

# ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT D'HYDRAULIQUE



## SUJET

**Contribution à la simulation des paramètres  
hydrométéorologiques (aléatoires et cycliques)  
à différents pas de temps (mensuel, journalier, horaire)  
par le modèle GESTOP.**

Proposé par :

M. N. DECHMI

Etudié par :

Mme A. HAMRICHE

JURY : M. BALI  
M. DECHMI  
M. AISSANI  
M. MOULOUDI  
M. ISSOLAH  
M. MOUKRAN  
M. OUABDESSALEM

Président  
Rapporteur  
Examineur  
Examineur  
Examineur  
Examineur  
Invité

PROMOTION  
97/98

E.N.P. 10, Avenue Hacén Badi - El Harrach - Alger

لإيجاد طريقة تسيير عقلانية للسدود طورنا نموذج رياضي «GESTOP» يمكن تطبيقه على الضواهر الجوية الدورية منها كالتبخر و العفوانية كالتدفق و حساب كمية الأمطار. و هذا بخنطرات متغيرة للوقت. (شهري، يومي، ساعي). يرتكز النموذج على تقنية التحليل بالمركبات الرئيسية. في ايطار هذا البحث، درسنا متغيرات مختلفة، تهدف أساسا للحفاظ على الخصائص الإحصائية و منها القيم القصوى، الشيء الذي لم تحققه النماذج الموجودة. النموذج «GESTOP» قد أعطى نتائج حسنة للضواهر الثلاث المدروسة في مختلف خطوات الوقت. قد أثريت هذه الاطروحة بمقارنة مع النماذج الموجودة، التي بينت قدرة النموذج «GESTOP» على التمثيل بطريقة موحدة ضواهر جوية مختلفة.

### Résumé

Pour une nouvelle approche d'une gestion rationnelle des barrages réservoirs on a élaboré un modèle de simulation appelé GESTOP, applicable aussi bien pour les paramètres hydrométéorologiques cycliques (ETP) qu'aléatoires (débit, pluviométrie) ; et ceci à différents pas de temps en allant du mensuel, le journalier et enfin le pas de temps horaire.

Le modèle est basé sur la technique de l'analyse en composantes principales. Dans le cadre de cette recherche plusieurs variantes ont été étudiées, visant principalement à reproduire les caractéristiques statistiques des données historiques entre autres les valeurs extrêmes, objectif que nombreux modèles existants n'ont pas atteint.

Le modèle GESTOP a donné de bons résultats pour les trois paramètres étudiés aux différents pas de temps. Cette étude a été enrichie avec une étude comparative qui a conclut sur la capacité du modèle élaboré a simuler d'une manière unificatrice les paramètres hydrométéorologiques.

### Abstract

For a new approach of dam management, we have elaborated a simulation model, called "GESTOP", applied both on the cyclical hydrometeorological phenomena (evaporation) and random ones (flow, rain). And this could be achieved in several time steps: monthly, daily and hourly as well.

The model is based on the analysis of the main components. In the following investigations, many variants have been studied aiming predominantly at reproducing statistical characteristics of historical data.

An attempt that the "GESTOP" has successfully achieved and realised on the three time steps presented below; for the three parameters given below.

This research has also been improved by comparative study, which conclude on the model ability to easily simulate the hydrometeorological phenomena.

# REMERCIEMENTS

المدرسة الوطنية المتعددة التقنيات  
BIBLIOTHEQUE — المكتبة  
Ecole Nationale Polytechnique

Ce travail de recherche a été réalisé dans le cadre d'une thèse de magister, au sein du département hydraulique de l'école nationale polytechnique.

Je tiens à remercier mon promoteur M. N. Dechmi maître de conférence au département hydraulique de l'école nationale polytechnique et directeur du laboratoire de recherche GESTOP, pour n'avoir ménagé aucun effort pour m'assurer de son appui. Qu'il trouve en ces quelques lignes l'expression de ma profonde gratitude et entière reconnaissance.

Je tiens également à remercier messieurs les membres de jury qui m'honorent de leur présence, d'avoir accepté d'examiner ce travail de recherche, et enrichir par leurs remarques la qualité et le contenu de cette thèse, en l'occurrence :

**M. BALI**

Professeur et maître de conférence à l'école nationale polytechnique

**M. OUABDESSALEM**

Professeur et maître de conférence à l'école nationale polytechnique

**M. AISSANI**

Professeur et maître de conférence à l'école nationale supérieure d'hydraulique (Blida)

**M. MOULOUDI**

Chargé de cours à l'école nationale polytechnique

**M. ISSOLAH**

Docteur chargé de cours à l'école nationale d'Agronomie

**M. MOKRAN**

Directeur à l'agence nationale des barrages (ANB)

Je n'omettrai pas de remercier M. E.Y. Bouzroua, M<sup>lle</sup> D. Souag ainsi que toute la sous direction des infrastructures d'assainissement de la direction des grands aménagements et infrastructures hydrauliques du ministère de l'équipement et de l'aménagement du territoire, et particulièrement à M. M. Medkour pour leur compréhension, leurs aides et leurs encouragements. Je remercie enfin tous ceux qui par un geste ou une parole m'ont aidé à mener à bien ce travail.

# LISTE DES FIGURES

## Chapitre I

Figure I-1	Représentation du processus de la ligne brisée	6
------------	--	---

## Chapitre II

Figure II-1	Organigramme de l'Analyse En Composantes principales	27
-------------	--	----

## Chapitre III

Figure III-1	Passage d'une fonction de probabilité F à une loi uniforme	31
Figure III-2	Tirage aléatoire à partir de la fonction de répartition	31
Figure III-3	Organigramme de simulation par le modèle GESTOP	34

## Chapitre IV

Figure IV-1	Evolution mensuelle des ETP (1965 - 1985)	42
Figure IV-2	Reconstitution de l'ETP (ACP normée)	44
Figure IV-3	Evolution mensuelle des Débits - Station Cheffia -	47
Figure IV-4	Reconstitution des Débits - ACP normée -	49
Figure IV-5	Evolution mensuelle de la pluviométrie (1953 - 1988 )	51
Figure IV-6	Reconstitution de la série de pluies - ACP normée -	53
Figure IV-7	Variation des CP (ETP journalières)	56
Figure IV-8	Variation des CP (ETP journalières transformées)	56
Figure IV-9	Tests des moyennes et écarts types des valeurs d'ETP annuelles	58
Figure IV-10	Ajustement des valeurs d'ETP annuelles observées et simulées	58
Figure IV-11	Test des moyennes des moyennes mensuelles d'ETP	59
Figure IV-12	Test des écarts types des moyennes mensuelles d'ETP	59
Figure IV-13	Ajustement des valeurs d'ETP maximales (Loi de Gumbel)	60
Figure IV-14	Variation des CP (données de débits brutes)	61
Figure IV-15	Variation des CP (données de débits transformées)	62
Figure IV-16	Tests des moyennes et écarts types des valeurs de débits annuelles (V1).	65
Figure IV-17	Ajustement des valeurs de débits annuelles observées et simulées (V1)	65
Figure IV-18	Test des moyennes des valeurs mensuelles de Débit (V1)	66
Figure IV-19	Test des écart types des valeurs mensuelles de Débits (V1)	66
Figure IV-20	Ajustement des valeurs de Débits maximales (Loi de Gumbel) (V1)	66
Figure IV-21	Tests des moyennes et écarts types des valeurs de débits annuelles (V2)	67
Figure IV-22	Ajustement des valeurs de débits annuelles simulées (V2)	67
Figure IV-23	Test des moyennes des valeurs mensuelles de Débit (V2)	68
Figure IV-24	Test des écart types des valeurs mensuelles de Débits (V2)	68
Figure IV-25	Ajustement des valeurs de Débits maximales simulées (Loi de Gumbel) (V2).	69
Figure IV-26	Tests des moyennes et écarts types des valeurs de débits annuelles (V3)	69
Figure IV-27	Ajustement des valeurs de débits annuelles observées et simulées (V3)	70
Figure IV-28	Test des moyennes des valeurs mensuelles de Débit (V3)	70
Figure IV-29	Test des écart types des valeurs mensuelles de Débits (V3)	70
Figure IV-30	Ajustement des valeurs de Débits maximales simulées (Loi de Gumbel) (V3)	71
Figure IV-31	Tests des moyennes et écarts types des valeurs de débits annuelles (V4)	71
Figure IV-32	Ajustement des valeurs de débits annuelles simulées (V4)	72
Figure IV-33	Test des moyennes des valeurs mensuelles de Débit (V4)	72
Figure IV-34	Test des écart types des valeurs mensuelles de Débits (V4)	73
Figure IV-35	Ajustement des valeurs de débits maximales simulées (Loi de Gumbel) (V4)	73
Figure IV-36	Tests des moyennes et écarts types des valeurs annuelles (V5)	74
Figure IV-37	Ajustement des valeurs de débits annuelles simulées (V5)	74
Figure IV-38	Test des moyennes des valeurs mensuelles de Débit (V5)	75
Figure IV-39	Test des écart types des valeurs mensuelles de Débits (V5)	75
Figure IV-40	Ajustement des valeurs de Débits maximales (Loi de Gumbel) (V5)	75

Figure IV-41	CP calculées sur la série de pluies brutes	77
Figure IV-42	CP calculées sur la série de pluies transformées	77
Figure IV-43	Tests des moyennes et écarts types des valeurs de pluies annuelles (V1)	79
Figure IV-44	Ajustement des valeurs annuelles de pluies observées et simulées (V1)	80
Figure IV-45	Test des moyennes des valeurs de Pluies mensuelles (V1)	80
Figure IV-46	Test des écart types des valeurs de Pluies mensuelles (V1)	80
Figure IV-47	Ajustement des pluies maximales simulées (Loi de Gumbel) (V1)	81
Figure IV-48	Tests des moyennes et écarts types des valeurs de pluies annuelles (V2)	81
Figure IV-49	Ajustement des valeurs annuelles de Pluies simulées (V2)	82
Figure IV-50	Test des moyennes des valeurs de Pluies mensuelles (V2)	82
Figure IV-51	Test des écart types des valeurs de Pluies mensuelles (V2)	82
Figure IV-52	Ajustement des pluies maximales simulées (Loi de Gumbel) (V2)	83
Figure IV-53	Tests des moyennes et écarts types des valeurs de pluies annuelles (V3)	83
Figure IV-54	Ajustement des valeurs annuelles de Pluies simulées (V3)	84
Figure IV-55	Test des moyennes des valeurs de Pluies mensuelles (V3)	84
Figure IV-56	Test des écart types des valeurs de Pluies mensuelles (V3)	84
Figure IV-57	Ajustement des pluies maximales simulées (Loi de Gumbel) (V3)	85
Figure IV-58	Tests des moyennes et écarts types des valeurs de pluies annuelles (V4)	85
Figure IV-59	Ajustement des valeurs annuelles de Pluies simulées (V4)	85
Figure IV-60	Test des moyennes des valeurs de Pluies mensuelles (V4)	86
Figure IV-61	Test des écart types des valeurs de Pluies mensuelles (V4)	86
Figure IV-62	Ajustement des pluies maximales observées et simulées (Loi de Gumbel) (V4)	86
Figure IV-63	Tests des moyennes et écarts types des valeurs de pluies annuelles (V5)	87
Figure IV-64	Ajustement des valeurs annuelles de Pluies observées et simulées (V5)	87
Figure IV-65	Test des moyennes des valeurs de Pluies mensuelles (V5)	88
Figure IV-66	Test des écart types des valeurs de Pluies mensuelles (V5)	88
Figure IV-67	Ajustement des pluies maximales simulées (Loi de Gumbel) (V5)	88

### Chapitre V

Figure V-1	Evolution de l'ETP horaire sur 72 heures	90
Figure V-2	Test de la moyenne et l'écart type des valeurs journalières (ETP)	91
Figure V-3	Test de la moyenne et l'écart type des valeurs journalières (ETP)	91
Figure V-4	génération de la seconde CP par la fonction de répartition	91
Figure V-5	Test de la moyenne et l'écart type des valeurs journalières (Pluies)	94
Figure V-6	Test de la moyenne et l'écart type des valeurs journalières (Pluies)	94
Figure V-7	Génération de la première CP par la loi de Goodrich	95

