

---

# ANALYSE FRÉQUENTIELLE HYDROLOGIQUE :

## LOGICIEL HYFRAN-PLUS (VERSION-V2.1)

---

Salaheddine EL ADLOUNI<sup>(1)</sup>

Bernard BOBÉE<sup>(2)</sup>

<sup>(1)</sup> Professeur, Département de Mathématiques  
et statistique, Université de Moncton

<sup>(2)</sup> Professeur Émérite  
INRS-ETE

Version du 13 novembre 2014

**Note : Les articles cités dans la liste des références (page 61) et marqués par \* sont disponibles lors de l'installation du logiciel**

Citation: El Adlouni, S. and B. Bobée (2014). Analyse Fréquentielle avec le logiciel HYFRAN-PLUS.  
Guide d'utilisateur disponible avec la version  
Démon. <http://www.wrpllc.com/books/HyfranPlus/indexhyfranplus3.html>

## TABLE DES MATIERES

---

<b>1. Menu Principal</b> .....	6
1.1. Interface .....	6
1.2. Fichier .....	7
1.3. Édition .....	8
1.4. Échantillon .....	8
1.5. Système d'aide à la décision (SAD) .....	9
1.6. Ajustement .....	11
1.7. Graphique .....	13
1.8. Affichage .....	14
1.9. Fenêtre .....	15
<b>2. Tutorial</b> .....	15
2.1. Saisie des données et étude des caractéristiques statistiques de l'échantillon aléatoire .....	16
2.1.1 Description .....	17
2.1.2. Données .....	18
a) <i>Tableur intégré au logiciel</i> .....	18
- <i>Données désactivées</i> .....	20
- <i>Tri des données</i> .....	22
- <i>Transformation des données</i> .....	22
b) <i>Importation des fichiers de données</i> .....	23
c) <i>Presse-papier</i> .....	24
2.1.3. Statistiques de base .....	24
2.1.4. Tests d'hypothèse .....	25
2.1.5. Graphiques .....	29
2.2. Système d'aide à la décision (SAD) .....	33
2.3. Ajustement d'une distribution statistique à une série de données .....	49
2.3.1. Choix de la distribution statistique la plus adéquate .....	49
<i>Comparaison graphique</i> .....	49
<i>Critères d'information</i> .....	50
2.3.2. Ajustement .....	52
<i>Résultats de l'ajustement</i> .....	52
<i>Graphique</i> .....	53
<i>Autre période de retour</i> .....	54
<i>Adéquation</i> .....	55

*Caractéristiques statistiques de la distribution ajustée* ..... 56  
*Discordance* ..... 57  
**Références** ..... 61  
**Annexe 1 : Séries d'observations** ..... 62  
**Appendice A : Intervalles de confiance asymptotiques des quantiles** ..... 70  
**Appendice B : Classification des distributions** ..... 71

### Liste des Tableaux (Numéros de page à ajouter)

Tableau A.1 : Données observées de la série 1	-----	62
Tableau A.2 : Données simulées de la série 2	-----	64
Tableau A.3 : Données transformées de la série 2	-----	66
Tableau A.4 : Données simulées de la série 3	-----	68

### Liste des Figures (Numéros de page à ajouter)

Figure 1 : Fenêtre d'ouverture du logiciel HYFRAN-PLUS	-----	6
Figure 2 : Menu «Fichier»	-----	7
Figure 3 : Menu «Édition»	-----	8
Figure 4 : Menu échantillon	-----	9
Figure 5 : Menu SAD	-----	10
Figure 6 : Diagramme du SAD	-----	11
Figure 7 : Menu Ajustement	-----	12
Figure 8 : Menu graphique	-----	13
Figure 9 : Menu Affichage	-----	14
Figure 10 : Menu Fenêtre	-----	15
Figure 11 : Création d'un nouveau projet	-----	16
Figure 12 : exemple de description du projet Harricana	-----	18
Figure 13 : Tableur intégré du logiciel (projet Harricana)	-----	19
Figure 14 : tableur intégré pour une donnée désactivée (ligne 14 de l'exemple du projet Harricana)	----	20
Figure 15 : exemple de graphique correspondant à une donnée désactivée (projet Harricana)	-----	21
Figure 16 : Décision d'inclure et d'exclure une donnée	-----	22
Figure 17 : transformation des données	-----	23
Figure 18 : Statistiques de base de l'échantillon (projet Harricana)	-----	25
Figure 19 : Test d'indépendance : projet Harricana	-----	26
Figure 20 : Test de stationnarité (projet Harricana)	-----	27
Figure 21 : Test d'homogénéité à l'échelle annuelle	-----	28
Figure 22 : Test d'homogénéité à l'échelle saisonnière (projet Harricana)	-----	29
Figure 23 : Observation sur papier de probabilité normal (projet Harricana)	-----	30
Figure 24 : Histogramme des observations classées par valeur (projet Harricana)	-----	31
Figure 25 : Histogramme des observations classées par mois (projet Harricana)	-----	32
Figure 26 : Courbe chronologique (données annuelles)	-----	32
Figure 27 : Diagramme (Cv, Cs) préalable à l'utilisation du test de log-normalité (série 1)	-----	34
Figure 28 : Décision sur l'utilisation du test de log-normalité (série 1)	-----	35
Figure 29 : Graphique log-log (série 1)	-----	36
Figure 30 : Décision pour le test log-log (série 1)	-----	37
Figure 31 : Rapport de Hill (série 1)	-----	38
Figure 32 : Statistique de Jackson (série 1)	-----	38
Figure 33 : Diagramme (Cv, Cs) préalable à l'utilisation du test de log-normalité (série 2)	-----	39
Figure 34 : Décision sur l'utilisation du test de log-normalité (série 2)	-----	40
Figure 35 : Décision du test de Jarque-Bera (série 2)	-----	41

Figure 36 : représentation graphique de l'ajustement de la série 2 par la distribution Log-Normale	----42
Figure 37 : Diagramme (Cv, Cs) préalable à l'utilisation du test de log-normalité (série 3)	-----43
Figure 38 : Décision sur l'utilisation du test log-normalité (série 3)	-----44
Figure 39 : Diagramme log-log (série 3)	-----45
Figure 40 : Décision du graphique log-log (série 3)	-----45
Figure 41 : Fonction Moyenne des Excès (série 3)	-----46
Figure 42 : Décision pour la Fonction Moyenne des Excès (série 3)	-----47
Figure 43: Rapport de Hill (série 3)	-----48
Figure 44 : Statistique de Jackson (série 3)	-----48
Figure 45 : Comparaison graphique des ajustements de la série 1	-----50
Figure 46 : Comparaison à l'aide de critères d'information des ajustements de la série 1	-----51
Figure 47 : Illustration graphique d'un ajustement par une distribution Gamma Inverse de la série 1	----53
Figure 48 : Résultats de l'ajustement par une distribution Gamma Inverse de la série 1	-----54
Figure 49: Illustration de l'ajout d'une autre période de retour	-----55
Figure 50 : Adéquation de l'ajustement par une distribution Gamma Inverse de la série 1	-----56
Figure 51 : Caractéristiques statistiques pour un ajustement par une distribution Gamma Inverse de la série 1	-----57
Figure 52 : Test de discordance de la plus grande observation dans le cas d'un ajustement par la distribution Log-normale (série 2)	-----58
Figure 53 : Ajustement de la loi Log-normale à la série 2 pour la mise en évidence des valeurs discordantes sous l'hypothèse de Log-normalité.	-----59
Figure 54 : Test de discordance pour un ajustement par la distribution Log-normale avec une nouvelle grande observation simulée (série 2).	-----60

## 1. Menu Principal

---

Le logiciel HYFRAN-PLUS a été conçu pour l'analyse fréquentielle en hydrologie particulièrement pour les valeurs extrêmes. Ainsi pour l'analyse des crues, on s'intéresse au débit maximum annuel. Cependant il peut être utilisé pour n'importe quelle série de valeurs extrêmes dans d'autres domaines avec différents pas de temps, à conditions que les observations soient Indépendantes et Identiquement Distribuées (Hypothèses IID, cf 2.1.4 et Bobée et El Adlouni, 2015).

### 1.1. Interface

---

Au démarrage du logiciel, la fenêtre suivante (Figure 1) apparaît avec la barre de menu composé des menus suivants :

- a) Fichier
- b) Edition
- c) Échantillon
- d) Système d'Aide à la Décision (SAD)
- e) Ajustement
- f) Graphique
- g) Affichage
- h) Fenêtre
- i) Aide (?)

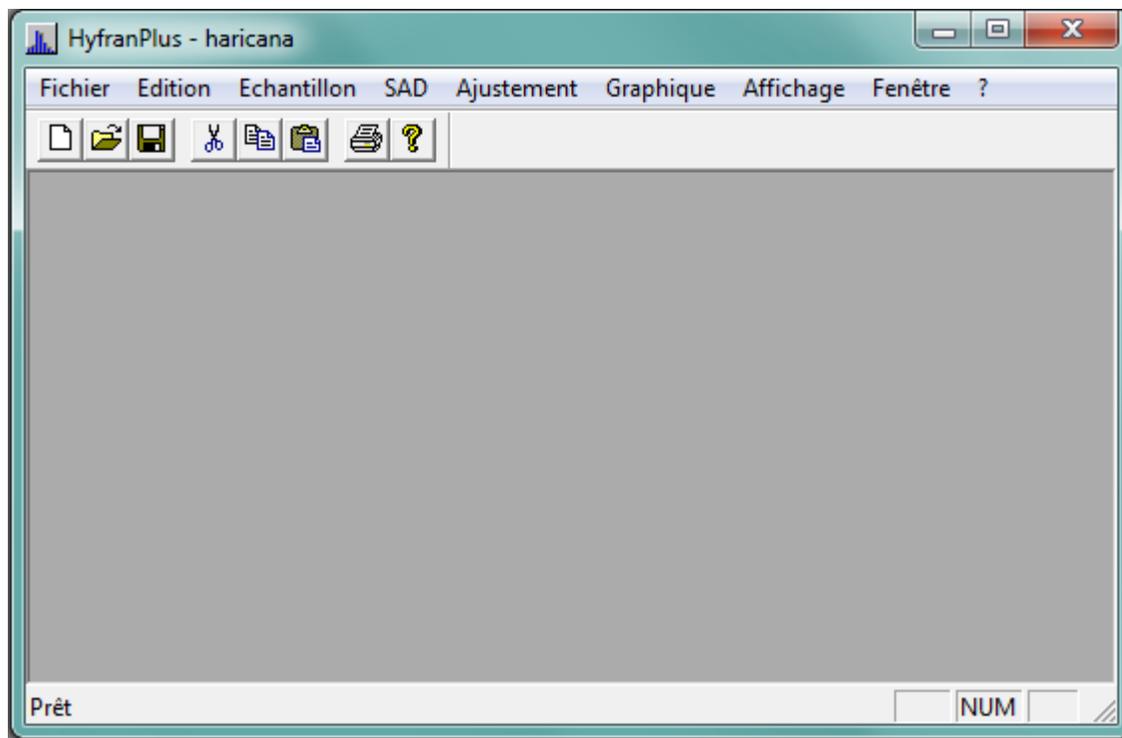


Figure 1 : Fenêtre d'ouverture du logiciel HYFRAN-PLUS

## 1.2. Fichier

Ce menu (Figure 2) contient les options de base pour la création, l'ouverture et l'enregistrement des échantillons ainsi que les options d'impression.

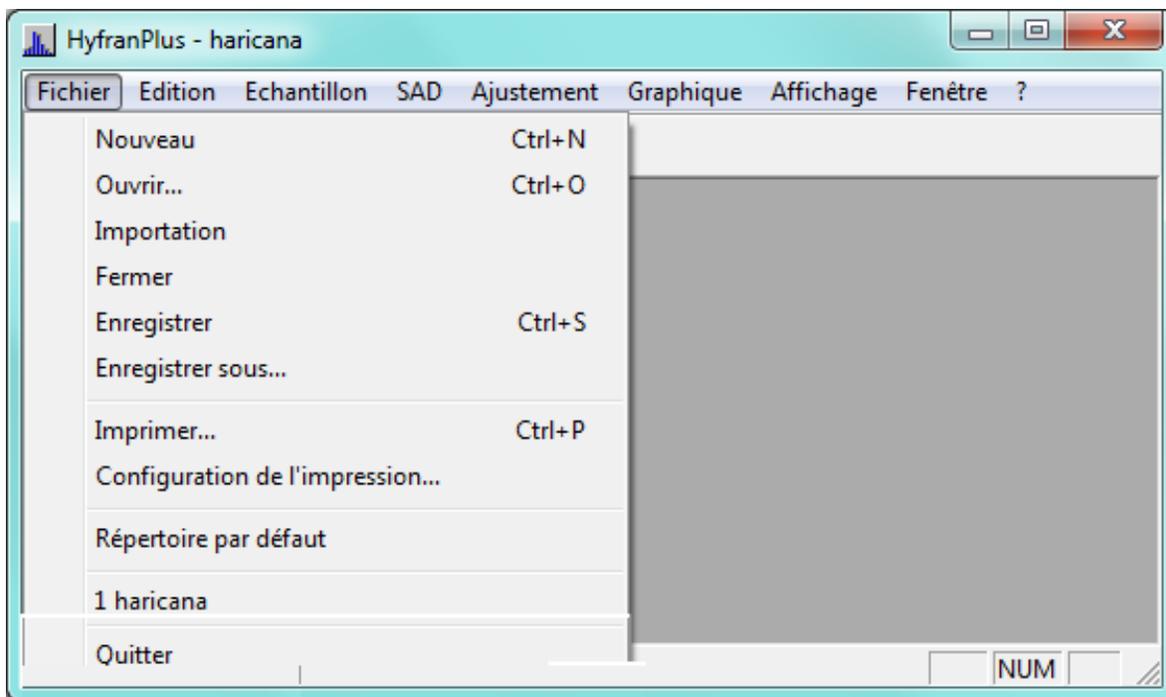


Figure 2 : Menu «Fichier»

- **Nouveau** : Cette option permet de créer un nouveau projet. On entre les données manuellement ou en utilisant le presse-papier.
- **Ouvrir** : Lors de la sélection de cette option, une boîte de dialogue permettant de sélectionner un fichier projet apparaît. On peut choisir entre ouvrir un fichier de type HYFRAN-PLUS (\*.hyf) ou importer différents types de fichier.
- **Importation** : Cette option permet d'importer des fichiers de formats variés. Pour utiliser cette option, il suffit de sélectionner le fichier à importer dans son répertoire et d'appuyer sur le bouton «ouvrir». Si le format du fichier est reconnu, le curseur sera positionné sur le bon format de fichier dans le nouvel écran qui s'affiche. Par la suite la conversion du fichier au format HYFRAN-PLUS s'effectuera après que l'utilisateur ait entré un nom pour ce nouveau fichier.
- **Enregistrer**: cette option permet de sauvegarder les modifications effectuées dans le fichier projet.
- **Enregistrer sous** : l'option «enregistrer sous» permet de changer le nom ou le chemin d'accès d'un fichier projet existant.

- **Imprimer** : L'ensemble des informations de toutes les fenêtres peuvent être imprimées peu importe le type de contenu (texte ou graphique).
- **Configuration de l'impression** : cette option permet de configurer l'imprimante (choix de l'imprimante, du papier, de l'orientation).
- **Répertoire par défaut** : Cette option permet de choisir un répertoire par défaut pour tout fichier de données.
- **Haricana** : La fenêtre correspondant au projet ouvert.
- **Quitter** : En choisissant cette option on quitte le logiciel HYFRAN-PLUS. Si des modifications sur l'échantillon sont effectuées et qu'elles n'ont pas été sauvegardées un avertissement apparaîtra.

### 1.3. Édition

---

Ce menu (Figure 3) contient les options de base pour l'édition, c'est-à-dire annuler, couper, copier et coller. On peut également exporter les graphiques dans une autre application en choisissant l'option «copier».

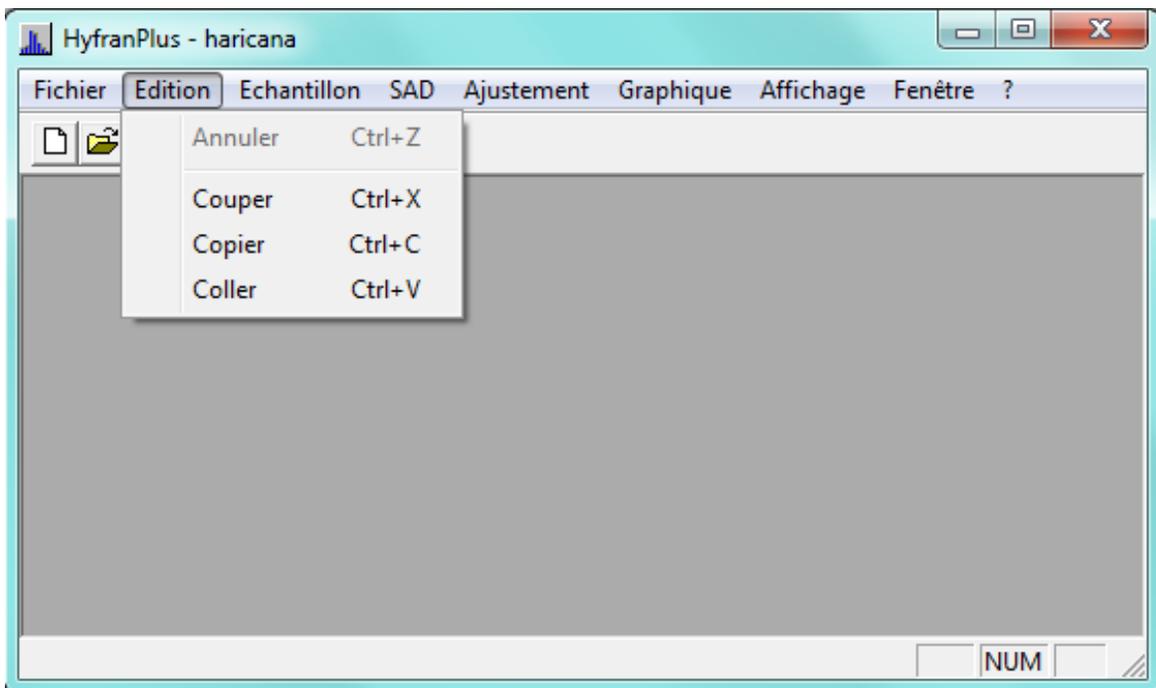


Figure 3 : Menu «Édition»

### 1.4. Échantillon

---

Ce menu (Figure 4) contient les options utiles pour l'entrée des données.

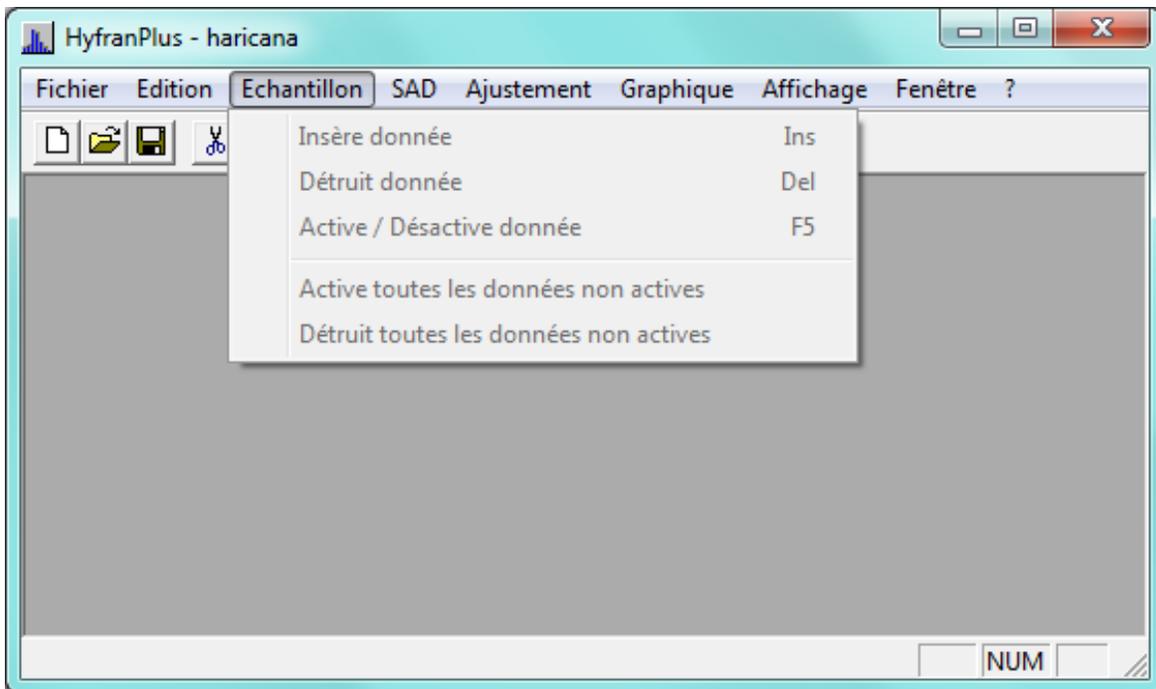


Figure 4 : Menu échantillon

- **Insère une donnée** : Cette option est sélectionnée pour ajouter une donnée.
- **Détruit une donnée** : À l'aide de cette option, on efface une ou plusieurs données.
- **Active/Désactive une donnée** : Cette option permet de désactiver une ou plusieurs observations qui sont actives ou inversement d'activer une ou plusieurs observations inactives (cf Section 2.1.2),
- **Active toutes les données non actives** : Afin d'activer toutes les données qui sont inactives, vous choisissez cette option,
- **Détruit toutes les données non actives** : Afin d'effacer toutes les données qui sont inactives, vous choisissez cette option.

### 1.5. Système d'aide à la décision (SAD)

---

Le SAD (Figure 5) est la principale addition à la version antérieure de HYFRAN-PLUS (Version 2.0); il s'agit dans HYFRAN-PLUS (Version 2.1), de l'ajout d'un Système d'Aide à la Décision pour permettre le choix de la classe de distributions la plus adéquate pour représenter une série de données. Les principaux éléments du SAD sont présentés en détail dans El Adlouni et Bobée (2011) et El Adlouni, Bobée et Samoud (2012).

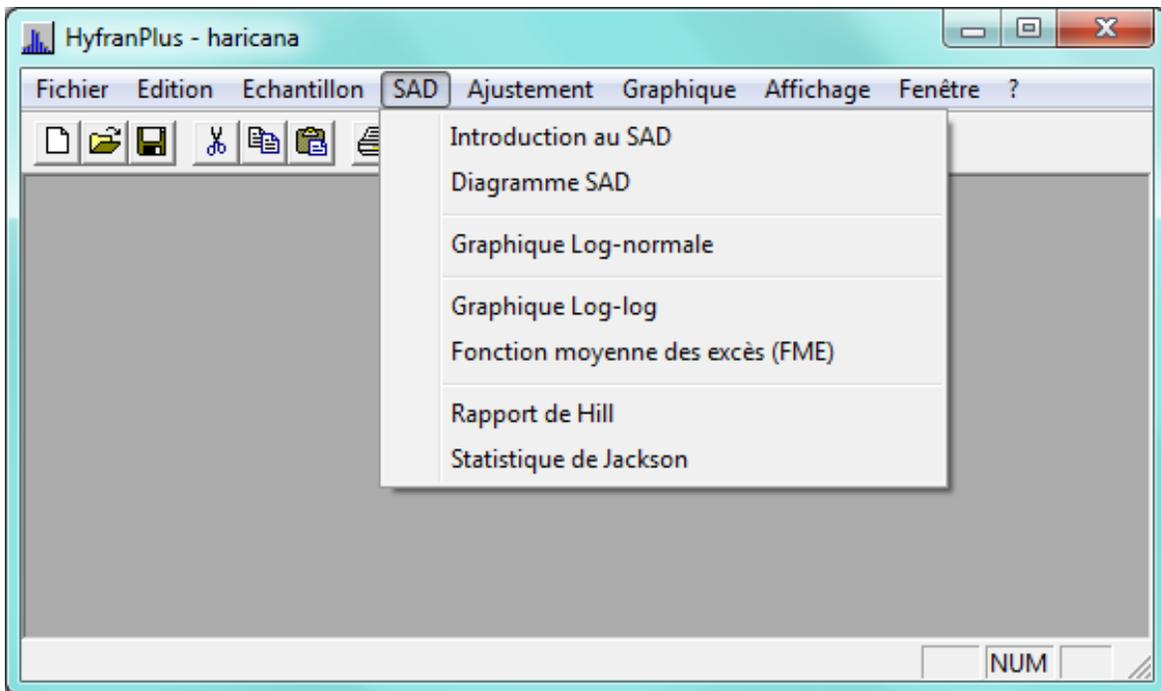


Figure 5 : Menu SAD

- **Introduction au SAD** : une présentation succincte du SAD est offerte ainsi qu'une référence aux articles publiés à propos du SAD.
- **Diagramme SAD** : Afin de mieux comprendre la méthodologie du SAD (Figure 5) une illustration graphique des étapes est fournie dans l'interface de HYFRAN-PLUS (Figure 6). Elle représente les différentes étapes du SAD.  
Le menu SAD (Figure 5) propose aussi les graphiques et les courbes sur lesquels repose le système d'aide à la décision tels que : le graphique log-normale, le graphique log-log, la fonction moyenne des excès (FME), le rapport de Hill et la statistique de Jackson (cf El Adlouni et Bobée, 2011 ; El Adlouni, Bobée et Samoud, 2012).
- **Graphique Log-normale (1)**: Pour tester la log-normalité avec le test de Jarque-Berra.
- **Graphique Log-log (1)** : Vérifier l'appartenance à la classe C (distributions à variations régulières).
- **Fonction moyenne des excès (FME) (1)**: Vérifier l'appartenance à la classe D (distributions sub-exponentielles).
- **Rapport de Hill et Statistique de Jackson (1)** : Analyse confirmatoire de la classe choisie.

**Note :** (1) Ces points sont développés et illustrés par des exemples de la section 2.2

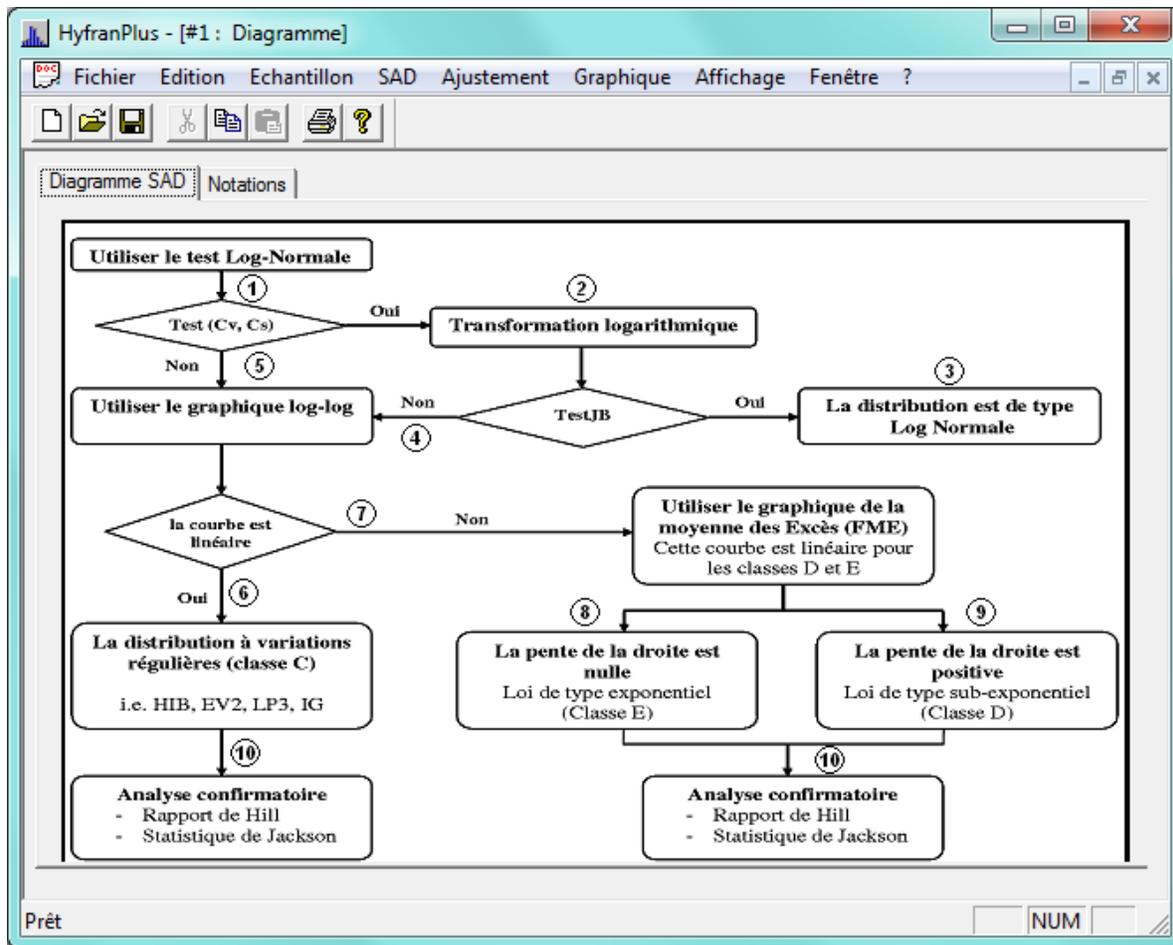


Figure 6 : Diagramme du SAD

## 1.6. Ajustement

HYFRAN-PLUS permet d'ajuster différentes distributions statistiques (Figure 7) qui y sont présentes (Compaoré, El Adlouni et Bobée, 2014) à un échantillon aléatoire respectant les conditions IID [données **I**ndépendantes et **I**dentiquement **D**istribuées, cf. la section 2.1.4] à l'aide de plusieurs méthodes d'estimation (Bobée et El Adlouni, 2015; Bobée et Askar 1991).

