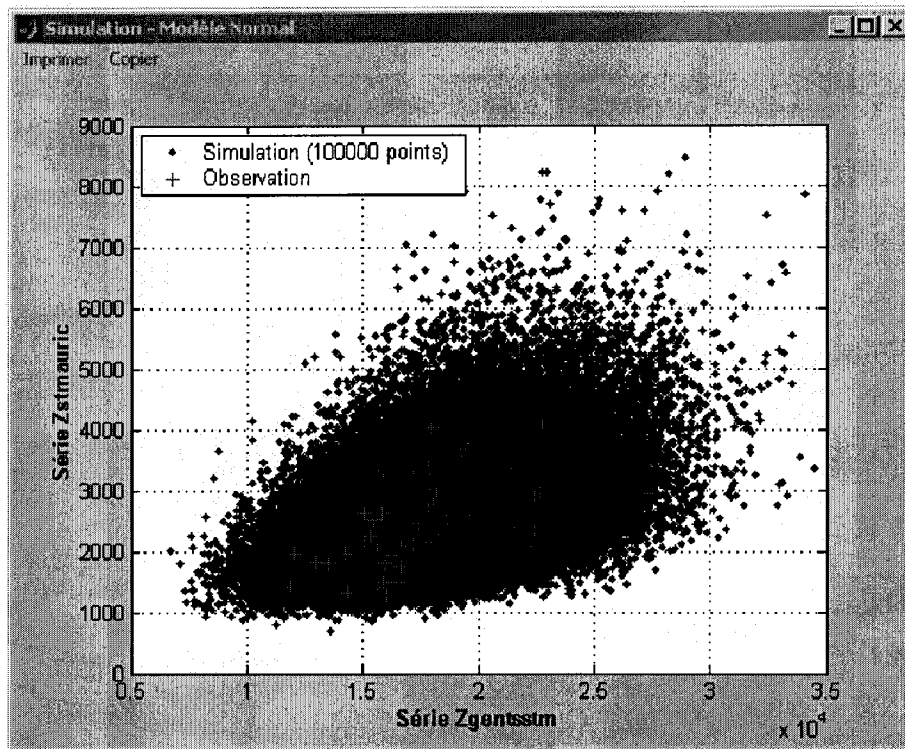


Figure 11- Comparaison des valeurs observées des débits avec des valeurs simulées.

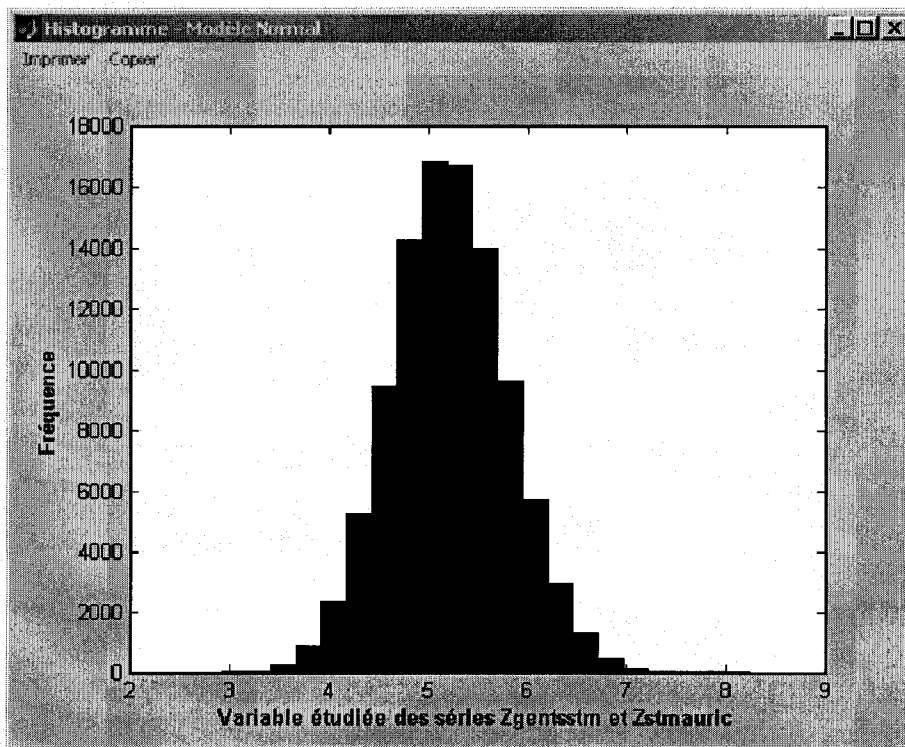


Figure 12- Histogramme de la variable étudiée z pour le modèle normal.

Statistiques - Modèle Normal						
Imprimer Copier		Série Zgentssm		Série Zstmauric		z
	Observation	Simulation	Observation	Simulation	Simulation	
Minimum	11899.2	6662.248	1266.6	703.762	3.049	
Maximum	27455.9	34514.624	5282.7	8448.474	8.107	
Moyenne	18241.889	18282.351	2672.94	2625.329	5.222	
Écart-type	2962.677	3364.098	943.19	776.529	0.584	
Asymétrie	0.231	0.268	0.926	1.011	0.116	
Aplatissement	3.197	3.092	3.597	4.875	3.038	
Coefficient de variation	16.241	18.401	35.287	29.578	11.178	

Figure 13- Statistiques pour les données observées et simulées – Modèle Normal.

Modèle Gamma

Pour faire l'analyse des risques combinés à partir de la distribution gamma bivariée, allez au menu Modèle Gamma. Ce menu vous offre les commandes suivantes : Ajustement, Probabilités Jointes et Simulation. Dans le menu Modèle Gamma, cliquez sur l'option Ajustement et ensuite sur l'option Première Série pour ouvrir la fenêtre 'Ajustement de la loi Gamma à une série'. Cette fenêtre vous offre les options suivantes: Probabilité, Quantile Quantile, Imprimer et Copier:

Probabilité: Cette option affiche la probabilité empirique cumulée et la probabilité théorique cumulée ainsi que les paramètres de la distribution Gamma (Figure 14 et 16).

Quantile Quantile: Cette option crée un graphique quantile-quantile (Figures 15 et 17).

Imprimer: Cette option vous permet d'imprimer la figure en cours.

Copier: Cette option vous permet de copier la figure en cours dans le presse-papier.

La Figure 14 présente les résultats de l'ajustement de la loi gamma au débit maximum annuel de la station fleuve St-Laurant à Gentilly. Cette figure montre que l'on peut considérer l'ajustement d'une distribution gamma à cette série. Le coefficient de détermination (test de Filliben) entre les quantiles théoriques et empiriques est 98% (Figure 15). La Figure 16 présente les résultats de l'ajustement de la loi gamma au débit maximum annuel de la station rivière St-Maurice. Cette figure montre que l'on peut aussi considérer l'ajustement d'une distribution gamma à cette série. Le coefficient de détermination (test de Filliben) entre les quantiles théoriques et empiriques est 98% (Figure 17).