### Chapitre IV

Figure VI-1	Tests sur les moyennes et écarts types des valeurs d'ETP annuelles	96
Figure VI-2	Ajustement des valeurs d'ETP annuelles	97
Figure VI-3	Tests sur les moyennes des valeurs mensuelles	97
Figure VI-4	Tests sur les écarts types des valeurs d'ETP mensuelle	97
Figure VI-5	Ajustement des valeurs d'ETP maximales journalières	98
Figure VI-6	Tests sur les moyennes et écarts types des valeurs de débits annuels	99
Figure VI-7	Ajustement des valeurs de débits annuels	99
Figure VI-8	Tests sur les moyennes des valeurs de débits mensuels	100
Figure VI-9	Tests sur les écarts types des valeurs de débits mensuels	100
Figure VI-10	Ajustement des valeurs des débits maximums journaliers	100
Figure VI-11	Tests des moyennes et écarts types des valeurs de pluies annuelles	101
Figure VI-12	Ajustement des valeurs de pluies annuelles	102
Figure VI-13	Test sur les moyennes des valeurs de pluies mensuelles	102
Figure VI-14	Test sur les écarts types des valeurs de pluies mensuelles	102
Figure VI-15	Ajustement des valeurs de pluies maximales journalières	102

## LISTE DES TABLEAUX

### Chapitre I

Tableau I.8-1	Matrice de passage d'ordre 1	13
Tableau I.8-2	Matrice de passage d'ordre 2	13

### Chapitre IV

Tableau IV.3-1	Paramètres statistiques de la série d'ETP	43
Tableau IV.3-2	CP calculées sur la série d'ETP (ACP normée)	43
Tableau IV.3-3	CP calculées sur la série d'ETP (ACP normée) données transformées (SQRT)	46
Tableau IV.3-4	Paramètres statistiques des Débits mensuels	48
Tableau IV.3-5	CP calculées sur la série d'apports de la Cheffia -ACP normée- Données brutes	48
Tableau IV.3-6	CP calculées sur les données de débits transformées (ACP normée)	50
Tableau IV.3-7	Paramètres statistiques de la série de Pluies d'Annaba (1953 - 1986 )	52
Tableau IV.3-8	Les CP calculées avec ACP normée sur la série de pluies	52
Tableau IV.3-9	Les CP calculées sur la série de pluie - ACP normée - Données transformées-	54
Tableau IV.4-1	Paramètres statistiques de la série d'ETP	55
Tableau IV.4-2	Paramètres statistiques de la série des débits historiques	61
Tableau IV.4-3	Les différentes variantes de simulation	64
Tableau IV.4-4	Résultats des simulations des débits	64
Tableau IV.4-5	Paramètres statistiques de la série de pluies historiques	76
Tableau IV.4-7	Résultats des simulations des pluies	79

### Chapitre V

Tableau V.1-1	Caractéristiques statistiques de la série d'ETP horaires	89
Tableau V.1-2	Les CP calculées sur la série d'ETP (ACP centrée, données transformées)	90
Tableau V.2-1	Caractéristiques statistiques de la série de pluies horaires	92
Tableau V.2-2	Les CP calculées sur la série de pluies (ACP centrée, données transformées)	92

### Chapitre VI

Tableau VI-1	Comparaison entre les modèles de simulation	103
--------------	---	-----

## SOMMAIRE

**LISTE DES TABLEAUX**  
**LISTE DES FIGURES**

<b>INTRODUCTION</b>	<b>1</b>
<b>CHAPITRE I LES MODÈLES DE SIMULATION EN HYDROLOGIE.</b>	<b>2</b>
<b>I.1. INTRODUCTION.</b>	<b>2</b>
<b>I.2. LE MODÈLE DE MONTE CARLO</b>	<b>3</b>
<b>I.3. LES MODÈLES AUTORÉGRESSIFS DE THOMAS ET FIERING</b>	<b>4</b>
<b>I.4. MODÈLE DU BRUIT GAUSSIEN FRACTIONNEL (FGN)</b>	<b>5</b>
<b>I.5. LE MODÈLE DE LA LIGNE BRISÉE (BROKEN LINE)</b>	<b>6</b>
<b>I.6. LES MODÈLES AUTORÉGRESSIFS</b>	<b>7</b>
I.6.1 FORMULATION MATHÉMATIQUE DU MODÈLE.	8
I.6.2 APPLICATION DES MODÈLES AUTORÉGRESSIFS	9
I.6.3 LIMITES DES MODÈLES AUTORÉGRESSIFS	9
<b>I.7. LES MODÈLES DE DÉSAGRÉGATION.</b>	<b>10</b>
<b>I.8. LES MODÈLES MARKOVIENS</b>	<b>12</b>
I.8.1 DÉFINITION	12
I.8.2 DÉDUCTION ANALYTIQUE DES PARAMÈTRES	12
I.8.3 PRINCIPE DE LA SIMULATION	14
<b>I.9. SIMULATION DE VARIABLES CYCLIQUES PAR DÉCOMPOSITION DE LA SÉRIE HISTORIQUE.</b>	<b>14</b>
I.9.1 ESTIMATION DES PARAMÈTRES DE LISSAGE	15
I.9.2 SIMULATION DES FLUCTUATIONS	15
I.9.3 SIMULATION DE LA PORTEUSE	16
<b>I.10. LE MODÈLE SVD (SINGULAR VALUE DÉCOMPOSITION)</b>	<b>16</b>
I.10.1 CALIBRATION	16
I.10.2 SIMULATION	17
<b>CHAPITRE II ANALYSE EN COMPOSANTES PRINCIPALES THÉORIE ET PROGRAMMATION</b>	<b>18</b>
<b>II.1. HISTORIQUE DE L'ANALYSE DES DONNÉES (ADD)</b>	<b>18</b>
<b>II.2. DÉFINITIONS</b>	<b>19</b>
II.2.1 NOTIONS D'INDIVIDU ET DE CARACTÈRE	19
II.2.2 DÉFINITION ALGÈBRIQUE	19
II.2.3 DÉFINITION GÉOMÉTRIQUE	20
<b>II.3. FORMULATION MATHÉMATIQUE DU PROBLÈME</b>	<b>20</b>
<b>II.4. CHOIX DE LA MÉTRIQUE</b>	<b>21</b>
<b>II.5. PROCÉDÉ D'APPLICATION DE L'ACP</b>	<b>22</b>
II.5.1 CALCUL DE LA MATRICE DES COVARIANCES	22
II.5.2 RECHERCHE DES AXES PRINCIPAUX	22
II.5.3 CALCUL DES COMPOSANTES PRINCIPALES	25
II.5.4 CONCLUSION	25
<b>II.6. ACP PROGRAMMATION</b>	<b>25</b>
II.6.1 FICHIER DE DONNÉES INPUT	26
II.6.2 ORGANIGRAMME DE CALCUL	26

**CHAPITRE III ELABORATION DU MODÈLE GESTOP DE SIMULATION** **28**

<b>III.1. CONCEPT DE SIMULATION</b>	<b>28</b>
<b>III.2. MODÈLE DE SIMULATION</b>	<b>28</b>
<b>III.3. MÉTHODES DE SIMULATION</b>	<b>30</b>
III.3.1 SIMULATION PAR LES FONCTIONS DE RÉPARTITION	30
III.3.2 SIMULATION PAR LES CHAÎNES DE MARKOV	31
III.3.3 SIMULATION PAR LES LOIS D'AJUSTEMENT	32
<b>III.4. ORGANIGRAMME</b>	<b>33</b>

**CHAPITRE IV APPLICATION DU MODÈLE GESTOP AU PARAMÈTRES HYDROMÉTÉOROLOGIQUES** **35**

<b>IV.1. CARACTÉRISTIQUES DE LA RÉGION</b>	<b>35</b>
IV.1.1 L'ÉVAPOTRANSPIRATION	35
IV.1.2 LES PRÉCIPITATIONS	35
IV.1.3 LES APPORTS	36
<b>IV.2. MÉTHODOLOGIE</b>	<b>37</b>
IV.2.1 FICHER DE DONNÉES	37
IV.2.2 CALCUL DES C.P.	37
IV.2.3 RECONSTITUTION	38
IV.2.4 SIMULATION	39
<b>IV.3. SIMULATION DES PARAMÈTRES HYDROMÉTÉOROLOGIQUES AU PAS DE TEMPS MENSUEL</b>	<b>42</b>
IV.3.1 SIMULATION DES ETP	42
IV.3.2 SIMULATION DES APPORTS	47
IV.3.3 SIMULATION DES PLUIES	51
<b>IV.4. SIMULATION DES PARAMÈTRES HYDROMÉTÉOROLOGIQUES AU PAS DE TEMPS JOURNALIER</b>	<b>55</b>
IV.4.1 SIMULATION DES VARIABLES CYCLIQUES (ETP)	55
IV.4.2 SIMULATION DES APPORTS	61
IV.4.3 SIMULATION DES PLUIES	76

**CHAPITRE V SIMULATION DES PARAMÈTRES HYDROMÉTÉOROLOGIQUES AU PAS DE TEMPS HORAIRE** **89**

<b>INTRODUCTION</b>	<b>89</b>
<b>V.1. SIMULATION DES ETP</b>	<b>89</b>
V.1.1 DONNÉES UTILISÉES	89
V.1.2 RÉSULTATS DE L'ACP	90
V.1.3 RÉSULTATS DE LA SIMULATION DES ETP	90
<b>V.2. SIMULATION DES PLUIES</b>	<b>93</b>
V.2.1 LES DONNÉES UTILISÉES	93
V.2.2 RÉSULTATS DE L'ACP	93
V.2.3 RÉSULTATS DE LA SIMULATION	94



<b>CHAPITRE VI ETUDE COMPARATIVE DES MODÈLES DE SIMULATION EN HYDROLOGIE</b>	<b>96</b>
<b>INTRODUCTION</b>	<b>96</b>
<b>VI.1. SIMULATION DES ETP</b>	<b>96</b>
<b>VI.2. SIMULATION DES DÉBITS</b>	<b>98</b>
<b>VI.3. SIMULATION DES PLUIES</b>	<b>101</b>
<b>CONCLUSION</b>	<b>104</b>

## INTRODUCTION

Une gestion rationnelle des barrages réservoirs ne peut se faire que par une confrontation des paramètres d'entrée (pluviométrie, Débit) qui constituent les apports au niveau du barrage, et de sortie représentés par la demande agricole, domestique et industrielle.

La demande agricole est fonction de l'évapotranspiration, qui elle-même est fonction de la température. Pour optimiser les volumes d'eau disponibles au niveau des barrages il faut une excellente connaissance des paramètres du bilan hydrique : apports et demande, et cela ne peut se faire que par le biais de la simulation.

Les modèles de simulation sont fonction du pas de temps (saisonnier, mensuel, hebdomadaire, journalier et horaire) et du type de phénomène (aléatoire ou cyclique); parmi cette panoplie de modèles on trouve les modèles de Monte Carlo, Thomas et Fiering, Broken Line, Fractional Gaussien Noise, les modèles autorégressifs (ARMA), les modèles de désagrégation, Les modèles Markoviens et les modèles de décomposition de la série historique.

Tout ces modèles exigent des contraintes bien spécifiques tel que la stationnarité, l'homogénéité et la chronologie, ce qui limite leurs application.

Le but de cette thèse de magister, et de développer une approche unificatrice de la simulation des paramètres hydrométéorologiques.

Le modèle GESTOP élaboré est basé sur l'Analyse en Composantes Principales, qui est usuellement utilisée pour l'étude des phénomènes spatiaux.

Dans le chapitre I on s'est intéressé aux différents modèles de simulation en hydrologie, la théorie de ces modèles a été développée en ressortant le type de paramètres étudié, ainsi que les contraintes imposées par chaque modèle.

La théorie de la technique de base utilisée (l'Analyse en Composantes Principales) pour le développement du modèle GESTOP a été développée au chapitre II.

Le chapitre III a été consacré à l'élaboration du modèle de simulation, et un organigramme des différentes combinaisons de simulation a été intégré .

Le modèle GESTOP a été appliqué à la simulation des paramètres hydrométéorologiques cycliques (ETP) et aléatoires (pluviométrie, Débit) au pas de temps mensuel, journalier et horaire, ceci a fait l'objet du chapitre IV.

Une étude comparative des modèles GESTOP, Singular Value Décomposition, Markov et Décomposition de la série historique a été développée au chapitre V afin de montrer les avantages et les inconvénients des modèles pris en considération.