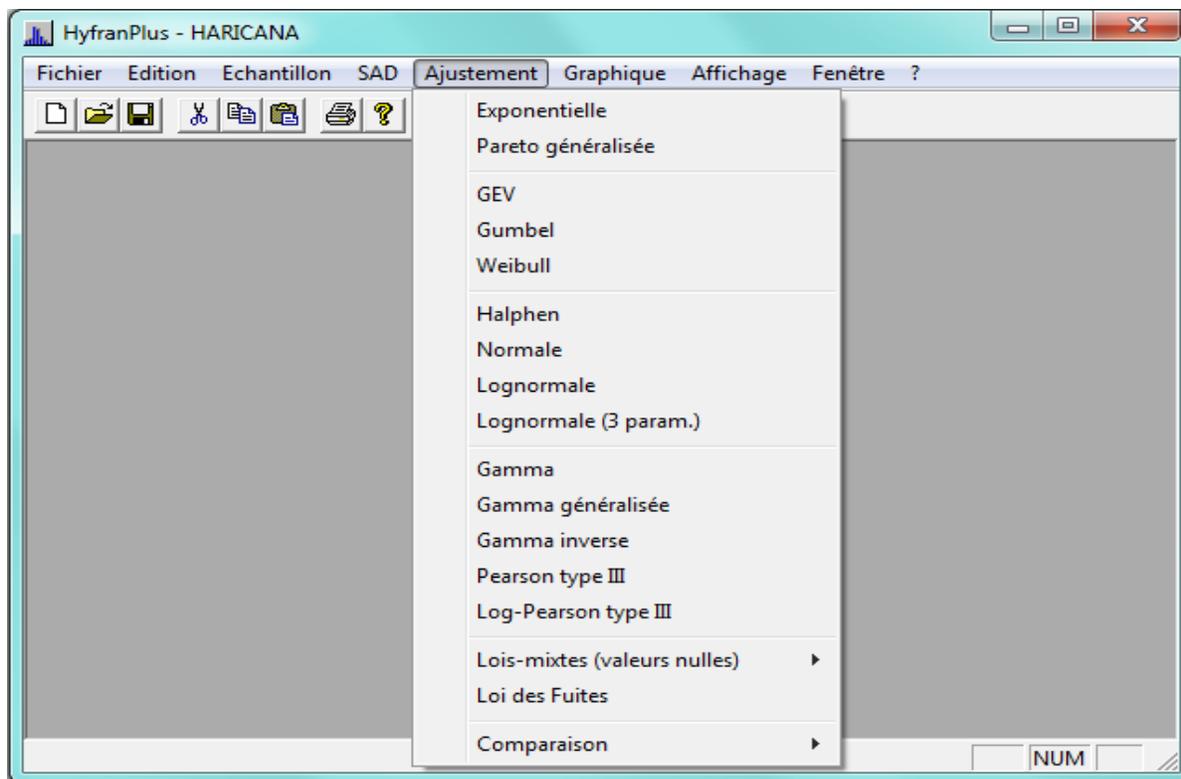


Figure 7 : Menu Ajustement

L'option «comparaison» permet de comparer plusieurs ajustements sélectionnés afin de pouvoir choisir lequel convient le mieux pour représenter la série de données considérée. On peut comparer les ajustements de manière graphique ou suivant différents critères :

- **Graphique** : il est possible de comparer les résultats de plusieurs ajustements différents (de 2 à 5) à l'aide d'un graphique sur papier de probabilité Normal ou Gumbel.
- **Critères** : deux critères sont disponibles, il s'agit des critères d'information d'Akaike (AIC) et Bayésien (BIC) (cf. Ehsanzadeh, El Adlouni and Bobée, 2010).

**Note :**

- L'onglet Halphen contient les trois lois de Halphen (type A, type B et type B Inverse) avec leurs lois limites (Gamma et Gamma Inverse) (Morlat, 1956). Cependant, dans l'ajustement le choix d'une des trois lois se fait automatiquement en fonction des caractéristiques de l'échantillon et des propriétés théoriques de la famille des lois de Halphen (Perreault, Bobée et Rasmussen, 1999).
- L'onglet GEV englobe les trois distributions Fréchet, Gumbel et Weibull. Le choix se fait automatiquement en fonction de l'estimation du paramètre de forme. Cependant, l'utilisateur peut choisir d'ajuster la loi Gumbel ou Weibull séparément (Figure 7).

## 1.7. Graphique

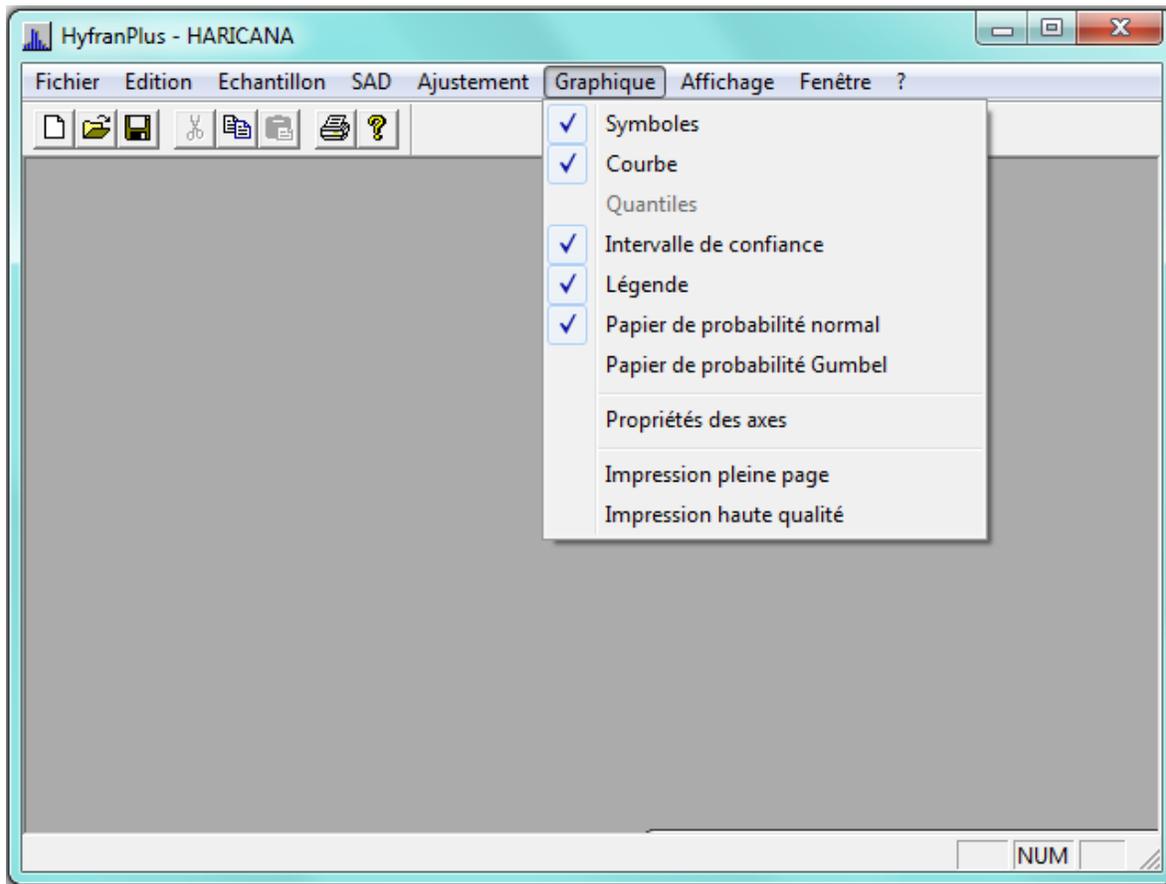


Figure 8: Menu graphique

Ce menu (Figure 8) contient des options qui permettent d'ajouter ou d'éliminer des éléments sur les graphiques. Il s'agit des options «symboles», «courbe», «quantiles», «intervalle de confiance» et «légende».

De plus il est possible de sélectionner le type de papier de probabilité sur lequel on désire visualiser les données et les courbes d'ajustement. Deux types de papier de probabilité sont disponibles : Normal et Gumbel. Une distribution normale (resp. Gumbel) serait représentée de manière linéaire sur un papier normal (resp. Gumbel).

L'option «propriétés des axes» permet de faire un zoom sur une zone particulière du graphique. En cas de recouvrement d'une partie du graphique par la légende, on peut remonter cette dernière en utilisant cette option.

L'option «Impression pleine page» permet d'imprimer un graphique en pleine page. Il faut sélectionner cette option (apparition d'une marque devant cette option du menu), et ensuite aller dans le menu Fichier pour imprimer. Par défaut, l'impression sera en format paysage. Cette option est conseillée lorsqu'on rencontre des problèmes d'impression avec certains modèles d'imprimantes.

L'option «Impression haute qualité» permet d'imprimer un graphique avec une plus grande précision au niveau des traits.

**Note :** Cette option fonctionne avec la majorité des imprimantes actuelles mais pas avec toutes.

## 1.8. Affichage

---

Ce menu (Figure 9) permet d'afficher ou non les barres d'outils et d'état.

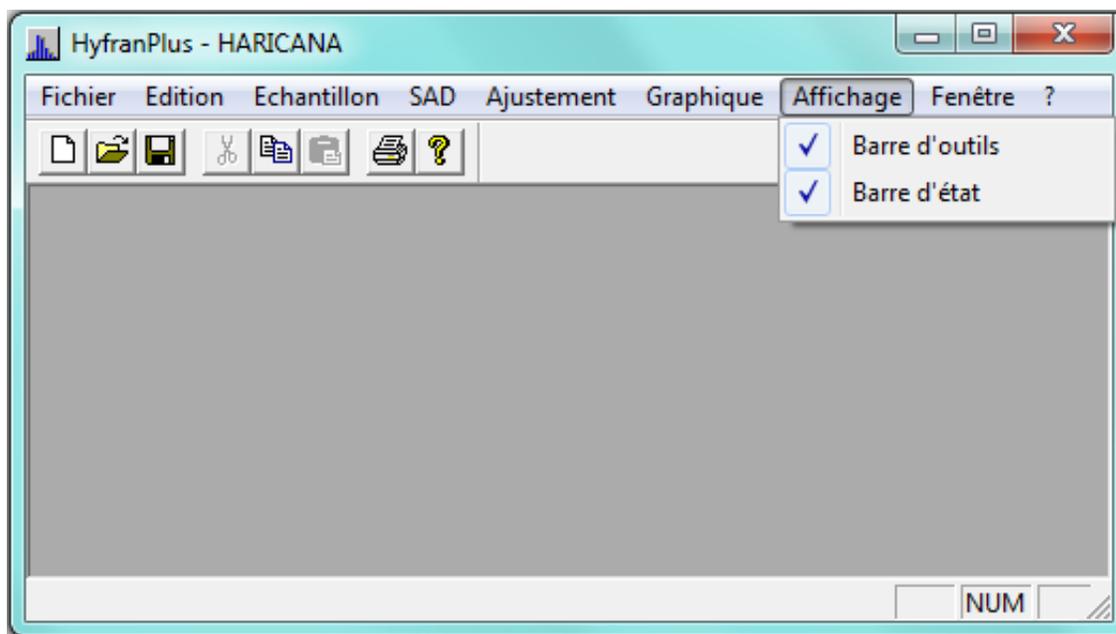


Figure 9 : Menu Affichage

## 1.9. Fenêtre

---

Ce menu (Figure 10) permet de positionner les différentes fenêtres ouvertes et de naviguer entre elles.

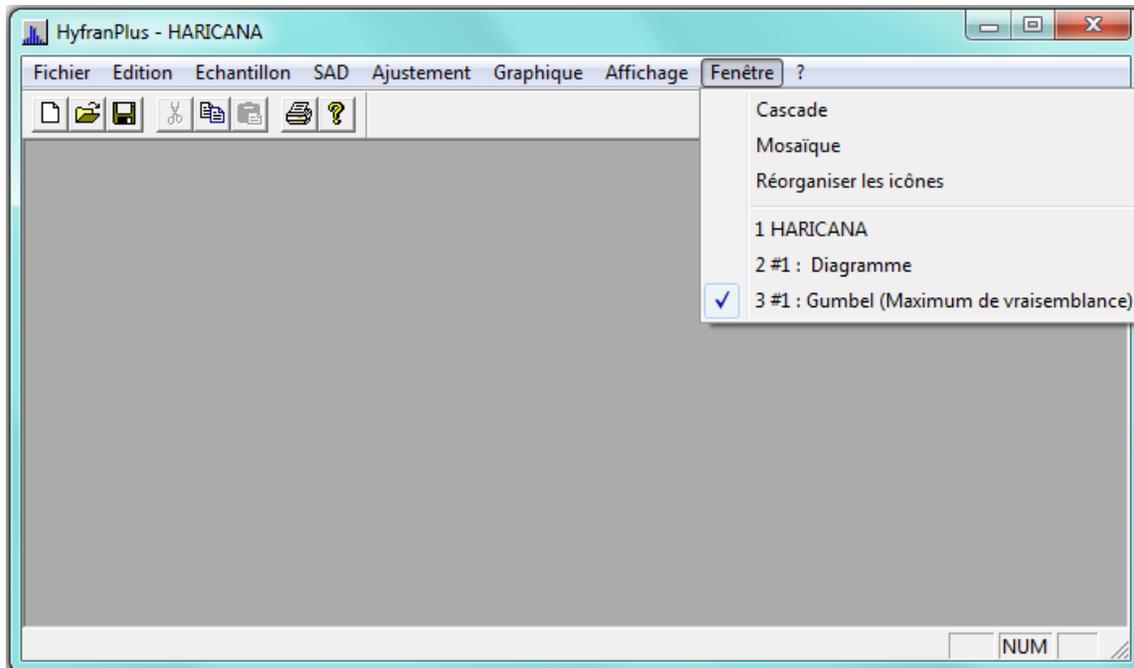


Figure 10 : Menu Fenêtre

## 2. Tutorial

---

Le logiciel HYFRAN-PLUS a été développé dans l'environnement du système d'exploitation Windows, il est conçu de façon à présenter clairement les étapes reliées à l'ajustement d'une distribution statistique à un échantillon aléatoire. Ces étapes peuvent être regroupées en deux catégories:

- Saisie des données et étude des caractéristiques statistiques de l'échantillon aléatoire (section 2.1)
- Procédures d'ajustement (section 2.2 et 2.3)

Le logiciel HYFRAN-PLUS contient un projet par défaut appelé Harricana qui sera utilisé dans certaines parties du tutoriel.

Une boîte à onglets munies d'options adaptées correspond à chacune des fonctions disponibles dans HYFRAN-PLUS. On présentera dans ce qui suit les différentes étapes d'une analyse fréquentielle en utilisant les outils statistiques et graphiques dans HYFRAN-PLUS.

## 2.1. Saisie des données et étude des caractéristiques statistiques de l'échantillon aléatoire

Lorsqu'on crée un nouveau projet (Figure 11) ou que l'on ouvre un projet existant à partir du menu «Fichier», on obtient une boîte à onglets qui permet de saisir certaines informations sur le projet, d'entrer ou de modifier des données, d'étudier certaines caractéristiques statistiques de l'échantillon aléatoire associé au projet, d'effectuer certains tests statistiques, ainsi que de produire plusieurs graphiques.

Pour effectuer ces différentes tâches on doit naviguer entre cinq onglets différents, nommés :

- 1)Description
- 2)Données
- 3)Statistiques de base
- 4)Tests d'hypothèse
- 5)Graphiques

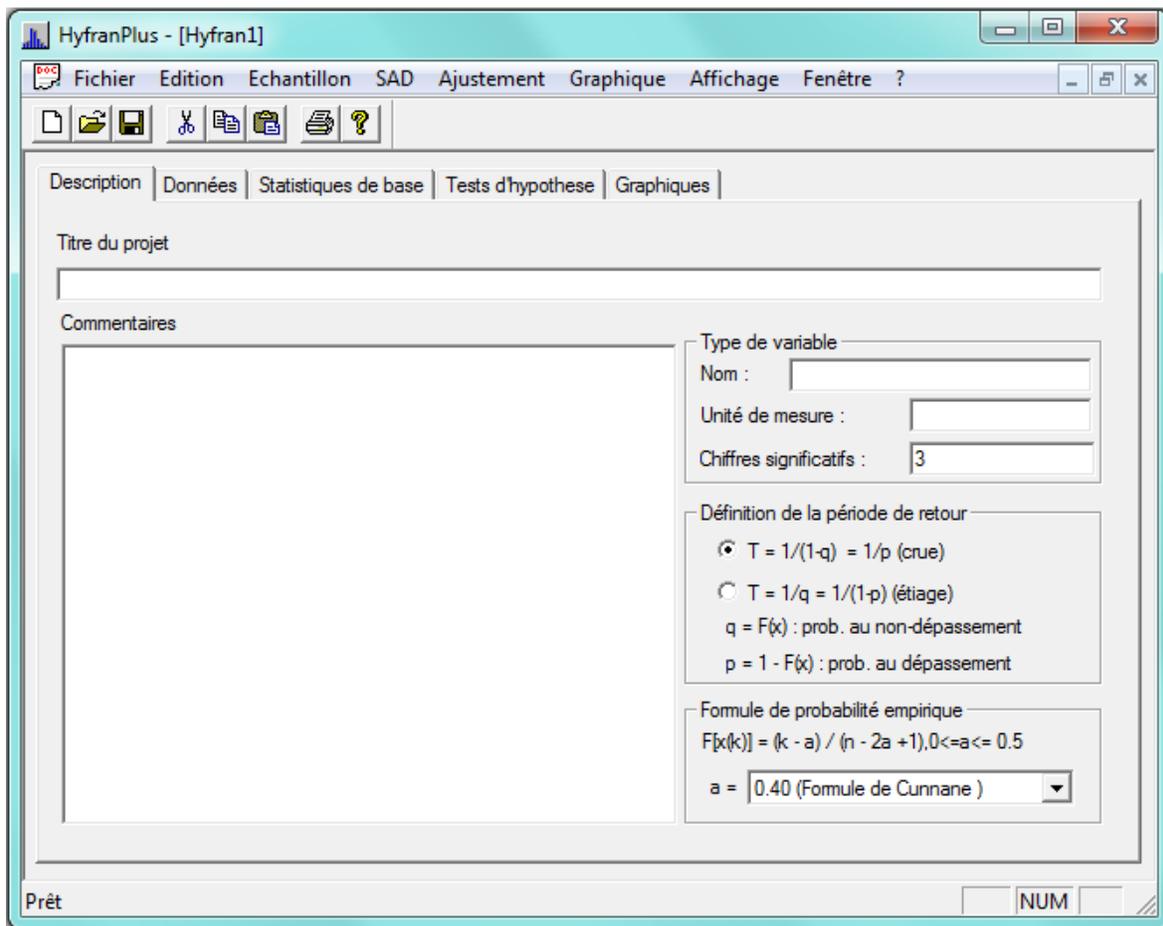


Figure 11: Création d'un nouveau projet

---

### 2.1.1 Description

---

Dans l'onglet «Description du projet» (Figure 12), on retrouve les différentes informations sur le projet qui sont utilisées pour construire les graphiques ainsi que pour présenter les tableaux de données et de résultats:

- a) On entre d'abord le titre du projet qui sera le titre des graphiques produits ensuite par HYFRAN-PLUS.
- b) On inscrit ensuite le nom de la variable qui apparaîtra comme le titre d'un des axes des graphiques (l'abscisse ou l'ordonnée, selon le graphique demandé).
- c) Puis on inscrit l'unité de mesure des observations qui sera indiqué à la suite du nom de la variable sur les graphiques.
- d) On peut aussi spécifier le nombre de chiffres significatifs des données (entre 1 et 18). Cette valeur est utilisée pour présenter les résultats avec un nombre correct de chiffres significatifs mais n'influence pas la précision des calculs.
- e) On peut choisir une définition pour le concept de période de retour. Il peut s'agir soit :
  - de l'inverse de la probabilité au dépassement pour l'étude des crues,
  - de l'inverse de la probabilité au non-dépassement pour l'étude des étiages (cf. Bobée et El Adlouni, 2015).
- f) On peut choisir une formule de probabilité empirique (Bobée et Ashkar, 1991), qui est ensuite utilisée pour tracer les observations sur un papier de probabilité. Dans HYFRAN-PLUS, les formules de probabilité empirique utilisées sont de la forme suivante:  $P_k = (k - a) / (n - 2a + 1)$  (cf. Bobée et El Adlouni, 2015; Bobée et Ashkar, 1991).  $P_k$  correspond à la probabilité au non-dépassement de l'observation  $X_k$  d'ordre k dans l'échantillon de taille n classé en ordre croissant.  
La formule de Cunnane ( $a = 0.4$ ) est utilisée par défaut dans le logiciel HYFRAN-PLUS mais les autres formules disponibles peuvent être utilisées selon le choix de l'utilisateur (Bobée et Ashkar, 1991 – Table 1.3 page 11).
- g) On peut finalement inscrire des commentaires quelconques sur le projet. Les informations inscrites dans la section «Commentaires» ne sont pas prises en compte par HYFRAN-PLUS mais peuvent être utiles à l'utilisateur pour décrire le projet plus en détail (Figure 12).

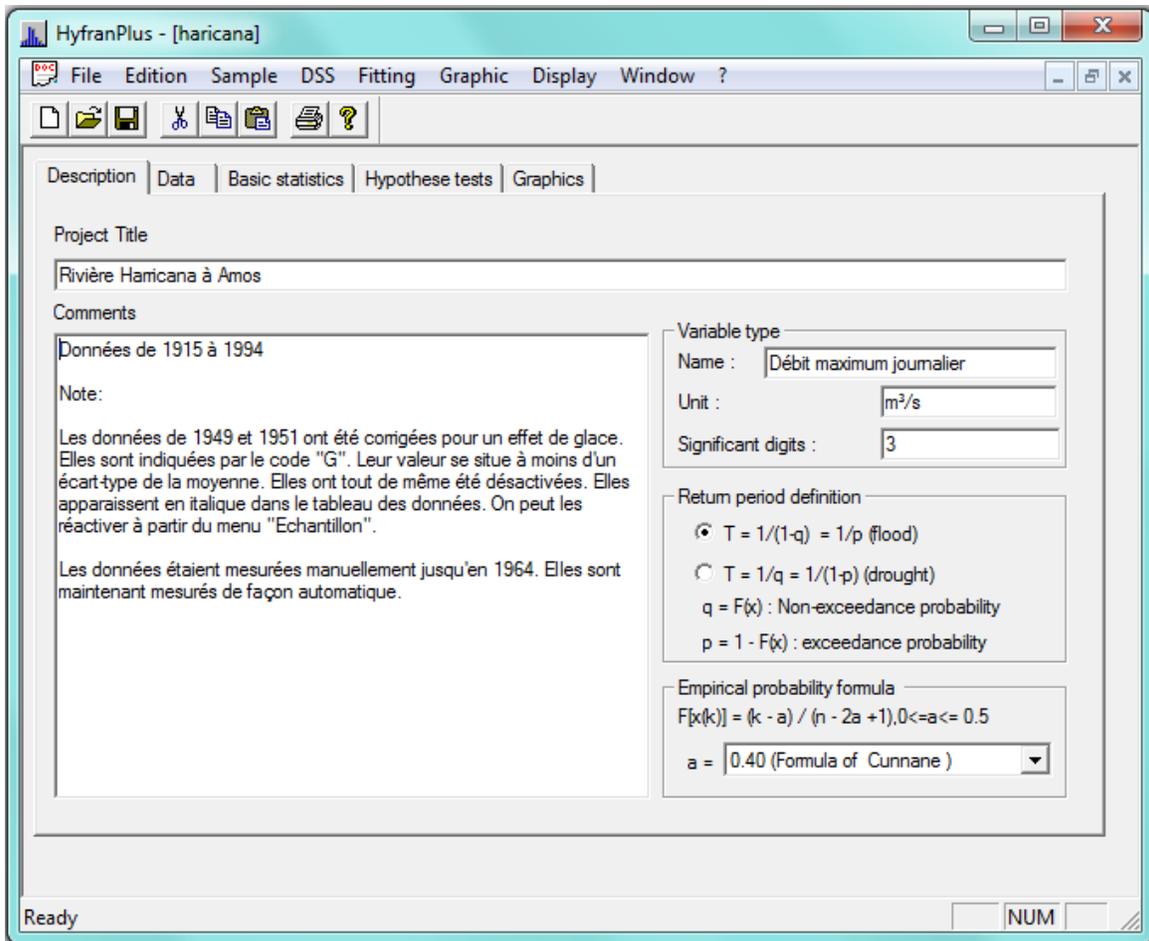


Figure 12 : Exemple de description du projet Harricana

### 2.1.2. Données

Il existe trois façons d'entrer les données dans le logiciel HYFRAN-PLUS :

- Utiliser le tableur intégré au logiciel (cf. Figure 13).
- Importer un fichier de données (pour les formats reconnus par HYFRAN-PLUS); ce qui est décrit dans l'une des sections suivantes.
- Utiliser le presse-papier

#### a) *Tableur intégré au logiciel*

HYFRAN-PLUS est muni d'un tableur qui permet de saisir et de visualiser les données. Chaque ligne correspond à une observation et est divisée en quatre colonnes (Figure 13):

- Observation:** la valeur numérique de chaque observation doit être inscrite dans cette colonne;
- Identificateur:** on peut associer un numéro séquentiel à chaque observation; si l'on veut inscrire une date, elle doit nécessairement être entrée dans le format AAAA-MM-JJ; on peut omettre le jour et le mois et entrer seulement l'année (AAAA);

3. Probabilité empirique: la probabilité empirique associée à chaque observation est calculée automatiquement et affichée dans cette colonne, en utilisant la formule spécifiée dans la page de description du projet (Figure 11);
4. Code: cette colonne peut être utilisée pour marquer d'un code certaines observations; ce code doit être composé d'un seul caractère; par exemple, on pourrait indiquer par un «M» les données qui ont été mesurées manuellement, et par un «R» les données qui ont été révisées.

HyfranPlus - [HARICANA]

Fichier Edition Echantillon SAD Ajustement Graphique Affichage Fenêtre ?