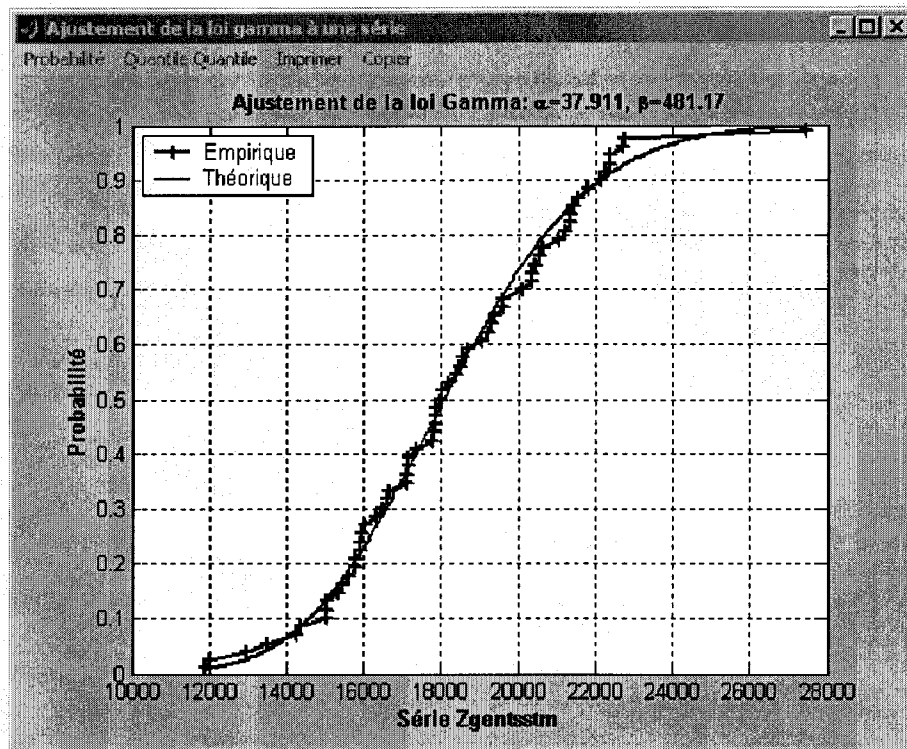


Figure 14- Ajustement de la loi gamma à la série du fleuve Saint-Laurent

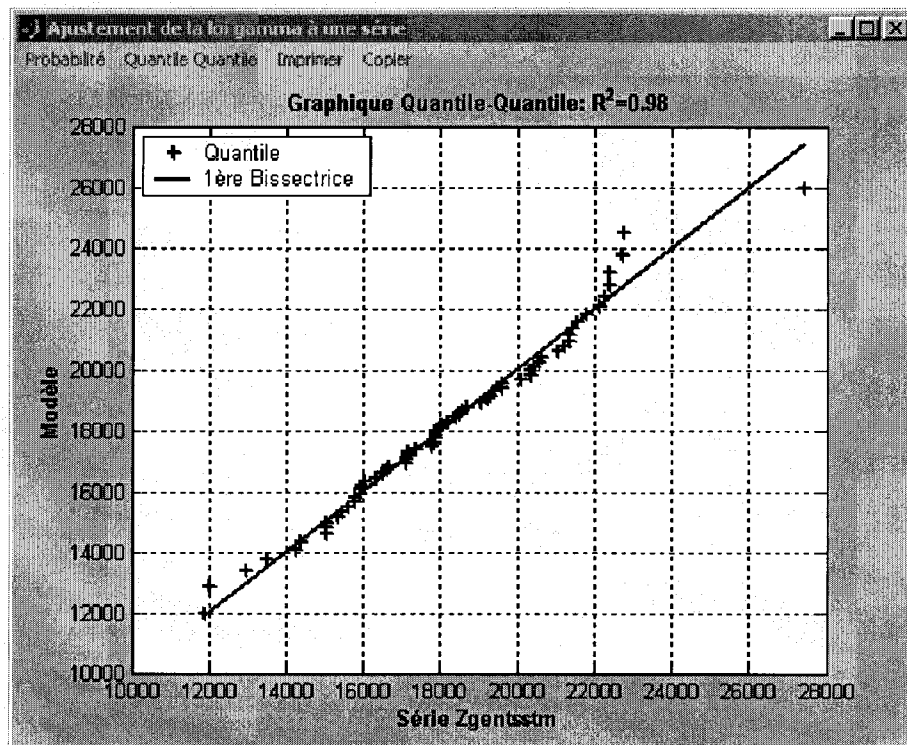


Figure 15- Graphique QQ pour les débits du Saint-Laurent

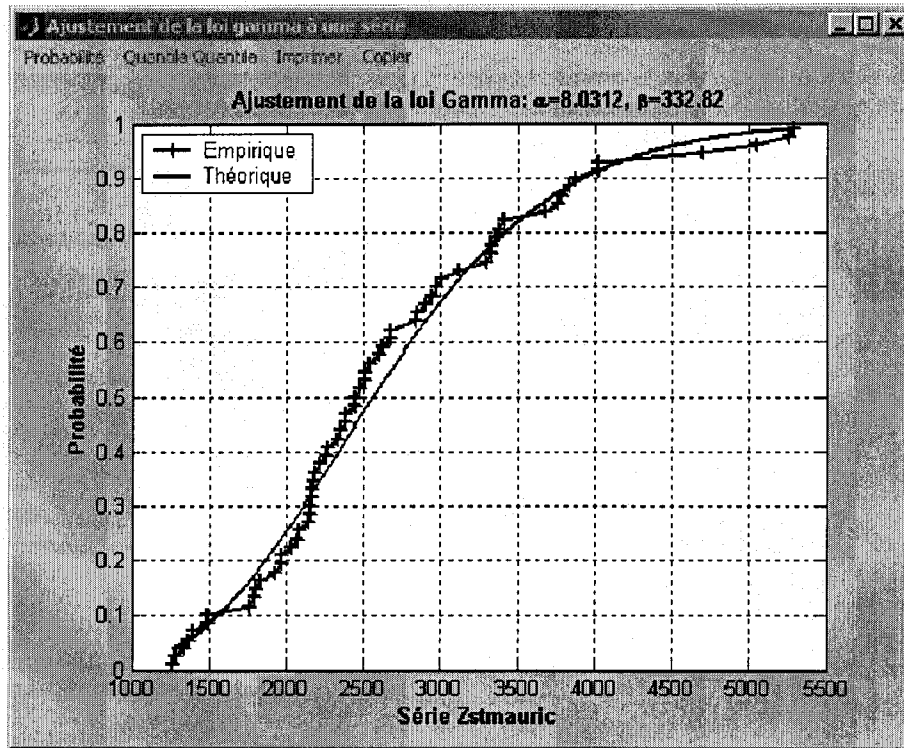


Figure 16- Ajustement de la loi gamma à la série de la rivière Saint-Maurice

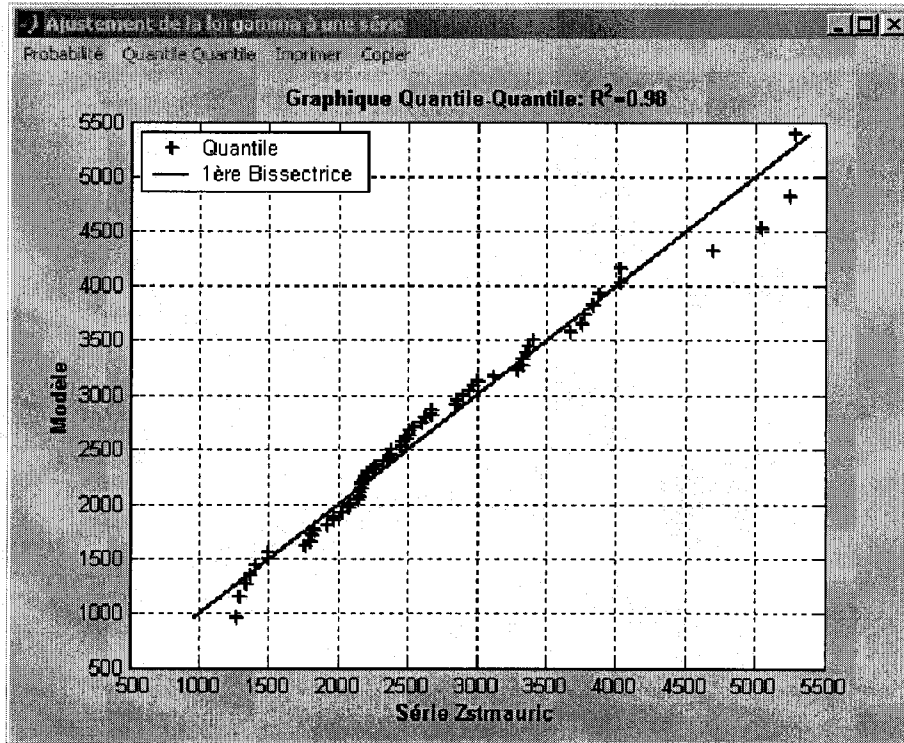


Figure 17- Graphique QQ pour les débits de la rivière Saint-Maurice

Après faire l'ajustement de la loi gamma aux deux séries, choisissez l'onglet Simulation du menu Modèle Gamma. Cet onglet est semblable à celui du menu Modèle Normal.

Pour les périodes de retour 2, 10, 50, 100, 1000 et 10000 ans, nous avons calculé les débits en provenance des stations de Gentilly et St-Maurice ainsi que les probabilités conjointes au non-dépassement en générant des vecteurs aléatoires provenant de copule de Clayton (Marshall et Olkin, 1988) de dimension 100000 (Figure 18).

Remarque
Le temps du calcul pour générer les vecteurs aléatoires provenant de copule de Clayton est généralement plus long que le temps du calcul pour générer les vecteurs aléatoires suivant une distribution Gaussienne bivariée.

En utilisant la relation entre les débits et le niveau, nous avons calculé par simulation les niveaux correspondants aux périodes de retour 2, 10, 50, 100, 1000 et 10000 ans (Figure 19).

L'option 'Graphiques' de l'onglet Simulation vous permet d'afficher les valeurs observées et simulées des débits ainsi que l'histogramme de la variable étudiée (Figures 20 et 21). En cliquant sur l'option 'Statistique', vous pouvez faire apparaître la fenêtre des statistiques pour les données observées et simulées (Figure 22).