**CHAPITRE I**  
**LES MODELES DE SIMULATION EN**  
**HYDROLOGIE**

## CHAPITRE I LES MODELES DE SIMULATION EN HYDROLOGIE.

### I.1. INTRODUCTION.

La simulation a fait l'objet de nombreux sujets de recherche, elle répond à la nécessité de générer des séries chronologiques futures similaires à l'historique, dans le sens où elles préservent les propriétés statistiques de la série historique. La génération de séries synthétiques trouve souvent son application dans la recherche de méthodes d'optimisation de la gestion des barrages réservoirs, et constitue aussi les entrées des modèles pluie - débit. En général la simulation est destinée au design, la planification, et les études opérationnelles des schémas de gestion de la ressource en eau.

Différents types de modèles stochastiques ont été proposés pour la modélisation des séries hydrologiques en général, et des séries d'apports en particulier. Ces modèles comprennent le modèle de Monte Carlo, le modèle autorégressif (Thomas et Fiering 1962), le modèle du bruit Gaussien fractionnel FGN (Mandelbrot et Wallis 1969), les modèles autorégressifs à moyenne mobile ARMA (Carlson et Al. 1970, O'Connel 1971, Hipel et Al. 1977), le modèle de la ligne brisée (Rodriguez et Al. 1972), le modèle de désagrégation (Mejia et Rousselle 1976), Les modèles ARMA-Markov (Lettenmaier et Burges 1977).

En général les modèles sont formulés tel que la valeur courante d'une variable est la somme pondérée des valeurs antécédentes et des nombres aléatoires représentant des effets inconnus. Il est important de noter les hypothèses sur lesquelles sont basées le concept et l'application des modèles stochastiques.

Aucune analyse pratique et utile ne peut être faite si la séquence enregistrée ne représente pas convenablement le processus. C'est pourquoi il est impératif que l'échantillon approche à un degré suffisant la population, ceci implique implicitement la vérification de la condition de stationnarité. Afin de contourner ce problème des modèles non stationnaires peuvent être formulés, tels que les modèles saisonniers, et les modèles ARIMA.

Il est nécessaire d'identifier et de rectifier les erreurs de mesures et d'enregistrements si elles existent, par une analyse préliminaire des données. De la même manière tout effet forcé sur les séries temporelles dues aux activités humaines (déviations des rivières, constructions de barrages), doivent être pris en considération pour que les séries représentent l'événement naturel tel qu'il s'est produit.

Le modèle doit être maniable dans sa forme, capable de reproduire les caractéristiques statistiques de la série historique, il doit être aussi en mesure de générer effectivement des séries synthétiques.

En pratique le choix du modèle est généralement dicté par le jugement, l'expérience ou une préférence personnelle. Dans la majorité des cas on utilise les propriétés statistiques des séries simulées ainsi que celles des échantillons pour identifier le modèle le plus approprié. Cependant une présélection doit être faite en se basant sur la stationnarité, ce qui a conduit à répartir les modèles suivant le type de séries traitées (annuelles, mensuelles, journalières et horaires).

## I.2. LE MODELE DE MONTE CARLO

La méthode de Monté Carlo est le fruit du mariage entre deux branches de la physique théorique. La théorie probabiliste des phénomènes aléatoires (mouvement Brownien), et la théorie du potentiel.

Le développement de la méthode de Monté Carlo, date des années quarante, et à été le résultat des travaux d'un groupe de chercheurs qui furent parmi les premiers à se pencher sur l'arme nucléaire, notamment Von Neumann et Ulma à Los Alamos ( USA).

Le nom de Monté Carlo attribué à la méthode provient de la célèbre ville de Monté Carlo, haut lieu du jeu de hasard.

La méthode de Monté Carlo peut être appliquée comme méthode de simulation, ou comme méthode d'échantillonnage

Pour la génération de nombre aléatoire par une méthode donnée, on est souvent amenés à considérer une fonction  $g(X)$ . Dans le cas de la méthode de Monté Carlo, cette fonction génère le nombre aléatoire suivant l'algorithme  $X_{k+1} = g(X_k)$ . En initialisant la série par une valeur  $X_0$ .

La forme générale est

$$[aX + c] \cdot [\text{mod } m]$$

$X_0$  : Valeur initiale positive

$a$  : Multiple positif

$c$  : Incrément positif

$m$  : Modulo

La séquence de nombre aléatoire est obtenue par :

$$X_{n+1} = [aX_n + c] \cdot [\text{mod } m]$$

Pour générer des séquences de périodes relativement importantes (ce qui est souvent nécessaire), il est proposé de choisir des valeurs assez grandes de  $m$ .

Pour générer des valeurs aléatoires de distribution uniforme  $U$  à l'intérieur de l'intervalle  $[0,1]$  :

$U = \frac{X_{n+1}}{m}$ ,  $U$  ne prendra que des valeurs comprises entre 0 et 1 appartenant à l'intervalle

$$\left\{ 0, \frac{1}{m}, \frac{2}{m}, \frac{3}{m}, \dots, \frac{m-1}{m}, 1 \right\}$$

Pour générer des nombres aléatoires appartenant à l'intervalle  $[a, b]$ , on utilisera la formule de transformation suivante :

$$X = a + (b - a)U$$

Comme indiqué ci-dessus, les nombres aléatoires ainsi obtenus seront de distribution uniforme. Cependant il est possible de générer des nombres aléatoires de distribution normale à partir de la transformation de Box et Muller :

$$\begin{cases} n_i = \sqrt{-2LnU_i} \cdot \cos(2\pi U_{i+1}) \\ n_{i+1} = \sqrt{-2LnU_i} \cdot \sin(2\pi U_{i+1}) \end{cases}$$

### I.3. LES MODELES AUTOREGRESSIFS DE THOMAS ET FIERING

En hydrologie les modèles saisonniers appliqués aux séries mensuelles étaient utilisés à l'origine par Thomas et Fiering (1962) et par Roesner et Yevjevich (1966), parmi les différents modèles saisonniers, le modèle alternatif de Box Jenkins, et le modèle de Thomas et Fiering. Le modèle de Thomas et Fiering est basé sur 12 équations de régression, où chaque mois  $j$  est régressé sur le mois précédent  $j-1$ , dans lesquelles il tient compte de la structure corrélationnelle mois à mois non stationnaire.

Dans sa forme la plus simple le modèle s'écrit comme suit:

$$q_{j+1} = \bar{q}_{j+1} + b_j(q_j - \bar{q}_j) + Z_j \cdot S_{j+1} \cdot \sqrt{1 - r_{(j,j+1)}^2}$$

$q_j$ : Débit du mois  $j$ .

$\bar{q}_j$ : Débit moyen mensuel interannuel du mois  $j$

$b_j$ : Coefficient de régression pour l'estimation du débit du mois  $j+1$  à partir du mois  $j$

$$b_j = r_{(j,j+1)} \cdot \frac{S_{j+1}}{S_j}$$

$Z_j$ : Déviation normale aléatoire  $N(0,1)$

$S_{j+1}$ : Déviation standard entre les débits au pas  $j$  et  $j+1$

$r_{(j,j+1)}$ : Coefficient de corrélation entre les mois  $j$  et  $j+1$ .

Si on considère la variable débit centrée réduite  $Q = \frac{q - \bar{q}}{S}$ , la formulation du modèle devient:

$$Q_{j+1} = r_{(j,j+1)} Q_j + Z_j \sqrt{1 - r_{(j,j+1)}^2}$$

Ce modèle peut être généralisé de deux manières :

- Faire une régression multiple au lieu d'une régression simple, ce qui revient à agrandir le temps de réponse.

Soit par exemple la régression multiple du second degré:

$$Q_{j+1} = r_{(j,j+1)} Q_j + r_{(j+1,j-1)} Q_{j-1} + Z_j \sqrt{1 - R_{(j,j+1)}^2}$$

où  $R$ : Coefficient de corrélation multiple.

- Introduire d'autres variables comme la hauteur de la pluie, le modèle de Thomas et Fiering a été utilisé dans ce sens par Colston et Wiggert (1968).

Le modèle de Thomas et Fiering a subi des modifications afin de pallier au problème de génération de séquences de débits pour des rivières complètement sèches durant l'été, l'approche est la suivante :

On suppose la disponibilité d'un relevé historique assez long de N années, sur les N années on relève pour chaque mois j le nombre d'années  $N_j$  pour lesquelles un écoulement fut relevé, et

$$\text{on évalue } P_j = \frac{N_j}{N}.$$

Pour le mois j, générer un nombre aléatoire y le comparer à  $P_j$  si  $y > P_j$  alors aucun écoulement ne peut être relevé pour ce mois ci.

Le modèle de Thomas et Fiering est parvenu à résoudre le problème de la génération multivariable (alors que les modèles statistiques et empiriques s'intéressent à la reproduction de la structure corrélacionnelle de chaque variable), il a réussi à préserver la structure inter-corrélacionnelle entre les différents variables prises deux à deux.

Le modèle de Thomas et Fiering mis au point en 1962 pour générer des séries mensuelles, présente l'avantage de résoudre le problème de la non stationnarité, ceci en introduisant des paramètres saisonniers au modèle AR, mais l'inconvénient de son utilisation reste la multiplication des paramètres, et par conséquent les problèmes d'estimation de ces derniers.

#### I.4. MODELE DU BRUIT GAUSSIEN FRACTIONNEL (FGN)

En 1965 Mandelbrot a construit (après les résultats des études de Hurst) une classe de processus aléatoires appelés FGN (Fractional Gaussian Noise).

Les processus FGN diffèrent des processus ARMA dans leurs structures autocorrélatives. La différence de base consiste dans leurs mesures relatives de la persistance, qui selon la définition de Taylor (1938), peut être vu comme l'intégral de la fonction covariance entre les bornes 0 et  $\infty$ , la persistance est une entité finie pour tous les types de processus ARMA, tandis que l'intégral du variogramme d'un processus FGN est infini.

Pour la génération hydrologique, les FGN tentent de préserver les caractéristiques naturelles et par la même occasion les longues périodes d'étiages ou de crues. Ceci est appelé effet Joseph, il découle directement de la propriété de la persistance infinie des processus de bruit. Les processus FGN tentent également de reproduire les valeurs extrêmes (élevées ou faibles), cette caractéristique est dite effet Noah.

La théorie des FGN a été approximée à partir des mouvements browniens fractionnels basés sur le coefficient de Hurst h et une mémoire M, ce type de variable est généré par:

$$X_t = (h - 0.5) \cdot \sum_{u=t-M}^{t-1} (t-u)^{h-0.5} \cdot \eta_u \quad 1/2 < h < 1$$

où  $\eta_u$  :  $u = 0..t-1$  sont des variables normales standards indépendantes.

$$\text{pour } t=1 \quad X_1 = (h - 0.5) \cdot (1^{1.5-h} \eta_0 + 2^{1.5-h} \eta_{-1} + 3^{1.5-h} \eta_{-2} + \dots + M^{1.5-h} \eta_{-M+1})$$

$$\text{pour } t=2 \quad X_2 = (h - 0.5) \cdot (1^{1.5-h} \eta_1 + 2^{1.5-h} \eta_0 + 3^{1.5-h} \eta_{-1} + \dots + M^{1.5-h} \eta_{-M+2})$$

De cette manière, des séries de variables dépendantes sont obtenues. On note que pour générer une seule variable FGN,  $M$  variables aléatoires  $\eta_u$  sont nécessaires, la valeur recommandée pour  $M$  est de l'ordre de 20000.

L'inconvénient majeur dans toutes les approximations FGN est le temps de calcul nécessaire en plus des incertitudes.

### I.5. LE MODELE DE LA LIGNE BRISEE (BROKEN LINE)

Le processus ligne brisée fut introduit par Ditlevsen (1971) pour vérifier par la méthode de Monté Carlo quelques-uns de ces résultats. Il a été utilisé principalement dans le domaine de la géophysique, son apport comme un modèle possible pour la simulation de séries temporelles est basé sur deux propriétés.

- Existence de la dérivée secondaire à l'origine de sa fonction de corrélation.
- La possibilité de modéliser de longues interdépendances parmi ces composantes de telle sorte que la simulation du procédé soit plus rapide que celle de n'importe quel autre modèle connus.

Le processus de la ligne brisée  $B_i(t)$  est représenté par la figure I-1.