Description Données Statistiques de base Tests d'hypothese Graphiques

	Observation	Identificateur	Probabilité empirique	Code
1	122	1915-05-13	0.0698	
2	244	1916-05-07	0.9052	
3	214	1917-05-26	0.7431	
4	173	1918-05-20	0.3815	
5	229	1919-05-23	0.7930	
6	156	1920-05-10	0.1945	
7	212	1921-05-03	0.7307	
8	263	1922-04-27	0.9551	
9	146	1923-05-15	0.1322	
10	183	1924-05-21	0.5062	
11	161	1925-05-11	0.2195	
12	205	1926-05-28	0.7057	
13	135	1927-05-12	0.1072	
14	331	1928-05-26	0.9800	

Inclure les données désactivées dans le calcul des probabilités empiriques

Prêt NUM

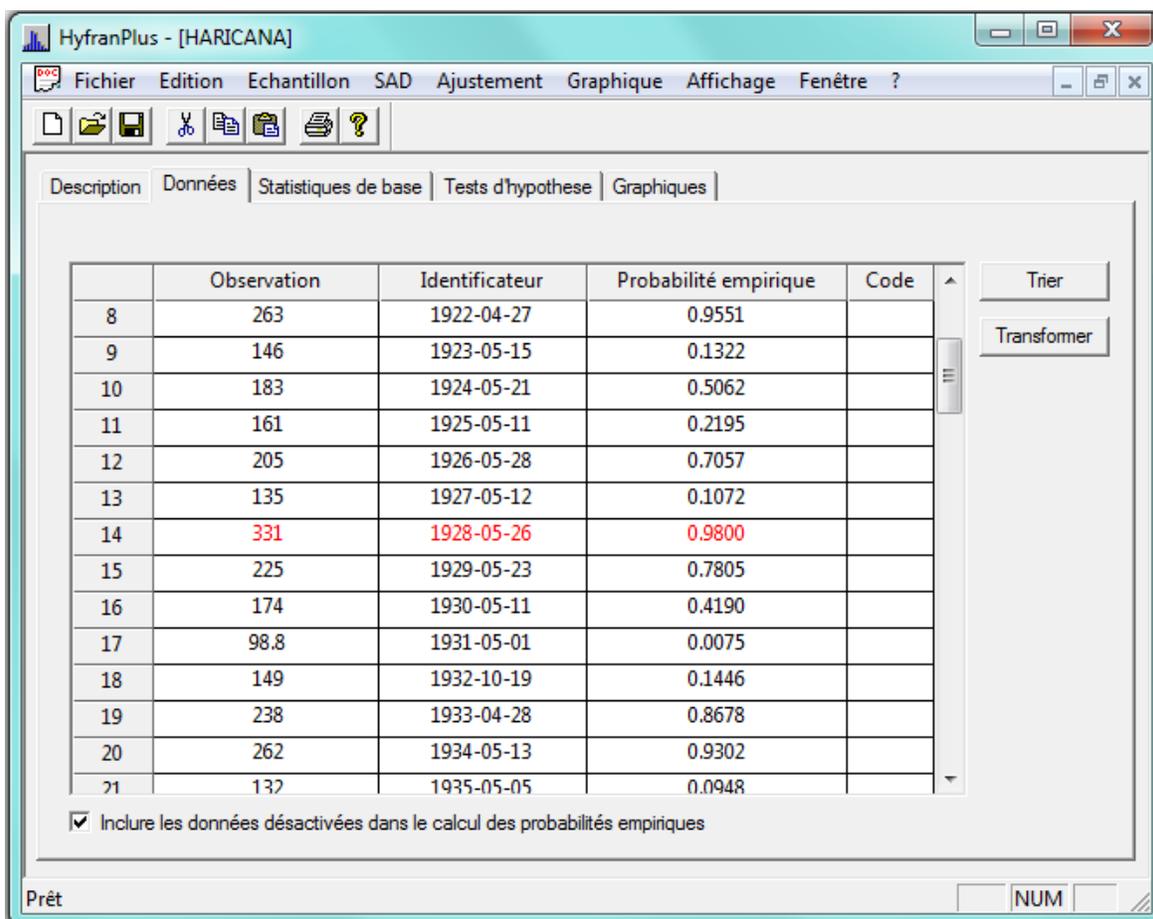
Figure 13 : Tableau intégré du logiciel (projet Harricana)

Pour insérer une nouvelle donnée (Figure 4), on utilise la touche «Insérer» du clavier ou l'option équivalente du menu «Échantillon». Pour détruire une ou plusieurs données, on sélectionne la ou les données (Figure 5), on utilise alors la touche «Del» du clavier ou l'option équivalente du menu "Échantillon" (Figure 4). Il est également possible de copier toutes les données du tableau de HYFRAN-PLUS dans le presse-papier en appuyant sur les touches **Ctrl-A** («Sélectionner tout» dans le menu «Édition»), ensuite les touches **Ctrl-C** («Copier» dans le menu «Édition») et enfin les touches **Ctrl-V** («Coller» dans le menu «Édition») [Figure 3]. De plus plusieurs options décrites dans ce qui suit (désactivation, tri et transformation des données) peuvent être utilisées à partir du tableau.

- *Données désactivées*

Un des avantages de HYFRAN-PLUS est de permettre de désactiver des données sans les détruire (Figure 14). Ceci permet entre autres d'étudier la sensibilité d'une analyse statistique aux données singulières. Ainsi une donnée extrême peut être représentée sur le graphique de la distribution ajustée mais n'avoir aucun poids dans l'ajustement de cette distribution.

Pour désactiver une donnée active ou activer une donnée désactivée, on peut appuyer sur «F5» ou choisir l'option correspondante du menu «Échantillon» (Figure 4).



	Observation	Identificateur	Probabilité empirique	Code
8	263	1922-04-27	0.9551	
9	146	1923-05-15	0.1322	
10	183	1924-05-21	0.5062	
11	161	1925-05-11	0.2195	
12	205	1926-05-28	0.7057	
13	135	1927-05-12	0.1072	
14	331	1928-05-26	0.9800	
15	225	1929-05-23	0.7805	
16	174	1930-05-11	0.4190	
17	98.8	1931-05-01	0.0075	
18	149	1932-10-19	0.1446	
19	238	1933-04-28	0.8678	
20	262	1934-05-13	0.9302	
21	132	1935-05-05	0.0948	

Inclure les données désactivées dans le calcul des probabilités empiriques

Figure 14 : Exemple de tableur intégré pour une donnée désactivée (ligne 14, projet Harricana)

Les données désactivées ne sont pas prises en compte pour effectuer les différents tests d'hypothèses et ajustements proposés par HYFRAN-PLUS. Cependant sur les graphiques elles sont affichées à l'aide d'un symbole différent (Figure 15).

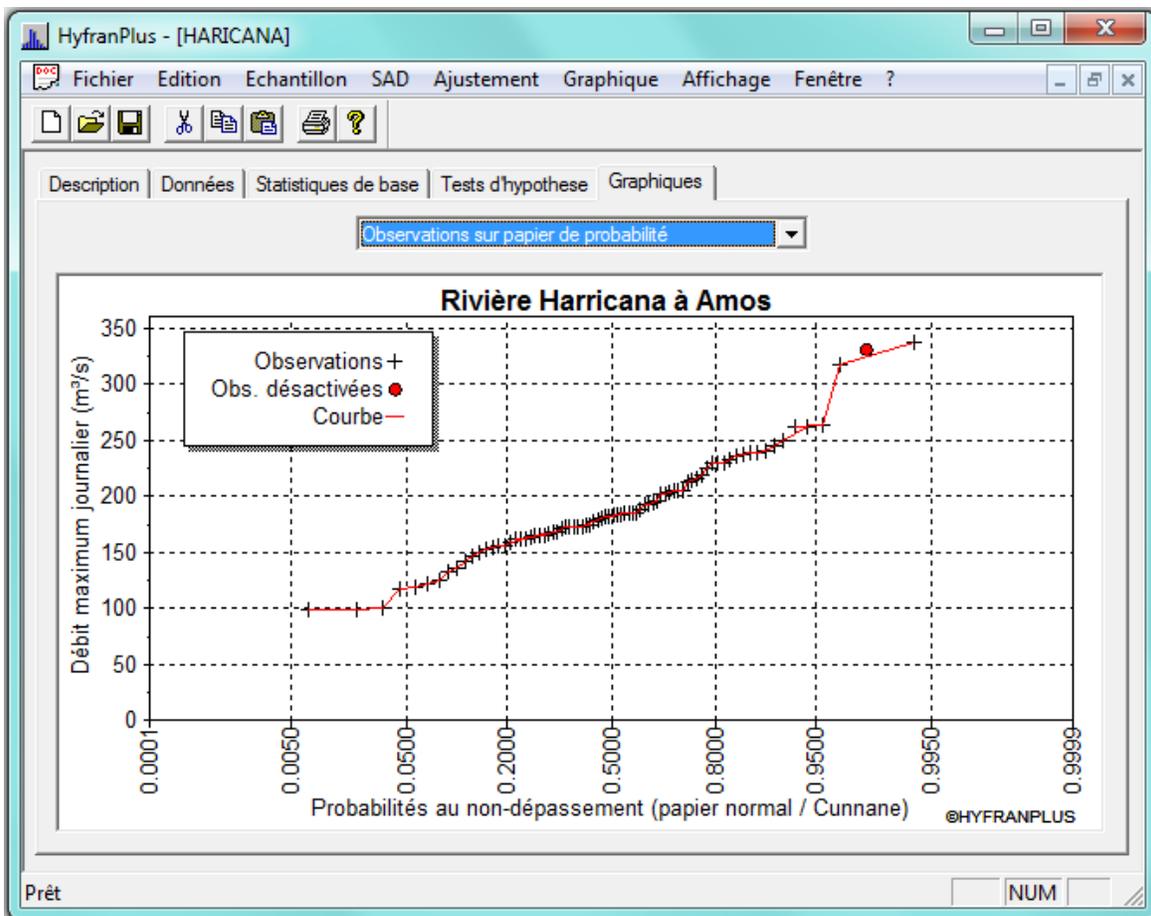


Figure 15 : exemple de graphique correspondant à une donnée désactivée (projet Harricana)

Pour le calcul des probabilités empiriques, il est possible d'inclure ou d'exclure les données désactivées (Figure 16). Si l'option «Inclure les données désactivées dans le calcul des probabilités empiriques» est cochée, toutes les observations (actives ou non) sont traitées de la même façon dans le calcul des probabilités empiriques. La probabilité empirique des observations désactivées est toujours calculée de la même façon, sur la base de l'échantillon complet. Dans le cas contraire, la probabilité empirique des observations actives est calculée sans tenir compte de l'existence des observations désactivées.

Pour activer toutes les données non actives on utilise l'option "Active toutes les données non actives" du menu "Échantillon" (Figure 4). La destruction des données non actives se fait à l'aide de l'option «Détruire toutes les données non actives» du menu "Échantillon".

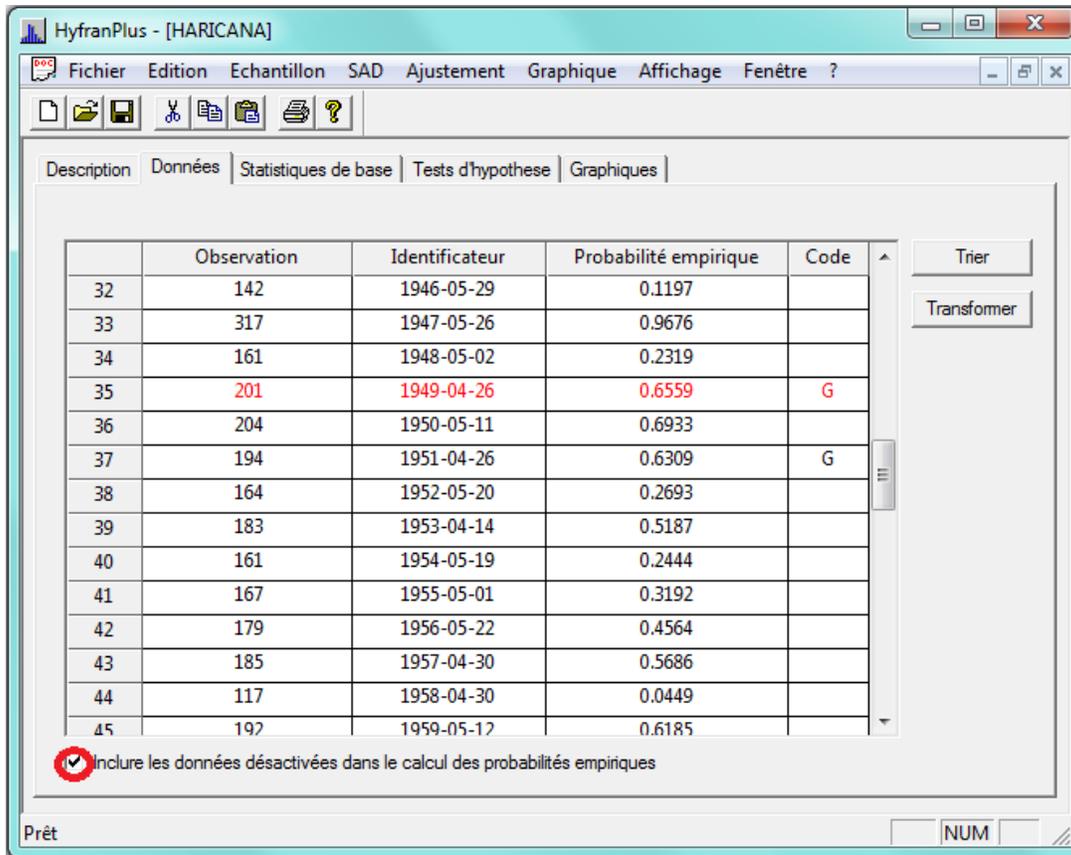


Figure 16 : Décision d'inclure et d'exclure une donnée

- *Tri des données*

Il est possible de trier les données par ordre croissant ou décroissant, et en fonction soit :

- de la valeur des observations (colonne 1)
- de l'identificateur (colonne2).

Pour trier les données, on appuie sur le bouton "Trier" (Figure 16).

- *Transformation des données*

Dans HYFRAN-PLUS on peut transformer les données à l'aide de plusieurs fonctions simples, ce qui permet d'effectuer l'ajustement d'une loi à la série transformée des données. Pour transformer les données, on doit appuyer sur le bouton "Transformer" (Figure 16). On ouvre alors une boîte de dialogue qui offre les transformations suivantes (Figure 17):

- Inverse:  $1/x$
- Opposé:  $-x$
- Valeur absolue:  $|x|$
- Exponentielle:  $a^x$
- Logarithmique:  $\log_a x$
- Échelle:  $a \times x$
- Position:  $a + x$
- Puissance:  $x^a$

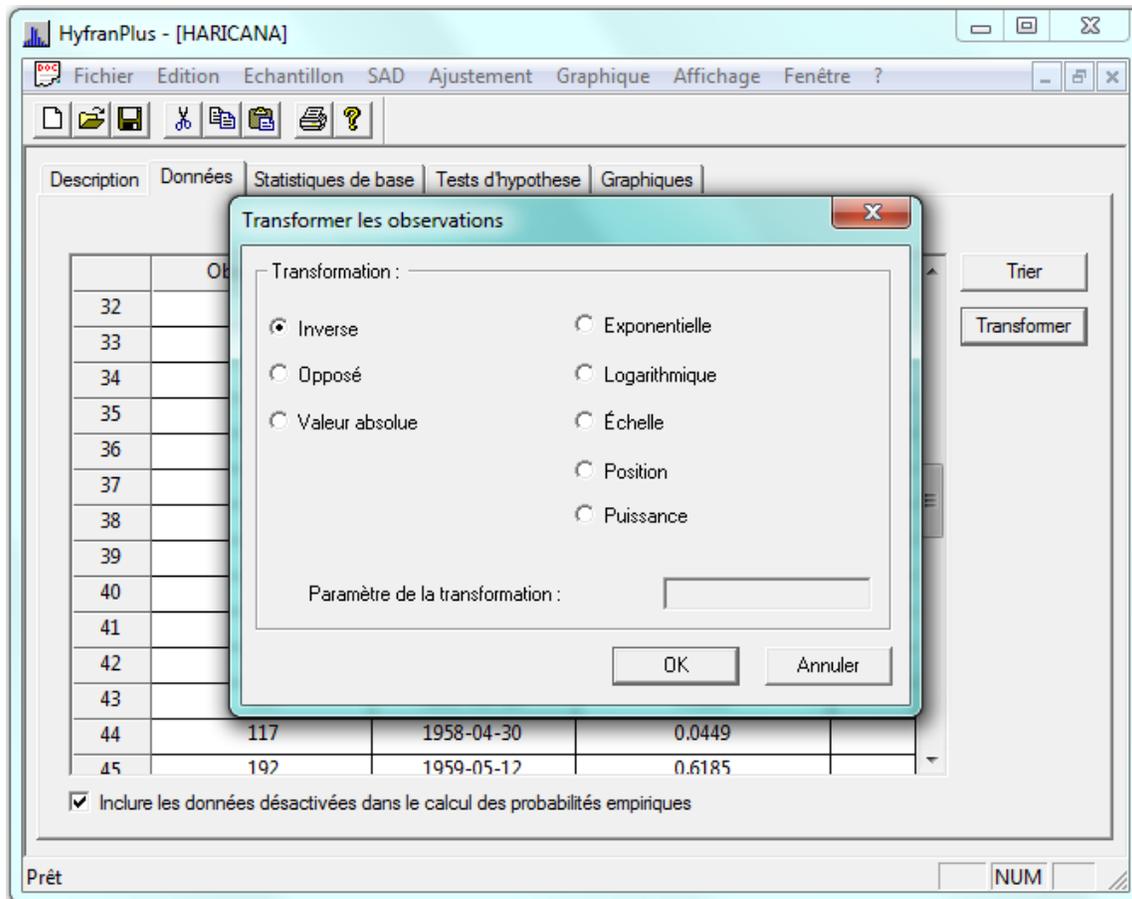


Figure 17 : transformation des données

**Note:** La valeur  $a$  correspond au paramètre de transformation, que l'on devrait spécifier dans cette même boîte de dialogue. Dans le cas d'une transformation logarithmique,  $a$  correspond à la base choisie. En pratique on considère  $a = 10$  (logarithme décimal) ou  $a = e$  (logarithme népérien).

#### *b) Importation des fichiers de données*

Le logiciel HYFRAN-PLUS permet d'importer divers types de format de données (Figure 2). Pour importer des données il faut aller dans le menu «Fichier» et cliquer sur l'onglet «Importation». Les différents fichiers importables sont cités ci-dessous.

- Formats HYDAT :
  - Export Extreme
  - Export Extreme Instantaneous
  - Export Mean
  - Print Extreme
  - Print Mean
- Format MATLAB : dans ce format texte, seules les données numériques sont présentes, les dates ne sont pas saisies. Les données sont séparées les unes des autres par un ou plusieurs espaces. Chaque ligne dans le fichier correspond à une station.

- Format Texte 1 colonne : les données numériques sont présentes sur une seule colonne; dans ce format, les dates ne sont pas entrées. L'utilisateur récupère l'ensemble de ces données dans un seul fichier.
- Format HCDN ASCII Annual Mean
- Format Excel avec le même ordre des colonnes que celui de HYFRAN-PLUS (Première colonne : Observations et Deuxième colonne l'identifiant qui est la date).
- Format libre : pour pouvoir importer ce type de données dans le logiciel HYFRAN-PLUS (incluant ou non des dates) le modèle suivant doit être respecté:
  - 1° ligne : on y entre un titre sur une seule ligne
  - 2° ligne : puis on y inscrit les mots «Format Libre» sans guillemets mais avec un espace entre Format et Libre
  - Lignes suivantes : elles ont consacrées aux données (1 donnée par ligne). Les données ne doivent pas contenir plus de 19 caractères en incluant le point décimal; elles contiennent un espace, une date ou un identifiant numérique (facultatif). Le format de la date est yyyy/mm/jj (année, mois puis jour). Le fichier doit être enregistré en format texte normal.

### c) *Presse-papier*

Le presse-papier peut être utilisé pour coller des données à partir d'une autre application Windows (traitement de texte, tableur, etc.). Il suffit d'utiliser les touches «Ctrl- C» («Copier» du menu «Édition») et «Ctrl-V» («Coller» du menu «Édition», Figure 3) pour insérer le contenu du presse-papier dans le tableur intégré à HYFRAN-PLUS.

---

### 2.1.3. Statistiques de base

---

La fenêtre des statistiques de base (Figure 18) présente en plus du titre éventuel de l'étude les statistiques suivantes de l'échantillon (cf. Bobée et El Adlouni, 2015):

- La taille de l'échantillon
- La valeur minimum
- La valeur maximum
- La moyenne
- L'écart-type
- La médiane
- Le coefficient de variation
- Le coefficient d'asymétrie
- Le coefficient d'aplatissement

Lorsqu'il y a des données non actives, les statistiques de base sont présentées en deux colonnes. La première colonne donne les statistiques de l'échantillon des données actives seulement et la deuxième colonne les statistiques de l'échantillon au complet (données actives et non actives).

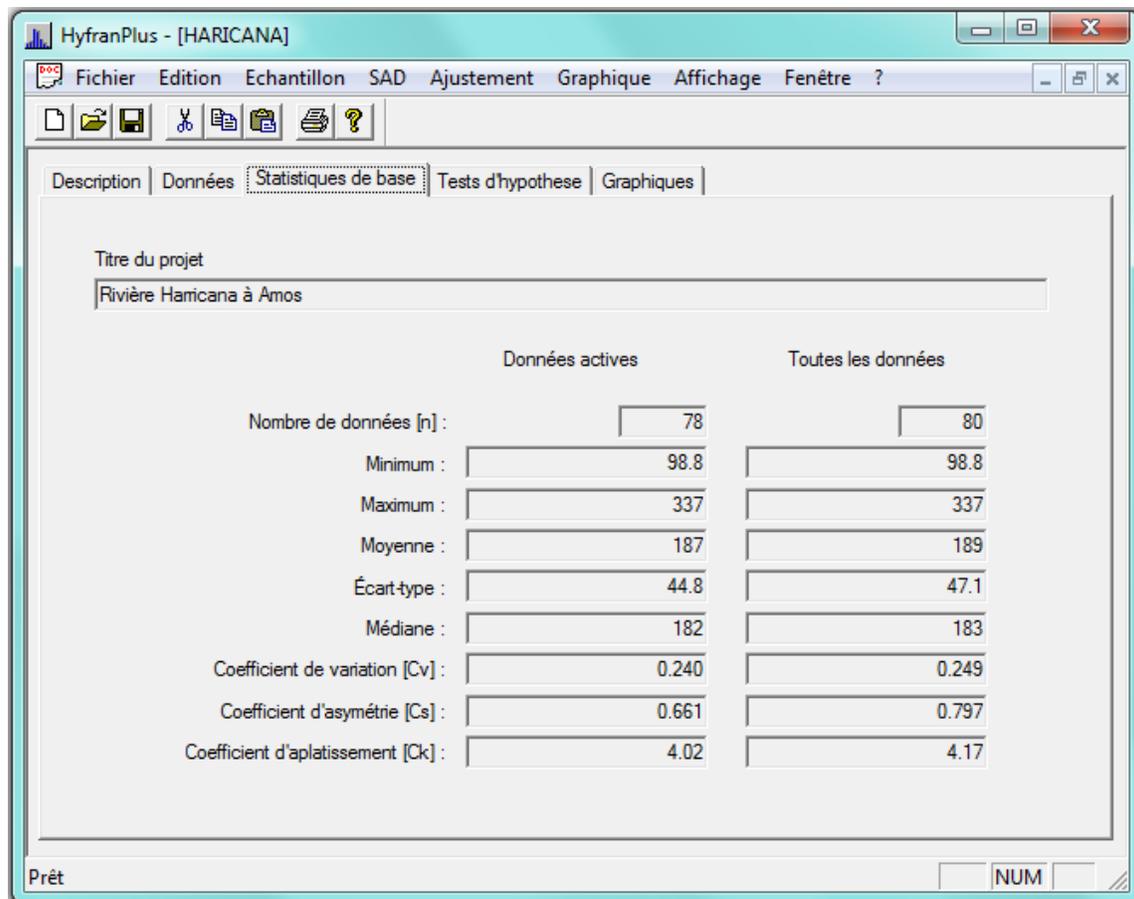


Figure 18 : Statistiques de base de l'échantillon (projet Harricana)

#### 2.1.4. Tests d'hypothèse

Avant de procéder à l'ajustement d'un échantillon à l'aide d'une distribution statistique, il est primordial de vérifier si les données sont indépendantes et identiquement distribuées (IID) (cf. Bobée et El Adlouni, 2015; Bobée et Ashkar, 1991). En effet, les observations doivent être des réalisations **I**ndépendantes de la même variable aléatoire et provenir de la même distribution statistique (i.e. **I**dentiquement **D**istribuées). Dans HYFRAN-PLUS, des tests statistiques sont disponibles pour vérifier les hypothèses d'indépendance, de stationnarité et d'homogénéité. Il s'agit des quatre tests d'hypothèse:

- a) Test d'indépendance (Wald-Wolfowitz)
- b) Test de stationnarité (Kendall);
- c) Test d'homogénéité à l'échelle annuelle (Wilcoxon);
- d) Test d'homogénéité à l'échelle saisonnière (Wilcoxon).

**Note :** Le test de Wilcoxon est connu aussi sous le nom du test de Mann-Whitney. Tous ces tests sont décrits dans Bobée et Ashkar (1991) et Bobée (2000).

Dans l'onglet «Tests d'hypothèse» (Figure 19) le titre du projet spécifié à la description du projet (Figure 12) est indiqué et on peut :

- en premier lieu choisir le type de test statistique à effectuer;
  - spécifier pour ce test l'hypothèse nulle ( $H_0$ ) et l'hypothèse alternative ( $H_1$ ).
  - Pour chaque test on obtient :
    - la valeur de la statistique du test et la  $p$ -value correspondante (probabilité au dépassement de la statistique),
    - la conclusion du test (obtenue à partir de la  $p$ -value), c'est-à-dire l'acceptation ou le rejet de l'hypothèse nulle à un niveau de signification de 5% ou de 1%.
- a) *Test d'indépendance (Test de Wald-Wolfowitz)* : Le test d'indépendance de Wald-Wolfowitz (Figure 19) permet de vérifier s'il existe une autocorrélation d'ordre 1, significative entre les observations.
- b) *Test de stationnarité (Test de Kendall)* : Le test de stationnarité de Kendall (Figure 20) permet de vérifier s'il existe une tendance significative dans la série.

HyfranPlus - [HARICANA]

Fichier Edition Echantillon SAD Ajustement Graphique Affichage Fenêtre ?

Description Données Statistiques de base **Tests d'hypothèse** Graphiques

Nom du test Test d'indépendance (Wald-Wolfowitz) Subdiviser l'échantillon

Titre du projet Rivière Haricana à Amos

Hypothèses

H0 Les observations sont indépendantes

H1 Les observations sont dépendantes (autocorrélation d'ordre 1)

Résultats

Valeur de la statistique	U  = 1.09
p-value	p = 0.275

Conclusion

Nous pouvons accepter H0 au niveau de signification de 5 %

Prêt NUM

Figure 19: Test d'indépendance : projet Harricana

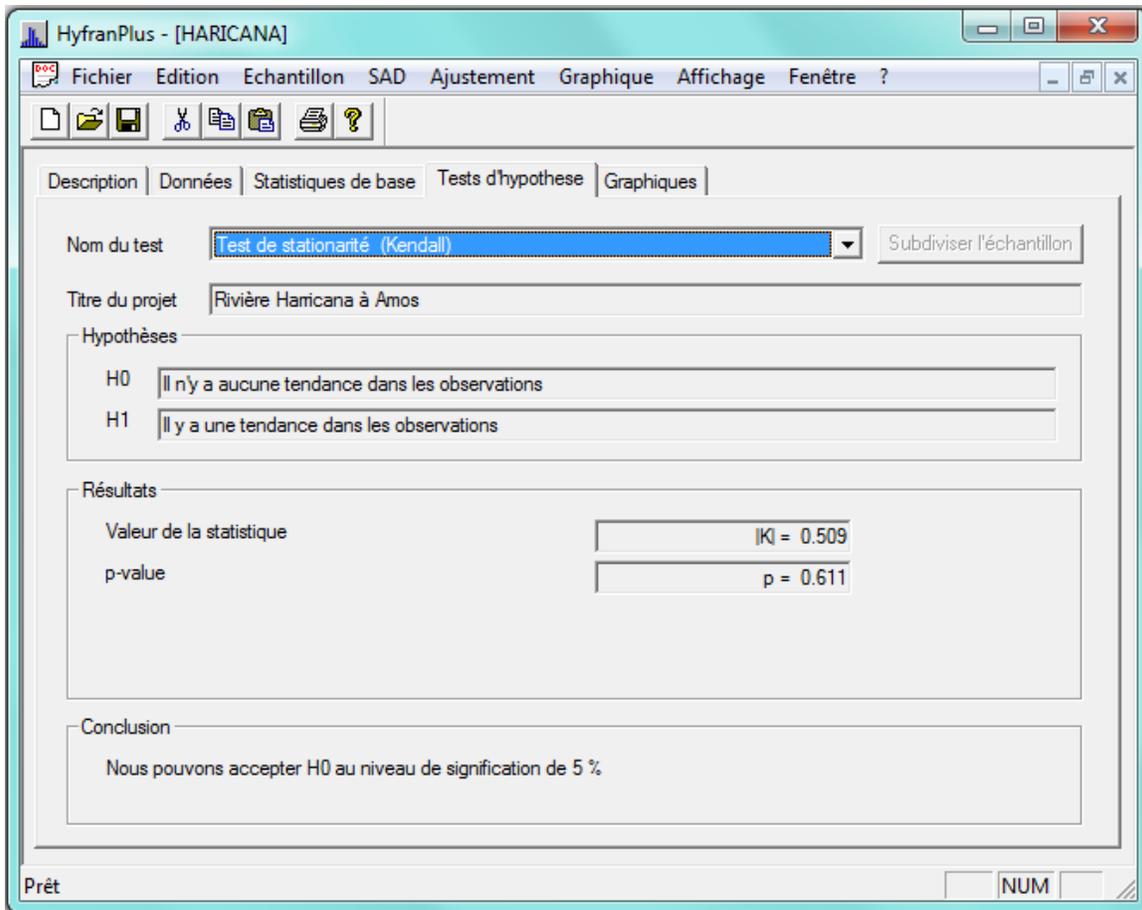


Figure 20 : Test de stationnarité (projet Harricana)

**Note :** Quand on choisit le test d'homogénéité (Figure 21) le bouton "Subdiviser l'échantillon" devient apparent pour les deux tests d'homogénéité considérés dans ce qui suit. Il est nécessaire dans ces cas de spécifier les deux sous-échantillons pour lesquels les moyennes vont être comparées.