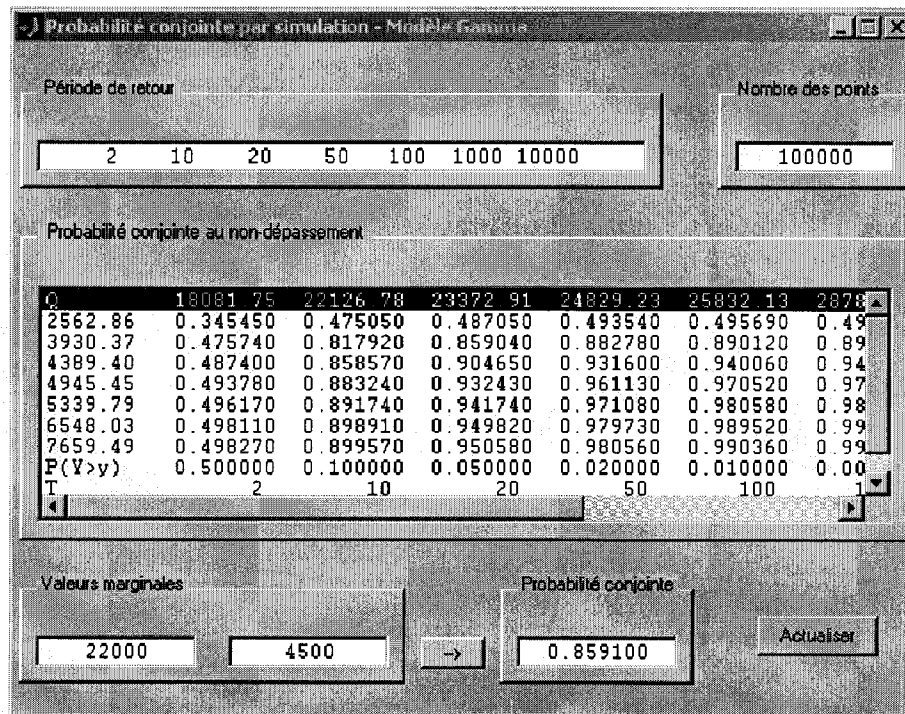


Figure 18- Fenêtre 'Probabilité conjointe' du logiciel RISCHYDRO pour le modèle gamma

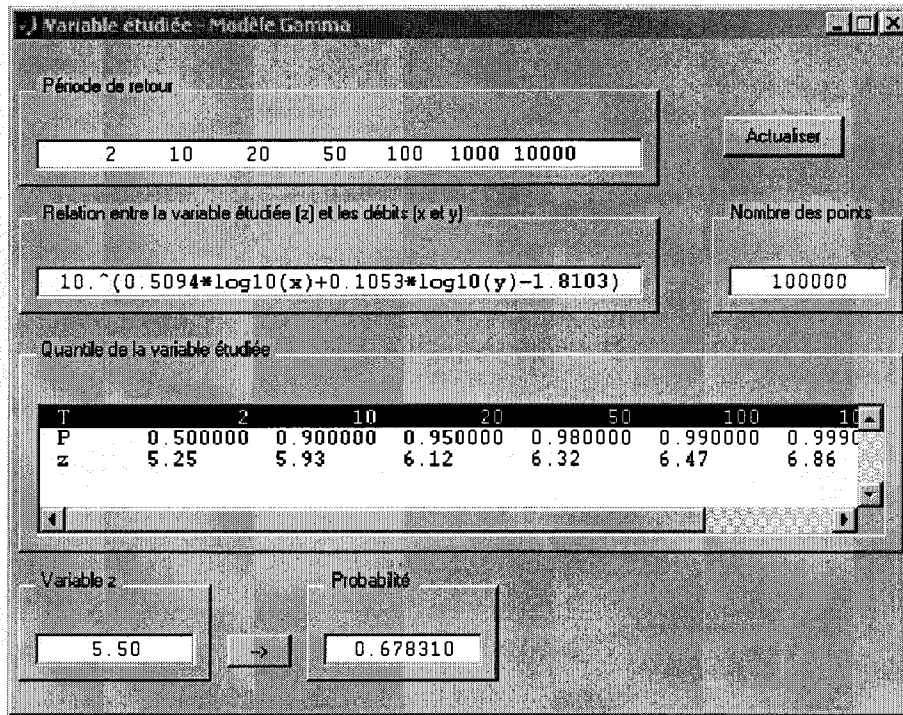


Figure 19- Fenêtre 'Variable étudiée' du logiciel RISCHYDRO pour le modèle gamma

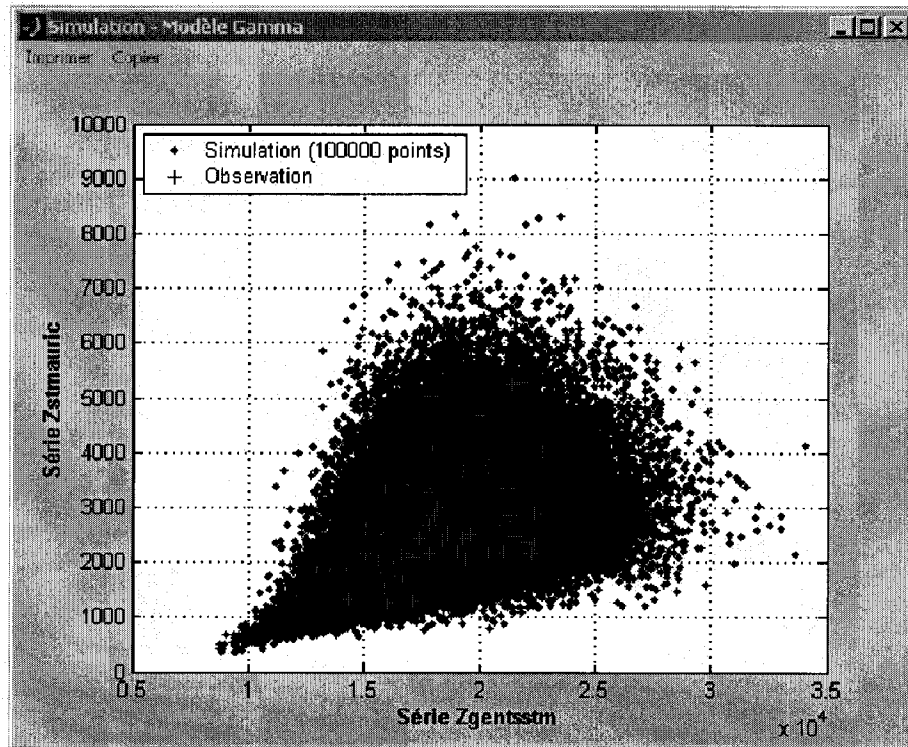


Figure 20- Comparaison des valeurs observées des débits avec des valeurs simulées.

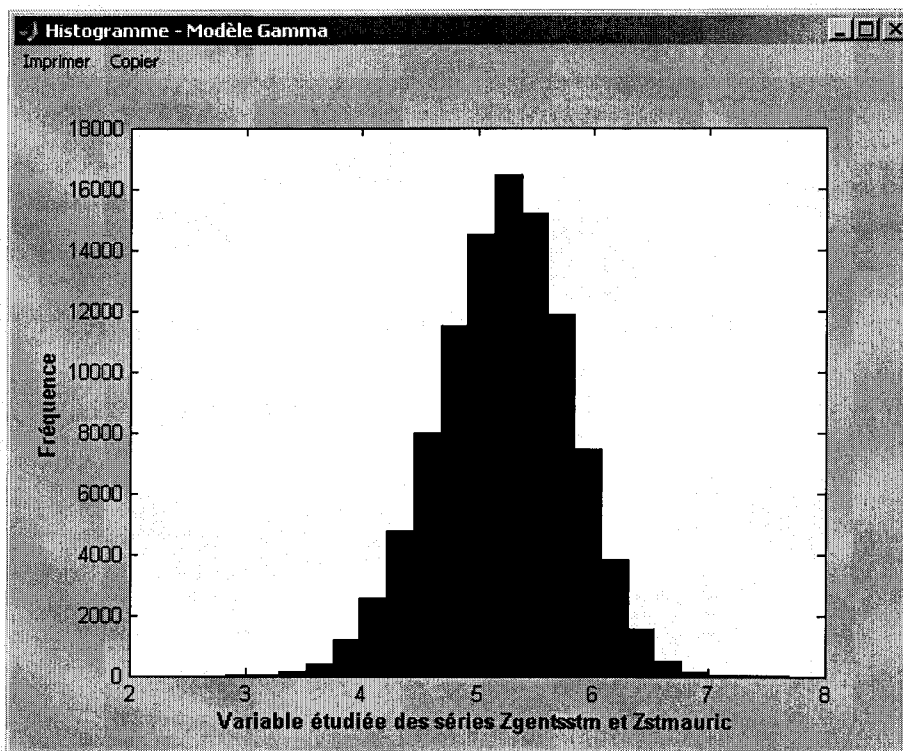


Figure 21- Histogramme de la variable étudiée z pour le modèle gamma.

	Série Zgentsstm		Série Zstmauric		z
	Observation	Simulation	Observation	Simulation	Simulation
Minimum	11899.2	8733.634	1266.6	354.103	2.949
Maximum	27455.9	34038.346	5282.7	9023.318	7.572
Moyenne	18241.889	18245.495	2672.94	2676.546	5.224
Écart-type	2962.677	2954.82	943.19	946.116	0.562
Asymétrie	0.231	0.315	0.926	0.701	-0.146
Aplatissement	3.197	3.121	3.597	3.704	2.933
Coefficient de variation	16.241	16.195	35.287	35.348	10.76

Figure 22- Statistiques pour les données observées et simulées – Modèle Gamma.

AJUSTEMENT DES LOIS MARGINALES

Pour ajuster une distribution de probabilités aux deux séries de données, allez au menu Lois marginales. Ce menu vous offre les commandes suivantes : Première Série et Deuxième Série (Figure 24). Après avoir cliquer sur l'onglet d'une des deux séries de données, une fenêtre intitulée « Lois_marginales » apparaît (Figure 25). Cette fenêtre est composée de 4 parties :

Lois sélectionnées : Cette section permet de choisir les lois statistiques que vous désirez ajuster à la série de données. Les lois disponibles dans RISCHYDRO sont : la loi normale, la loi gamma, la loi lognormale, la loi Gumbel (EV1 ou GEV avec $k=0$), la loi GEV et la loi Burr XII qui selon ses paramètres peut aussi être une loi Weibull ou Pareto.

Critère : Cette section permet de sélectionner le critère à utiliser pour comparer l'ajustement des différentes distributions. Les quatre critères présents dans RISCHYDRO sont : le critère d'information d'Akaike (AIC), le critère d'information bayésien (BIC), le critère d'Anderson-Darling et le critère de Cramer von-Mises.