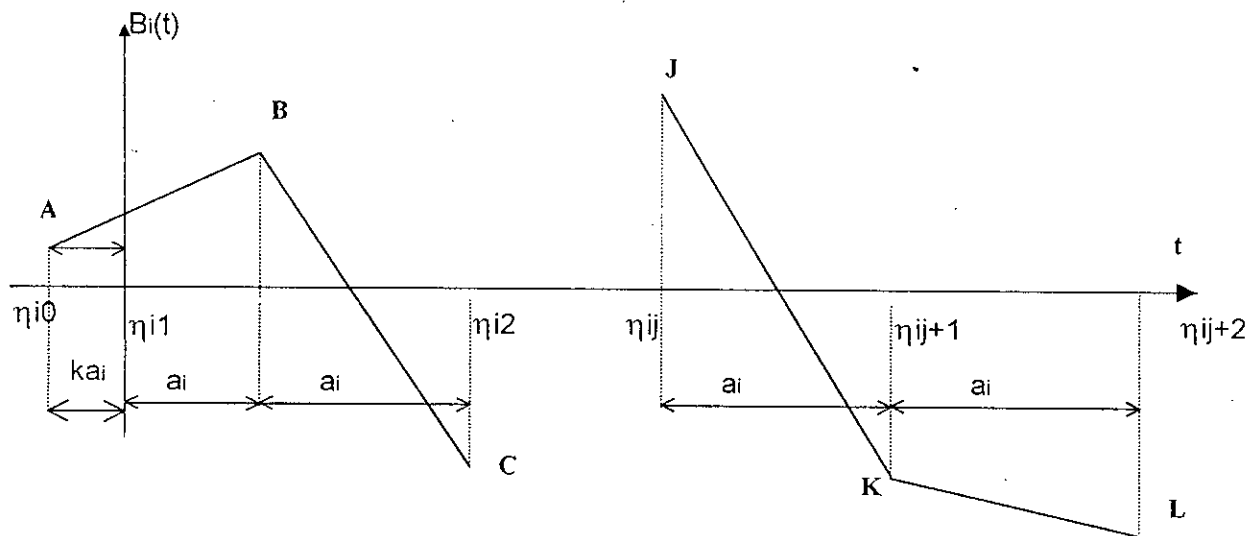


Figure I-1 : Représentation du processus de la ligne brisée

Les points A, B, C, ..., J, K, L ont des projections égales de longueurs  $a_i$  sur l'axe  $t$ , et leurs ordonnées sont distribuées normalement (chacun avec une moyenne et une variance  $\sigma_i^2$ ). Le point A se situe avant l'origine de l'axe des temps d'une quantité  $ka_i$ , où  $k$  est uniformément distribué dans l'intervalle  $[0, 1]$ .

Il s'ensuit qu'en général



$$\beta_i(t) = \eta_{ij} + \frac{t + k \cdot a_i - j \cdot a_i}{a_i} \cdot (\eta_{ij+1} - \eta_{ij})$$

$$j \cdot a_i - k \cdot a_i \leq t \leq (j+1)a_i - k \cdot a_i$$

On démontre que

$$\text{Var}\{\beta_i(t)\} = \frac{2}{3} \cdot \sigma_\varepsilon \cdot \sigma_i^2$$

La fonction de covariance est donnée par un processus  $\xi_t$  de lignes brisées, et est formée, (pour des pas de temps discrets) de la somme de n processus simples de lignes brisées  $\beta_i(t)$ ,  $i=1\dots n$ , donc:

$$\xi_t = \sum_{i=1}^n \beta_i(t)$$

L'application de ce modèle, nécessite la détermination de n,  $a_i$ , et  $\sigma_i$  pour  $i=1\dots n$ , totalisant  $2n+1$  paramètres.

La méthode a été jugée utile au moins dans certains cas, et flexible, le grand inconvénient reste toujours l'estimation de nombreux paramètres.

## I.6. LES MODELES AUTOREGRESSIFS

Les modèles autorégressifs ont été utilisés depuis le début des années soixante en hydrologie pour la modélisation des séries temporelles annuelles et périodiques, l'application de ces modèles a été intéressante en hydrologie du fait que la forme autorégressive est une caractéristique des séries étudiées (séries autocorrélées) en plus de leur simplicité d'utilisation.

Le développement historique des modèles AR en hydrologie peut être divisé en deux étapes. La décennie de 1960 a été initiée principalement par les travaux de Thomas et Fiering (1962) et Yevjevich (1963), et la décennie de 1970 motivée par la publication du livre de Box et Jenkins (1971).

Durant la première période, la procédure utilisée pour l'estimation des paramètres des modèles était basée sur la méthode des moments, et les tests d'adéquation des modèles étaient basés sur l'analyse du corrélogramme.

Depuis 1970 les chercheurs en hydrologie ont utilisé des méthodologies plus raffinées proposées par Box et Jenkins et autres, spécialement pour la vérification des hypothèses du modèle et la sélection du modèle adéquat.

## I.6.1 Formulation mathématique du modèle.

### I.6.1.1 Modèle Autorégressif à paramètres constants.

Soit  $Y$  une série temporelle normalement distribuée de moyenne  $\mu$  et d'écart type  $\sigma$  avec une structure de dépendance (variable autocorrélée). Le modèle autorégressif d'ordre  $p$  noté AR( $p$ ) représente la variable  $Y_t$  s'écrit généralement comme suit:

$$Y_t = \mu + \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \dots + \phi_p(Y_{t-p} - \mu)$$

$$Y_t = \mu + \sum_{j=1}^p \phi_j(Y_{t-j} - \mu) + \varepsilon_t$$

Où

$\varepsilon_t$ : Série indépendante normalement distribuée de moyenne nulle et d'écart type  $\sigma_\varepsilon$ .

$\phi_1 \dots \phi_p$ : Coefficients autorégressifs.

Les paramètres du modèle estimés à partir des données sont  $\mu$ ,  $\sigma_\varepsilon^2$ ,  $\phi_1 \dots \phi_p$  et  $\sigma_\varepsilon$ .

### I.6.1.2 Les modèles autorégressifs à paramètres périodiques.

Les modèles autorégressifs à paramètres périodiques sont des modèles où tous les paramètres ou une partie de ces derniers varient d'une manière périodique au cours de l'année.

La périodicité peut être dans la moyenne, la variance et (ou) les coefficients autorégressifs

- Le modèle autorégressif AR( $p$ ) à moyenne et variance périodiques et des coefficients autorégressifs constants, peut être représenté par

$$Y_{v,\tau} = Y_\tau + \sigma_\tau Z_{v,\tau}$$

et

$$Z_{v,\tau} = \sum_{j=1}^p \phi_j \cdot Z_{v,\tau-j} + \varepsilon_{v,\tau}$$

Où :

$Y_{v,\tau}$ : Variable dépendante pour l'année  $v$  et le temps  $\tau$  (mois, semaine, etc.)

Avec:  $\tau = 1, \dots, \omega$  et  $\omega$  est le nombre d'intervalles de temps.

$Y_{v,\tau}$  est supposé normalement distribué de moyenne  $\nu_\tau$  et de variance  $\sigma_\varepsilon$ , et a une structure autorégressif d'ordre  $p$ .

$Z_{v,\tau}$ : Variable dépendante normalement distribuée de moyenne nulle et de variance unité.

$\varepsilon_{v,\tau}$ : Variable indépendante de moyenne nulle et de variance  $\sigma_\varepsilon^2$ .

Les paramètres du modèle sont:  $\mu$ ,  $\sigma_\varepsilon^2$ ,  $\phi_1 \dots \phi_p$  et  $\sigma_\varepsilon$

Un modèle autorégressif d'ordre  $p$  à coefficients périodiques est donné par

$$Z_{v,\tau} = \sum_{j=1}^p \phi_{j,\tau} \cdot Z_{v,\tau-j} + \sigma_{\varepsilon\tau} \xi_{v,\tau}$$

où:

$\phi_{j,\tau}$ :  $j^{\text{ième}}$  coefficient autorégressif périodique au temps  $\tau$ .

$\sigma_{\varepsilon\tau}$ : Coefficient périodique.

$\xi_{v,\tau}$ : Variable indépendante standardisée.

Les paramètres du modèle sont  $\mu_\tau, \sigma_\tau^2, \phi_{1\tau}, \dots, \phi_{p\tau}$  et,  $\tau = 1, \dots, \omega$ .

## I.6.2 Application des modèles autorégressifs

### I.6.2.1 Les modèles autorégressifs annuels

Les modèles autorégressifs annuels sont utilisés pour la génération des séries annuelles pour la planification des systèmes de ressources hydriques avec une régularisation interannuelle.

La gestion des systèmes de ressources hydriques à régularisation interannuelle requiert des études opérationnelles à un niveau saisonnier, dans de tels cas, on procède à une simulation de séries annuelles autorégressives puis une désagrégation en valeurs saisonnières.

Quelques problèmes de gestion de système de ressources hydriques consistant en un ensemble de réservoirs, requièrent l'agrégation des différents réservoirs équivalent (le débit entrant au réservoir est l'agrégation des débits entrants aux petits réservoirs) dans ce cas des modèles AR univariés peuvent être utilisés pour la génération de séries synthétiques de débits agrégés.

### I.6.2.2 Modèles autorégressifs périodiques

Les modèles AR à coefficients périodiques sont utilisés pour la génération de séries synthétiques périodiques et pour la prévision. Ils sont nécessaires à la planification et à la conception des systèmes de ressources hydriques utilisant des procédés de régularisation saisonnière qui nécessitent des données au pas de temps mensuel décadaire ou même journalier.

## I.6.3 Limites des modèles autorégressifs

### I.6.3.1 Modèles autorégressifs annuels

Il est généralement supposé que les modèles AR(p) préservent au sens statistique la moyenne, l'écart type, et les  $p$  premiers coefficients historiques. La moyenne en général ne pose aucun problème, mais pour les coefficients d'autocorrélation (vrai dans le cas où les paramètres seraient estimés par la méthode des moments), la seule ressemblance entre les corellogrammes historique et simulés et dans l'amplitude et la forme.

Il existe des cas où les modèles AR peuvent ne pas présenter une ressemblance des corellogrammes que pour quelques premiers pas, ceci est dû au fait que les modèles AR sont généralement des modèles à mémoire courte.

Un autre facteur important dans les applications pratiques des modèles AR est relatif à la modélisation des séries asymétriques, pour lesquelles trois approches ont été proposées :

- Normaliser les données asymétriques par une transformation appropriée.
- Appliquer le modèle AR aux données asymétriques originales, et trouver la distribution des résidus.
- Utiliser une relation (si elle existe) entre les deux premiers moments des séries originales asymétriques et ceux des séries normales, tels que les moments des séries originales soient préservés.

Chaque approche à ses avantages et ses inconvénients, et c'est à la base d'un jugement personnel que se fera le choix de telle ou telle approche.

Finalement, il reste à ne pas négliger l'aspect physique des séries temporelles, qui conduit à des sous estimations, et par conséquent à des décisions fausses.

### 1.6.3.2 Modèles AR périodiques

L'identification des modèles AR périodiques requiert de :

Résoudre le problème de la non normalité des données.

Décider si les séries de Fourier seront utilisées pour l'estimation des paramètres.

Choisir le modèle AR à coefficients constants ou à coefficients périodiques.

Les solutions à ces problèmes, et les décisions à prendre sont basées en grande partie sur la taille des séries historiques, ainsi que sur l'expérience et le jugement personnel du modélisateur.

## 1.7. LES MODELES DE DESAGREGATION.

Les modèles de désagrégation ont été introduit pour la première fois pour la modélisation en hydrologie par Harms et Campbell (1966), mais le premier modèle qui fut accepté est celui présenté par Valencia et Shaake (1973), ce modèle est la base des techniques de désagrégation dans la simulation des séries hydrologiques.

Mejia et Rouselle ont proposé des modifications du modèle de base pour surmonter les défauts de ce dernier, Lane (1979) a aussi proposé quelques améliorations supplémentaires.

La plupart des applications de la désagrégation ont été dans le domaine temporel, Lane (1979) a appliqué les mêmes principes dans le domaine spatial.

Le modèle de désagrégation est un processus par lequel des séries temporelles sont générées dépendant des séries déjà disponibles. Les séries de départ sont générées par un modèle stochastique choisi, puis désagrégé en sous séries.

Généralement dans le cas de désagrégation saisonnière (données saisonnières) la somme des sous séries doit donner exactement le série originale (données annuelles) où série clé.

Les deux formes de bases des modèles de désagrégation sont spatiales (désagrégation des données annuelles en données mensuelles et décadaires) et temporelles (désagrégation des débits d'une rivière en débit individuel des affluents).

Il est à noter que la désagrégation n'est pas limitée à un seul niveau, autrement dit, une série annuelle peut être désagrégée en série semi-annuelle en première étape, qui elle-même peut être désagrégée en séries mensuelles.

Modèle général de désagrégation

$$Y = A \cdot X + B\varepsilon$$

Y : Série d'observations à simuler (sous série).

X : Série clé.

$\varepsilon$  : Terme stochastique.