- c) *Test d'homogénéité à l'échelle annuelle (Test de Wilcoxon ou de Mann-Whitney)*: Le test d'homogénéité à l'échelle annuelle (Figure 21) permet de vérifier si la moyenne d'un premier sous-échantillon est significativement différente de celle d'un deuxième sous-échantillon. Ce test peut, par exemple, être utilisé dans le cas du déplacement d'une station de mesure si on veut comparer la moyenne des données annuelles avant et après la date du déplacement c.à.d. vérifier si les données appartiennent à une population statistique de même moyenne. Le premier sous-échantillon est composé des observations de l'année de début des observations à une année de coupure. Le deuxième sous-échantillon est formé des observations de l'année suivant l'année de coupure à l'année de fin des observations. L'année de coupure est spécifiée après avoir appuyé sur le bouton «Subdiviser l'échantillon».

**Note :** Pour effectuer ce test, il faut que l'année soit spécifiée pour chaque observation lors de l'entrée ou de l'importation des données (Figure 13).

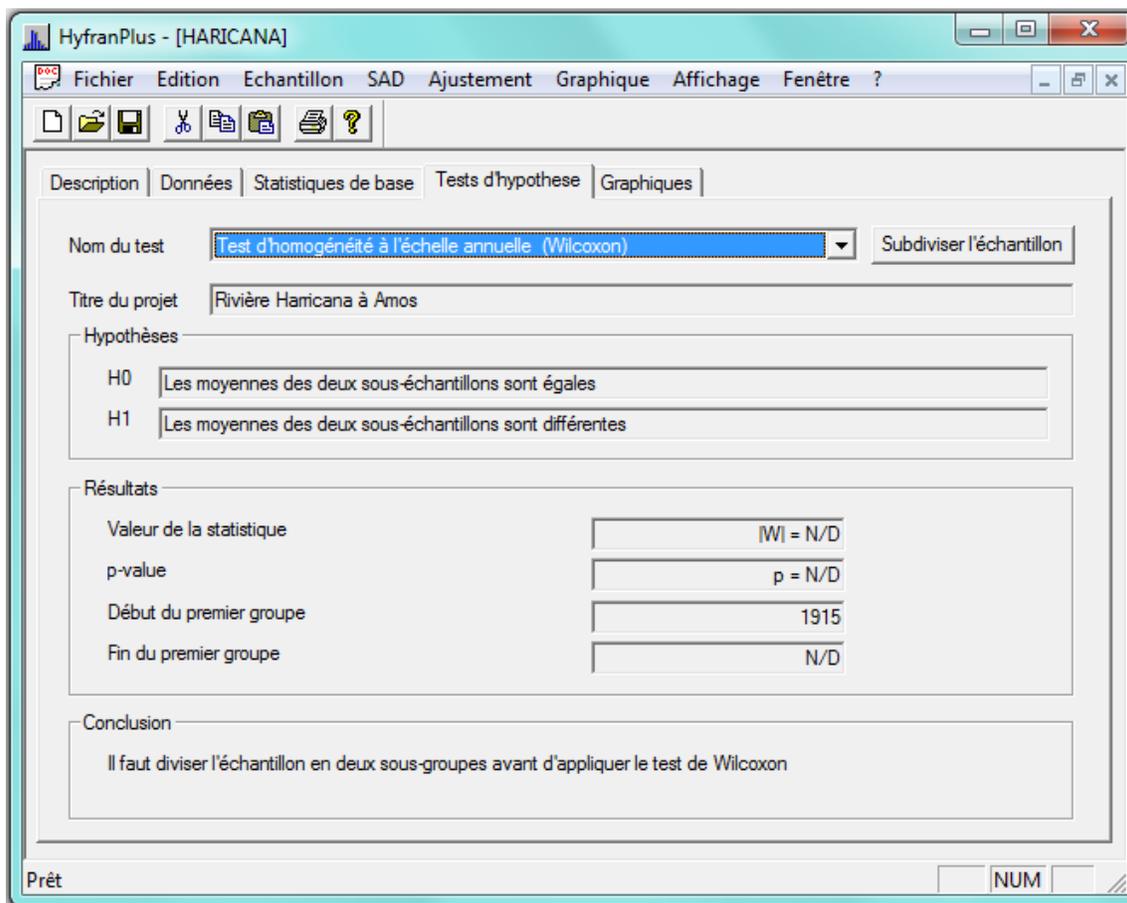


Figure 21 : Test d'homogénéité à l'échelle annuelle

d) *Test d'homogénéité à l'échelle saisonnière (Test de Wilcoxon ou de Mann-Whitney)*: Dans le test d'homogénéité à l'échelle saisonnière (Figure 22), tout comme le test d'homogénéité à l'échelle saisonnière le premier sous-échantillon est composé des observations comprises dans la période allant du mois m (début de la première saison) au mois n (fin de la première saison). Le deuxième sous-échantillon est formé des observations inclus dans la période allant du mois n+1 (début de la deuxième saison) au m-1 (fin de la deuxième saison). Les mois m et n sont spécifiés après avoir appuyé sur le bouton «Subdiviser l'échantillon» (Figure 19). Ce test peut être utilisé, par exemple, pour vérifier l'homogénéité des crues printanières, dues à la fonte des neiges (janvier-juin), et des crues automnales, dues aux précipitations juillet-décembre), afin de savoir si on peut les regrouper dans le même échantillon.

**Note :** Pour effectuer ce test, il faut que le mois soit spécifié pour chaque observation lors de l'entrée ou l'importation des données (Figure 13).

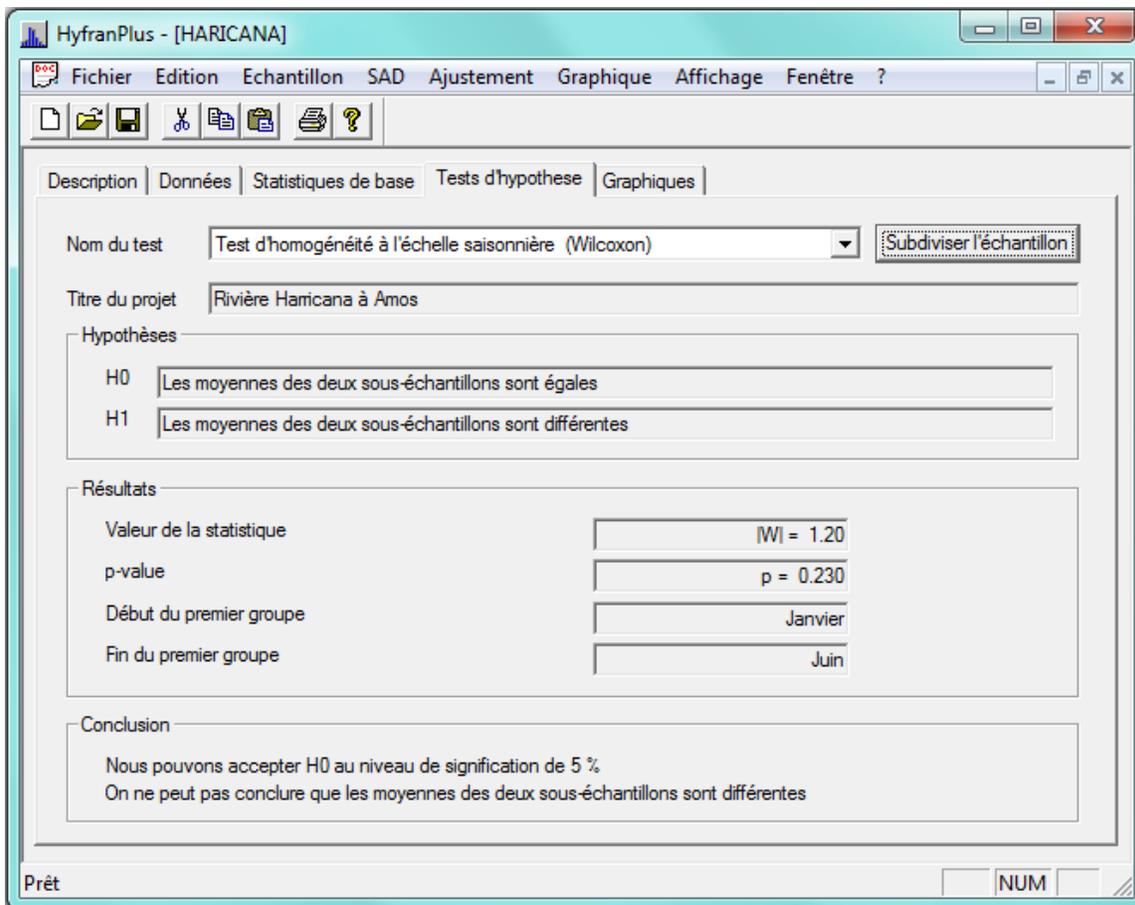


Figure 22 : Test d'homogénéité à l'échelle saisonnière (projet Harricana)

### 2.1.5. Graphiques

HYFRAN-PLUS offre la possibilité à partir de l'onglet «Graphiques» (Figure 15) de visualiser les données de différentes manières:

- Observations sur papier de probabilité (normal ou Gumbel);
- Histogramme des observations classées par valeur;
- Histogramme des observations classées par mois;
- Courbe chronologique.

Dans ce qui suit, ces quatre points seront détaillés et illustrés par des exemples.

- Observations sur papier de probabilité (Figure 23) : Ce graphique présente, sur papier de probabilité normal ou Gumbel, les observations en fonction des probabilités empiriques de l'échantillon. La sélection du type de papier de probabilité est effectuée dans le menu "Graphique" (Figure 8) où l'on retrouve les deux options "Papier de probabilité normal" et "Papier de probabilité Gumbel".

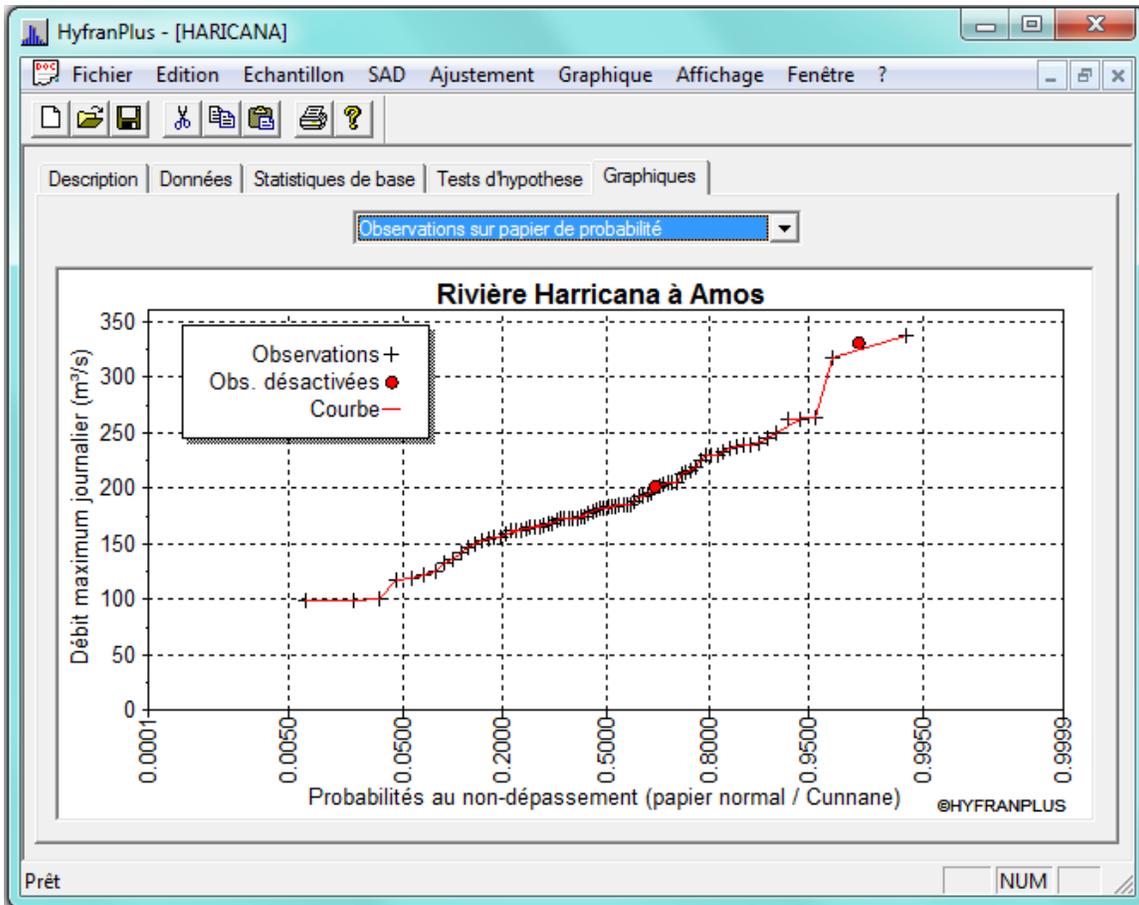


Figure 23 : Observation sur papier de probabilité normal (projet Harricana)

**Note :** Les points en rouge correspondent aux valeurs désactivées.

- b) *Histogramme des observations classées par valeur* (Figure 24): l'histogramme des observations classées par valeur est un graphique permettant de visualiser le nombre d'observations par classe de valeur. Ce type de graphique donne une image de la fonction de densité de probabilité des données.

Le nombre de classes équidistantes est calculé approximativement selon l'équation suivante:

$$\text{Nombre de classes} = [5 \times \log(n)]$$

Où  $n$  est la taille de l'échantillon et  $[.]$  représente la partie entière.

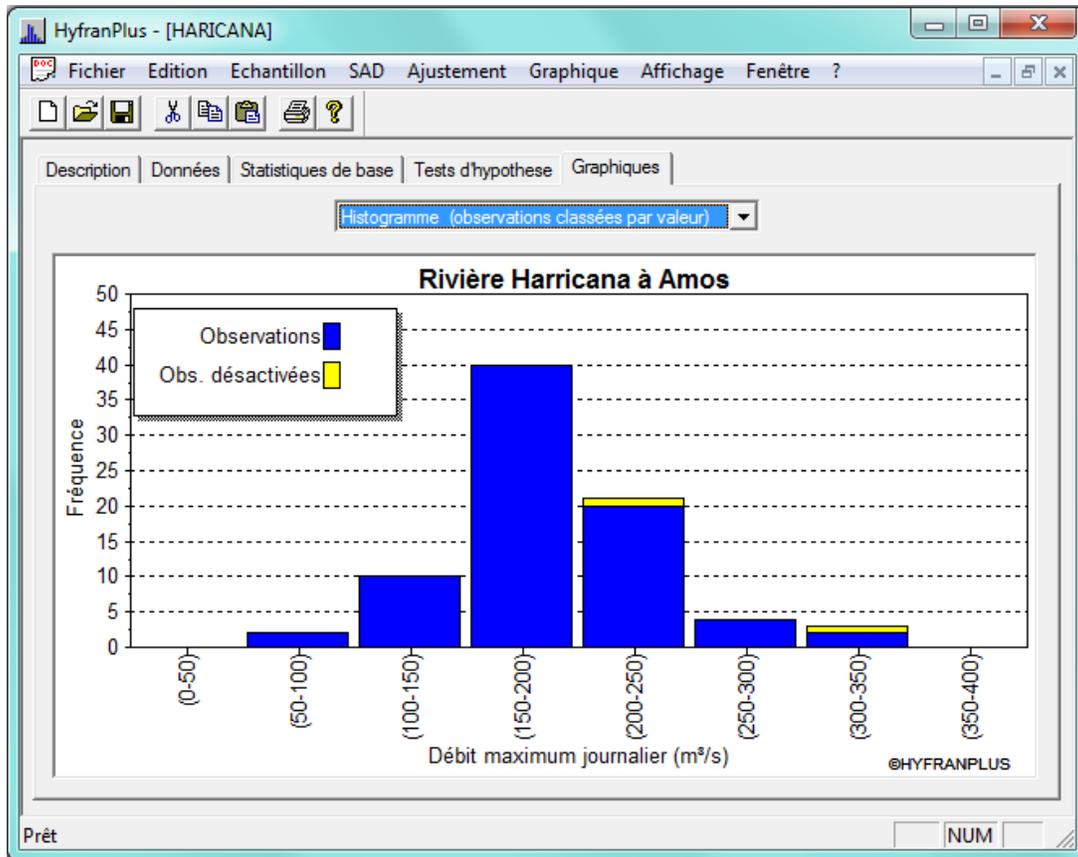


Figure 24 : Histogramme des observations classées par valeur (projet Harricana)

**Note :** Les valeurs désactivées sont également prises en compte dans la représentation sous forme d’histogramme par une couleur différente (Figures 24 et 25).

- c) Histogramme des observations classées par mois (Figure 25) : l’histogramme des observations classées par mois est un graphique permettant de visualiser le nombre d’observation par mois.

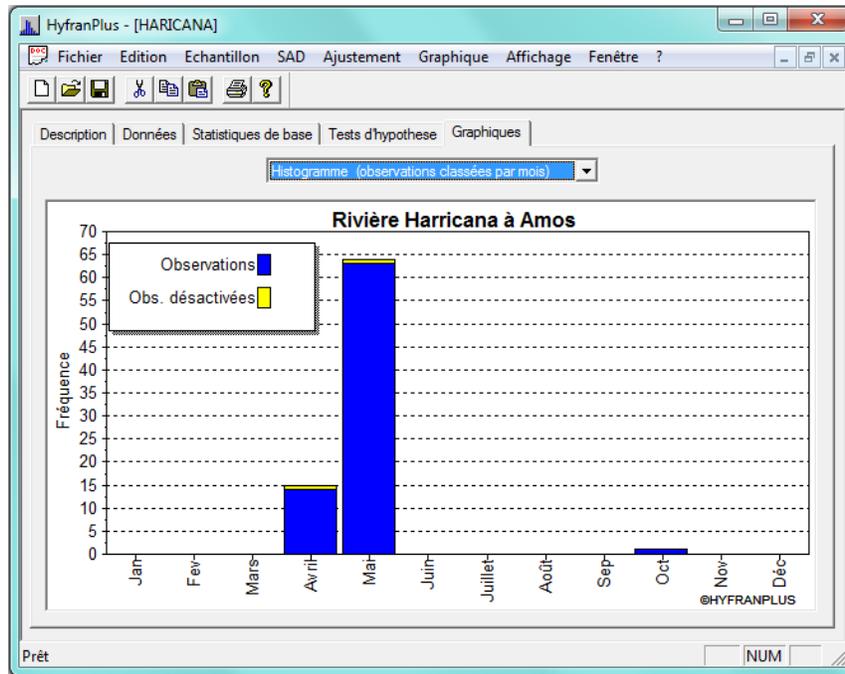


Figure 25 : histogramme des observations classées par mois (projet Harricana)

d) Courbe chronologique (Figure 26) : Ce graphique présente les observations en fonction du temps (année).

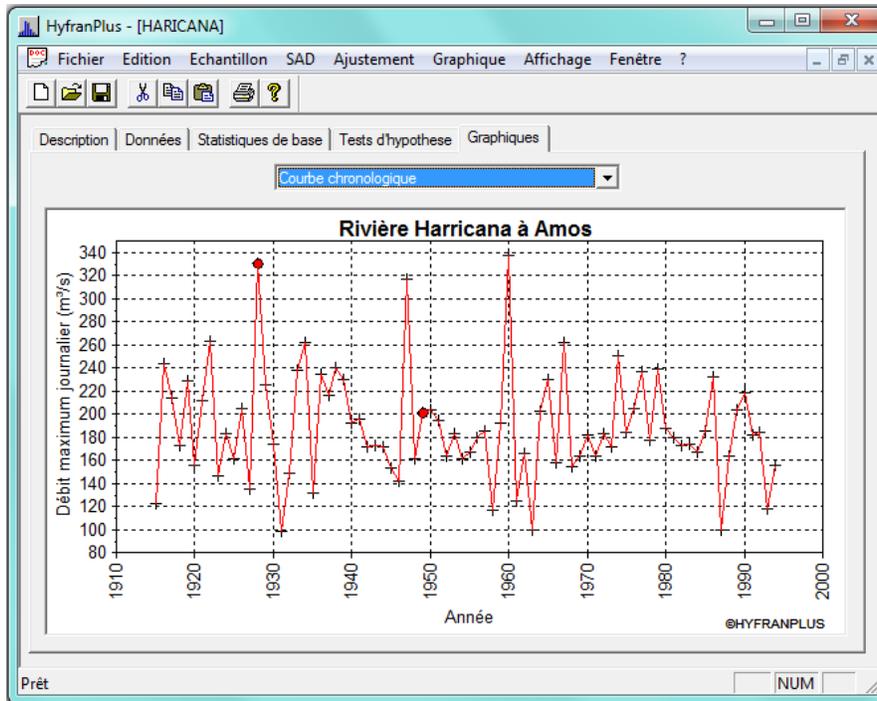


Figure 26 : courbe chronologique des données annuelles (projet Harricana)

**Note :** Les points en rouge correspondent aux valeurs désactivées.

## 2.2. Système d'aide à la décision (SAD)

---

Avant de passer à l'ajustement d'une distribution statistique à une série de données il est souhaitable de déterminer en premier lieu à quelle classe cette distribution appartient (El Adlouni, Bobée et Ouarda, 2008). Pour des fins d'illustration on utilisera 3 séries (voir Annexe 1) :

- La série correspondant au projet Harricana (Tableau A1; Série 1) est disponible à l'installation de HYFRAN-PLUS. On montrera, dans ce qui suit (Figure 30), que la distribution qui s'ajuste le mieux à cette série appartient à la classe C,
- Une série simulée à partir de la loi Log-normale (Tableau A2; Série 2) puis transformée<sup>(1)</sup> en loi normale (Tableau A3; Série 2 transformée). La loi Log-normale est un cas limite des classes C et D (El Adlouni, Bobée et Ouarda, 2008), et,
- Une série simulée à partir d'une loi Gamma (classe D) (Tableau A4; Série 3).

**Note :** <sup>(1)</sup> En effet, si  $X = \ln(Y)$  suit une loi Normale alors  $Y = e^X$  suit une loi Log-normale.

Dans tous ces cas on utilise le SAD comme décrit dans El Adlouni et Bobée (2011) ; El Adlouni, Bobée et Samoud (2012). La Figure B-1 de l'appendice B représente la classification des distributions les plus utilisées, en hydrologie, pour l'ajustement des valeurs extrêmes, par rapport à leurs queues droites.

### a) **Série 1 (Annexe 1 – Tableau A.1) :**

On effectue en premier lieu le test (Cv, Cs) (étape 1 du diagramme SAD, Figure 6) pour décider si on peut ensuite tester la log-normalité. Ceci est effectué en utilisant l'option « diagramme log-normale » disponible dans le menu SAD du logiciel HYFRAN-PLUS (voir Figure 5). Cette option permet de naviguer entre 4 onglets :

- «Graphiques» pour pouvoir observer le diagramme (Cv, Cs) (Figure 27),
- «Décision» qui donne les conclusions du test (Cv, Cs) (Figure 28),
- «Test JB» qui représente les conclusions du test JB, quand on peut l'utiliser,
- «Aide» qui contient l'information utile concernant les tests de log-normalité et de Jarque-Bera (Martel, El Adlouni and Bobée, 2012).

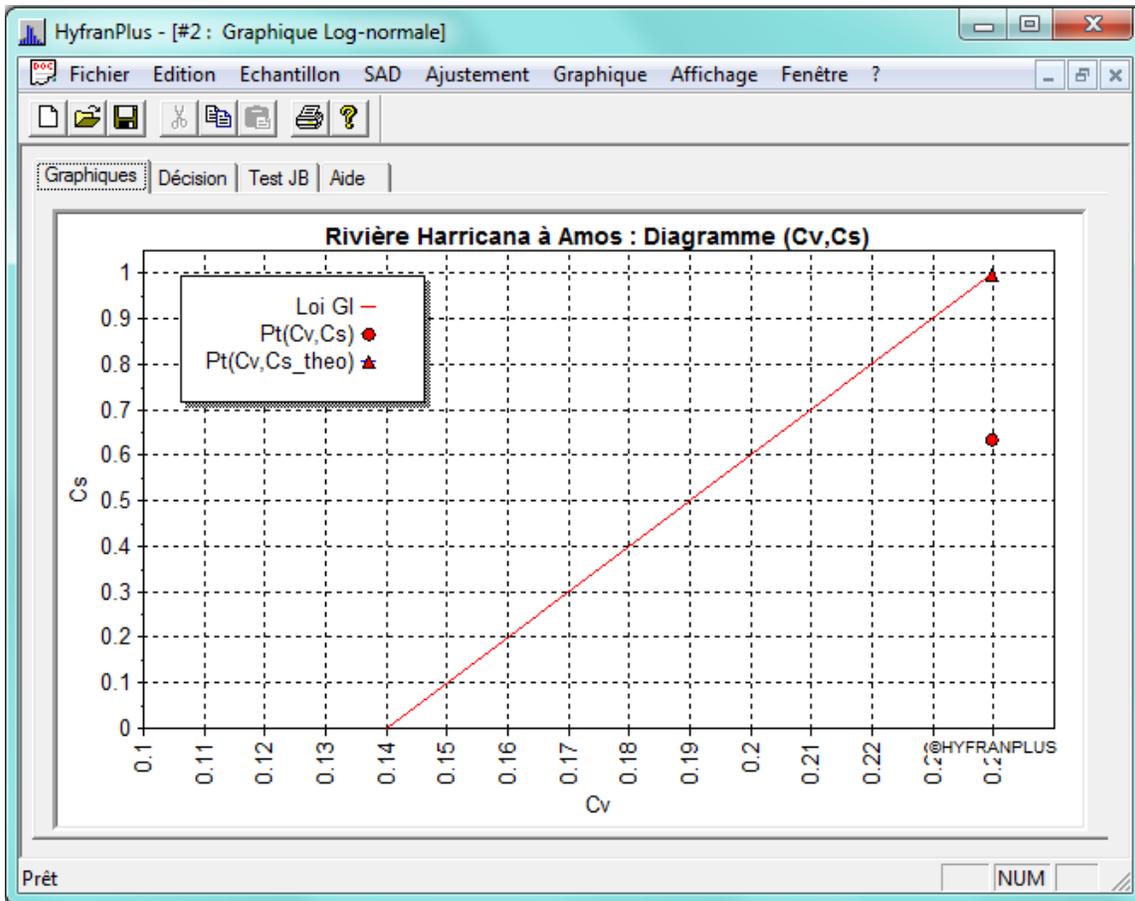


Figure 27 : diagramme (Cv, Cs) préalable à l'utilisation du test de log-normalité (série 1)

- Le point ( $C_v, C_s$ ) observé (•) est situé en dessous de la droite (Figure 27), on déduit donc (Figure 28) qu'on ne peut pas utiliser le test de log-normalité de Jarque-Bera (cf. Martel, El Adlouni and Bobée, 2012).