Best Fit ! : Ce bouton permet d'obtenir les résultats de l'ajustement et des comparaisons. La fenêtre « Résultat » indique la loi sélectionnée par le critère choisi et ses paramètres (Figure 26). Cliquez sur le bouton « Ok » pour fermer la fenêtre « Résultat ».

Fermer : Ce bouton ferme la fenêtre « Lois_marginales ».

Les figures 25 et 26 illustrent un exemple d'ajustement de distributions de probabilités. À la figure 25, on désire comparer l'ajustement des lois normales, gamma, lognormale, Gumbel et GEV aux débits de la série de la station de Gentilly à l'aide du critère d'information bayésien (BIC). En cliquant sur le bouton « Best Fit ! », on obtient les résultats de cette analyse. La figure 26 indique que la loi qui s'ajuste le mieux aux données est la loi gamma de paramètres $a=38,1$ et $b=478,76$.

Remarque

Pour ajuster une distribution de probabilités en particulier, sélectionnez seulement cette distribution dans la section « Lois sélectionnées » et choisissez n'importe lequel des critères de comparaison. De cette façon, en cliquant sur le bouton « Best Fit ! » la loi sélectionnée par RISCHYDRO sera celle que vous aurez choisie.

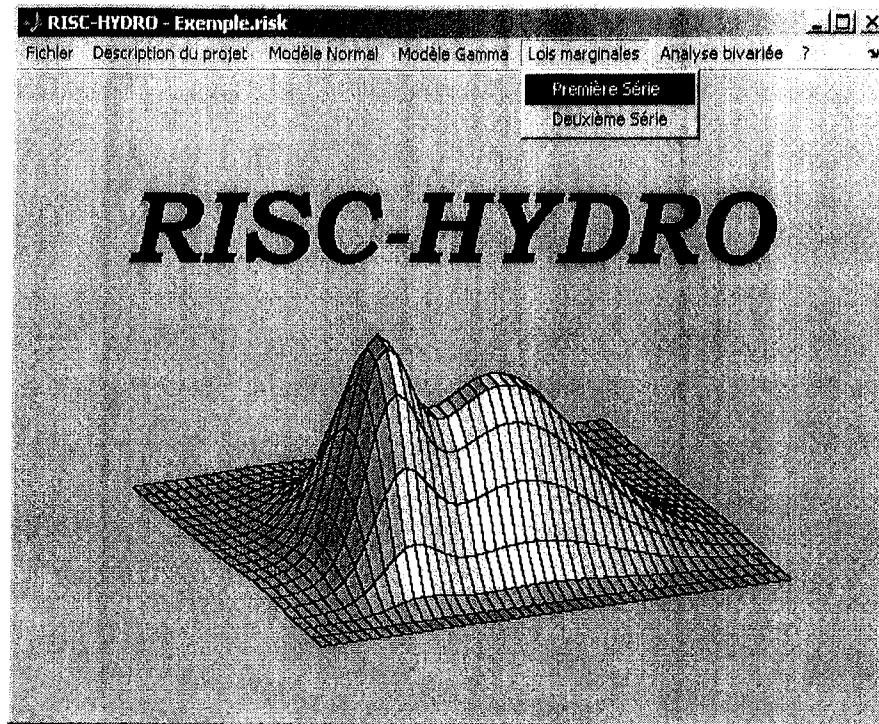


Figure 24 – Menu « Lois marginales » du logiciel RISCHYDRO

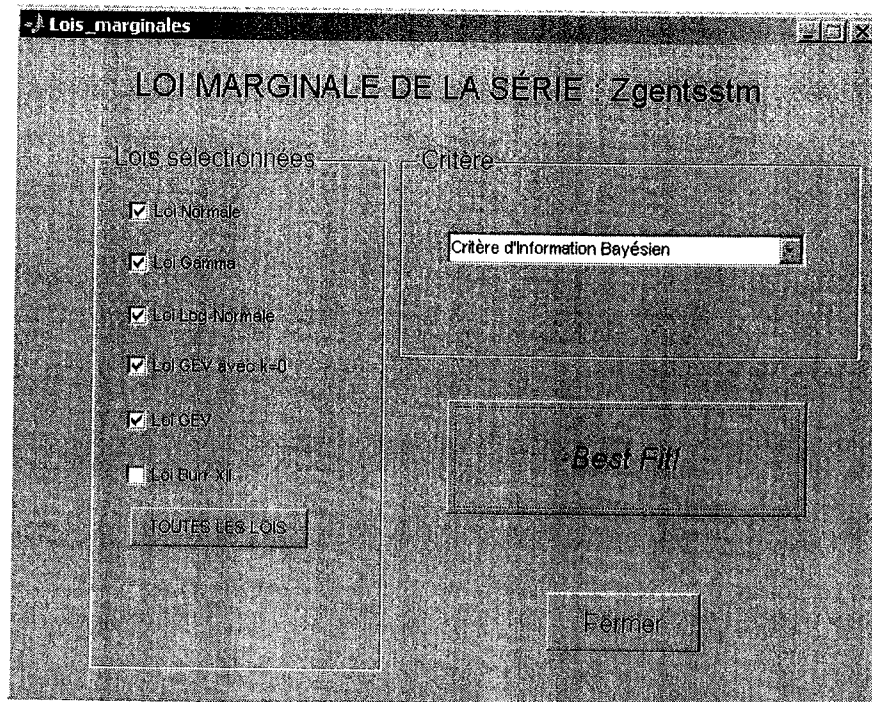


Figure 25 – Fenêtre « Lois_marginales » du logiciel RISCHYDRO

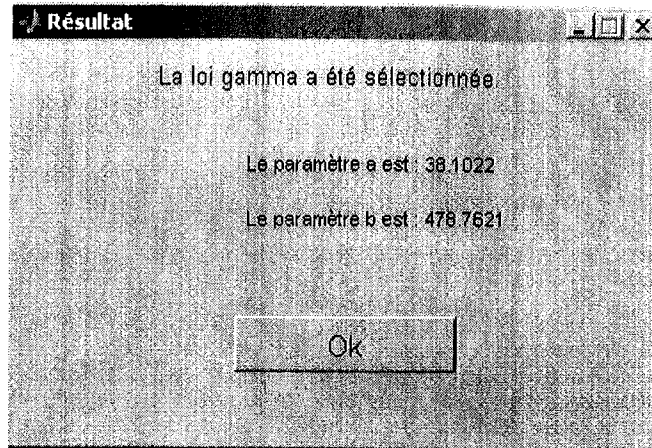


Figure 26 – Résultat de l'ajustement des lois marginales

MODÈLE BIVARIÉ

Le menu « Analyse bivariée » permet de faire l'analyse des risques combinés à partir de copules. Ce menu vous offre les commandes suivantes : Analyse bivariée, Graphiques, Probabilités conjointes et Périodes de retour (Figure 27).

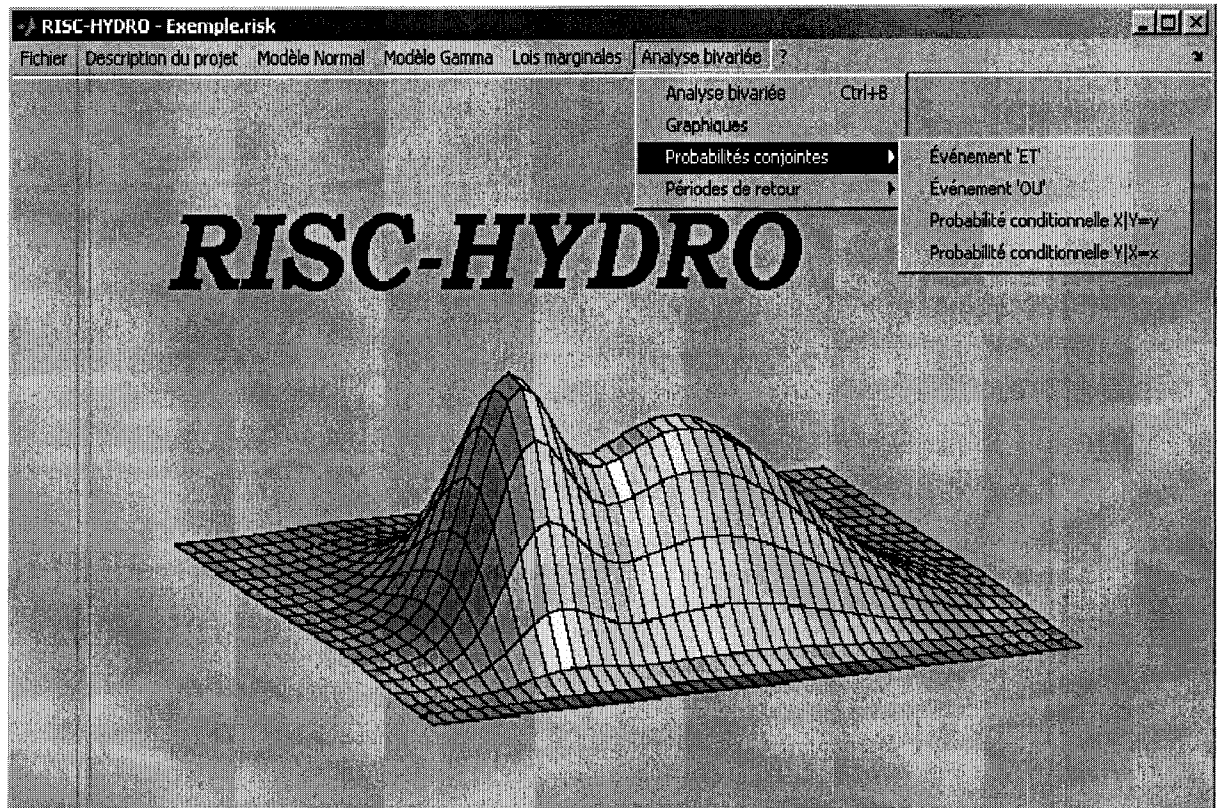


Figure 27 - Menu « Analyse bivariée » du logiciel RISCHYDRO

Dans le menu « Analyse bivariée », cliquez d'abord sur l'onglet « Analyse bivariée » pour ouvrir la fenêtre « Analyse_bivariee » (Figure 28). Cette fenêtre se compose des sections et des boutons suivants :

Type de copules : Cette section permet de choisir les copules à ajuster et à comparer. Les copules disponibles dans RISCHYDRO sont la copule de Gumbel, la copule de Clayton et la copule de Frank.