Modèle temporel de désagrégation à site unique: il existe trois formes du modèle de désagrégation temporel à site unique.

1- Le modèle de base: c'est celui de Valencia et Shaake proposé (1978), il est de la forme :

$$Y = A \cdot X + B\varepsilon$$

X : Série mère.

Y : Série des valeurs saisonnières .

A,B : Paramètres du modèle.

C'est le modèle le plus simple, sa structure permet de préserver la covariance entre les valeurs annuelles, et les valeurs saisonnières, et de préserver la variance et la covariance entre valeurs saisonnières.

2- Le modèle étendu.

Développé par Mejia et Rousselle (1979), c'est une extension du modèle de base, un terme supplémentaire a été inclus pour préserver les covariances saisonnières entre saisons successives, il a la forme :

$$Y = A \cdot X + B \cdot Y + C \cdot Z$$

Z : Matrice colonne contenant des valeurs de l'année précédente.

C : Matrice des paramètres supplémentaires.

Le nombre de paramètres du modèle est excessif, et ceci rend l'estimation de ces paramètres plus compliquée.

3- Le modèle condensé

Lane (1979) a développé une approche qui annule chaque paramètre du modèle étendu qui n'est pas importante, c'est le modèle condensé, le nombre de paramètres estimés, ainsi que le nombre de moments préservés est réduit, le modèle utilisé est le modèle étendu mais sur une base saisonnière, il est donné par

$$Y = A_{\tau} X + B_{\tau} \varepsilon + C_{\tau} Y_{\tau-1}$$

$\tau$  : Saison à générer, donc il y a W saisons, donc W équations individuelles suivant la forme de l'équation et donc W ensembles de paramètres  $A_{\tau}, B_{\tau}, C_{\tau}$ .

Le principal avantage de ce modèle est la réduction du nombre de paramètres.

## I.8. LES MODELES MARKOVIENS

### I.8.1 Définition

Une chaîne de Markov est un système qui subit au cours du temps des changements d'états à transition aléatoire, et qui sans être dénué de mémoire ne garde de son passé que le souvenir le plus récent.

Une chaîne de Markov est dite d'ordre  $k$  si l'équation ci-dessous est vérifiée:

$$P\{X_t, X_{t-1}, \dots, X_{t-k}, X_{t-k-1}\} = P\{X_t, X_{t-1}, \dots, X_{t-k+1}, X_{t-k}\}$$

Où  $X$  : Variable aléatoire.

Pour l'ordre  $k$ , la réalisation d'un état donné ne dépend que des  $k$  réalisations précédentes, dans ce qui suit la pluviométrie (phénomène aléatoire par excellence) est prise comme exemple.

### I.8.2 Dédution analytique des paramètres

Les précipitations sont décrites par le processus stochastique suivant:

$$Z_t = X_t \cdot Y_t$$

$X_t$ : Caractérise l'état au jour  $t$  ( $X_t=0$ : état sec,  $X_t=1$ : état pluvieux)

$Y_t$ : Hauteur de la pluie (si  $X_t=1$ )

Pour la détermination de la matrice de passage, on s'intéresse seulement à la variable  $X_t$  qui décrit le passage d'un état  $E_i$  à un état  $E_j$ .

#### I.8.2.1 Processus d'ordre 1

Soit  $X$  la variable aléatoire caractérisant l'état,  $X$  peut donc prendre deux valeurs (système binaire), zéro pour l'état sec et un pour l'état pluvieux.

Ce processus est caractérisé par les probabilités conditionnelles de passage de l'état de la veille à l'état d'aujourd'hui.

Ainsi si l'état au jour  $k$ , ne dépend que de l'état au jour  $k-1$ , on note :

$$\text{prob}(X_k = j / X_{k-1} = i) = L_{ij}$$

$L_{ij}$ : Représente la probabilité d'obtenir au jour  $k$  l'état  $j$ , sachant qu'au jour  $k-1$ , on avait l'état  $i$ .

En considérant deux états possibles, on obtient une matrice de passage P présentée dans le tableau I-1.

Etat au jour k-1	Etat au jour k	
	0	1
0	$L_{00}$	$L_{01}$
1	$L_{10}$	$L_{11}$

Tableau I.8-1: Matrice de passage d'ordre 1

La matrice P ayant les propriétés d'une matrice stochastique implique la relation :

$$L_{i0} + L_{i1} = 1 \quad \text{avec } i=0 \text{ ou } 1$$

Les coefficients de la matrice P sont obtenus par le décompte des jours caractérisés par les états (0,1).

La probabilité marginale (ou inconditionnelle)  $P_0$  peut s'écrire sous la forme suivante :

$$P_0 = \frac{L_{10}}{(1 - L_{00} + L_{10})}$$

### I.8.2.2 Processus d'ordre 2

On cherche la probabilité de l'état de la classe 0 ou 1 d'aujourd'hui en fonction des états connus d'hier et d'avant hier.

La matrice de transition est représentée par le tableau I-2.

Etats au jours k-1 et k-2	Etats au jours k-1 et k			
	00	01	10	11
00	$\beta_{000}$	$\beta_{001}$	-	-
01	-	-	$\beta_{010}$	$\beta_{011}$
10	$\beta_{100}$	$\beta_{101}$	-	-
11	-	-	$\beta_{110}$	$\beta_{111}$

Tableau I.8-2: Matrice de passage d'ordre 2

$\beta_{ijk}$  : Représente la probabilité conditionnelle d'obtenir un doublet de classe (i,j).

Du fait qu'il y a chevauchement de deux couples sur la veille, il est nécessaire d'avoir l'égalité de la classe de transition dans les deux couples. Pour cet ordre, l'état au jour k, dépend de l'état aux jours k-1 et k-2. Lors de la définition de la matrice de passage exprimant les probabilités de couples successifs, il y a impossibilité de certaines combinaisons, vu la succession de certains doublets. Si la matrice de transition d'ordre 1, élevée à l'exposant n converge à l'infini vers une matrice asymptotique, le système est alors stable en probabilité, on dit qu'il a la propriété érgodique, cette dernière permet de déterminer l'ordre maximal avec lequel on peut travailler.

### I.8.3 Principe de la simulation

La simulation aboutit à l'obtention de séries synthétiques à partir d'une série historique plus ou moins longue.

Comme on a vu précédemment la variable  $Z_t$  est décrite de la manière suivante:

$$Z_t = X_t \cdot Y_t$$

Pour l'ordre 1, à partir des probabilités conditionnelles  $L_{00}$  et  $L_{10}$ , on génère l'état journalier des séries synthétiques après tirage d'un nombre aléatoire compris entre 0 et 1 (constituant une probabilité) qu'on compare à  $L_{i0}$  ( $i=0$  ou  $1$ ), le jour simulé peut être sec ou pluvieux.

Le principe de simulation est similaire pour les ordres supérieurs, dans ce cas la probabilité obtenue par tirage aléatoire sera comparée aux probabilités conditionnelles de la matrice de passage correspondant à l'ordre du processus de simulation.

S'il s'avère que le jour simulé est pluvieux, on lui attribue une hauteur de pluie qu'on tire des fonctions de répartition conditionnelles.

En tenant compte de la position du jour humide dans un épisode donné, on donne à chaque jour simulé pluvieux une hauteur de pluie, cela se fait par le biais d'un deuxième tirage d'un nombre aléatoire qui sera confronté à la fonction de répartition considérée.

### I.9. SIMULATION DE VARIABLES CYCLIQUES PAR DECOMPOSITION DE LA SERIE HISTORIQUE.

L'examen de la représentation graphique des séries hydrométéorologiques cycliques permet de constater la vraisemblance de certaines composantes fondamentales à savoir :

- Un mouvement cyclique qui désigne des oscillations de périodes et d'amplitudes plus ou moins irrégulières.
- Un élément aléatoire qui échappe à toute précision, et qui entraîne des variations accidentelles plus au moins sporadiques qualifiées de variations résiduelles.

Donc on considère les données brutes comme la superposition de deux séries indépendantes, la première prenant en compte les variations saisonnières (porteuse), la seconde les variations résiduelles ou fluctuations (composante fluctuante).

Pour la séparation de ces deux composantes, on utilise une technique de lissage, les séries seront décomposées de la manière suivante :

$$X_t = V_{it} + W_{it}$$

- Avec  $X_{it}$  :Données brutes  
 $V_{it}$  : Composante saisonnière ou porteuse  
 $W_{it}$  : Composante fluctuante  
 $t= 1$  à  $n$  ( $n=$  nombre d'années de la série historique).



La décomposition en séries de Fourier de la série initiale permet d'extraire la composante saisonnière, la composante fluctuante est donnée par:

$$V_{it} = X_{it} - W_{it}$$

La simulation des composantes saisonnières et fluctuantes permet d'obtenir des séries synthétiques par superposition de ces dernières.

### 1.9.1 Estimation des paramètres de lissage

Sachant que la composante saisonnière peut se mettre sous la forme:

$$V_t = A_0 + a_1 \sin t + b_1 \cos t + a_2 \sin t + b_2 \cos t \dots$$

En considérant  $m$  harmoniques le lissage sera donné par:

$$V_t = A_0 + \sum_{j=1}^m \sum_{t=1}^T a_j \cos \frac{2\pi j t}{T} + b_j \sin \frac{2\pi j t}{T}$$

$A_0$  : Moyenne annuelle de la série brute.

$a_j, b_j$  : Coefficient de la série de Fourier.

$M$  : Nombre d'harmoniques.

$T$  : Nombre de jours par année.

La porteuse est définie par une moyenne annuelle à laquelle s'ajoutent les variations périodiques que caractérisent les coefficients  $a_j$  et  $b_j$ , d'amplitude  $\sqrt{a_j^2 + b_j^2}$ , de période  $(T/j)$  et d'angle de phase  $\text{Arctg}(a_j/b_j)$ , de façon que :

$$X_t = V_t + W_t$$

$$\sum W_t = 0, \quad A_0 = \frac{1}{T} \cdot \sum_{i=1}^T X_t$$

Les termes  $a_j$  et  $b_j$  sont estimés par la méthode des moindres carrés.

$$a_j = (2/T) \cdot \sum_{t=1}^T X_t \cdot \cos(2\pi j t / T)$$

$$b_j = (2/T) \cdot \sum_{t=1}^T X_t \cdot \sin(2\pi j t / T)$$

### 1.9.2 Simulation des Fluctuations

Les fluctuations sont obtenues en retranchant à la série des données brutes, la composante saisonnière (obtenue par la décomposition en série de Fourier de la série historique).

Les fluctuations sont de moyenne nulle, et leurs distributions suivent une loi normale.

Les fluctuations sont simulées par les chaînes de Markov, dans ce cas on considère deux états (état 0 si la fluctuation est négative ou nulle, et l'état 1 si celle ci est positive).

### I.9.3 Simulation de la porteuse

L'analyse harmonique permet de ressortir le mouvement saisonnier ou «porteuse» qui s'écrit :

$$V_i^t = X_{i0} + \sum_{j=1}^m \sum_{t=1}^{\pi} (a_{ij} \cdot \cos(\frac{2\pi jt}{T}) + b_{ij} \cdot \sin(\frac{2\pi jt}{T}))$$

Pour le procédé de simulation, la porteuse sera vue comme une courbe moyenne à laquelle on a superposé pour chaque année des fluctuations d'amplitudes et de phases bien définies.

La simulation se fera donc pour chaque année de la manière suivante :

Simulation de la moyenne interannuelle.

Simulation des coefficients  $a_{ij}$  et  $b_{ij}$  avec  $j$  variant de 1 à  $m$  ( $m$  étant le nombre d'harmoniques).

La simulation des coefficients  $a_{ij}$  et  $b_{ij}$  se fait de deux manières différentes, soit en tenant compte des corrélations qui existent entre ces variables ou bien en les ignorant.

### I.10. LE MODELE SVD (SINGULAR VALUE DECOMPOSITION)

Les modèles utilisés usuellement pour la simulation des débits requièrent un nombre élevé de paramètres à estimer. Le modèle SVD est apparu en alternative à ces derniers (Cavadias et Al. 1980-1985; Haché et Al. 1994) il est basé sur le théorème de décomposition en valeurs singulières. Dans sa forme originale le modèle SVD permet de générer des séries de débits. Le but de cette technique est la préservation des moyennes, variances et corrélations entre les paires de débits de différentes périodes (semaine, mois) de la même année.

$$[X] = [C] \cdot [\lambda]^{1/2} [A]^t$$

C: CP standardisées de X

$\lambda$  : Matrice diagonale des valeurs propres

A: Matrice des vecteurs propres

$$[C] = X \cdot [A]^t \cdot [\lambda]^{1/2}$$

La procédure de calcul se fait en deux étapes : **la calibration** et **la simulation**

#### I.10.1 Calibration

La calibration consiste en un premier lieu à associer à la variable d'origine Y, une variable X tel que  $x_{ij} = f_j(y_{ij})$ , où  $f$  est une fonction destinée à normaliser les débits saisonniers, la seconde étape est le calcul de la matrice de corrélation de X qu'on notera  $R_x$ . La troisième étape est la détermination de la matrice  $C_s$  des  $r$  composantes principales standardisées.