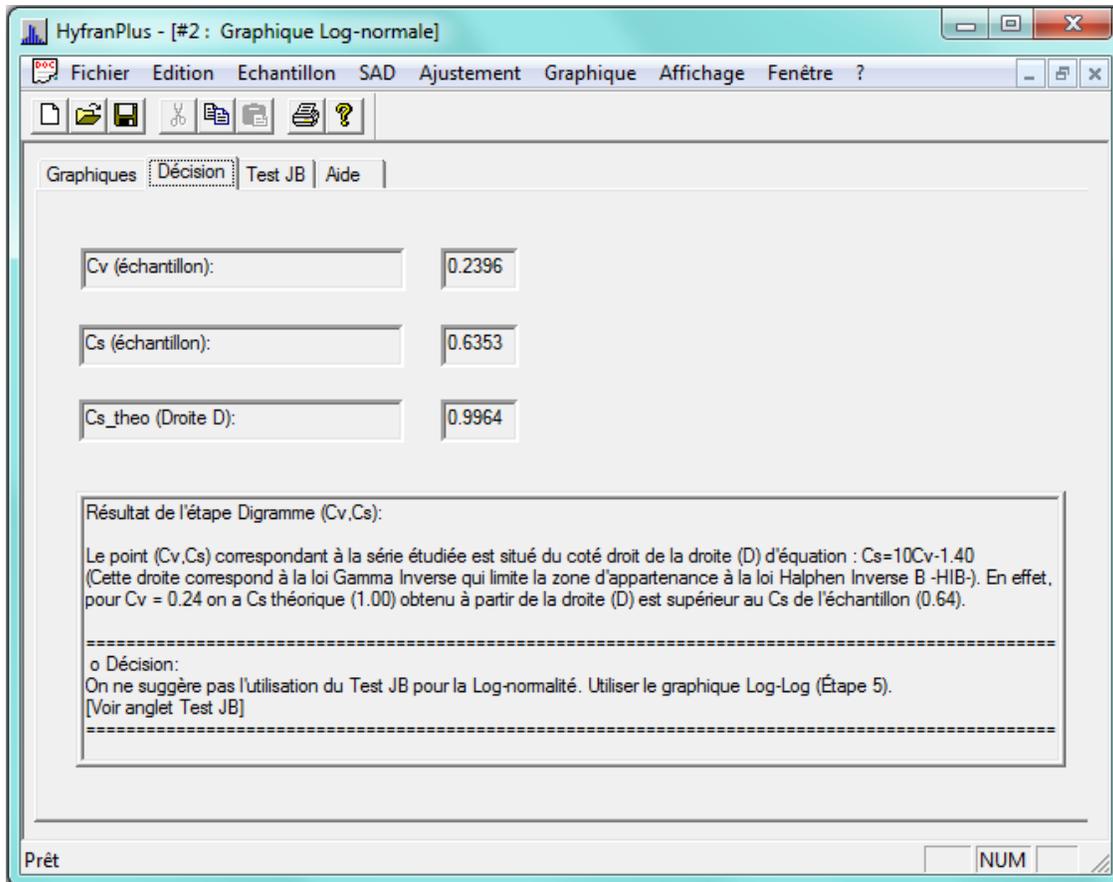


Figure 28 : décision sur l'utilisation du test de log-normalité (série 1)

- Donc on considère le graphique log-log (étape 5 du diagramme SAD, Figure 6) en utilisant l'option «graphique log-log» située dans le menu «SAD» (Figure 5). L'option choisie nous permet de naviguer entre 3 onglets :
  - o «Graphique» qui permet d'afficher le tracé du diagramme log-log pour le projet étudié (Figure 29),
  - o «Décision» qui permet de déterminer la classe à laquelle la série de donnée appartient (Figure 30) et,
  - o «Aide» qui contient des informations utiles relatives au test.

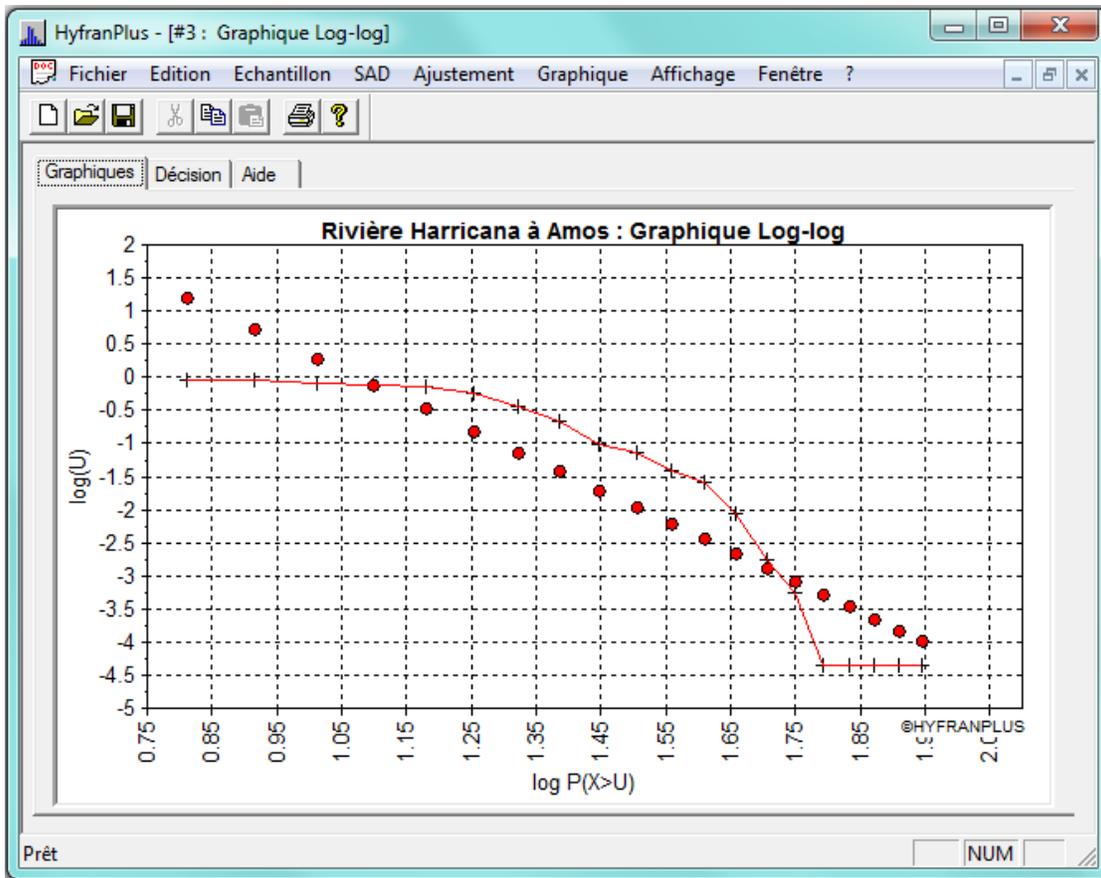


Figure 29 : Graphique log-log (série 1)

La Figure 29 montre que la linéarité de la courbe est acceptable, on en déduit donc que la série appartient à la classe C (étape 6 du diagramme SAD, Figure 6). En effet (Figure 30), le coefficient de corrélation observé étant supérieur à la valeur critique, il n'est donc pas significativement différent de 1. On accepte donc l'hypothèse  $H_0$  : la courbe est linéaire (cf. El Adlouni, Bobée et Samoud, 2012).

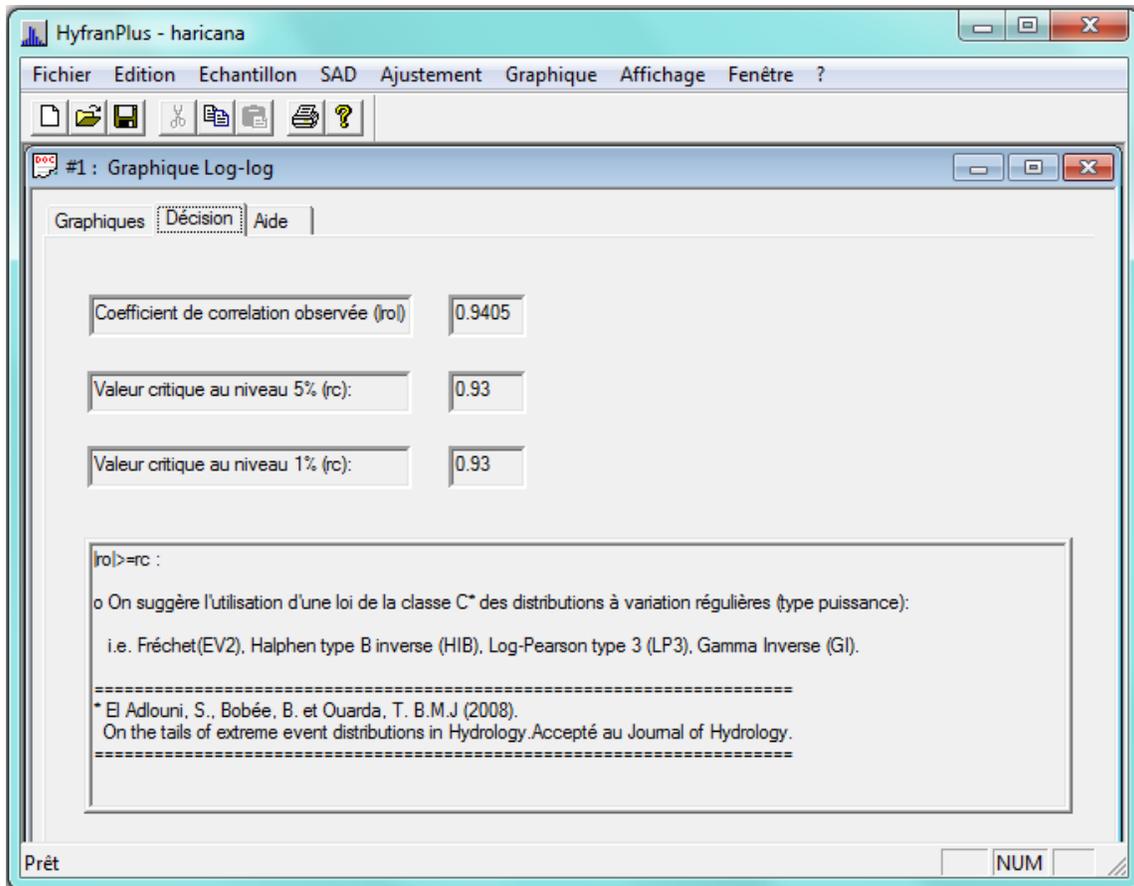


Figure 30 : Décision pour le test log-log (série 1)

**Note :** Les valeurs critiques pour l'étude de la linéarité du diagramme log-log (Figure 30), ont été obtenues par simulation (El Adlouni et Bobée 2011). Nous avons remarqué que les valeurs correspondant aux deux seuils 1% et 5% sont identiques même en considérant des simulations sur 10.000 échantillons.

- On utilise le rapport de Hill et la statistique de Jackson (Figure 5) pour confirmer le choix de la classe C (étape 10 du diagramme SAD, Figure 6). La Figure 31 montre que la statistique de Hill converge vers une valeur constante différente de zéro et la Figure 32 correspond à la statistique de Jackson qui dans ce cas converge vers 2. Ces résultats confirment donc (El Adlouni et Bobée, 2011) que la série 1 peut être représentée par une distribution de la classe C : Fréchet (nommée aussi EV2), Gamma Inverse, Halphen type B Inverse,... (Figure B-1, Appendice B).

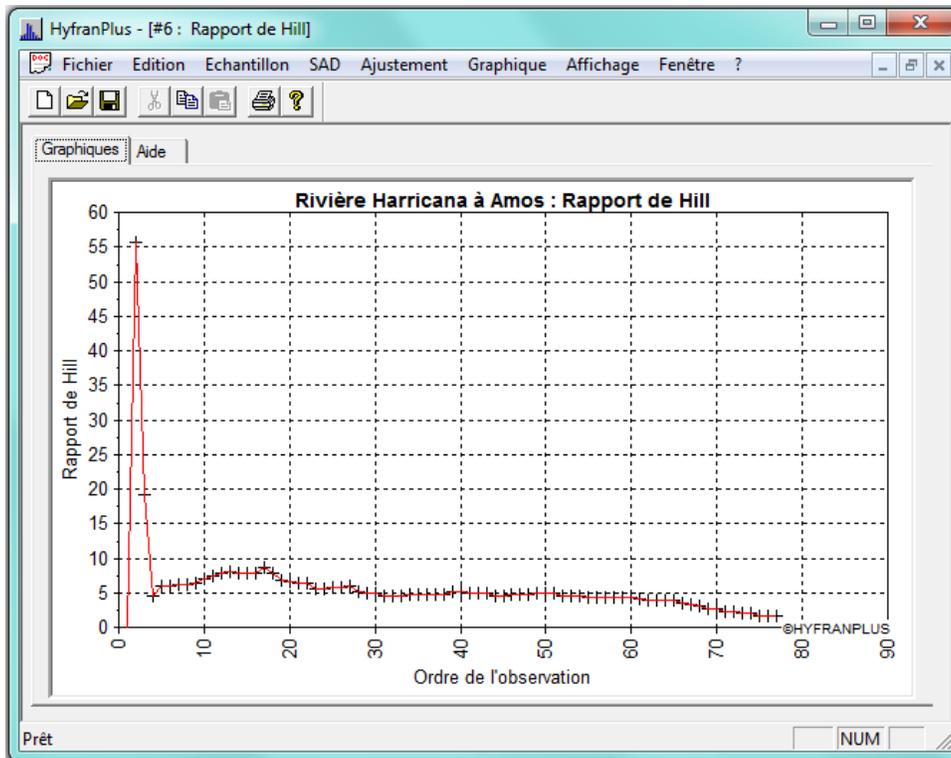


Figure 31 : Rapport de Hill (série 1)

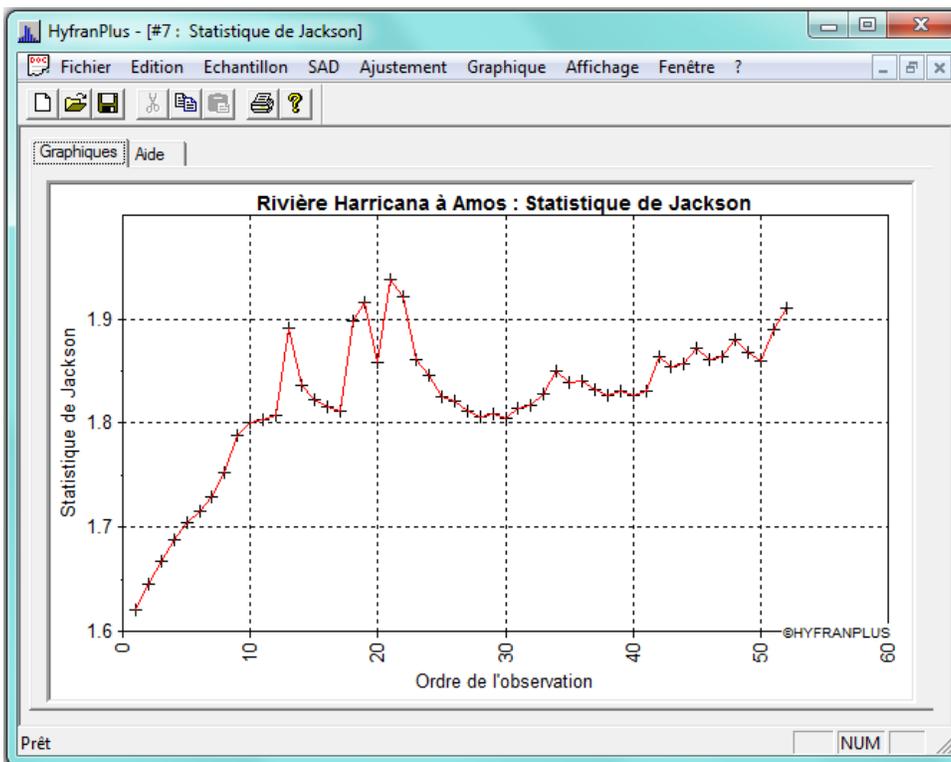


Figure 32 : Statistique de Jackson (série 1)

b) **Série 2 (Annexe 1 – Tableau A.2):**

- On trace le diagramme (Cv, Cs) pour savoir si on doit utiliser le test de log-normalité (étape 1 du diagramme SAD, Figure 6).

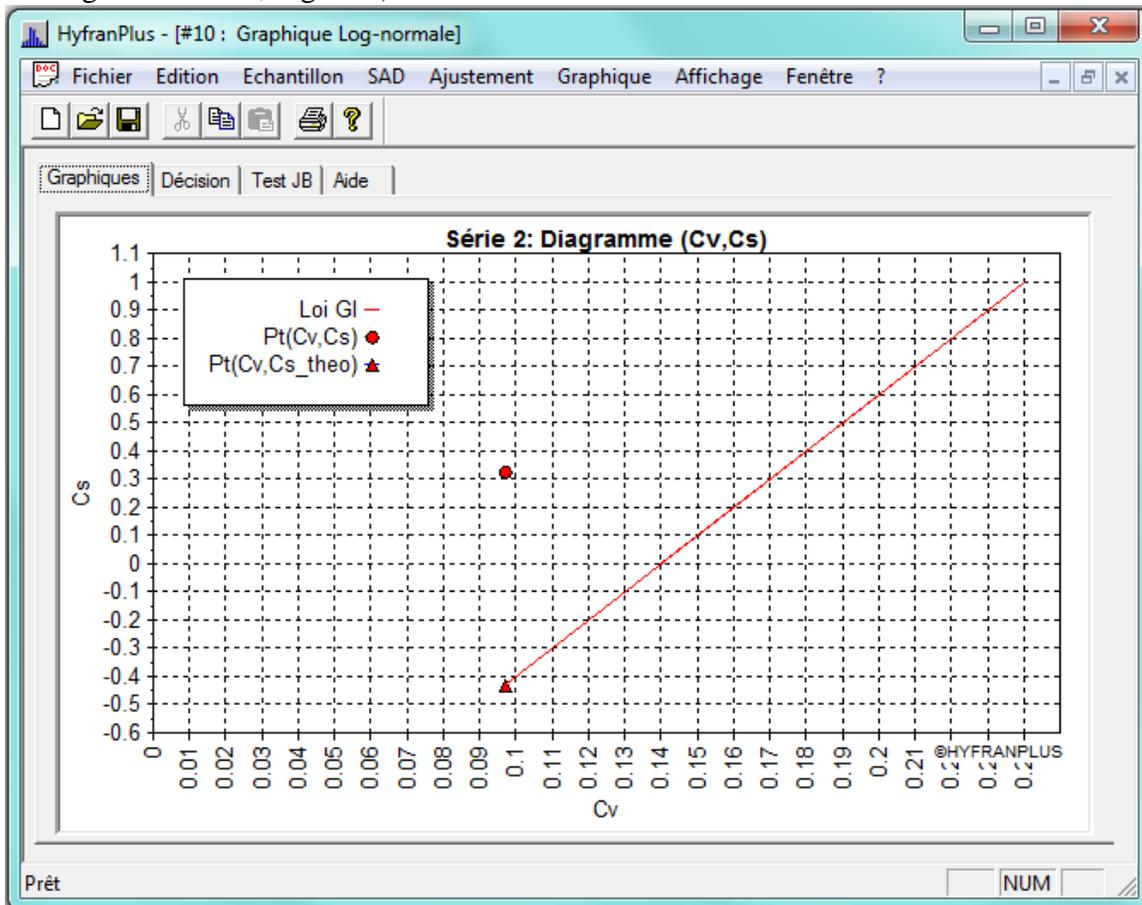


Figure 33 : diagramme (Cv, Cs) préalable à l'utilisation du test de log-normalité (série 2)

Étant donné que le point (Cv, Cs) observé (•) est au-dessus de la droite (Figure 33) et donc appartient à la zone HIB (cf. Martel, El Adlouni and Bobée, 2012), on en déduit donc (Figure 34) que le test de log-normalité est applicable (étape 2 du diagramme SAD, Figure 6). On effectue donc la transformation logarithmique (Figure 17) des données du tableau A.2 (Annexe 1) pour obtenir le Tableau A.3.

- La décision du diagramme (Cv, Cs) permet de tester la log-normalité (on applique donc le test de Jarque-Bera (JB)).

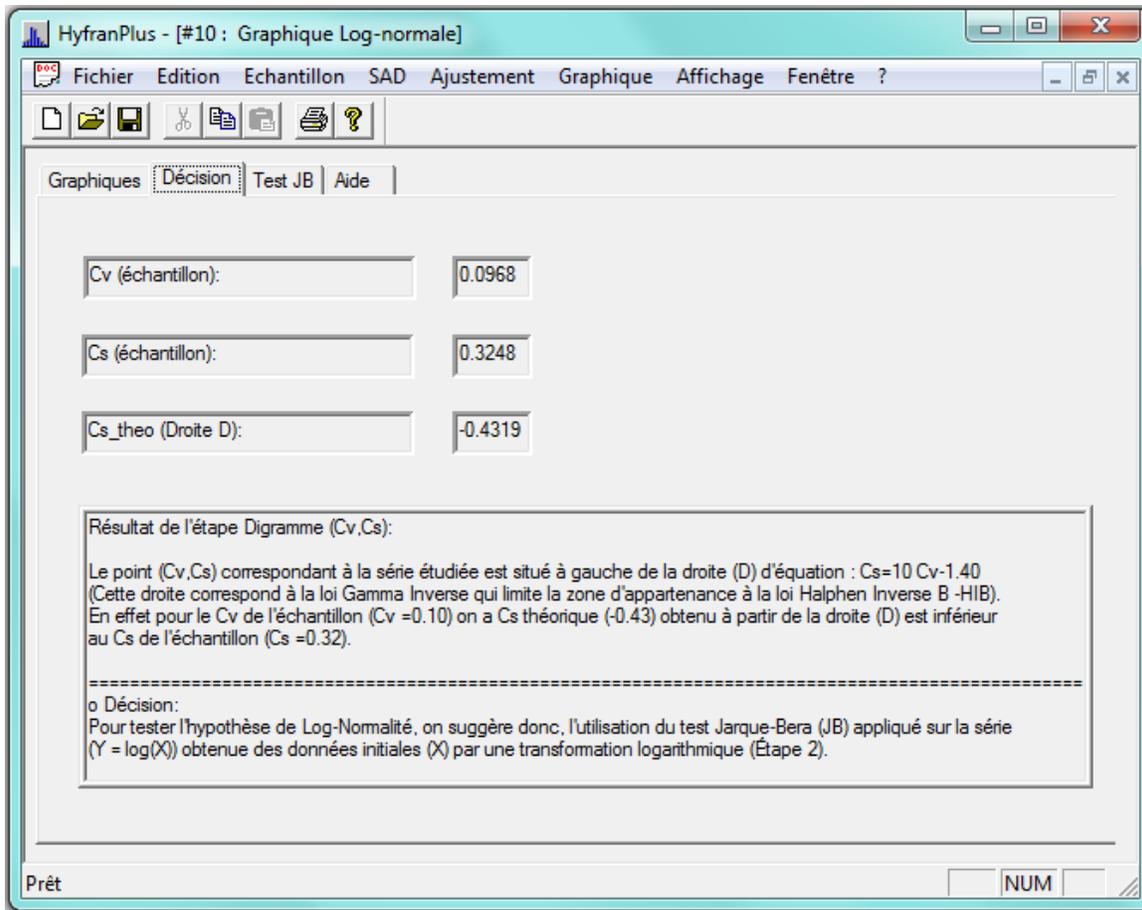


Figure 34 : Décision sur l'utilisation du test de log-normalité (série 2)

Le test de JB (Figure 35) démontre que l'hypothèse de la log-normalité est satisfaisante, on suggère donc l'emploi d'une distribution log-normale pour la représentation de la série 2 des données (étape 3 du diagramme SAD, Figure 6).

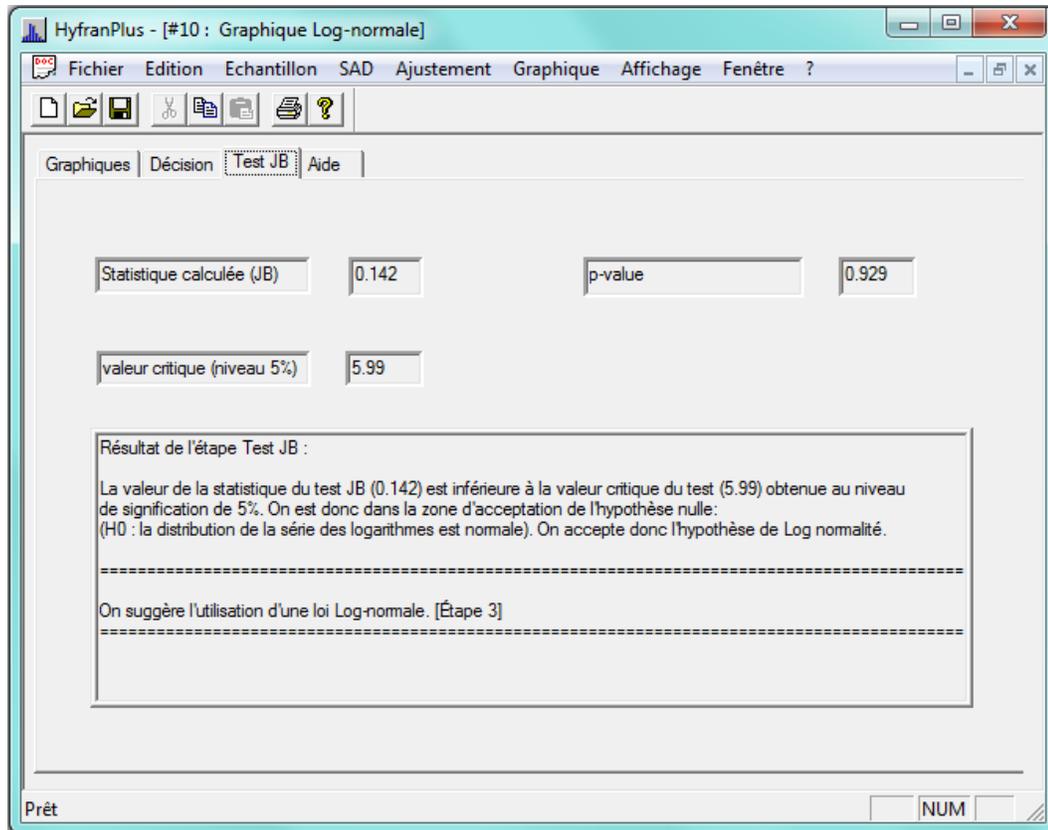


Figure 35 : Décision du test de Jarque-Bera (série 2)

On remarque (Figure 36) que si l'on choisit de représenter les données de la série 2 transformée logarithmiquement) par une distribution normale, l'ajustement correspondra à une droite sur du papier de probabilité normal. Cela confirme la validité de la loi Log-normale pour la série 2. En effet, Si  $Y=\ln(X)$  suit une loi normale alors  $X$  suit une loi Log-normale.

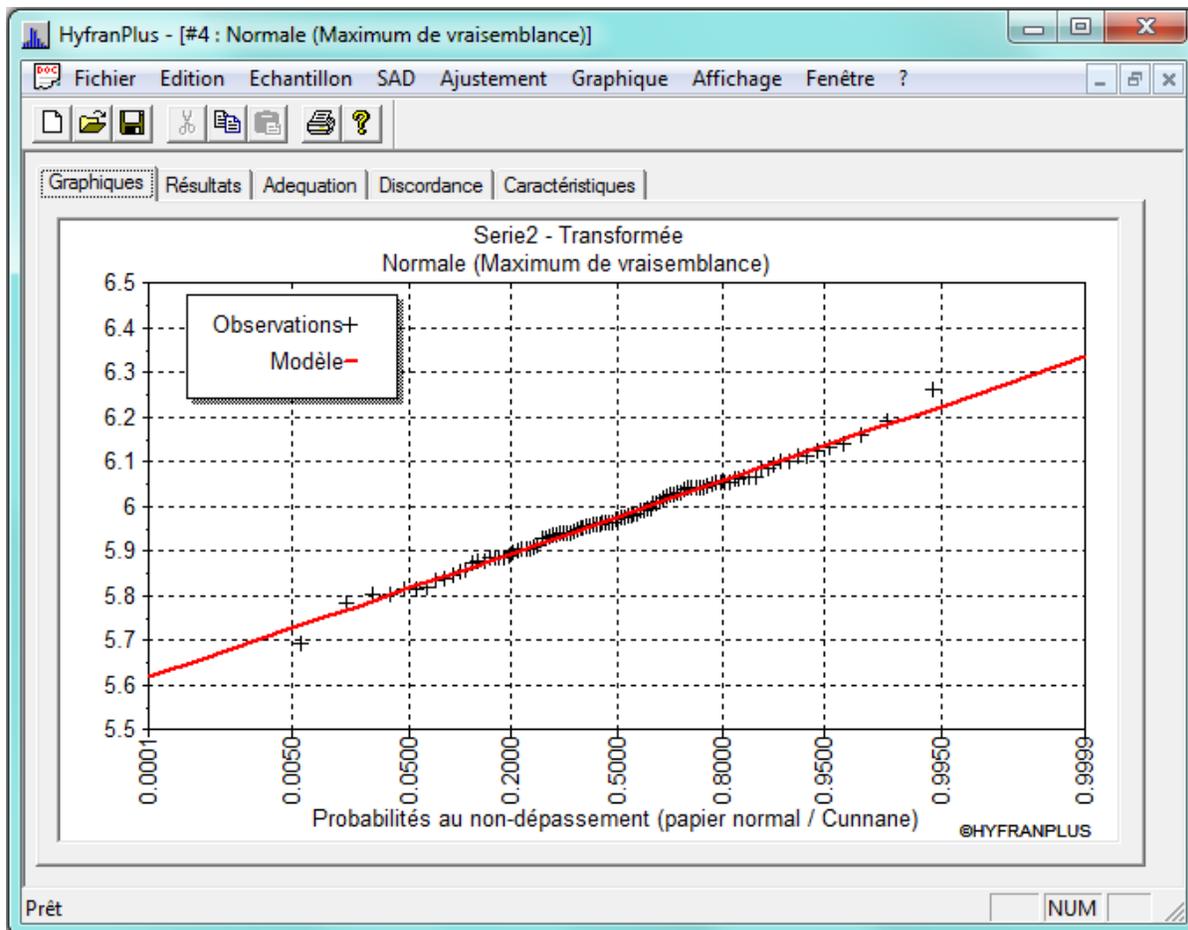


Figure 36 : Représentation graphique de l'ajustement de la série 2 transformée, par la distribution Normale (échelle logarithmique)

c) **Série 3 (Annexe 1 – Tableau A.4):**

- On effectue le test ( $C_v$ ,  $C_s$ ) (Figure 37) pour déterminer si on peut ensuite tester la log-normalité (étape 1 du diagramme SAD, Figure 6).