Critère à utiliser : Cette section permet de choisir le critère à utiliser pour comparer l'ajustement des copules sélectionnées dans la section précédente. Deux critères peuvent être employés : la fonction K et la statistique S.

Best Fit! : Ce bouton permet de comparer les copules sélectionnées à l'aide du critère indiqué. Les résultats du choix de la copule qui s'ajuste le mieux aux données s'affichent dans la fenêtre « Résultat » (Figure 29). Les résultats sont la meilleure copule et son paramètre.

Plot : Ce bouton donne le graphique des probabilités cumulatives des données de la deuxième série en fonction des probabilités cumulatives des données de la première série (Figure 30).

Fermer : Ce bouton ferme la fenêtre « Analyse_bivariee ».

Les figures 28 et 29 représentent les résultats pour notre exemple avec les stations de Gentilly et du St-Maurice. On compare l'ajustement des copules de Gumbel, Clayton et Frank à l'aide de la statistique S à la figure 28. En cliquant sur le bouton « Best Fit! », on obtient que la copule qui s'ajuste le mieux aux données est la copule de Clayton avec paramètre $\alpha=1,2602$ comme le montre la figure 29.

Remarque

Pour ajuster une copule en particulier, sélectionnez seulement cette copule dans la section « Type de copules » et choisissez n'importe lequel des critères de comparaison. De cette façon, en cliquant sur le bouton « Best Fit! » la copule sélectionnée par RISCHYDRO sera celle que vous aurez choisie.

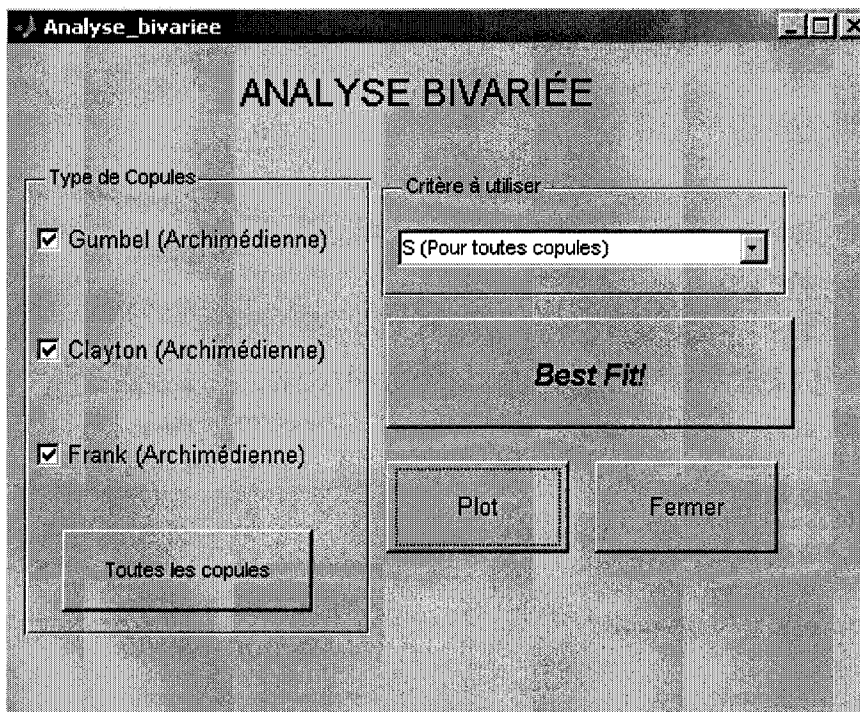


Figure 28 – Ajustement des copules de Gumbel, Clayton et Frank à l'aide de la statistique S

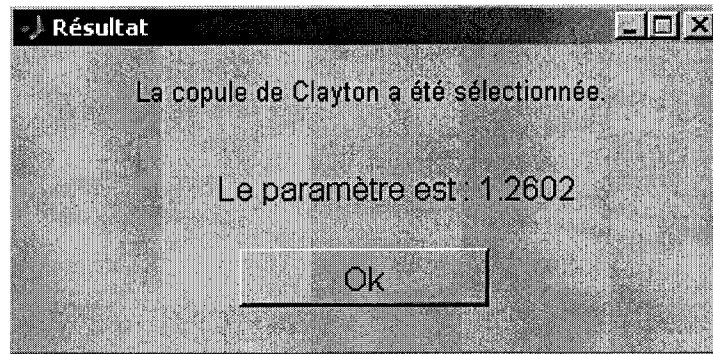


Figure 29 – Résultat de l'ajustement des copules

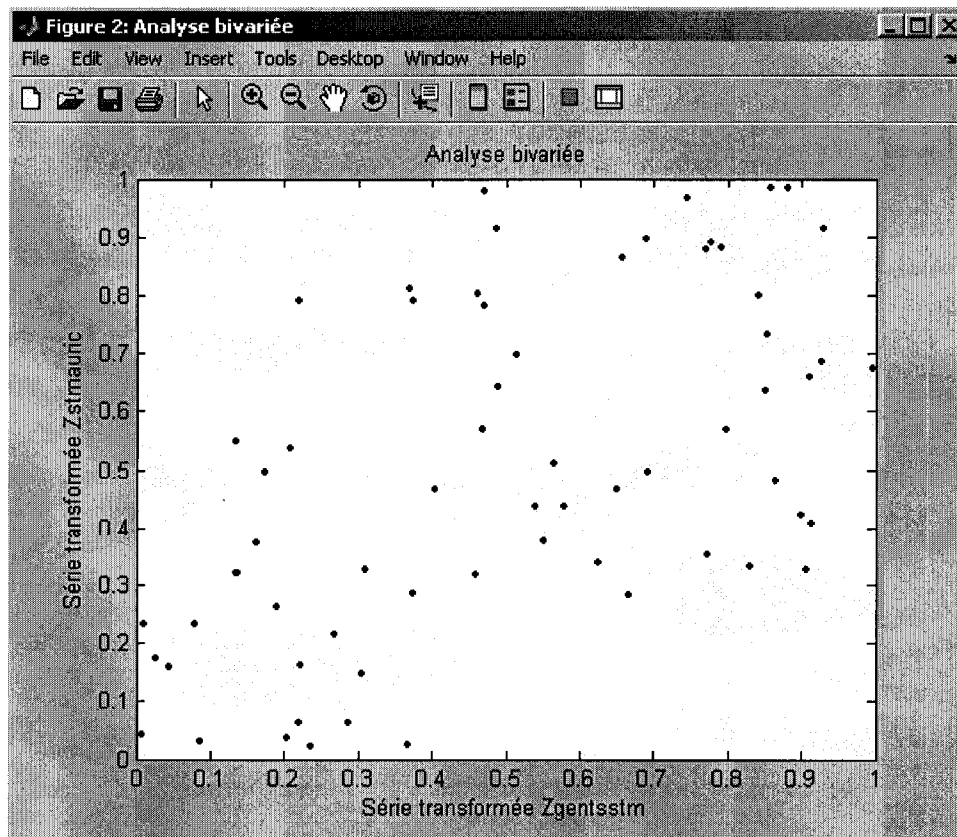


Figure 30 – Probabilités cumulatives des données de la deuxième série en fonction des probabilités cumulatives des données de la première série

Après avoir ajuster une copule, cliquez sur l'onglet « Graphiques » du menu « Analyse bivariable ». La fenêtre « Graphiques – Copules » (Figure 31) vous rappelle d'abord quelle est la copule qui s'ajuste le mieux aux données étudiées. Ensuite, les trois options suivantes s'offrent à vous :

Simulation : Cette option vous permet de simuler la copule de votre choix ajustée à vos données (Frank ou Clayton avec leur paramètre respectif) et de superposer sur un graphique les résultats de cette simulation (5000 points) et les données observées.

Fonction CDF : Cette option vous permet de tracer graphiquement la fonction de distribution cumulative (et ses courbes de niveau) de la copule de votre choix ajustée à vos données (Gumbel, Clayton ou Frank avec leur paramètre respectif).

Fonction PDF : Cette option vous permet de tracer graphiquement la fonction de distribution de probabilités (et ses courbes de niveau) de la copule de votre choix ajustée à vos données (Gumbel, Clayton ou Frank avec leur paramètre respectif).