En dernier lieu, on ajuste des distributions de probabilités aux éléments de chaque colonne de  $C_s$ .

### I.10.2 Simulation

La simulation par le modèle SVD consiste à suivre les étapes suivantes:

1- Génération de  $m$  valeurs pour chaque colonne de  $C_S$  pour obtenir une matrice de CP simulées qu'on notera  $\tilde{C}_{S1}$ ,  $\tilde{C}_{S1}$  est approximativement orthonormale en raison du caractère aléatoire du générateur.

2- Transformation de  $\tilde{C}_{S1}$  en une matrice orthonormale notée  $\tilde{C}_{S2}$ , en utilisant le théorème de décomposition en valeurs singulières, ceci afin de reproduire exactement la matrice de corrélation  $R_x$ .

$$\tilde{C}_{S2} = \tilde{C}_{S1} \cdot A_{Z-1} \cdot \lambda_{Z-1}^{1/2}$$

Calcul des débits simulés

$$\tilde{X}_S = \tilde{C}_{S2} \cdot A_{Z-1} \cdot \lambda_{Z-1}^{1/2}$$

Le modèle SVD dans sa version originale reproduit d'une manière presque identique (en terme de moyennes, écarts types, coefficients de corrélation) l'historique, cette option ne satisfait pas une des caractéristiques naturelles des données qui est l'aléa, d'où la nécessité de réexaminer l'algorithme de calcul afin d'introduire une variabilité dans les caractéristiques de l'échantillonnage des débits simulés, pour cela trois variantes ont été étudiées:

- Réduction du nombre de CP de 50% dans la phase de calibration.
- Elimination de la standardisation des CP (le passage de  $\tilde{C}_{S1}$  à  $\tilde{C}_{S2}$ )
- Elimination de la standardisation des CP et des débits simulés

Avec le modèle SVD une variabilité a put être introduite dans les caractéristiques statistiques, pour améliorer la qualité des simulations, l'application de cette variante aux débits observés a permis d'examiner de près le problème de la transformation des données qui influe directement sur la qualité de la simulation.

## **CHAPITRE II**

# **ANALYSE EN COMPOSANTES PRINCIPALES THEORIE ET PROGRAMMATION**

## CHAPITRE II ANALYSE EN COMPOSANTES PRINCIPALES THEORIE ET PROGRAMMATION

### II.1. HISTORIQUE DE L'ANALYSE DES DONNEES (ADD)

Les méthodes d'analyse des données ont largement démontré leur efficacité dans l'étude de grandes masses d'informations. Ce sont des méthodes multidimensionnelles, en opposition aux méthodes de statistiques descriptives simples (qui ne traitent qu'une ou deux variables à la fois). Elles permettent la confrontation d'un ensemble d'informations, ce qui est infiniment plus riche que leur examen séparé. Les représentations simplifiées de grands tableaux de données que ces méthodes permettent d'obtenir, s'avèrent un outil de synthèse remarquable.

L'analyse des données (ADD) a été définie (J.P.Benzecri 1982) comme étant un outil pour dégager de la langue des données le pur diamant de la vérité de nature. Parmi les méthodes utilisées, on étudiera en particulier, celles de l'analyse factorielle à partir de laquelle plusieurs variantes se sont développées :

L'analyse factorielle des correspondances (AFC), méthode de description des données qualitatives, proposée pour l'étude des tableaux de contingence (croisement de deux caractères nominaux).

L'analyse discriminante (AD), possédant des motivations descriptives et décisionnelles, cherche à déterminer les axes donnant une meilleure discrimination du nuage de points.

La variable à expliquer est nominale (qualitative).

Les variables explicatives sont métriques (mesurables).

L'analyse de variance (AV), applicable dans le cas inverse de l'analyse discriminante c'est à dire :

Variable à expliquer mesurable.

Variables explicatives nominales.

L'analyse des proximités, se base sur les tableaux de distance (tableau carré symétrique), cette méthode est dite aussi procédé non métrique.

L'analyse en composantes principales (ACP), permet de visualiser l'information contenue dans un tableau de données quantitatives.

Les premières analyses à plusieurs variables remontent au tout début du XX<sup>ième</sup> siècle (Pearson 1901, Spearman 1904, Burt 1909). Les bases théoriques de presque toutes celles utilisées actuellement ont été établies, de façon quasi définitive vers 1930 (Mahalanobis 1927, Fisher 1928, Hottelling 1931-1933, Barcette 1933) avec des compléments importants après 1950 (Rao, Anderson 1958).

Cependant faute de moyens de calcul suffisants, ces méthodes n'ont pu être utilisées que de façon extrêmement limitée.

Jusqu'aux années soixante, ces méthodes étaient perfectionnées et s'enrichissaient de variantes, mais toutes restaient inabordables pour les praticiens car elles nécessitaient une masse considérable de calcul.

C'est l'apparition, puis l'extraordinaire développement des ordinateurs qui permirent la vulgarisation des techniques statistiques d'analyse des données.

## II.2. DEFINITIONS

Dans le domaine de l'hydrologie, et devant la quantité importante d'informations, l'étude des phénomènes hydrométéorologiques constitue une tâche très complexe, c'est ainsi que l'analyse en composantes principales (ACP), par sa capacité de concentrer le maximum d'information dans un nombre réduit de nouvelles variables, s'est imposée en tant que technique efficace d'analyse des données.

### II.2.1 Notions d'individu et de caractère

On distingue généralement deux ensembles : les individus et les caractères relatifs à ces individus. Le terme "individus" peut désigner selon les cas : une année d'observation ou une autre unité de temps. L'ensemble des individus peut provenir d'un échantillonnage dans une population ou il peut s'agir de la population toute entière (cas rare surtout en hydrologie).

L'individu "i" est décrit par le vecteur appartenant à  $R^P$

$$X_i = \{X_{ij} / j = 1 \text{ à } P\}$$

Sur un individu, on relève un certain nombre de caractères (dits aussi variables) désignant en général un paramètre intervenant dans un phénomène complexe à étudié (la répartition spatiale ou temporelle des précipitations dans un bassin versant).

Le caractère (ou variable) "j" est décrit par le vecteur de  $R^N$

$$X_j = \{X_{ij} / i = 1 \text{ à } N\}$$

Si l'ensemble des individus doit être homogène, l'ensemble des variables peut être hétérogène.

### II.2.2 Définition algébrique

L'ACP considère P variables pour lesquelles on dispose de N individus, donc la matrice [X] résultant du croisement «NxP» est la matrice de données [X] tel que :

$$[X] = \begin{vmatrix} X_{11} & X_{12} & \dots & X_{1j} & \dots & X_{1P} \\ X_{21} & X_{22} & \dots & X_{2j} & \dots & X_{2P} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_{i1} & X_{i2} & \dots & X_{ij} & \dots & X_{iP} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_{N1} & X_{N2} & \dots & X_{Nj} & \dots & X_{NP} \end{vmatrix}$$

Le tableau [X] est remplacé par un tableau [Y] (individus X nouvelles variables) en réduisant le nombre de variables nécessaires pour décrire les individus, avec une perte minimale d'informations, ces nouvelles variables sont appelées composantes principales (ou CP).

$$\begin{cases} C_1 = a_{10} + a_{11}X_1 + \dots + a_{1N}X_N \\ C_2 = a_{20} + a_{21}X_1 + \dots + a_{2N}X_N \\ \dots \\ C_i = a_{i0} + a_{i1}X_1 + \dots + a_{iN}X_N \\ \dots \\ C_N = a_{N0} + a_{N1}X_1 + \dots + a_{NN}X_N \end{cases}$$

$C_j$ :  $j^{\text{ième}}$  Composante Principale  
 $X_j$ : Vecteur variable initiale  
 $a_{jk}$ : Coefficient du système

On note au passage que les termes  $a_{j0}$  désignent le vecteur permettant la translation de l'origine de l'ancien repère vers le centre de gravité du nuage de points. Un centrage des données initiales annule les coefficients  $a_{j0}$ .

### II.2.3 Définition géométrique

L'Analyse en Composantes Principales est puissante par son support géométrique, la méthode consiste à rechercher un premier axe qui soit le plus proche possible de tous les points au sens des moindres carrés, tel que la somme des carrés des distances des N points à cet axe serait minimale ou encore la projection de ces derniers sur cet axe possède une dispersion minimale. Cet axe est appelé "axe factoriel".

Un second axe est obtenu après projection des N points sur un hyperplan orthogonal au premier axe, tel que la dispersion des projections des N points sur celui-ci serait toujours maximale, et le procédé se réitère P fois.

## II.3. FORMULATION MATHÉMATIQUE DU PROBLÈME

La recherche des composantes principales est faite sous deux contraintes :

Elles doivent être indépendantes, c'est à dire, prise deux à deux, elles présentent obligatoirement des corrélations nulles.

Les axes factoriels doivent être déterminés par ordre d'importance décroissante, le premier axe expliquera le maximum de la variance totale tandis que le second expliquera le maximum de la variance résiduelle non expliquée par le premier, jusqu'au dernier axe. Mais l'expérience a montré qu'un nombre Q d'axes nettement inférieur à P suffit pour donner le maximum d'informations.

L'objectif de l'ACP étant de maximiser la variance, la formulation matricielle du problème est la suivante :

Soit les matrices colonnes  $V_1, V_2, \dots, V_q$  de dimension  $(N \times 1)$  représentant l'hyperplan formé par les axes principaux vérifiant les conditions de normalité et d'orthogonalité :

$$\begin{cases} \vec{V}_i \cdot \vec{V}_i = 1 \\ \vec{V}_i \cdot \vec{V}_j = 0 \end{cases} \rightarrow \begin{cases} i = 1 \rightarrow Q \\ j = 1 \rightarrow Q \end{cases}$$

on veut maximiser la quantité  $\sum Var(C_j)$

Sachant que:  $\sum Var(C_j) = V_j \cdot [M] \cdot [R] \cdot [M] \cdot V_j$

$C_j$  : Composante principale d'ordre j

$[R]$  : Matrice des covariances des variables  $(X_1, X_2, \dots, X_p)$

$[M]$  : Métrique définissant le produit scalaire sur l'espace  $R^p$

#### II.4. CHOIX DE LA METRIQUE

La métrique  $[M]$  possède deux options classiques :

$$[M] = I : \text{Matrice identité}$$

La covariance sera utilisée afin de quantifier les relations inter variables, on parlera alors d'une *ACP canonique*.

$$[M] = D_{\%}^2$$

Tel que :

$$D_{\%}^2 = \begin{vmatrix} \frac{1}{\sigma_1^2} & 0 & & & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 & & \\ & 0 & \cdot & & \\ & & 0 & \frac{1}{\sigma_1^2} & 0 \\ & & & \cdot & 0 \\ 0 & & & 0 & \frac{1}{\sigma_p^2} \end{vmatrix}$$

On utilise généralement cette métrique pour pallier au problème de l'hétérogénéité des caractères (variables), et éviter l'influence du choix d'unité des variables. Dans ce cas on parlera d'*ACP normée*, elle est équivalente à une ACP canonique effectuée sur des variables centrées réduites.

Les données ainsi transformées se présentent sous forme d'une matrice dont toutes les variables sont de moyenne nulle et d'écart type unité.



## II.5. PROCÉDE D'APPLICATION DE L'ACP

### II.5.1 Calcul de la matrice des covariances

La matrice des covariances, notée  $[R]$  est la base de l'ACP, elle est obtenue en appliquant la relation suivante :

$$[R] = (1/n)[X]'.[M].[X]$$

$[R]$  : Matrice de covariance de dimension (P x P).

$[X]$  : Matrice de données.

$[X]'$  : Matrice transposée de  $[X]$ .

$[M]$  : Métrique.

### II.5.2 Recherche des axes principaux

Le but est de construire un nouveau système d'axes avec un minimum de variables assurant un maximum de variance.