Le point ( $C_v$ ,  $C_s$ ) observé (•) est situé en dessous de la droite (Figure 37), on déduit donc (Figure 38) qu'on ne peut pas utiliser le test de log-normalité de Jarque-Bera (cf. Martel, El Adlouni and Bobée, 2012), comme dans le cas de la série 1.

On considère alors, le graphique log-log (étape 5 du diagramme SAD, Figure 6) en utilisant l'option «graphique log-log» située dans le menu «SAD» (Figure 5).

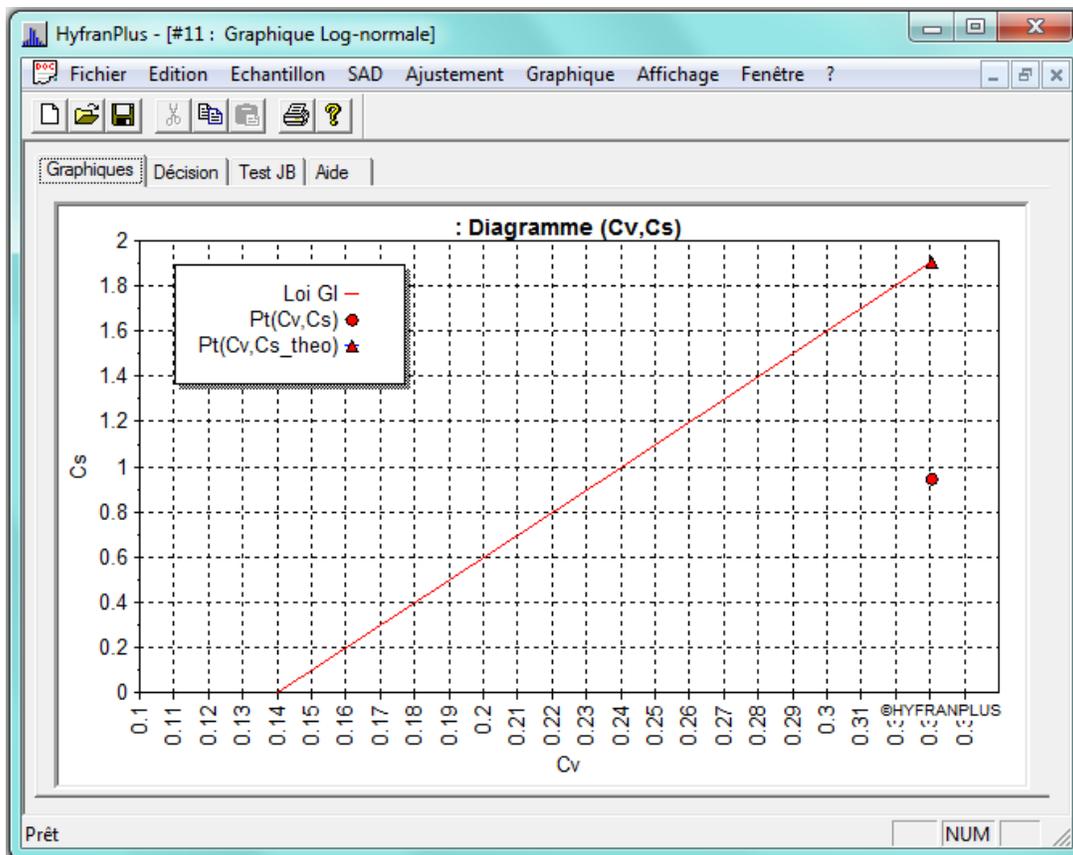


Figure 37 : Diagramme (Cv, Cs) préalable à l'utilisation du test de log-normalité (série 3)

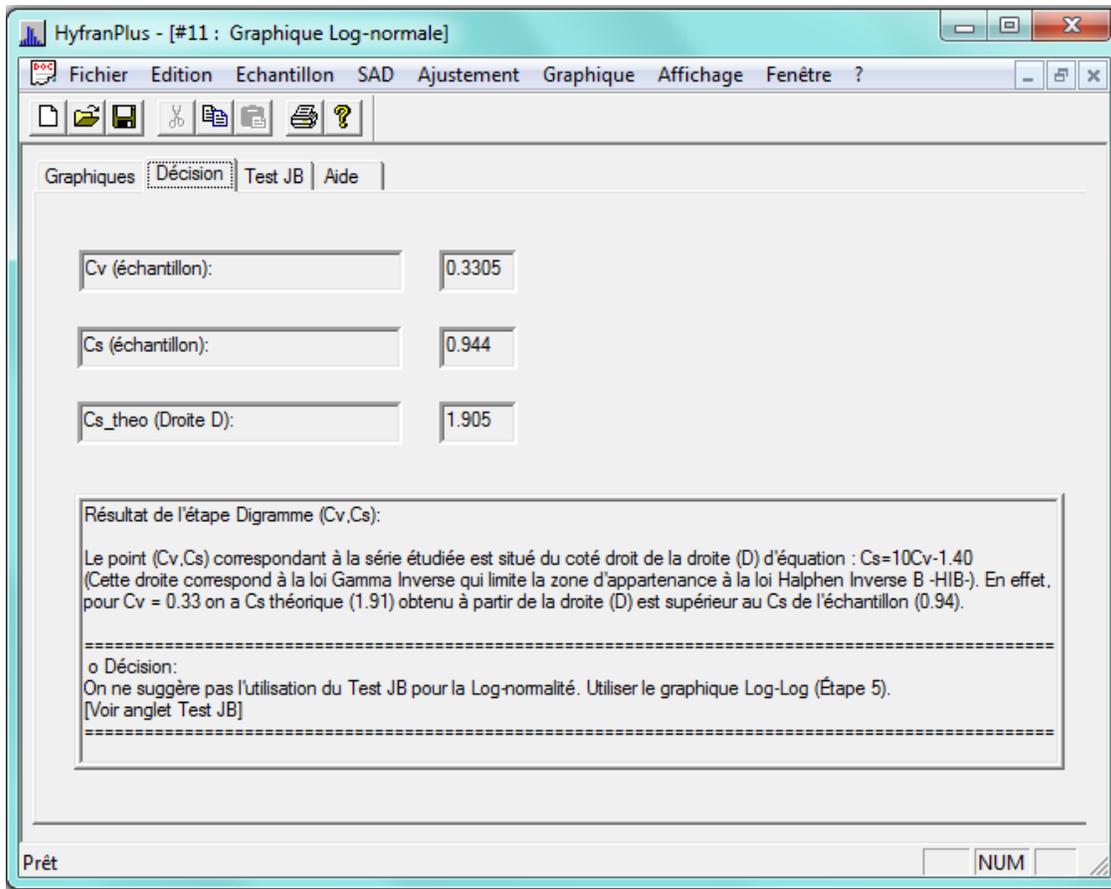


Figure 38 : Décision sur l'utilisation du test log-normalité (série 3)

Nous pouvons constater, en considérant le digramme log-log (Figure 39), que la courbe n'est pas linéaire (El Adlouni, Bobée et Samoud, 2012); on en déduit donc (Figure 40) qu'il faut utiliser le graphique de la Fonction Moyenne des Excès (FME) (étape 7 du diagramme SAD, Figure 6).

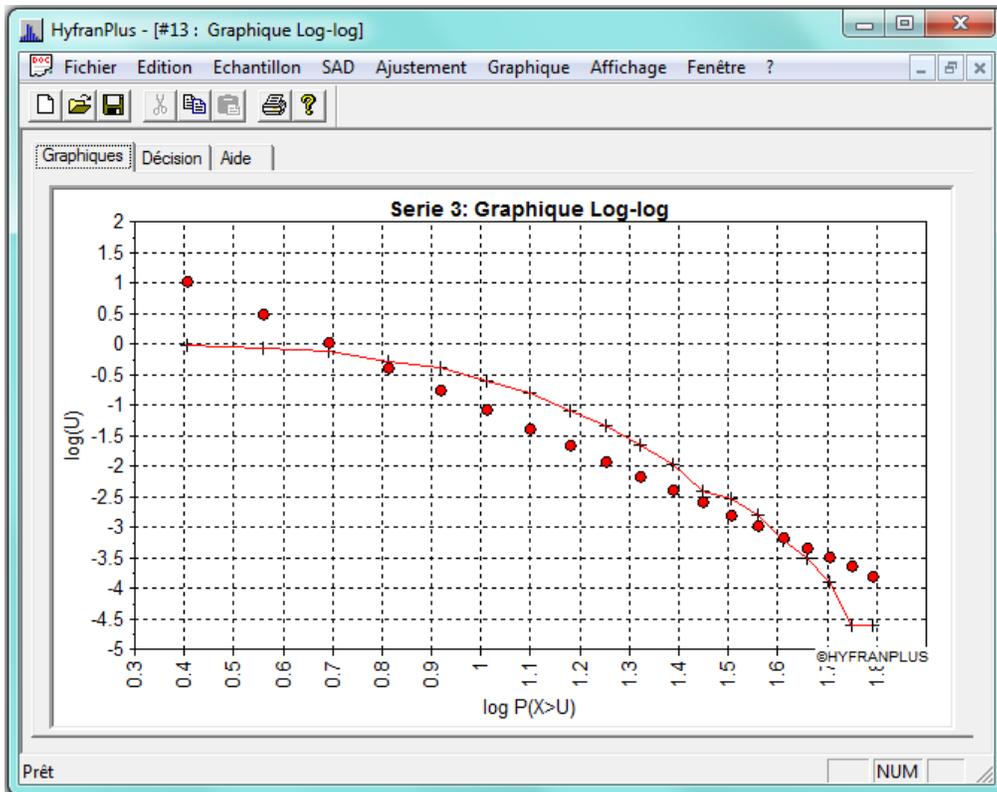


Figure 39 : Graphique log-log (série 3)

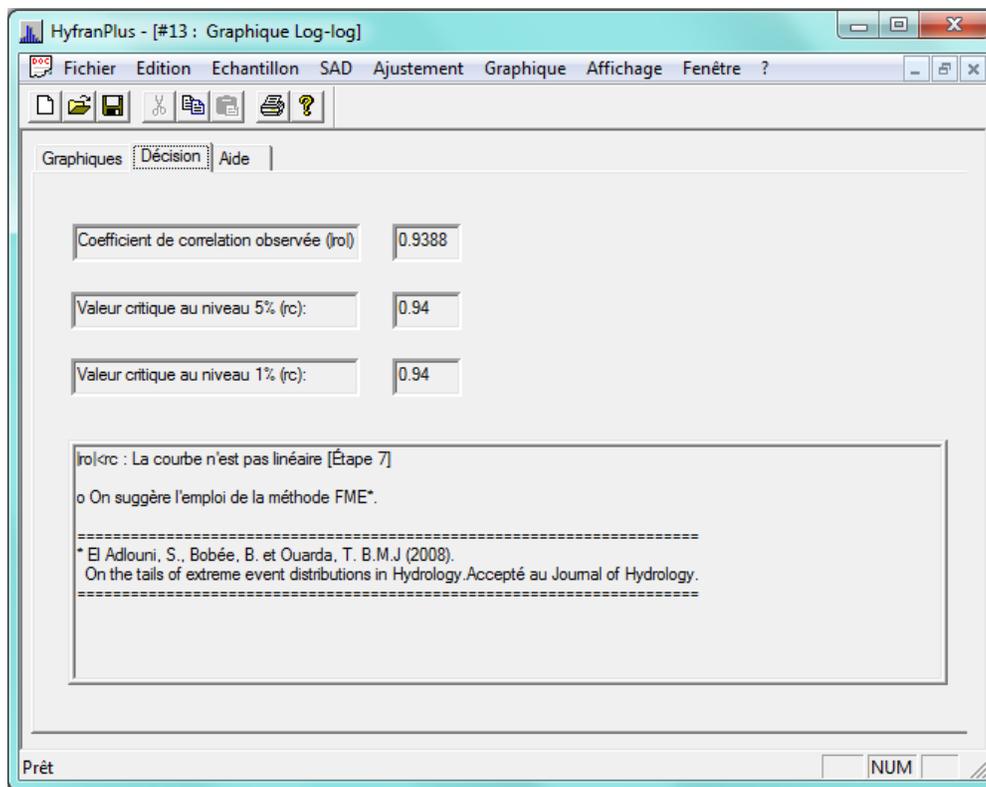


Figure 40 : Décision du graphique log-log (série 3)

- Pour effectuer le test de la FME on sélectionne l'option Fonction Moyenne des Excès que l'on retrouve dans le menu SAD (Figure 5). Cette option permet de naviguer entre 3 onglets :
  - o «Graphique » pour observer le tracé de la FME (Figure 41),
  - o «Décision» qui permet d'obtenir les conclusions du test de la FME (Figure 42) et,
  - o «Aide» qui contient l'information utile concernant le graphique e la FME.

On remarque que la pente est positive (Figure 41), on peut en déduire (El Adlouni et Bobée, 2011) que les distributions de la classe D peuvent être utilisées pour représenter les données de la série 3 (étape 9 du diagramme SAD, Figure 6; Figure B-1, Appendice B). Ceci est confirmé par la Figure 42. En effet, la pente observée  $ao$  étant supérieure à la valeur critique  $ac$ , on rejette donc l'hypothèse  $H_0$  que la pente est nulle (Étape 9 du diagramme SAD, Figure 6) (El Adlouni, Bobée et Samoud, 2012).

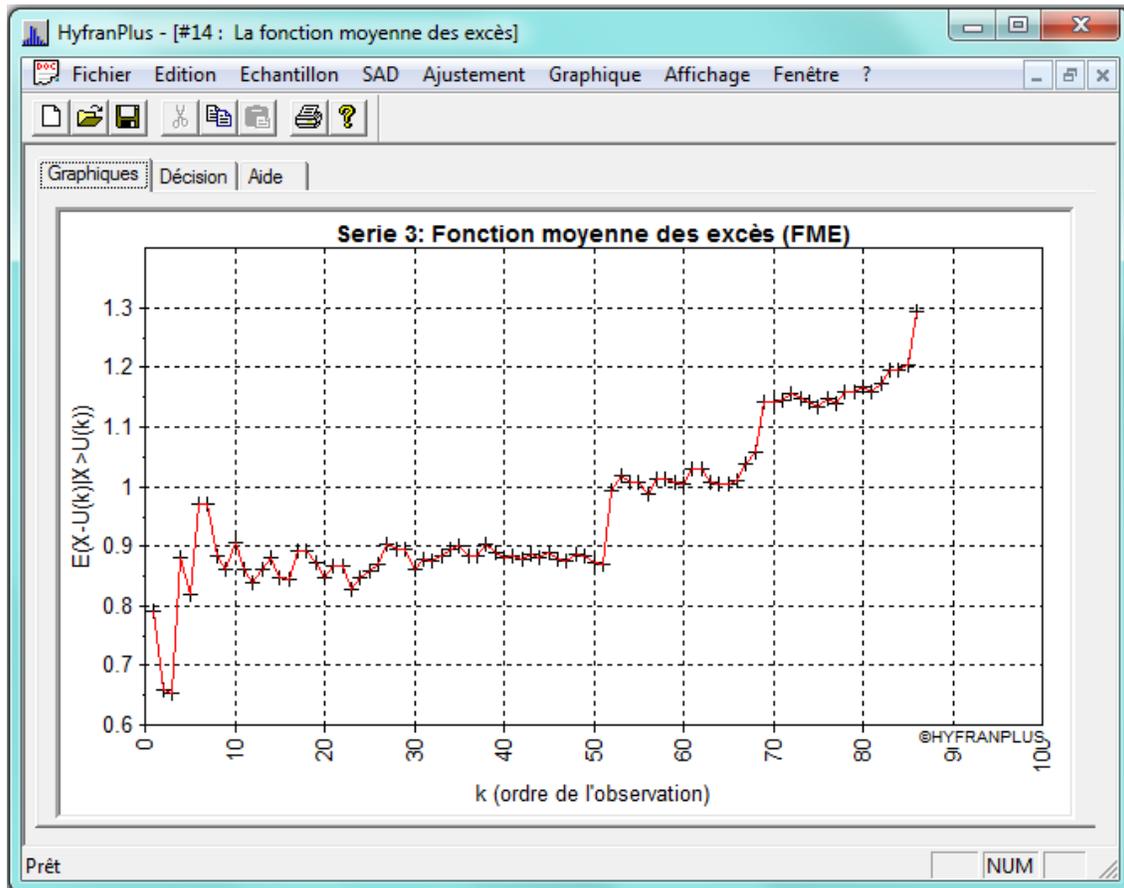


Figure 41 : Fonction Moyenne des Excès (série 3)

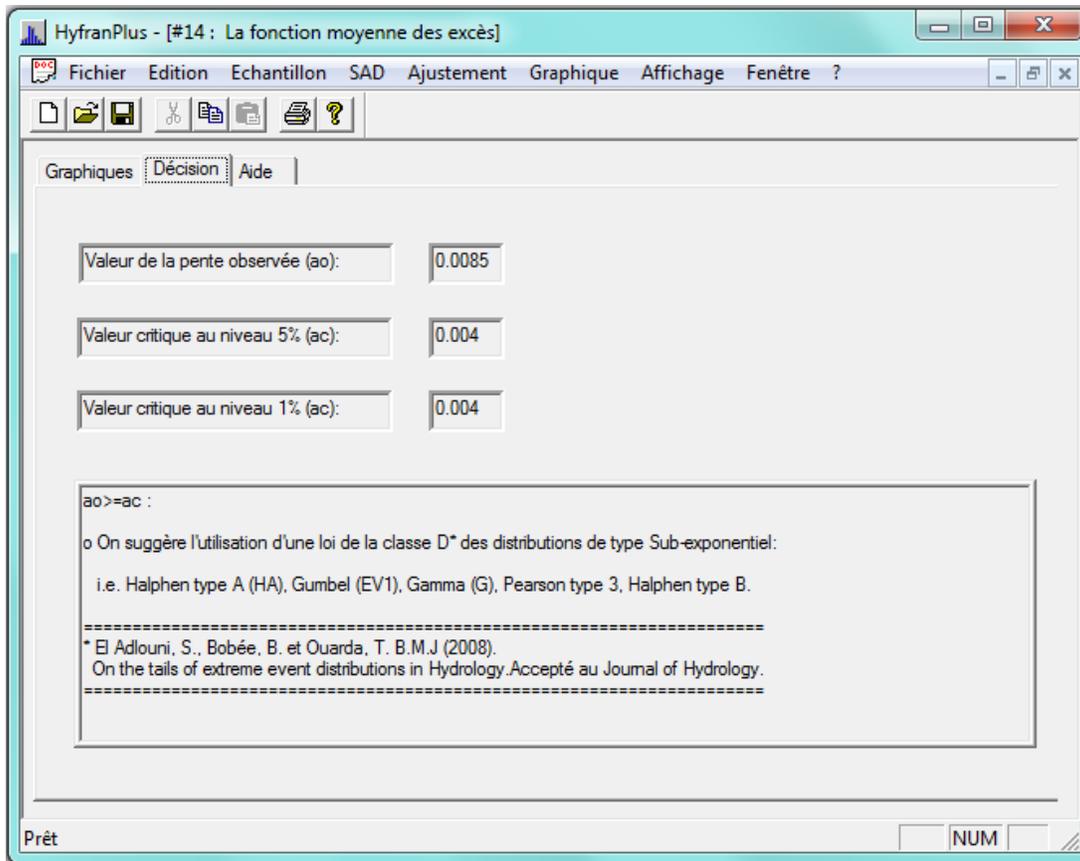


Figure 42 : Décision pour la Fonction Moyenne des Excès (série 3)

**Note :** Les valeurs critiques pour le diagramme FME (Figure 42), ont été obtenues par simulation (El Adlouni et Bobée 2011). Nous avons remarqué que les valeurs correspondant aux deux seuils 1% et 5% sont identiques même en considérant des simulations avec 10.000 échantillons.

- On utilise le rapport de Hill et la statistique de Jackson (Figure 5) pour confirmer le choix de la classe D (étape 10 du diagramme SAD, Figure 6). La Figure 43 montre que la statistique de Hill converge vers une valeur nulle et la Figure 44 qui correspond à la statistique de Jackson présente des irrégularités et ne converge pas vers 2. On en déduit donc (El Adlouni, Bobée et Samoud, 2012) que la série 3 peut être représentée par une distribution de la classe D (Figure B-1, Appendice B), de type sub-exponentiel : Gumbel, Halphen type B, Halphen type A,....

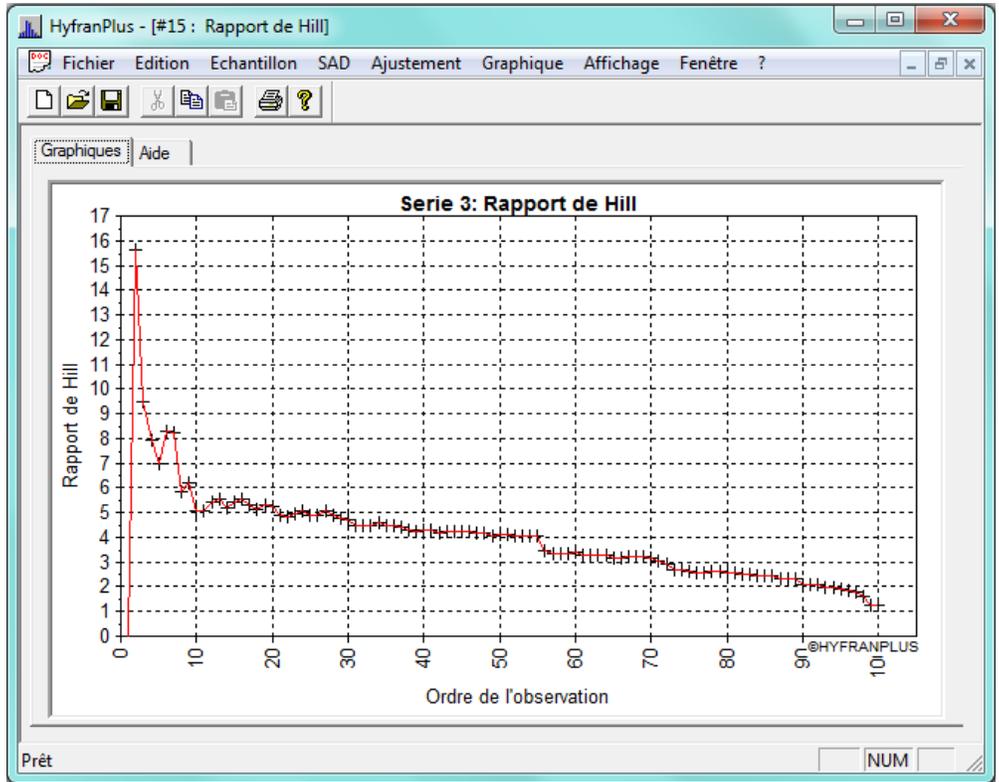


Figure 43: Rapport de Hill (série 3)

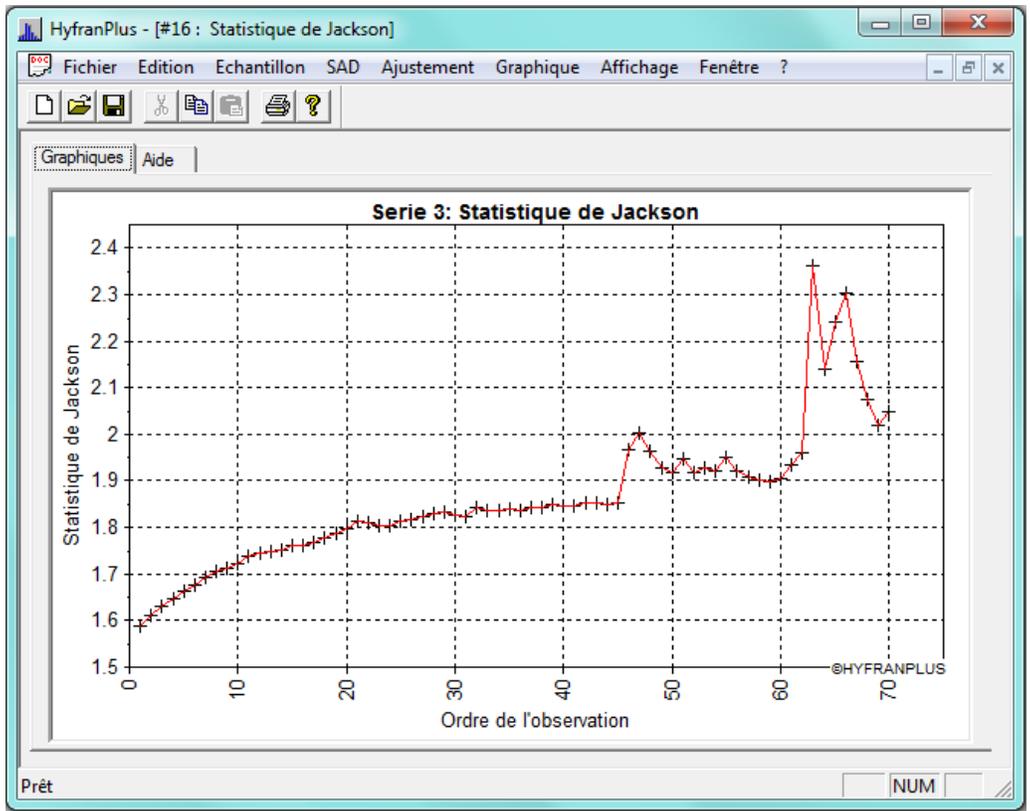


Figure 44 : Statistique de Jackson (série 3)

## 2.3. Ajustement d'une distribution statistique à une série de données

---

Dans ce qui précède, on a utilisé le SAD pour choisir la classe qui représente le mieux la forme de la distribution empirique. On procède ensuite à l'ajustement des distributions de la classe retenue à la série des observations. Les classes C et D contiennent plusieurs distributions (Figure B-1, Appendice B). Le choix de la distribution la plus adéquate peut être effectué par le biais de visualisation graphique ou en utilisant les critères d'information (AIC et BIC).

Dans ce qui suit, on utilisera, à titre d'illustration, la série 1 qui peut être représentée par une distribution de la classe C (cf. Figure 30). HYFRAN-PLUS permet d'ajuster différentes distributions statistiques à un échantillon aléatoire d'observations IID en considérant différentes méthodes d'estimation.

Pour effectuer un ajustement on suit les étapes suivantes:

- Choisir, tour à tour, chacune des distributions de la classe choisie (Classe C dans le cas de la série 1) dans le menu "Ajustement" (Figure 7);
- Choisir, pour chaque distribution, une méthode d'estimation dans la fenêtre contextuelle qui apparaît ensuite;
- Appuyer sur le bouton «OK».

---

### 2.3.1. Choix de la distribution statistique la plus adéquate

---

On a déjà démontré (Figure 30) que la série 1 pouvait être représentée par une distribution de la classe C (Fréchet, Gamma Inverse, Halphen type Inverse B, Log-Pearson type 3) (Figure B-1, Appendice B). Or le logiciel HYFRAN-PLUS permet de comparer l'ajustement de plusieurs distributions afin de choisir le modèle le plus adéquat pour représenter la série de données. Dans ce qui suit on présentera les méthodes (graphique et critères d'information) permettant de discriminer les distributions. On survolera l'ajustement pour les besoins de cette section.

#### *Comparaison graphique*

On peut comparer les résultats de plusieurs ajustements différents à l'aide d'un graphique sur papier de probabilité normal ou Gumbel. Il faut premièrement effectuer consécutivement l'ajustement des distributions (cinq au maximum) à comparer et ensuite sélectionner l'option «Comparaison» du menu «Ajustement» (Figure 7). Une boîte de dialogue qui permet de sélectionner les distributions apparaîtra.

Après avoir sélectionné les distributions d'intérêt on obtient un graphique représentant les distributions choisies ainsi que les probabilités empiriques des données de la série 1 (Figure 45).

**Note :** On ne peut pas effectuer l'ajustement de la distribution Halphen de type B Inverse pour cette série de données car le système d'équations du maximum de vraisemblance ne converge pas dans ce cas (Perreault, Bobée et Rasmussen, 1999).

On remarque sur la Figure 45 que la distribution Gamma Inverse (en bleu) représente mieux la série de données que les distributions Fréchet (en rouge) et Log-Pearson 3 (en vert).