La figure 32 représente les résultats d'une simulation de la copule de Clayton avec $\alpha=1,2602$ et de la superposition de cette simulation et des données observées. La fonction de distribution cumulative de la copule de Clayton avec $\alpha=1,2602$ et ses courbes de niveau sont illustrées aux figures 33 et 34. La fonction de distribution de probabilités de la copule de Clayton avec $\alpha=1,2602$ et ses courbes de niveau sont illustrées aux figures 35 et 36.

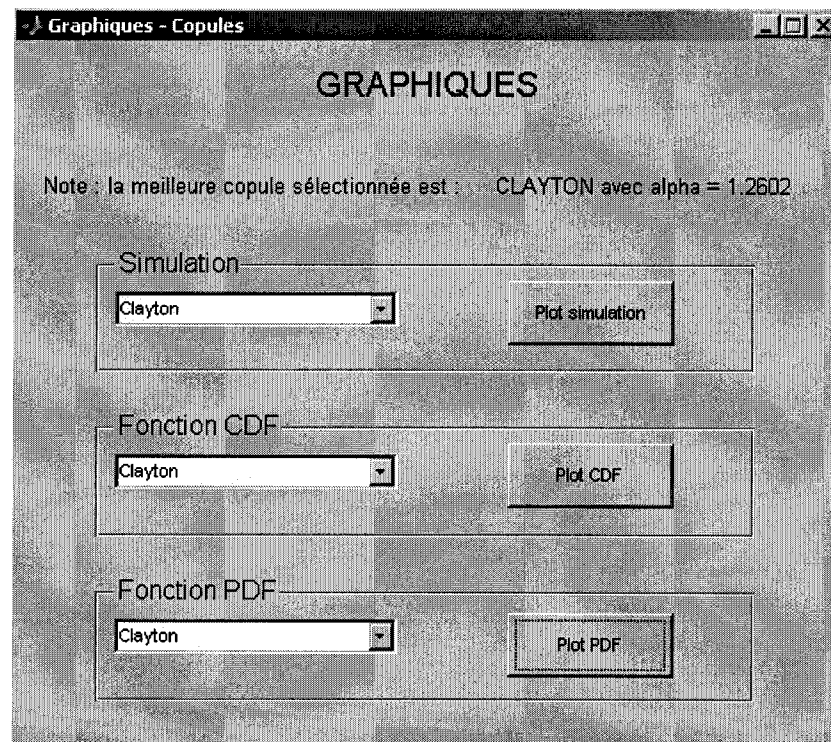


Figure 31 – Fenêtre « Graphiques – Copules » du logiciel RISCHYDRO

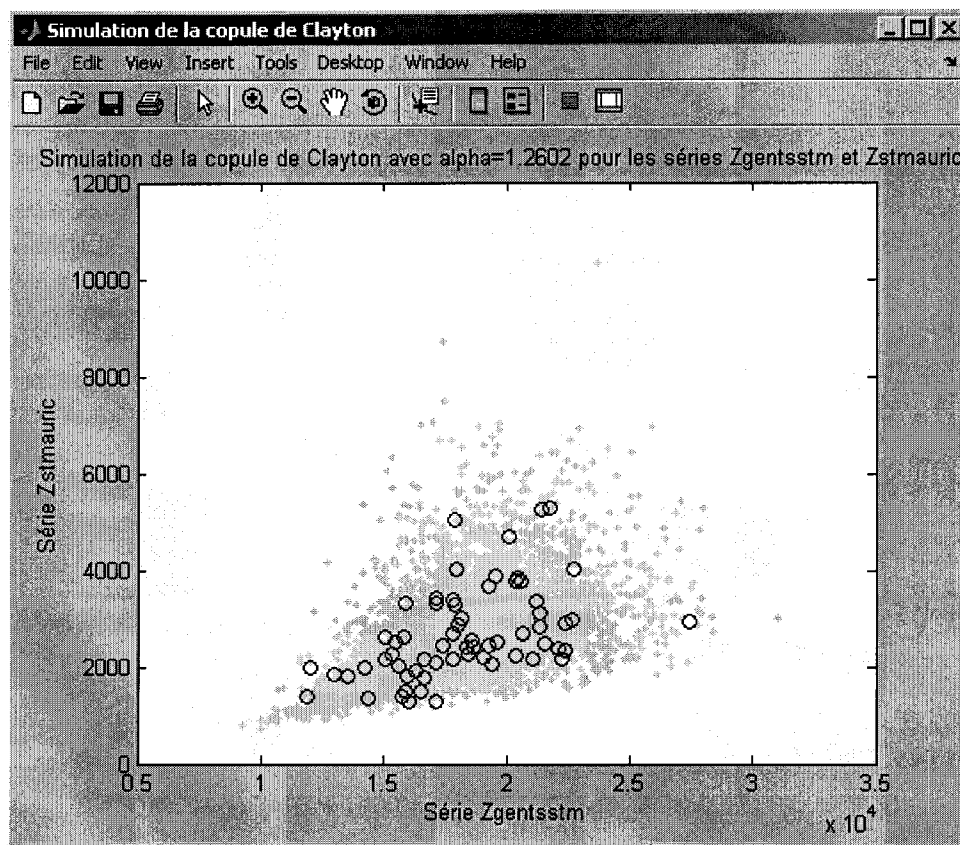


Figure 32 – Superposition de la simulation de la copule de Clayton avec $\alpha=1,2602$ et des données observées

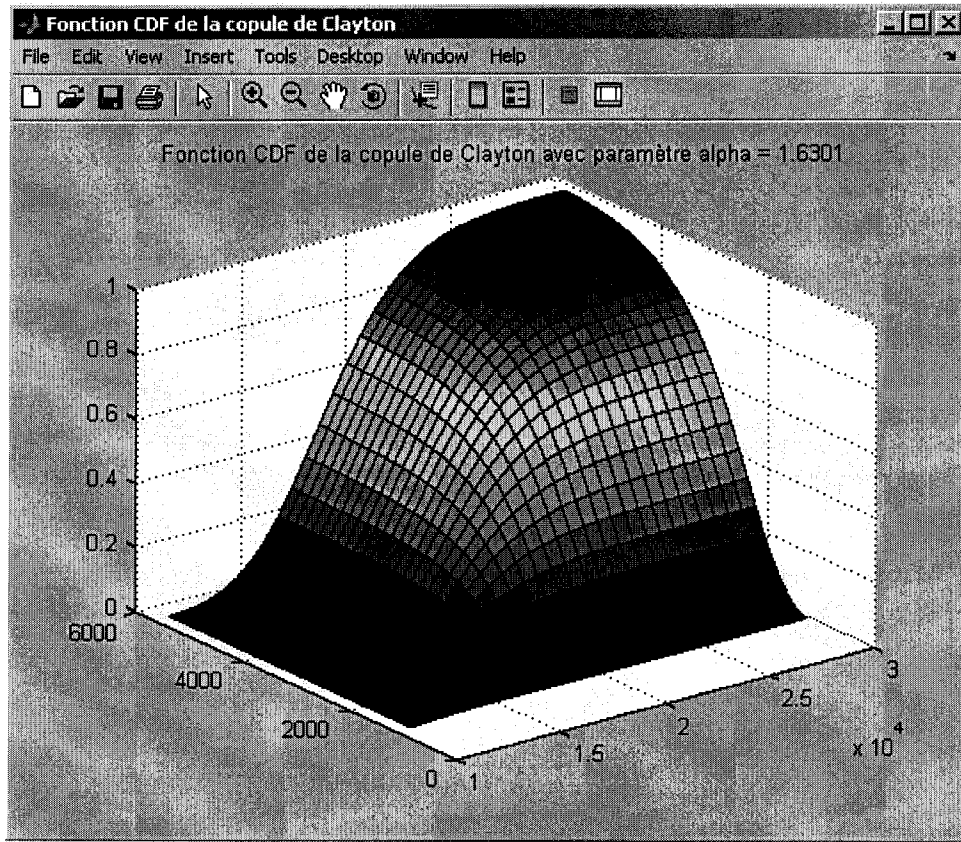


Figure 33 – Fonction de distribution cumulative de la copule de Clayton avec $\alpha=1,2602$