#### II.5.2.1 Recherche du premier axe

Tel que mentionné précédemment, la contribution maximale est donnée par le premier axe principal. On doit donc maximiser la variance relative à celui-ci, la recherche du premier axe principal consiste à résoudre le problème

$$\begin{cases} \text{MaxVar}(C_1) \\ V_1'.[M].V_1 = 1 \end{cases}$$

On peut déterminer la variance de  $C_1$  à l'aide de la matrice des covariances  $[R]$  du vecteur :

$$X = (X_1, X_2, \dots, X_j, \dots, X_p)$$

$$\text{Var}(C_1) = V_1'.[M].[R].[M].V_1$$

En utilisant la méthode du multiplicateur de LAGRANGE on peut écrire :

$$L = V_1'.[M].[R].[M].V_1 - \lambda_1(V_1'.[M].V_1 - 1)$$

La dérivée par rapport à V est nécessairement nulle :

$$\partial L / \partial V_1 = 2.[M].[R].[M].V_1 - 2\lambda_1.[M].V_1 = 0$$

Puisque la matrice  $[M]$  est inversible :

$$[R].[M].V_1 = \lambda_1.V_1$$

Donc  $V_1$  est le vecteur propre de la matrice  $[R] \cdot [M]$ . Il suffit de choisir comme vecteur  $V_1$  le vecteur propre associé à la plus grande valeur propre  $\lambda_1$  de la matrice  $[R] \cdot [M]$  pour maximiser la variance de  $C_1$ .

### II.5.2.2 Recherche du second axe

On cherche à déterminer le vecteur unitaire  $V_2$  de telle façon que la composante  $C_2$  soit de variance maximale et non corrélée à  $C_1$ .

Sachant que :

$$\begin{aligned} \text{Var}(C_2) &= V_2' \cdot [M] \cdot [R] \cdot [M] \cdot V_2 \\ V_2' \cdot [M] \cdot V_2 &= 1 \\ \text{COV}(C_1, C_2) &= 0 \end{aligned}$$

L'expression de  $\text{COV}(C_1, C_2)$  est donnée par :

$$\text{COV}(C_1, C_2) = V_1' \cdot [M] \cdot [R] \cdot [M] \cdot V_2$$

Comme la covariance ne tient pas compte de l'ordre on a :

$$\text{COV}(C_1, C_2) = \text{COV}(C_2, C_1) = V_2' \cdot [M] \cdot [R] \cdot [M] \cdot V_1$$

Or on sait que  $V_1$  est un vecteur propre de  $[R] \cdot [M]$  associé à la valeur propre  $\lambda_1$ . On en déduit que :

$$\text{COV}(C_1, C_2) = \lambda_1 V_2' \cdot V_1 = 0$$

Une covariance nulle entre  $C_1$  et  $C_2$  est équivalente à l'orthogonalité des vecteurs  $V_1$  et  $V_2$  :

$$\text{COV}(C_1, C_2) = 0 \Leftrightarrow V_1 \cdot V_2 = 0$$

En appliquant la même méthode pour la recherche du deuxième axe, on aura :

$$L = V_2' \cdot [M] \cdot [R] \cdot [M] \cdot V_2 - \lambda_2 (V_2' \cdot [M] \cdot V_2 - 1) - \mu (V_2' \cdot [M] \cdot V_1)$$

$$\partial L / \partial V_2 = 2 \cdot [M] \cdot [R] \cdot [M] \cdot V_2 - 2\lambda_2 \cdot [M] \cdot V_2 - \mu \cdot [M] \cdot V_1 = 0$$

En simplifiant par  $[M]$  on obtient :

$$2[R] \cdot [M] \cdot V_2 - 2\lambda_2 \cdot V_2 - \mu \cdot V_1 = 0$$

On multiplie à gauche par  $V_1' \cdot [M]$

on obtient :

$$2 \cdot V_1' \cdot [M] \cdot [R] \cdot [M] \cdot V_2 - 2\lambda_2 \cdot V_1' \cdot [M] \cdot V_2 - \mu \cdot V_1' \cdot [M] \cdot V_1 = 0$$

Or

$$V_1' \cdot [M] \cdot V_2 = 0 \quad (\text{par hypothèse})$$

Donc:

$$V_1' \cdot [M] \cdot [R] \cdot V_2 = V_2' \cdot [M] \cdot [R] \cdot [M] \cdot V_1 = \lambda_1 \cdot V_2' \cdot [M] \cdot V_1 = 0$$

Puisque le vecteur  $V_1$  est unitaire:

$$V_1' \cdot [M] \cdot V_1 = 0$$

Le multiplicateur de Lagrange  $\mu$  est donc nul, et on est ramené au problème précédent. On peut donc énoncer la définition suivante: Le second axe est défini par le vecteur  $V_2$ , vecteur propre unitaire de la matrice  $[R] \cdot [M]$  orthogonal à  $V_1$  et associé à la plus grande valeur propre  $\lambda_2$  inférieure ou égale à  $\lambda_1$ .

### II.5.2.3 Recherche des autres axes

En répétant le procédé, on détermine donc les valeurs propres et les vecteurs propres de la matrice  $[R] \cdot [M]$  pour obtenir la  $i^{\text{ème}}$  composante principale  $C_i$ .

Le vecteur propre unitaire de la matrice  $[R] \cdot [M]$  définit le  $i^{\text{ème}}$  axe orthogonal à  $(V_1, V_2, \dots, V_{i-1})$  et associé à la  $i^{\text{ème}}$  plus grande valeur propre  $\lambda_i$ . On constate que la mise en équation de ces règles aboutit aux résultats suivants :

- On appelle  $i^{\text{ème}}$  vecteur principal : le vecteur propre unitaire  $V_i$  de la matrice  $[R] \cdot [M]$  associée, qui fournit les coefficients qui pondéreront les variables initiales pour le calcul des composantes principales.
- On appelle  $i^{\text{ème}}$  axe principal, la droite engendrée par le  $i^{\text{ème}}$  vecteur principal.
- Chaque composante  $C_k$  est portée par le  $k^{\text{ème}}$  axe principal.
- La dispersion des projections des variables sur la composante  $C_k$  est mesurée par la valeur propre  $\lambda_k$ .
- Les valeurs sont rangées par ordre décroissant :  $\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_Q$ .
- Les CP sont rangées de 1 à  $Q$  dans l'ordre des valeurs propres.
- La moyenne de chaque CP est nulle.

Si on veut normer les CP (imposer à chacune d'entre elles d'avoir un écart type unité) il suffit de diviser chacune d'elle par la variance expliquée  $(\lambda)^{1/2}$  correspondante.

### II.5.3 Calcul des composantes principales

On désigne par CP, la projection du nuage de points initiale sur le nouveau système d'axes fournis par les vecteurs propres.

Notons  $[C]$  la matrice des CP

$$[C] = [A]'[X]$$

$[X]$ : Matrice des données initiales.

$[A]'$ : Matrice transposée de la matrice composée par les vecteurs propres.

### II.5.4 Conclusion

L'Analyse en Composante Principales, dans son aspect qualitative, permet d'établir deux sortes de bilan :

- Un bilan de liaisons entre les variables pouvant nous renseigner sur les points suivants:  
Quelles sont les variables qui sont liées positivement, et celles qui s'opposent (liées négativement)?

Existe-t-il une répartition en groupes des variables inter-corrélées ?

- Un bilan de ressemblance entre individus répondant aux questions suivantes :

Quels sont les individus qui se ressemblent et ceux qui diffèrent ?

Existe-t-il des groupes homogènes d'individus ?

En dehors de l'analyse descriptive citée plus haut, l'ACP ouvre les portes sur un domaine beaucoup plus intéressant qui est l'analyse opérationnelle, cette dernière traite des problèmes de prévision, d'optimisation et de simulation.

## II.6. ACP PROGRAMMATION

Dans ce paragraphe nous traiterons de la programmation de l'Analyse en Composantes Principales. Le logiciel que nous avons mis au point, est constitué d'une suite de sous programmes traitant chacun une partie bien spécifique du problème. Le modèle élaboré est très simple à utiliser et affranchit l'utilisateur du format des fichiers INPUT, vu que la lecture de ceux-ci se fait sous format libre.

Nous expliciteront alors les différentes phases du programme par un organigramme.

### II.6.1 Fichier de données INPUT

Le fichier des données doit être constituer comme suite :

```

N
P
X11 X12 ..... X1j ..... X1p
X21 X22 ..... X2j ..... X2p
.....
XN1 XN2 ..... XNj ..... XNP

```

<EOF>

N, P : N nombres entiers désignant respectivement le nombre d'individus et le nombre de variables.

X<sub>ij</sub> : Données à introduire.

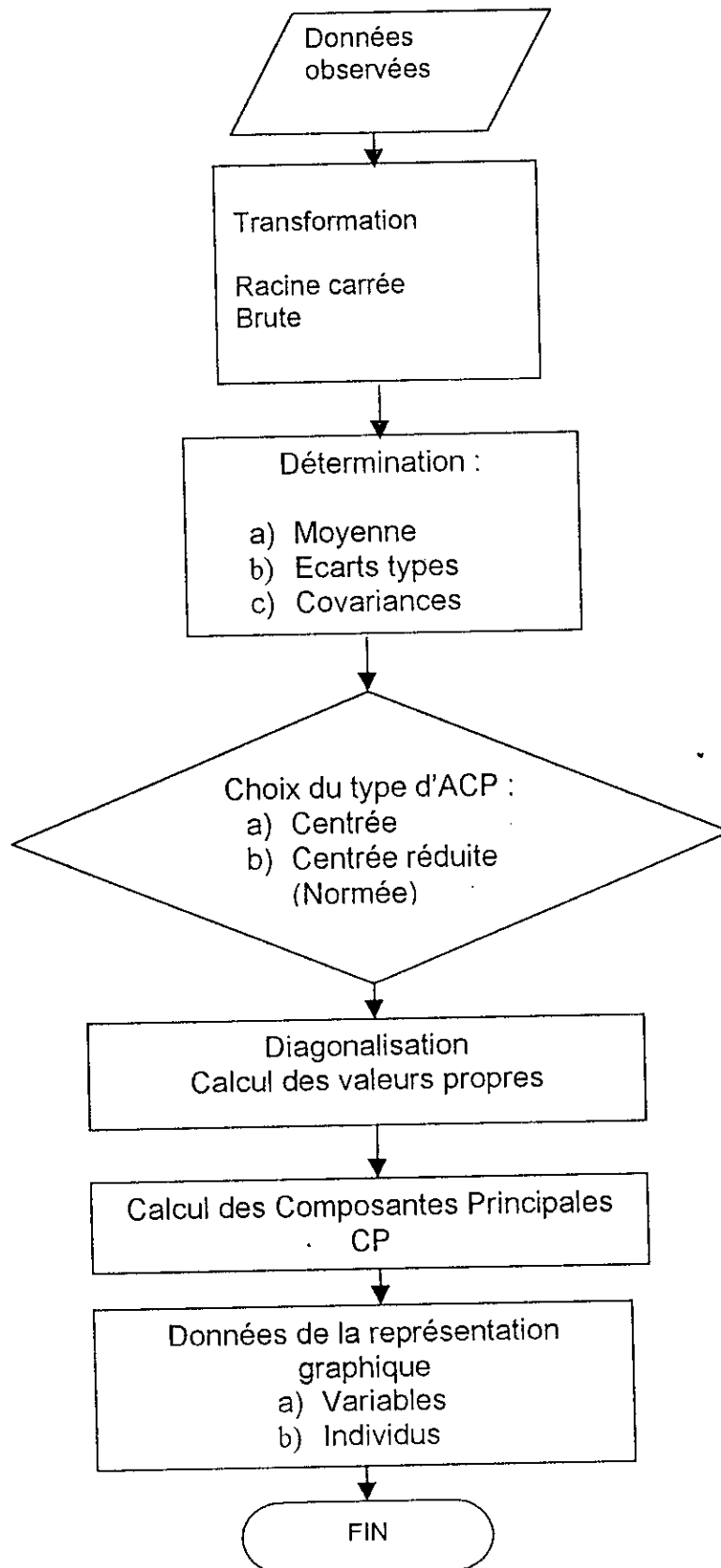
L'introduction des données se fait ligne par ligne, dont chacune d'elle représente un individu.