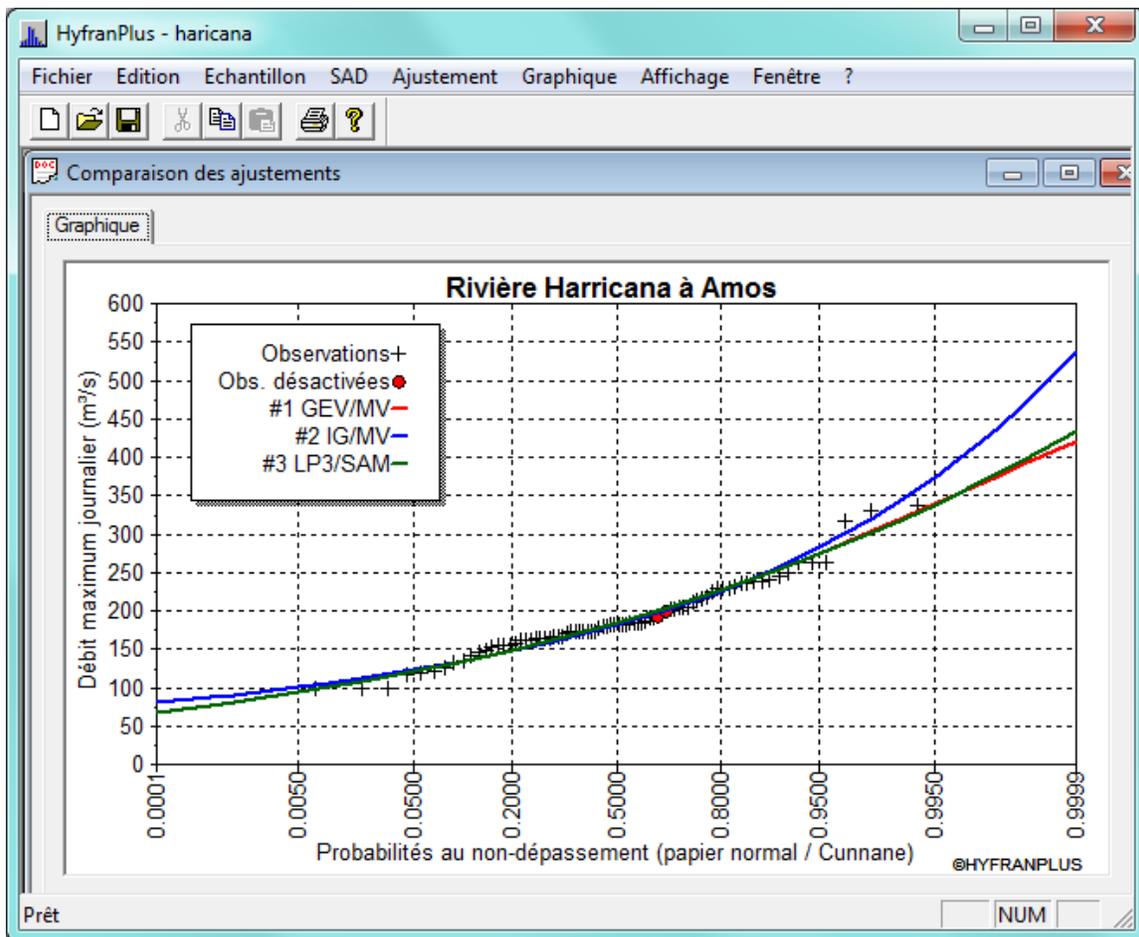


Figure 45 : comparaison graphique des ajustements de la série 1

**Note :** La distribution de Fréchet (EV2) correspond à un cas particulier de la loi GEV (Generalized Extreme Value).

#### Critères d'information

Pour discriminer les différents ajustements on peut en outre utiliser les deux critères suivants disponibles dans HYFRAN-PLUS (Ehsanzadeh, El Adlouni and Bobée, 2010):

- Critère d'information d'Akaike (AIC)
- Critère d'information bayésien (BIC)

Ces deux critères utilisés consécutivement (AIC puis BIC) permettent de construire un classement de modèles statistiques tenant compte du principe de parcimonie. Les meilleurs ajustements correspondent aux plus faibles valeurs de ces critères (Ehsanzadeh, El Adlouni and Bobée, 2010).

Pour faire la comparaison il faut premièrement effectuer consécutivement l'ajustement des distributions de la même classe à comparer et ensuite sélectionner l'option «Comparaison» du menu «Ajustement» (Figure 7). Une boîte de dialogue qui permet de sélectionner les distributions (cinq au

maximum) apparaîtra. Dans le cas du BIC, après avoir sélectionné les distributions d'intérêt, on spécifie le rapport des probabilités *a priori* (i.e.  $P(M_i)$  le poids donné *a priori* à chaque distribution) ainsi que la période de retour. La probabilité *a posteriori*  $P(M_i|x)$  est ensuite déterminée, pour chaque distribution, en tenant compte de la série des observations pour en déduire le BIC (voir Ehsanzadeh, El Adlouni et Bobée, 2010). Cette procédure conduit à une fenêtre qui contient toutes les informations relatives aux distributions comparées (Figure 46).

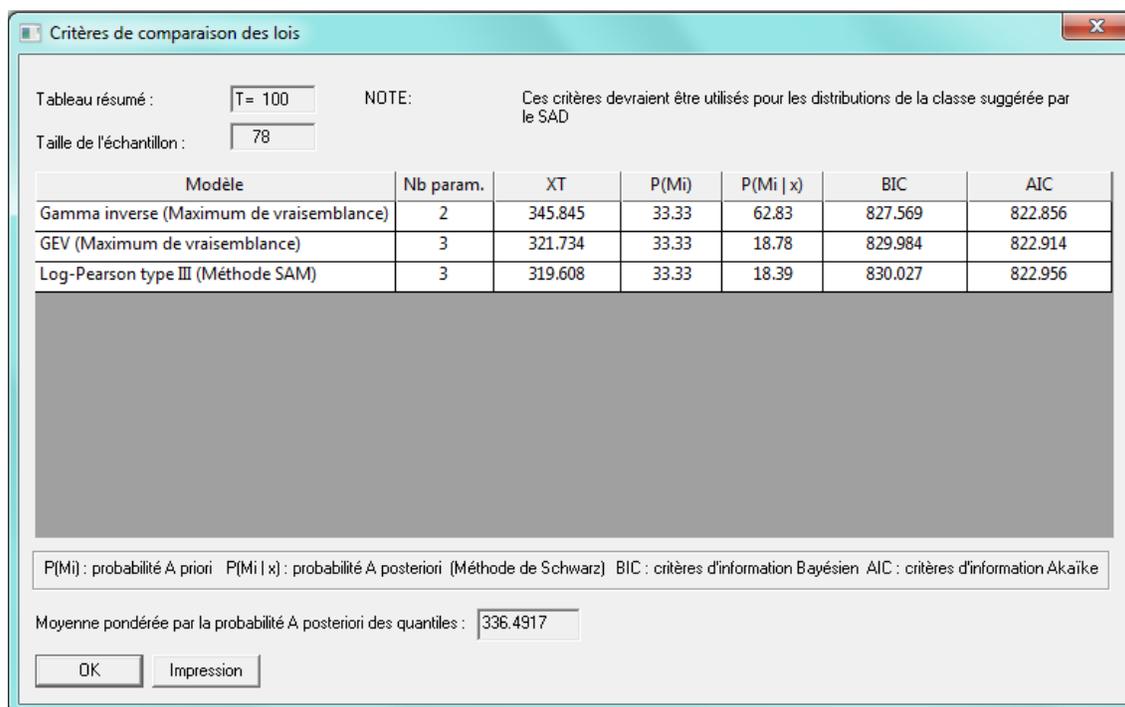


Figure 46 : comparaison à l'aide de critères d'information des ajustements de la série 1

Les critères AIC (Critère d'Information d'Akaike) et BIC (Critère d'Information Bayésien) ainsi que les notations du tableau de la Figure 46 sont expliqués en détail dans Ehsanzadeh, El Adlouni et Bobée (2010).

**Note :** Les critères AIC et BIC favorisent l'utilisation de distributions ayant le plus petit nombre de paramètres (principe de parcimonie). Cependant, de manière générale pour la modélisation des séries de maxima annuels en hydrologie, Morlat (1956) recommande l'utilisation des distributions à trois paramètres pour prendre en compte la forme (asymétrie des distributions). En effet, dans le cas de distributions à 2 paramètres la forme est fixée (par exemple, le coefficient d'asymétrie pour la loi de Gumbel est 1.137)

---

### 2.3.2. Ajustement

---

En accord avec la comparaison graphique (Figure 45) et les critères AIC et BIC (Figure 46) la distribution Gamma inverse est la plus adéquate pour représenter la série 1 (Rivière Harricana à Amos entre 1915-1994, Tableau A.1, Annexe 1). On représentera donc cette série en utilisant la distribution Gamma Inverse pour la partie du tutoriel concernant l'ajustement.

Lorsqu'on effectue l'ajustement d'une distribution statistique à un échantillon, une boîte à onglets apparaît (Figures 2 et 7), présentant les résultats de l'ajustement ainsi que les informations connexes.

Pour accéder aux différents aspects de l'ajustement on navigue entre les onglets suivants (Figure 47):

- Résultats
- Graphique
- Adéquation
- Caractéristiques de la population
- Discordance (apparaît seulement pour les lois Normale et Log-normale)

#### *Résultats de l'ajustement*

La fenêtre des résultats de l'ajustement (Figure 47) présente les options suivantes:

- Le projet, autrement dit le nom du fichier qui contient l'échantillon et son chemin d'accès;
- Le titre du projet;
- La taille de l'échantillon;
- La valeur des paramètres estimés de la loi Gamma Inverse;
- Les quantiles  $x_T$  pour 21 périodes de retour prédéfinies. Dans l'ordre de gauche à droite on a :
  - la période de retour ( $T = 1/p$ ), où  $p$  est la probabilité au dépassement de  $x_T$ ;
  - la probabilité au non-dépassement ( $q = 1 - p$ ),
  - la valeur du quantile  $x_T$  correspondant, l'écart-type du quantile  $\sigma_{x_T}$  et enfin,
  - l'intervalle de confiance  $IC$  du quantile à un niveau de confiance donné (La démarche théorique pour déterminer le quantile  $x_T$ , l'écart-type du quantile  $\sigma_{x_T}$  et l'intervalle de confiance associé sont disponibles dans l'Appendice A, Bobée et El Adlouni, 2015);
- Un bouton «Autre période de retour»;
- Le niveau de confiance (95% par défaut) de l'intervalle de confiance que l'on peut modifier.

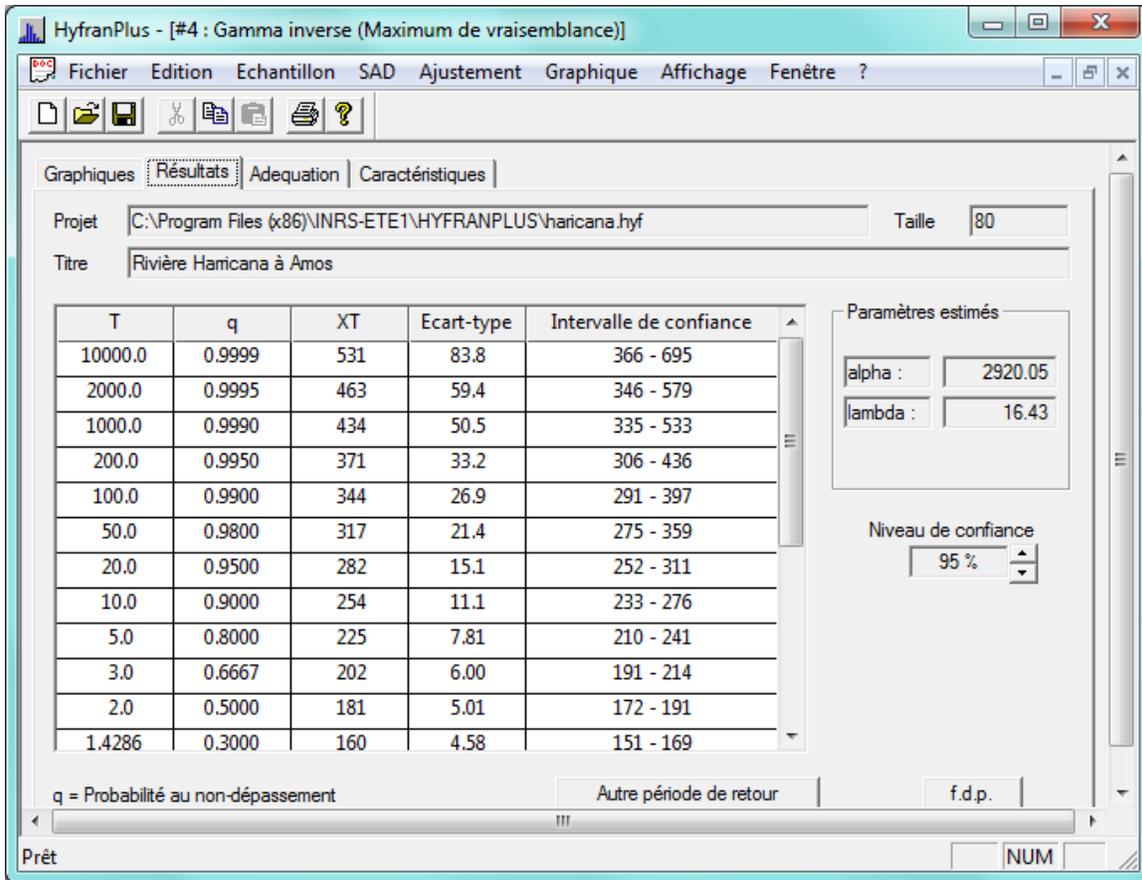


Figure 47 : Résultats de l'ajustement par une distribution Gamma Inverse de la série 1

### Graphique

Il est approprié d'analyser l'adéquation d'un ajustement à l'aide d'un graphique. HYFRAN-PLUS permet la visualisation de l'ajustement sur papier de probabilité normal ou Gumbel (le choix du papier est effectué dans le menu "Graphique", Figure 8).

Par défaut, le graphique (Figure 48) présente les données de l'échantillon, la courbe théorique de l'ajustement (ligne rouge), l'intervalle de confiance (lignes bleues) dont le niveau est spécifié à l'onglet "Résultats" et la légende. Il est toutefois possible de retirer ou d'ajouter certains des éléments sur la Figure à partir du menu «Graphique» que l'on retrouve à la barre des menus (Figure 8).

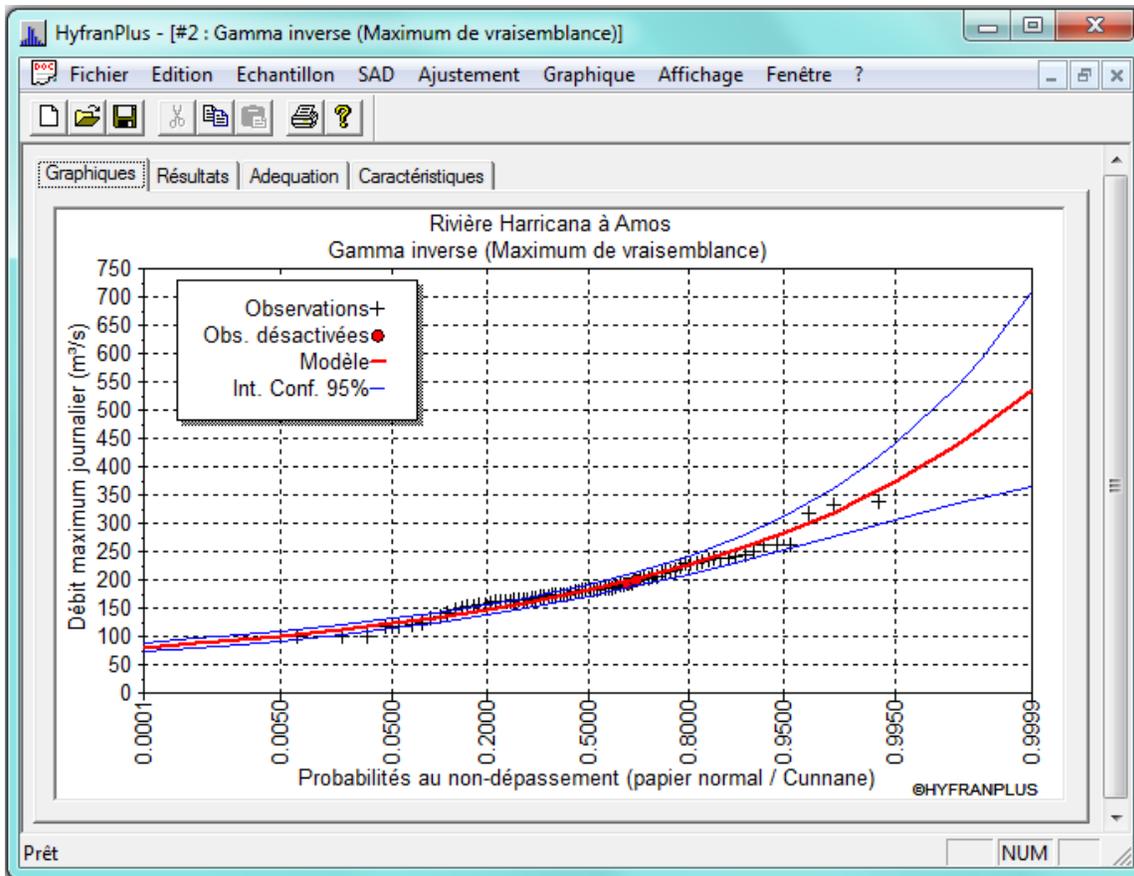


Figure 48 : illustration graphique d'un ajustement par une distribution Gamma Inverse de la série 1

#### Autre période de retour

Lors d'un ajustement, HYFRAN-PLUS présente les résultats pour 21 périodes de retour, choisies de manière à répondre à la plupart des problèmes étudiés. Toutefois, il peut être nécessaire dans le cadre de certains projets, de connaître la valeur de la variable pour une ou plusieurs autres périodes de retour.

Pour ajouter une période de retour, à partir de l'écran des résultats de l'ajustement il faut procéder comme suit :

- Appuyez sur le bouton «Autre période de retour» (Figure 47);
- La boîte de dialogue «Autre période de retour» apparaît immédiatement (Figure 49);
- Dans la zone de texte «Valeur de T» entrez la nouvelle période de retour que vous désirez ajouter.
- Appuyez sur le bouton «OK». La nouvelle période de retour est alors insérée dans la liste.

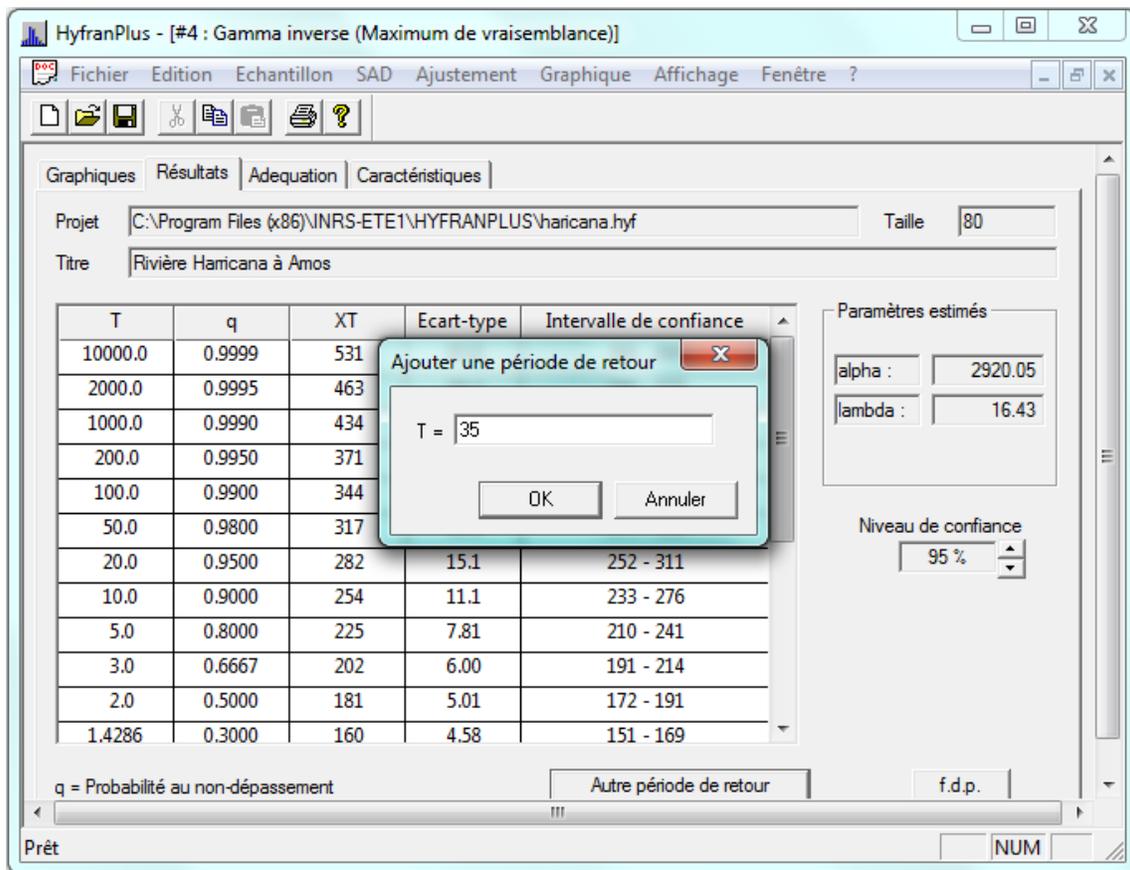


Figure 49: illustration de l'ajout d'une autre période de retour

- Modification du niveau de l'intervalle de confiance

Par défaut, le niveau de l'intervalle de confiance est de 95%. Pour le modifier il faut utiliser les flèches situées à droite de la zone de texte «Niveau de confiance». Le niveau de confiance varie de 1% à 99 %. HYFRAN-PLUS affiche immédiatement les résultats pour le nouveau niveau de confiance.

### Adéquation

Afin de juger d'une manière objective de la qualité de l'ajustement aux données, il existe divers tests statistiques d'adéquation. Selon la distribution utilisée et la taille d'échantillon, HYFRAN-PLUS présente les résultats obtenus pour certains des tests suivants (cf Bobée et El Adlouni, 2015 et Compaoré, El Adlouni et Bobée, 2013):

- Test du khi-carré (applicable pour toutes les distributions statistiques);
- Test sur les moments empiriques (applicable uniquement pour les distributions normale et log-normale).

Dans l'onglet "Adéquation" (Figure 50), on retrouve les informations suivantes:

- Le projet c.à.d. le nom du fichier qui contient l'échantillon et son chemin d'accès (Figure 11);
- Le titre du projet spécifié à l'onglet «Description du projet» est indiqué (Figure 12);

- L'hypothèse nulle ( $H_0$ ) et l'hypothèse alternative ( $H_1$ );
- Les résultats du test : i.e. la valeur de la statistique du test et la p-value de la statistique (cf Bobée et El Adlouni, 2015 et Compaoré, El Adlouni et Bobée, 2013);
- Enfin la conclusion du test c.à.d. l'acceptation ou le rejet de l'hypothèse nulle à un niveau de signification de 5% ou de 1%.

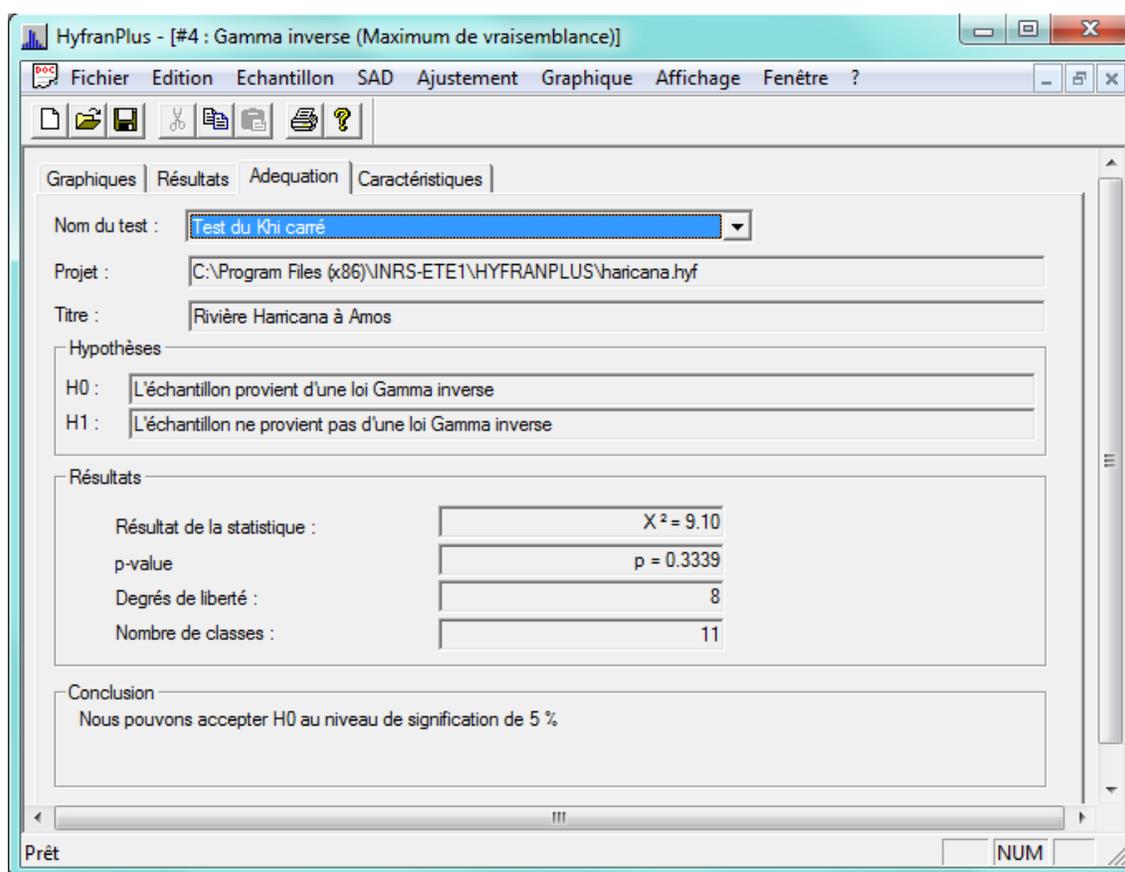


Figure 50 : Adéquation de l'ajustement par une distribution Gamma Inverse de la série 1

#### Caractéristiques statistiques de la distribution ajustée

La fenêtre des «caractéristiques de la population» (Figure 51) présente les statistiques pour la distribution utilisée. Dans la première colonne on retrouve les caractéristiques de la population et dans la seconde colonne on retrouve les caractéristiques de l'échantillon.

La taille de l'échantillon est donnée à la colonne des caractéristiques de l'échantillon. Le projet et le titre du projet apparaissent aussi dans cette fenêtre.

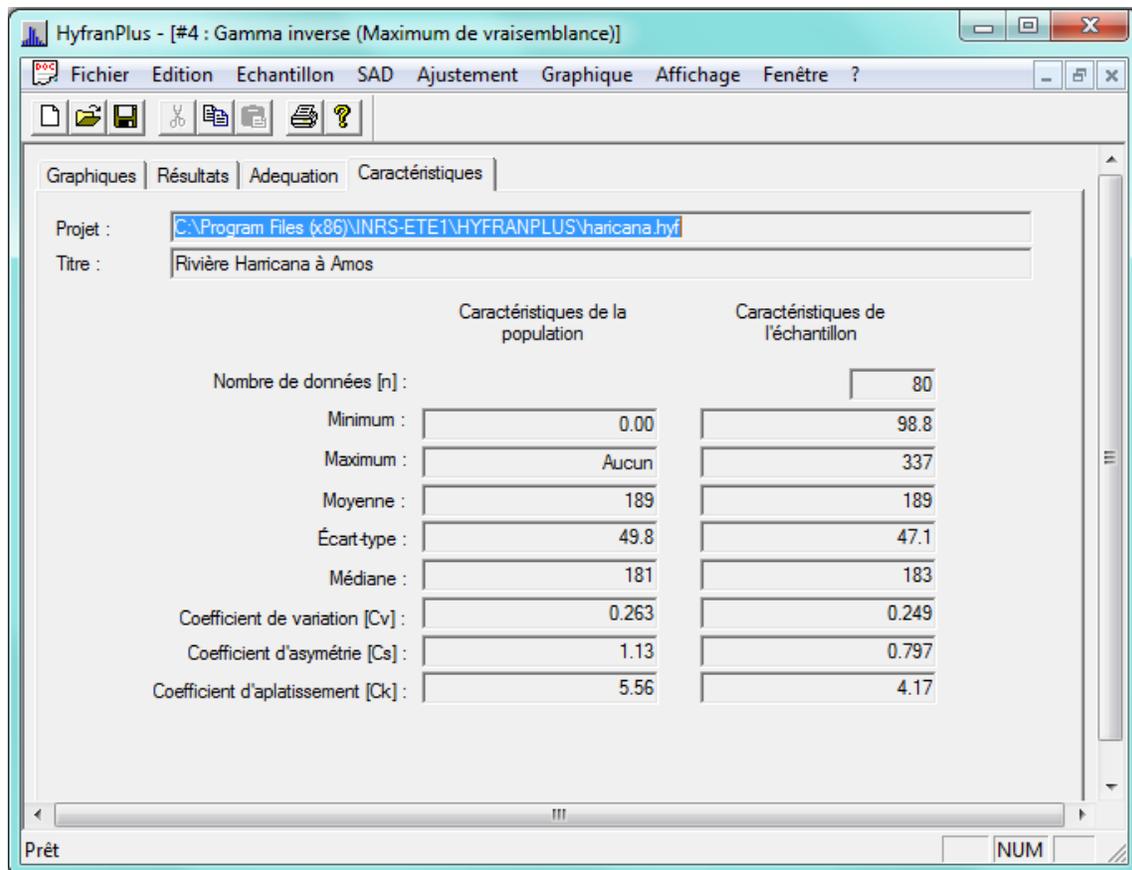


Figure 51 : Caractéristiques statistiques pour un ajustement par une distribution Gamma Inverse de la série 1

### Discordance

Cet onglet permet de vérifier, dans le cas de loi Normale et Log-normale, si l'échantillon comporte des données singulières (outliers), c'est-à-dire des observations qui semblent ne pas provenir de la distribution considérée. Pour vérifier une telle éventualité, on utilise le test de Grubbs-Beck pour la détection de données singulières adapté pour les distributions normale et log-normale (Bobée et El Adlouni, 2015).