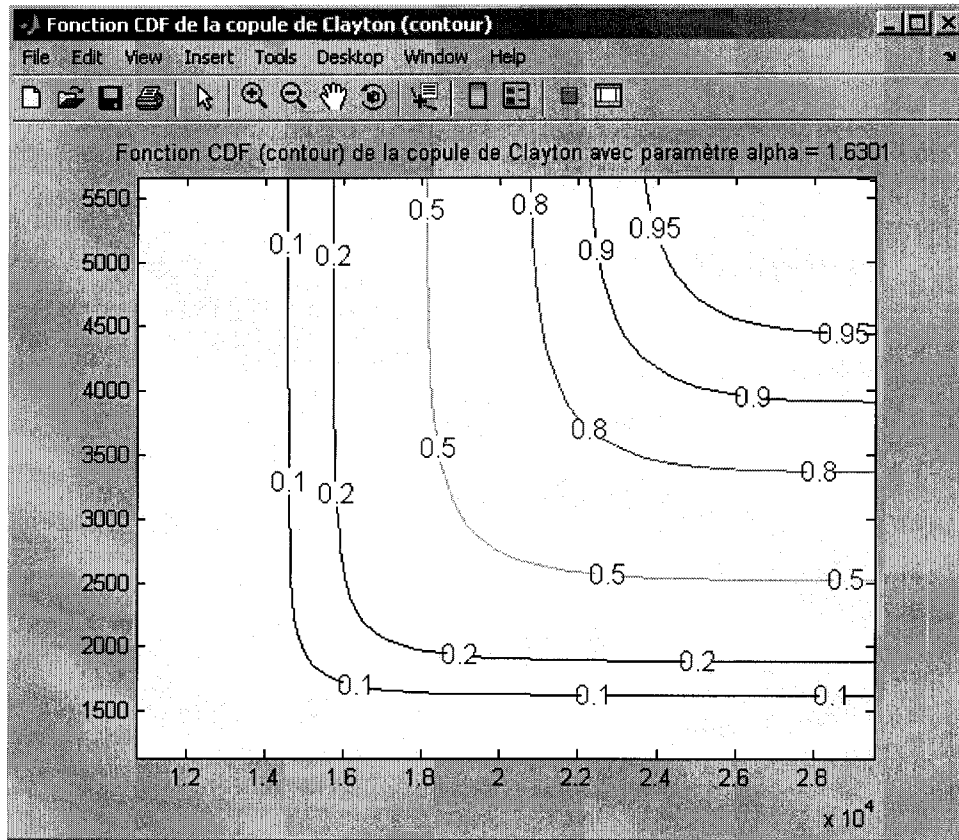


Figure 34 – Courbes de niveau de la fonction de distribution cumulative de la copule de Clayton avec $\alpha=1,2602$

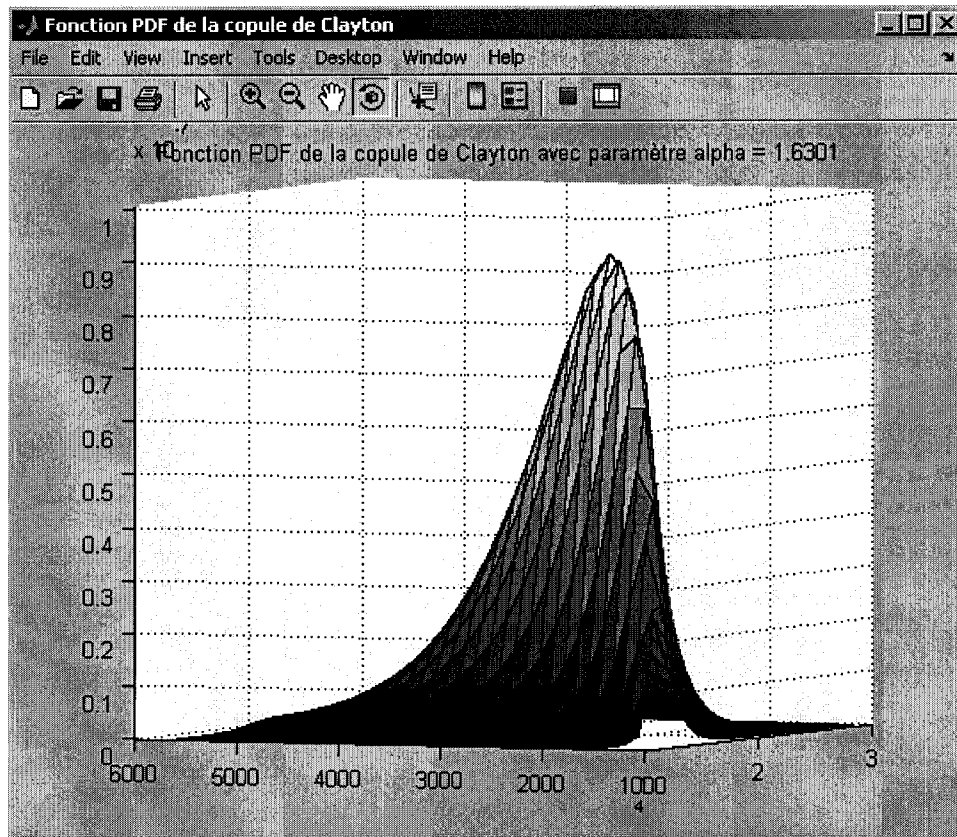


Figure 35 – Fonction de distribution de probabilités de la copule de Clayton avec $\alpha=1,2602$

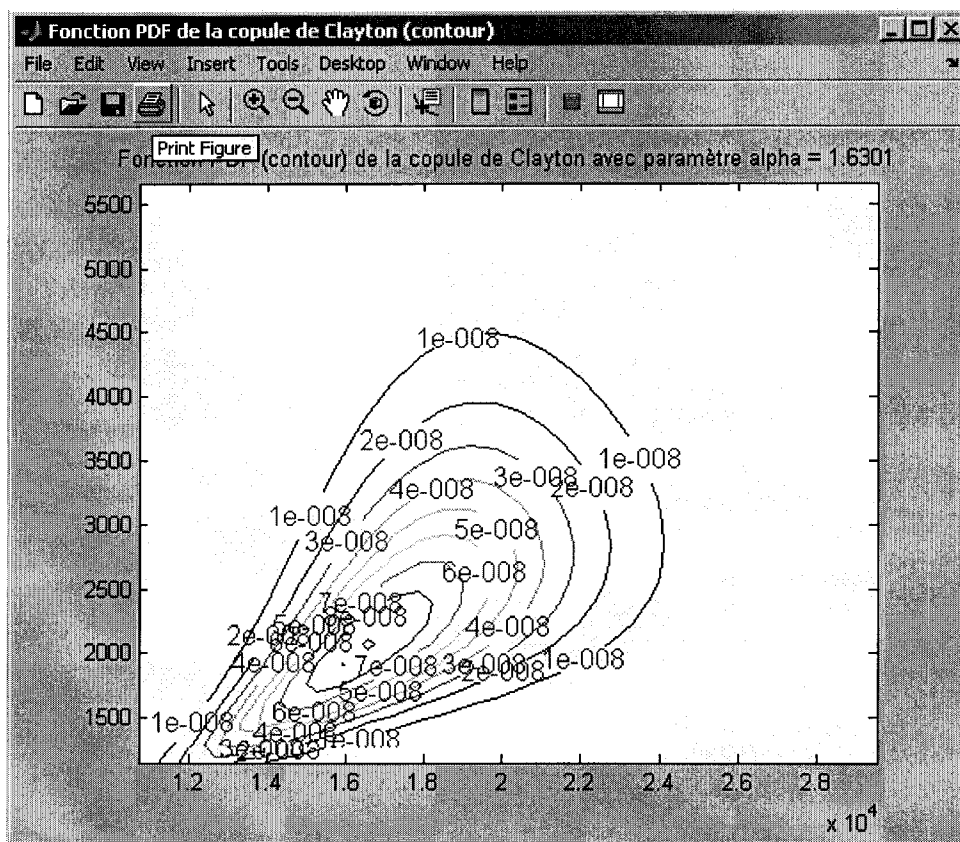


Figure 36 – Courbes de niveau de la fonction de distribution de probabilités de la copule de Clayton avec $\alpha=1,2602$

Pour obtenir les probabilités conjointes et conditionnelles de différents scénarios, cliquez sur l'onglet « Probabilités conjointes » puis sur l'onglet du scénario de votre choix. En cliquant sur ces onglets, on obtient les tableaux des probabilités conjointes au dépassement pour l'événement sélectionné ou des probabilités conditionnelles pour cet événement pour des périodes de retour de 2, 5, 10, 20, 50, 100 et 1000 ans (Figure 37). Il est également possible, dans ces tableaux, de calculer ces probabilités pour des valeurs de X et de Y que vous spécifierez. Il suffit d'entrer les valeurs souhaitées dans la section « Valeurs marginales » et d'appuyer ensuite sur le bouton « => ». Le résultat s'affichera dans la section « Probabilité conditionnelle ». Les scénarios pouvant être étudiés dans RISCHYDRO sont :

Événement 'ET' : Ce scénario calcule les probabilités conjointes $P(X>x \text{ et } Y>y)$.

Événement 'OU' : Ce scénario calcule les probabilités conjointes $P(X>x \text{ ou } Y>y)$.

Événement 'X|Y=y' : Ce scénario calcule les probabilités conditionnelles $P(X>x | Y=y)$.

Événement 'Y|X=x' : Ce scénario calcule les probabilités conditionnelles $P(Y>y | X=x)$.

La figure 37 représente les probabilités conditionnelles $P(X > x | Y = y)$ de notre exemple où X et Y sont correspondent respectivement à la première et à la deuxième série de données.

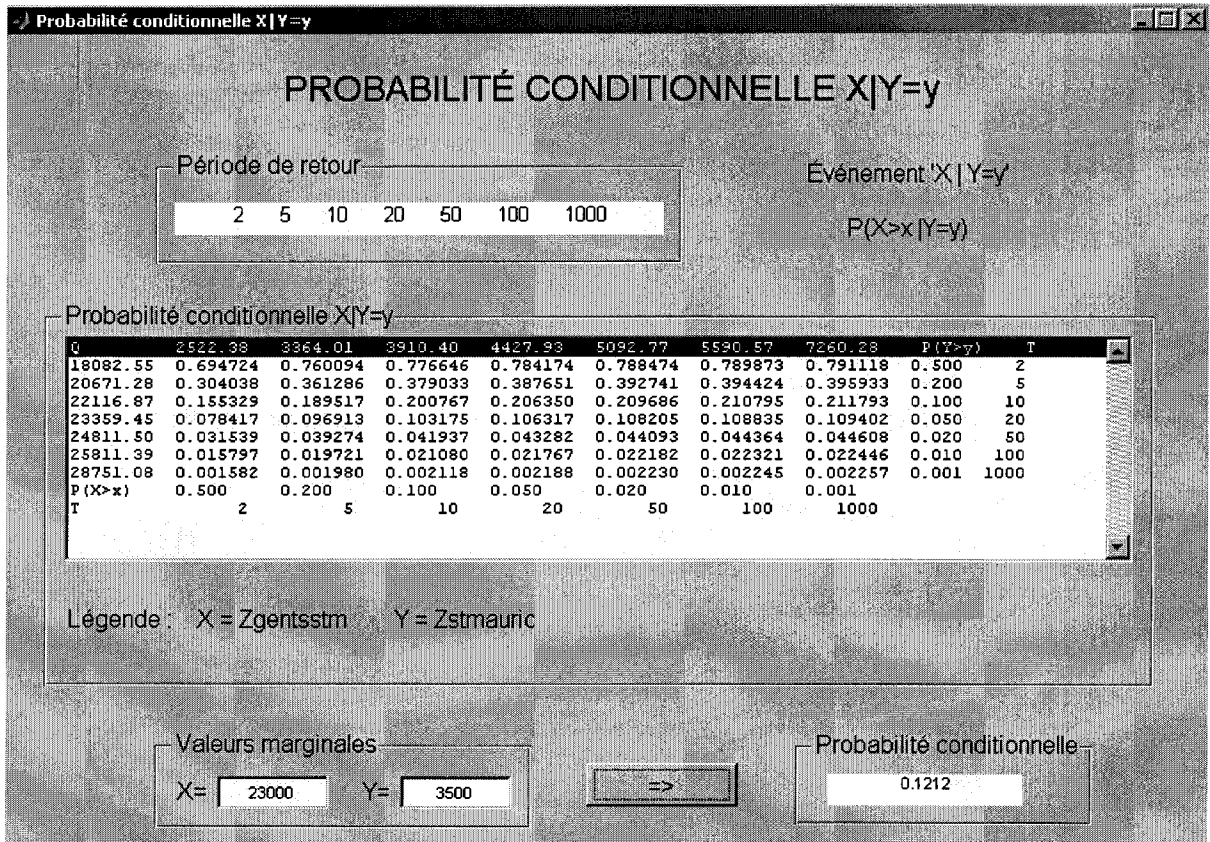


Figure 37 – Probabilités conditionnelles $P(X > x | Y = y)$

Pour obtenir les périodes de retour conditionnelles de l'événement 'X|Y=y' ou de l'événement 'Y|X=x', cliquez sur l'onglet « Périodes de retour » puis sur l'onglet représentant l'événement de votre choix du menu « Analyse bivariée ». Les périodes de retour conditionnelles voulues apparaissent dans un tableau pour des périodes de retour de 2, 5, 10, 20, 50, 100 et 1000 ans (Figure 38). Il est également possible de calculer une période de retour conditionnelle pour des valeurs de X et de Y que vous spécifierez. Il suffit d'entrer les valeurs souhaitées dans la section « Valeurs marginales » et d'appuyer ensuite sur le bouton « => ». Le résultat s'affichera dans la section « Période de retour cond. ».

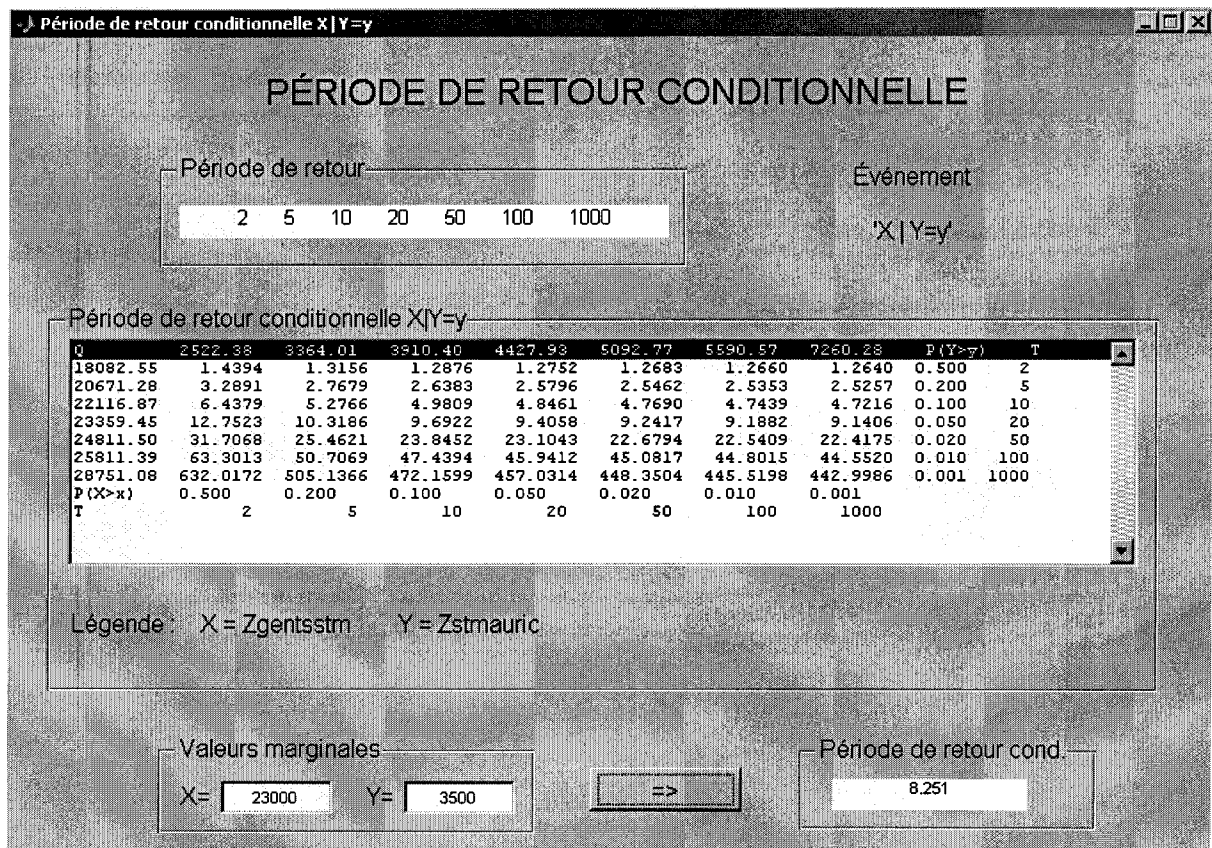


Figure 38 – Périodes de retour conditionnelles de l'événement 'X|Y=y'

Enregistrer un projet

Cliquez sur Enregistrer dans le menu Fichier pour enregistrer le projet sur lequel vous travaillez. Cette action fait apparaître la fenêtre Enregistrer. Tapez le nom de projet et cliquez sur le bouton Enregistrer. L'option 'Enregistrer sous' vous permet d'enregistrer une copie d'un projet. Pour ce faire, cliquez sur 'Enregistrer sous' dans le menu Fichier. Dans la zone 'nom de fichier', tapez un nouveau nom pour le projet et cliquez sur Enregistrer.

Ouvrir un projet

Dans le menu Fichier, cliquez sur Ouvrir. La fenêtre Ouvrir permet d'afficher la boîte de dialogue d'ouverture de projet. Seuls les fichiers comportant l'extension .risk apparaissent dans la fenêtre. Choisissez un fichier et cliquez sur le bouton ouvrir. Quand vous ouvrez un projet, le projet précédent sera automatiquement fermé.

Exportation des résultats

Dans le menu Fichier, cliquez sur Exporter. Dans la zone 'Nom de fichier', tapez un nom de fichier, puis cliquez sur Enregistrer. Cette action produit un document au format fichier texte qui contient les résultats des analyses de RISCHYDRO.

Quitter RISCHYDRO

Pour quitter RISCHYDRO, cliquez sur Quitter dans le menu Fichier ou appuyez sur Ctrl-Q.

Références

- Bobée, B., et Ashkar, F., 1991. The Gamma family and derived distributions applied in hydrology. Water Resources Publications, USA.
- Box, G.E.P., et Cox, D.R., 1964, An analysis of transformations. Journal of the Royal Statistical Society, Ser. B, 127: 211-252.
- Bruneau, P., Ashkar, F., et Bobée, B., 1994. SMPLNORM: un modèle simple pour obtenir les probabilités conjointes de deux débits et le niveau qui en dépend. Revue canadienne de génie civil, 21(5), 883-895.
- Hald, A., 1967. Statistical theory with engineering applications, John Wiley & Sons, Inc., New York, 602.
- Genest,C. et Rivest,L., 1993, Statistical inference procedures for bivariate Archimedean copulas, Journal of the American Statistical Association, vol.88, pp.1034-1043.
- Laio, F., 2004. Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters, Water Resources Research, 40.
- Linnet, K., 1988. Testing normality of transformed data. Applied Statistics, 37, 180-186.
- Salvadori, G., 2004, Bivariate return period via 2-Copulas, Statistical Methodology 1,129-144.