**Note :** La majorité des tests de discordance sont basés sur l'hypothèse de normalité ou dans le cas d'une loi Log-normal en considérant une transformation logarithmique. En effet, si  $Y \sim LN$  alors  $X = \ln(Y)$  suit une loi normale.

Il existe deux tests, un pour la plus petite valeur et un autre pour la plus grande valeur de l'échantillon. On vérifie donc si la plus petite ou la plus grande observation est discordante par rapport au modèle envisagé (loi Normale ou Log-normale).

Dans l'onglet «Discordance», on retrouve les informations suivantes (Figure 52):

- Le choix du type de test : la plus petite ou la plus grande observation.

- Le projet c.à.d. le nom du fichier qui contient l'échantillon et son chemin d'accès
- Le titre du projet
- L'hypothèse nulle ( $H_0$ ) et l'hypothèse alternative ( $H_1$ )
- Les résultats du test : on retrouve dans ce cadre la valeur de la statistique du test et la  $p$ -value de la statistique.
- Enfin, la conclusion du test c.à.d. l'acceptation ou le rejet de l'hypothèse nulle à un niveau de signification de 5% ou de 1%

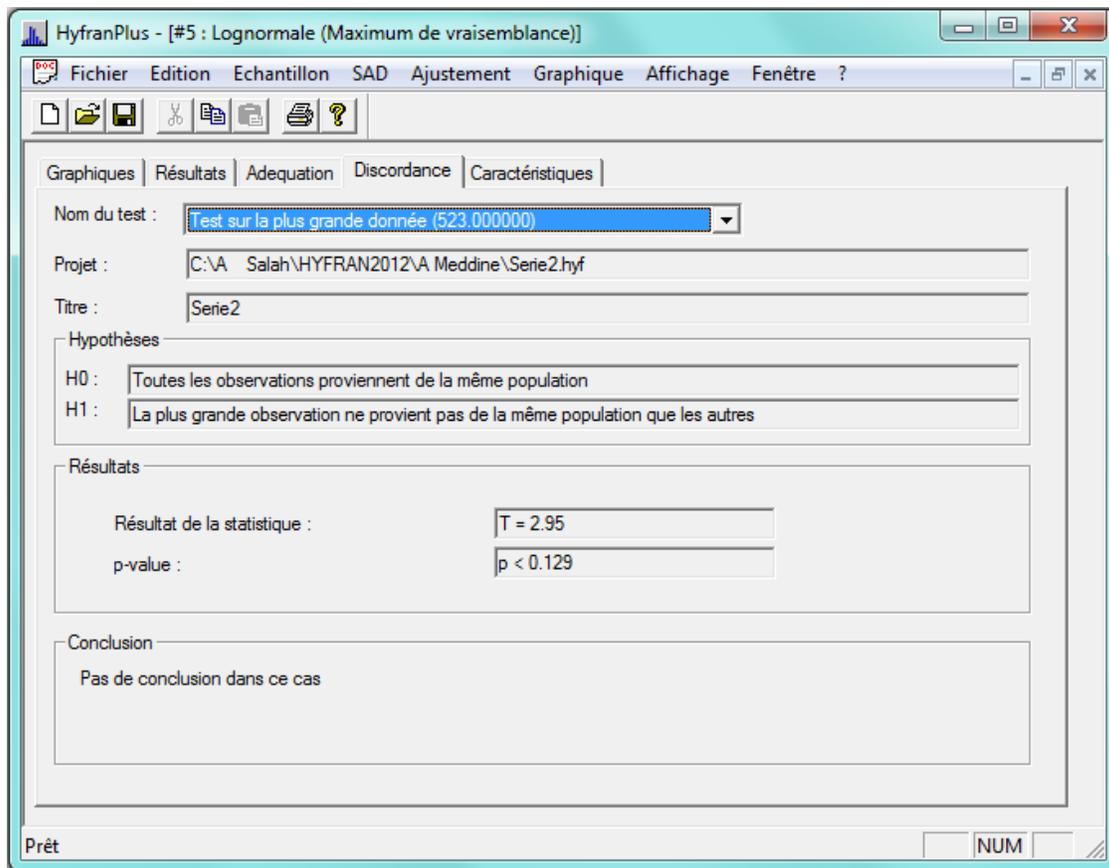


Figure 52 : Test de discordance de la plus grande observation dans le cas d'un ajustement par la distribution Log-normale (série 2)

Cas de la série 2 (Générée à partir d'une loi Log-normale)

Lorsqu'on ajuste une loi Log-normale à la série 2, on remarque que la plus grande observation est sur la ligne correspondant à la borne supérieure de l'intervalle de confiance des quantiles (Figure 53). On peut examiner si cette observation appartient à la population.

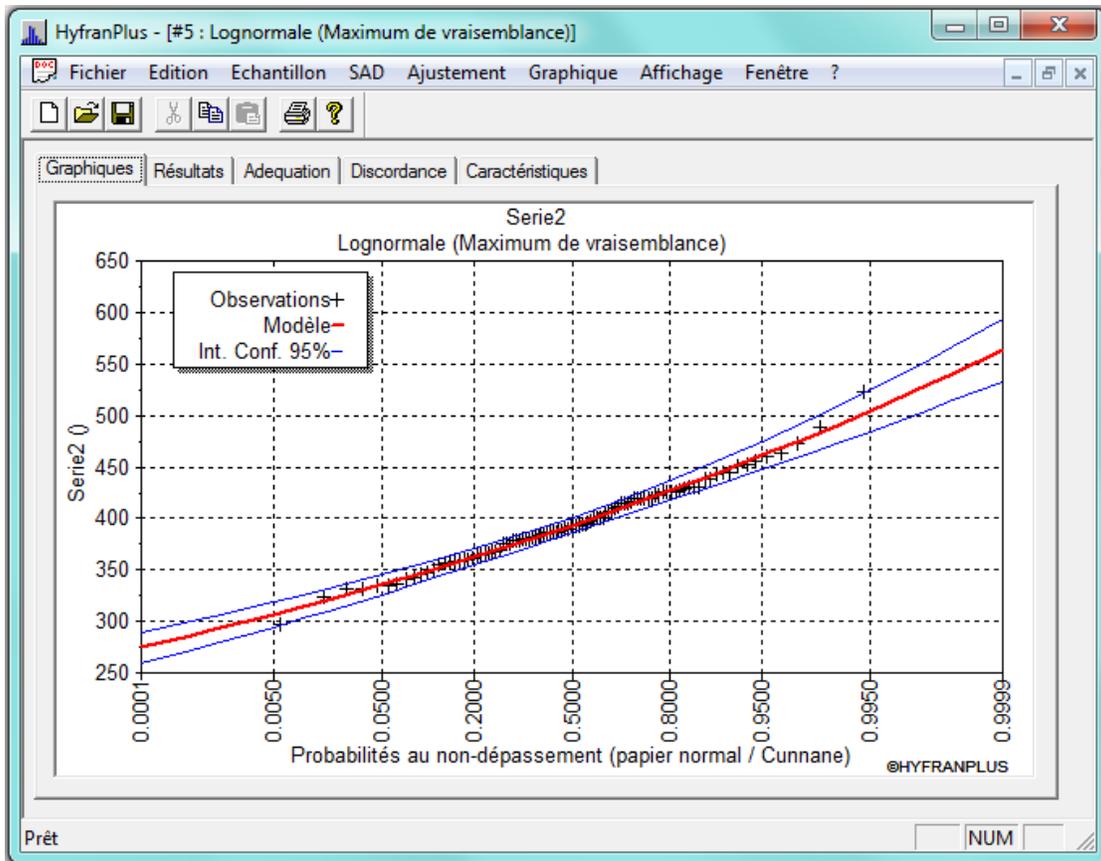


Figure 53 : Ajustement d'une loi Log-normale à la série 2 pour la mise en évidence des valeurs discordantes sous l'hypothèse de Log-normalité.

L'application du test de discordance à la série 2 montre que lorsqu'on considère toute la série, le test ne permet pas de conclure sur la Log-normalité de toute la série et donc sur l'appartenance de la plus grande observation 523 à la population Log-normale (Figure 52). En effet, dans ce cas (Bobée et El Adlouni, 2015) on sait seulement (Figure 52) que la p-value de  $T = 2.95$  est telle que p est plus petit que 0.129; mais elle peut alors être supérieure ou inférieure à la valeur critique de 5%.

Lorsqu'on ajoute une nouvelle valeur de l'ordre de 10% de plus que la plus grande valeur (523), soit 575, l'utilisation de la commande « Discordance » montre que cette nouvelle valeur est discordante avec l'hypothèse de Log-normalité de la distribution. En effet, dans ce cas (Figure 54), la p-value correspondant à  $T = 3.64$  est telle que  $p < 0.008$  qui est inférieure à l'erreur de première espèce de 0.05; et même à 0.01. Ce qui traduit une forte discordance.

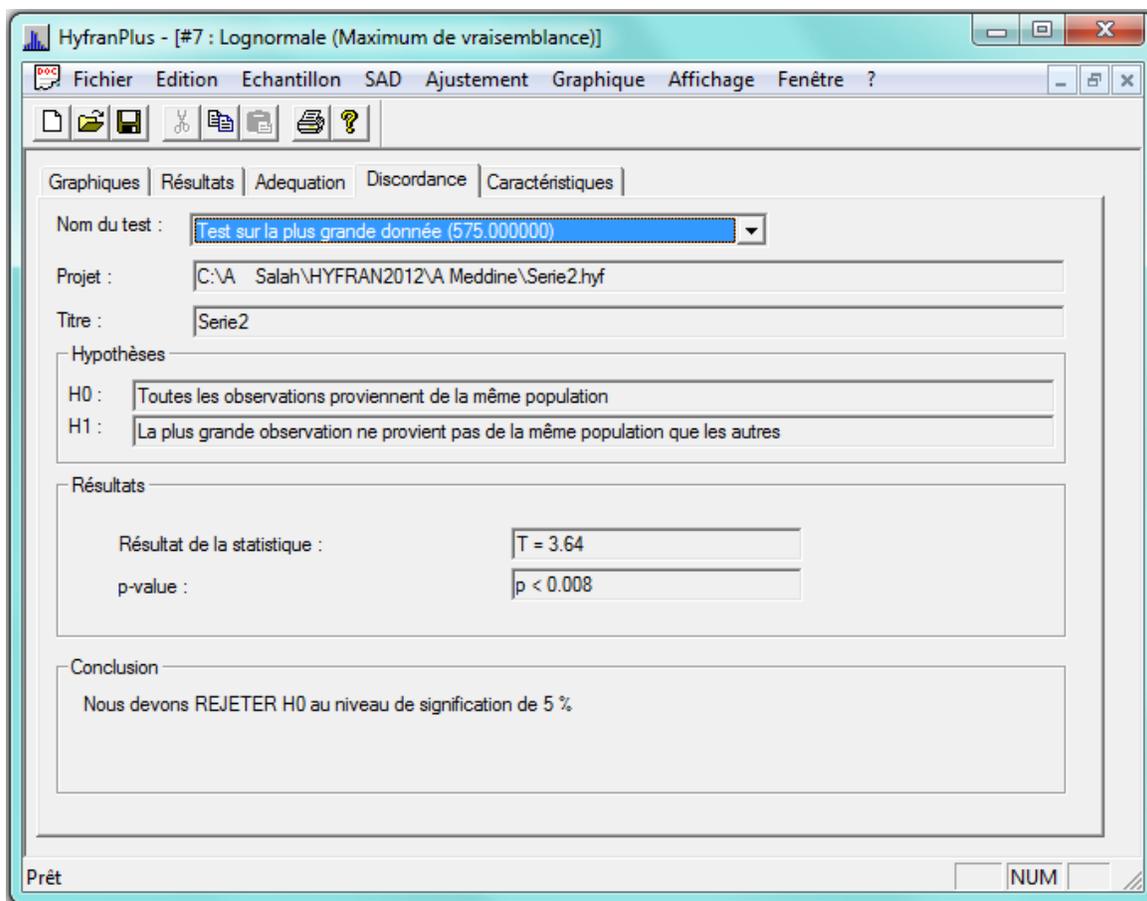


Figure 54 : Test de discordance pour un ajustement par la distribution Log-normale avec une nouvelle grande observation simulée (série 2)

La Figure 54 présente le résultat de ce test pour la nouvelle plus grande observation. La conclusion du test est le rejet de l'Hypothèse « H0 : Toutes les observations proviennent de la même population ». En conclusion, le test montre clairement (avec une p-value inférieure à 0.008) que la valeur 575 ne fait pas partie de la population sous l'hypothèse de Log-normalité.

**Note :** Lorsqu'une valeur singulière est détectée, on doit vérifier s'il s'agit :

- Une valeur aberrante (exemple : erreur de mesure ou de saisie) que l'on peut éliminer ou,
- Une valeur extrême réelle et donc très importante à conserver. Cette validation doit être effectuée à partir du contexte hydro-météorologique.

## Références

---

- \*Bobée B. et S. El Adlouni (2015). *Éléments d'Analyse Fréquentielle*. Institut National de la Recherche Scientifique (INRS-ETE). En cours de révision.
- Bobée, B., et F. Ashkar (1991). *The Gamma Family and Derived Distributions Applied in Hydrology*, Water Resources Publications, Littleton, Colorado, 203 pages.
- \*Compaore C., S. El Adlouni et B. Bobée (2013). *Analyse fréquentielle en hydrologie : Logiciel HYFRAN-PLUS*. Stage-Coop de l'Université de Moncton.
- Ehsanzadeh E., S. El Adlouni et B. Bobée (2010). Frequency analysis incorporating a Decision Support System (DSS) for hydro-climatic variables. *Journal of Hydrologic Engineering*, Vol 15:11, pp. 861-881.
- \*El Adlouni, S. et B. Bobée (2011). *Decision Support System for Flood Risk Assessment*. Volume 1-2, p. 11-26.
- \*El Adlouni, S., B. Bobée et T.B.M.J. Ouarda (2008). On the tails of extreme event distributions. *Journal of Hydrology*, 355, 16-33.
- \*El Adlouni S., B. Bobée et O. Samoud (2012). *Guide pour les méthodes du système d'aide à la décision (SAD) DE HYFRAN-PLUS*. Water Resources Publication (<http://www.wrpllc.com/books/HyfranPlus>).
- \*Martel, B., S. El Adlouni et B. Bobée (2012). Comparison of the power of Log-Normality tests with different right tail alternative distributions. *Journal of Hydrologic Engineering*, (ASCE). 18(1), 1-9. doi: 10.1061/(ASCE)HE.1943-5584.0000595.
- Morlat (1956). Les lois de probabilité de Halphen, *Revue de statistiques appliquée*, 4(3) : pp 21-46.
- Perreault L., B. Bobée et PF Rasmussen (1999). Halphen distribution system. I: Mathematical and statistical properties. *Journal of Hydrologic Engineering* 4 (3), 189-199.

## Annexe 1 : Séries d'observations

---

Nous présenterons ci-dessous les données des séries 1, 2, 3 et 4 utilisées pour le tutoriel de HYFRAN-PLUS.

Il faut noter que la série 1 correspond à la série par défaut du projet Harricana contenue dans le logiciel HYFRAN-PLUS pour exemple d'application.

Les séries 2 et 4 sont des données simulées à partir du logiciel Matlab.

Série 1			
Observation	Probabilité empirique	Observation	Probabilité empirique
122	0.0698	167	0.3192
244	0.9052	179	0.4564
214	0.7431	185	0.5686
173	0.3815	117	0.0449
229	0.7930	192	0.6185
156	0.1945	337	0.9925
212	0.7307	125	0.0823
263	0.9551	166	0.3067
146	0.1322	99.1	0.0200
183	0.5062	202	0.6683
161	0.2195	230	0.8180
205	0.7057	158	0.2070
135	0.1072	262	0.9426
331	0.9800	154	0.1696
225	0.7805	164	0.2818
174	0.4190	182	0.4813
98	0.0075	164	0.2943
149	0.1446	183	0.5312
238	0.8678	171	0.3441
262	0.9302	250	0.9177
132	0.0948	184	0.5436
235	0.8429	205	0.7182
216	0.7556	237	0.8554
240	0.8928	177	0.4439
230	0.8055	239	0.8803
192	0.6060	187	0.5935
195	0.6434	180	0.4688
172	0.3566	173	0.4065
173	0.3940	174	0.4314
172	0.3691	167	0.3317
153	0.1571	185	0.5810
142	0.1197	232	0.8304

317	0.9676	100	0.0324
161	0.2319	163	0.2569
201	0.6559	203	0.6808
204	0.6933	219	0.7681
194	0.6309	182	0.4938
164	0.2693	184	0.5561
183	0.5187	118	0.0574
161	0.2444	155	0.1820

Tableau A.1 : Données observées de la série 1

Série 2			
Observation	probabilité empirique	Observation	probabilité empirique
426	0.8044	350	0.1158
346	0.0958	375	0.2854
395	0.5449	452	0.9341
384	0.3952	428	0.8244
419	0.7046	355	0.1257
420	0.7545	324	0.0160
420	0.7246	381	0.3653
389	0.4750	412	0.6547
380	0.3353	443	0.8942
380	0.3553	407	0.6347
439	0.8842	361	0.1956
335	0.0459	416	0.6946
395	0.5349	359	0.1557
414	0.6747	366	0.2355
378	0.3154	378	0.3054
423	0.7745	357	0.1457
401	0.5948	393	0.5150
367	0.2455	369	0.2655
410	0.6447	392	0.5050
393	0.5250	385	0.4052
387	0.4351	387	0.4451
376	0.2954	384	0.3852
406	0.6248	456	0.9441
335	0.0559	429	0.8343
388	0.4551	406	0.6148
382	0.3752	348	0.1058
368	0.2555	343	0.0858
431	0.8643	331	0.0359
375	0.2754	523	0.9940
426	0.7844	445	0.9042
445	0.9142	414	0.6647
397	0.5649	366	0.2255
415	0.6846	360	0.1756
430	0.8543	426	0.8144
400	0.5848	472	0.9741
426	0.7944	341	0.0758
356	0.1357	386	0.4152

451	0.9242	400	0.5749
365	0.2156	331	0.0259
336	0.0659	439	0.8743
463	0.9641	387	0.4251
401	0.6048	488	0.9840
422	0.7645	388	0.4651
389	0.4850	420	0.7445
364	0.2056	360	0.1856
297	0.0060	379	0.3253
430	0.8443	359	0.1657
392	0.4950	420	0.7146
396	0.5549	460	0.9541
420	0.7345	380	0.3453

Tableau A.2 : Données de la série 2, simulées à partir d'une loi Log-normale

Série 2-transformée			
Observation	Probabilité transformée	Observation	Probabilité transformée
6.05	0.8044	5.86	0.1158
5.85	0.0958	5.93	0.2854
5.98	0.5449	6.11	0.9341
5.95	0.3952	6.06	0.8244
6.04	0.7046	5.87	0.1257
6.04	0.7545	5.78	0.0160
6.04	0.7246	5.94	0.3653
5.96	0.4750	6.02	0.6547
5.94	0.3353	6.09	0.8942
5.94	0.3553	6.01	0.6347
6.08	0.8842	5.89	0.1956
5.81	0.0459	6.03	0.6946
5.98	0.5349	5.88	0.1557
6.03	0.6747	5.9	0.2355
5.93	0.3154	5.93	0.3054
6.05	0.7745	5.88	0.1457
5.99	0.5948	5.97	0.5150
5.91	0.2455	5.91	0.2655
6.02	0.6447	5.97	0.5050
5.97	0.5250	5.95	0.4052
5.96	0.4351	5.96	0.4451
5.93	0.2954	5.95	0.3852
6.01	0.6248	6.12	0.9441
5.81	0.0559	6.06	0.8343
5.96	0.4551	6.01	0.6148
5.95	0.3752	5.85	0.1058
5.91	0.2555	5.84	0.0858
6.07	0.8643	5.8	0.0359
5.93	0.2754	6.26	0.9940
6.05	0.7844	6.1	0.9042
6.1	0.9142	6.03	0.6647
5.98	0.5649	5.9	0.2255
6.03	0.6846	5.89	0.1756
6.06	0.8543	6.05	0.8144
5.99	0.5848	6.16	0.9741
6.05	0.7944	5.83	0.0758
5.87	0.1357	5.96	0.4152
6.11	0.9242	5.99	0.5749

5.9	0.2156	5.8	0.0259
5.82	0.0659	6.08	0.8743
6.14	0.9641	5.96	0.4251
5.99	0.6048	6.19	0.9840
6.05	0.7645	5.96	0.4651
5.96	0.4850	6.04	0.7445
5.9	0.2056	5.89	0.1856
5.69	0.0060	5.94	0.3253
6.06	0.8443	5.88	0.1657
5.97	0.4950	6.04	0.7146
5.98	0.5549	6.13	0.9541
6.04	0.7345	5.94	0.3453

Tableau A.3 : Données de la série 2 transformée logarithmiquement (série normale)

<b>Série 3</b>			
Observation	Probabilité empirique	Observation	Probabilité empirique
494	0.7146	572	0.8443
330	0.2754	580	0.8643
358	0.3154	467	0.6547
330	0.2655	549	0.8144
479	0.6946	524	0.7844
319	0.2255	360	0.3253
293	0.1357	459	0.6347
476	0.6747	369	0.3653
355	0.3054	308	0.1756
450	0.6148	488	0.7046
363	0.3553	412	0.4651
437	0.5649	381	0.4052
187	0.0060	440	0.5848
464	0.6447	543	0.8044
704	0.9341	454	0.6248
551	0.8244	428	0.5349
530	0.7944	439	0.5749
375	0.3952	410	0.4551
426	0.5250	320	0.2355
434	0.5549	511	0.7445
415	0.4850	616	0.8942
314	0.2056	344	0.2854
363	0.3453	719	0.9441
383	0.4251	383	0.4152
349	0.2954	375	0.3852
519	0.7645	308	0.1856
864	0.9840	318	0.2156
271	0.0958	303	0.1657
419	0.4950	449	0.6048
262	0.0758	651	0.9142
596	0.8743	473	0.6647
616	0.9042	521	0.7745

245	0.0359	603	0.8842
723	0.9541	191	0.0160
291	0.1257	384	0.4351
432	0.5449	449	0.5948
561	0.8343	500	0.7246
422	0.5150	273	0.1058
803	0.9741	298	0.1457
921	0.9940	511	0.7545
421	0.5050	390	0.4451
228	0.0259	476	0.6846
362	0.3353	302	0.1557
313	0.1956	247	0.0459
321	0.2455	507	0.7345
288	0.1158	577	0.8543
413	0.4750	658	0.9242
759	0.9641	255	0.0559
325	0.2555	270	0.0858
371	0.3752	259	0.0659

Tableau A.4 : Données de la série 3, simulées à partir d'une loi Gamma

## Appendice A : Intervalles de confiance asymptotiques des quantiles

En AFH l'ajustement d'une distribution statistique quelconque  $D$ , par la méthode  $M$ , à un échantillon d'observations de débit maximum annuel de crue, a pour but de déterminer l'estimateur du quantile  $\hat{X}_T = F^{-1}(1-1/T; \hat{\theta})$ , où  $\hat{\theta}$  (estimation du vecteur des paramètres  $\theta$ ) et  $\hat{X}_T$ , sont des variables aléatoires.

En général la distribution exacte du quantile  $\hat{X}_T$  est inconnue sauf pour certaines distributions : exponentielle, normale ou log-normale. Mais lorsque  $N$  grand on admet que  $\hat{X}_T$  est distribuée asymptotiquement selon une loi normale :

- de moyenne  $X_T$  (vraie valeur inconnue)
- de variance  $\text{var } \hat{X}_T$

$$\hat{X}_T \sim N(X_T; \text{var } \hat{X}_T) \Rightarrow u = \frac{\hat{X}_T - X_T}{\sqrt{\text{var } \hat{X}_T}} \sim N(0, 1)$$

Pour une loi donnée  $\text{var } \hat{X}_T$  dépend de la méthode  $M$  d'estimation utilisée (cf. Appendice D, Bobée et Ashkar, (1991)) où l'on retrouve des détails de calculs pour les méthodes MM et MV. On peut déduire l'intervalle de confiance au niveau  $(1-\alpha)$  de la vraie valeur  $X_T$  à partir de l'équation de  $\hat{X}_T$ .

Pour un ensemble distribution  $\times$  méthode ( $D \times M$ ) on peut en déduire l'intervalle de confiance (IC) $(1-\alpha)$  de la vraie valeur inconnue  $X_T$  au niveau de confiance  $(1-\alpha)$ . On a (cf. Appendice E, Bobée et Ashkar, (1991)):

$$P \left[ -u_{\alpha/2} \leq \frac{\hat{X}_T - X_T}{\sqrt{\text{var } \hat{X}_T}} \leq u_{\alpha/2} \right] = 1 - \alpha$$

Ou encore :

$$P \left[ \hat{X}_T - u_{\alpha/2} \sqrt{\text{var } \hat{X}_T} \leq X_T \leq \hat{X}_T + u_{\alpha/2} \sqrt{\text{var } \hat{X}_T} \right] = 1 - \alpha$$

$u_{\alpha/2}$  correspond au quantile de la loi normale centré-réduite de probabilité au dépassement  $\alpha/2$ .

## Appendice B : Classification des distributions

---

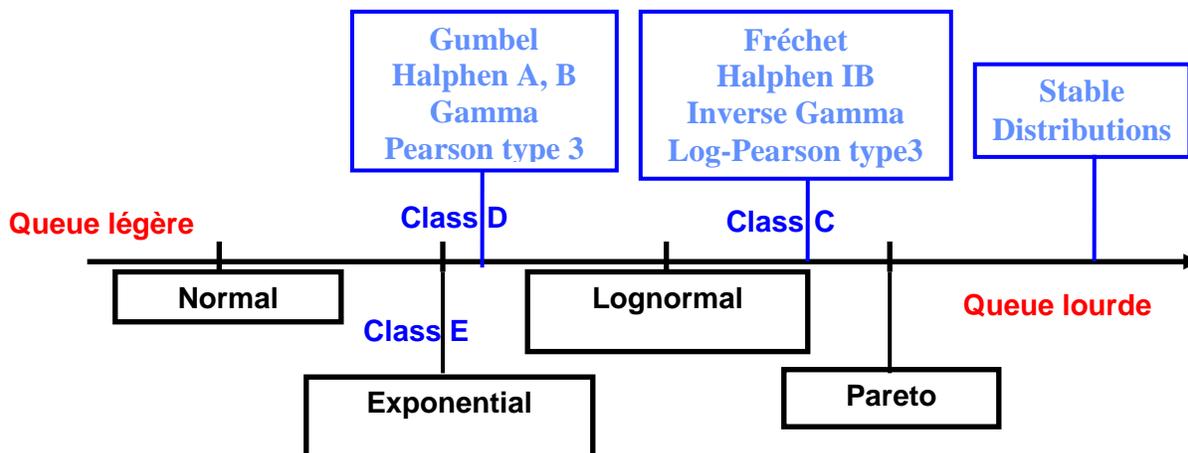


Figure B-1: Distributions ordonnées par rapport à la queue droite

(Tirée de El Adlouni, Bobée et Ouarda, 2008).