### II.6.2 Organigramme de calcul

Nous avons élaboré un programme, constitué de quatre sous programmes principaux (Figure II-1) effectuant les taches suivantes :

- 1- Transformation des données observées
- 2- Calcul de la matrice de covariance  $[R]$  ainsi que les paramètres statistiques
- 3- Diagonalisation de  $[R]$  (recherche des vecteurs propres)
- 4- Calcul des composantes principales

**Figure II-1: Organigramme de l'Analyse En Composantes Principales**



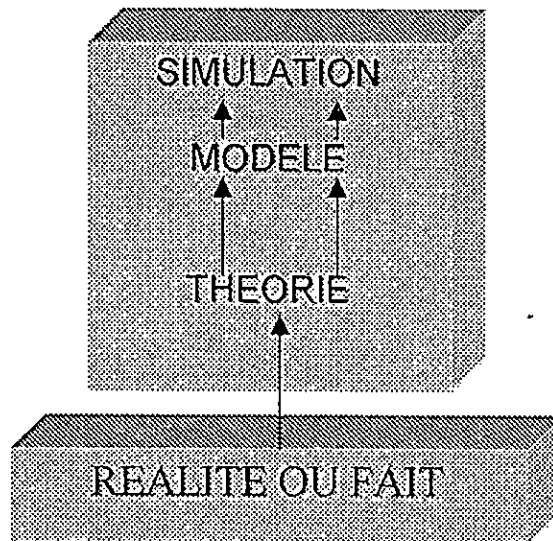
**CHAPITRE III**  
**ELABORATION DU MODELE GESTOP DE**  
**SIMULATION**

## CHAPITRE III ELABORATION DU MODÈLE GESTOP DE SIMULATION

### III.1. CONCEPT DE SIMULATION

La simulation est une technique qui ressemble à l'expérimentation scientifique utilisant des modèles dont toutes les caractéristiques sont connues.

La figure ci-dessous illustre le mécanisme de la simulation, à sa base, la figure repose sur la réalité ou les faits.



Selon les situations, il est possible de construire des modèles pour tester ou représenter une théorie. Ainsi, la simulation réside dans l'utilisation d'un modèle qui vise à identifier ou découvrir le comportement d'un procédé ou d'un système. Elle constitue finalement une approche de résolution de problèmes basée sur l'essai et l'erreur, elle aide à prendre de meilleures décisions et représente donc un excellent outil de planification.

Les méthodes actuelles de génération présentent une difficulté d'interprétation de "causes à effets", souvent attribuée au manque d'indépendance entre les variables utilisées ; or, les besoins actuels en Hydrologie nous orientent vers la recherche de nouvelles méthodes de génération pour que l'interprétation "causes à effets" soit facilitée. Il y aurait avantage à ce que ces méthodes utilisent des variables indépendantes (D. Tessier, 1983).

### III.2. MODELE DE SIMULATION

Par définition les Composantes Principales sont des combinaisons linéaires des variables initiales

$$C_i = \sum_{j=1}^P a_{ij} \cdot Y_j \quad (1)$$

$C_i$  :  $i^{\text{ème}}$  C.P.

$Y_j$  : Variable d'ordre  $j$

$a_{ij}$  : Cosinus directeur, élément de la matrice  $[A]$  des vecteurs propres.



Soit la composante principale normée  $C'_1$  tel que :

$$C'_1 = \frac{C_1}{\lambda_1^{1/2}} \quad (2)$$

$\lambda_i$ : Valeur propre correspondant à  $C_i$

En divisant l'équation (1) par  $\lambda_1^{1/2}$  on obtient :

$$C'_1 = \frac{C_1}{\lambda_1^{1/2}} = \frac{1}{\sqrt{\lambda_1}} \sum_{j=1}^p a_{1j} \cdot Y_j \quad (3)$$

Sous forme matricielle l'expression (3) devient :

$$[C'] = \lambda^{-1/2} [A]^t [Y] \quad (4)$$

Avec

$\lambda$  : Vecteur dont les composantes sont les valeurs propres.

$[A]$  : Matrice des cosinus directeurs, formée par les vecteurs propres.

$[A]^t$  : Matrice transposée de  $[A]$ .

$[Y]$  : Matrice des variables initiales centrées réduites.

Réciproquement on peut exprimer la variable  $Y_j$  en fonction des C.P. normées.

En transposant la matrice des vecteurs propres  $[A]$  on obtient :

$$[Y] = \lambda^{1/2} [A]^t [C'] \quad (5)$$

Sachant qu'on ne considère que les  $Q$  premières C.P. et qu'on ne perd pas de vue la variance non expliquée prise en compte par le terme résiduel noté  $\varepsilon_j$ .

En explicitant la formule (5) on aura :

$$Y_j = \sum_{i=1}^Q \lambda_i^{1/2} a_{ji} C'_i + \varepsilon_j \quad (6)$$

$\varepsilon_j$ : Variable résiduelle de moyenne nulle et d'écart type :

$$\sigma_{\varepsilon_j} = \sqrt{1 - \sum_{i=1}^Q a_{ji}^2 \lambda_i} \quad (7)$$

Parmi les résultats de l'ACP, le tableau des covariances des variables et des axes principaux

$$COV(Y_j, C'_i) = \sum_{i=1}^N Y_{ij} \cdot C'_{ij} \quad (8)$$

Pour déterminer la quantité  $COV(C'_i, Y_j)$ , on tire à partir de l'équation (5) la variable  $Y_j$  :

$$Y_{ij} = \sum_{l=1}^q \sqrt{\lambda_l} \cdot a_{ij} C'_{il} \quad (9)$$

Puisqu'on considère la covariance de Y avec la  $i^{\text{ième}}$  CP, on prend  $k=i$ , on obtient donc :

$$COV(Y_j, C'_i) = \sum_{i=1}^N C'_{ii} \cdot \sqrt{\lambda_i} \cdot a_{ij} C'_{ii} \quad (8)$$

$$COV(Y_j, C'_i) = \sqrt{\lambda_i} \cdot a_{ij} \cdot \sum_{i=1}^N C'_{ii} \cdot C'_{ii} \quad (9)$$

La quantité  $\sum_{i=1}^N C'_{ii} \cdot C'_{ii}$  représente la norme du vecteur  $C'_i$ , et par conséquent elle est égale à un.

Donc:

$$COV(Y_j, C'_i) = \lambda_i^{1/2} a_{ji} \quad (10)$$

En intégrant ce résultat dans l'équation (6) on obtient l'expression finale du modèle GESTOP :

$$Y_j = \sum_{l=1}^q COV(Y_j, C'_l) \cdot C'_l + \varepsilon_j \quad (11)$$

En se basant sur la structure du modèle ci-dessus, différentes possibilités de simulation sont à envisager.

On désignera dans ce qui suit par  $\tilde{Y}_j$  la variable simulée, et on remarque bien que c'est une combinaison linéaire des CP qui sont indépendantes par construction.

### III.3. METHODES DE SIMULATION

Pour simuler des grandeurs mesurant des phénomènes naturels (tels que la pluie ou le débit), il ne faut pas perdre de vue que cette notion est basée essentiellement sur le concept du hasard. Ce dernier n'a de sens scientifique que si l'on se réfère au mécanisme de génération.

Toutes les méthodes de simulation se basent sur un tirage de nombres aléatoires, la différence entre ces dernières réside dans le choix de l'échantillon sur lequel on opère, et la relation entre le paramètre à simuler et le nombre tiré au hasard.

#### III.3.1 Simulation par les fonctions de répartition

Si l'on considère la variable aléatoire X de fonction de répartition F(x), la nouvelle variable  $Y = F(x)$  (Figure III-1) est uniformément répartie sur le segment [0,1].

$$\text{Prob}(Y < y) = \text{Prob}[F(x) < y] = y$$

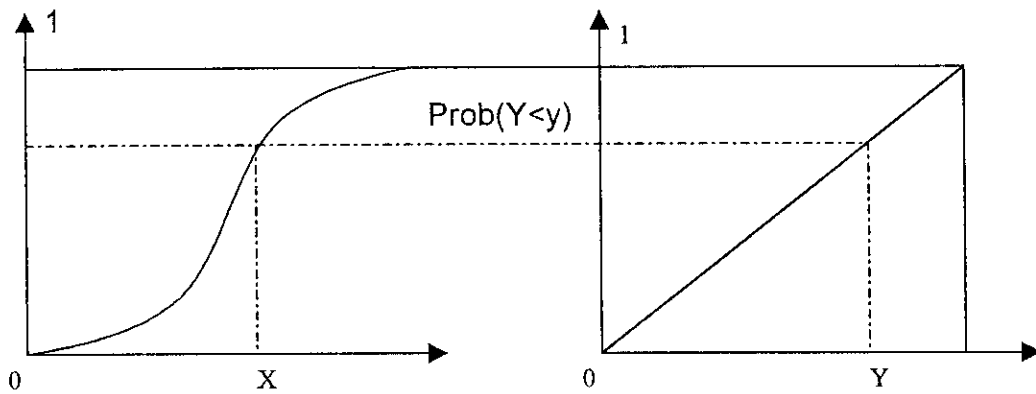


Figure III-1: Passage d'une fonction de probabilité F à une loi uniforme

Quelle que soit la variable aléatoire, on peut toujours la transformer en une variable uniformément distribuée sur le segment [0,1].

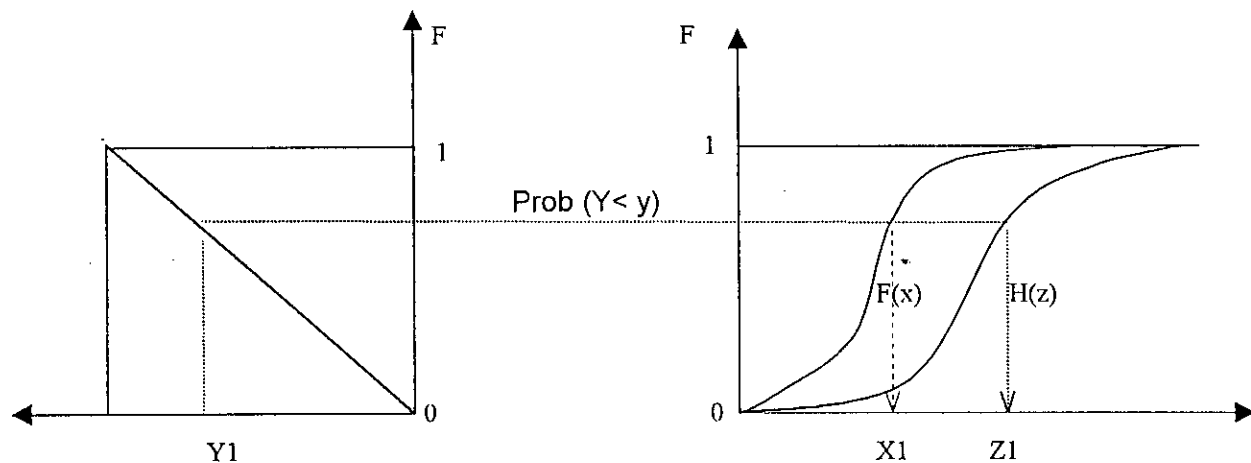


Figure III-2: Tirage aléatoire à partir de la fonction de répartition

La figure III-2 représente le tirage au hasard :

- D'une valeur de la variable aléatoire X définie par la loi de probabilité F(x).
- D'une valeur de la variable aléatoire Z définie par la loi de probabilité H(z).

On met la loi de probabilité F et/ou H dans l'une des urnes des nombres aléatoires, et on tire au hasard un nombre  $Y_1$ , qui après transformation fournit la valeur  $x_1$  et/ou  $z_1$ .

### III.3.2 Simulation par les chaînes de Markov

Un processus stochastique est un processus dont la valeur à l'instant t dépend des valeurs antérieures, celui-ci est dit Markovien, si la valeur à l'instant t ne dépend en fait que de la valeur la plus récente, c'est à dire :

$$P \{ X(t) = x / X(t_1) = x_1, X(t_2) = x_2, \dots, X(t_n) = x_n \} = x_n = P \{ X(t) = x / X(t_n) = x_n \}$$

Pour simuler une série donnée par les chaînes de MARKOV, il existe deux types de modèles : les modèles "binaires" dont le principe est basé sur l'existence de deux états (modèle choisie pour l'étude), et les modèles "multiclasses" où la série est subdivisée en plusieurs classes.

### III.3.2.1 Construction de la chaîne de Markov

On définit en premier lieu l'état 0, et l'état 1, sur lesquels on se base pour transformer la série donnée en système binaire (00111010..). Pour obtenir une telle suite, il suffit d'effectuer un comptage des éléments appartenant respectivement à l'état 0 et 1. On calcule ensuite les probabilités inconditionnelles  $P_0$  et  $P_1$ .

$$P_0 = \frac{\text{Nombre.d'éléments.appartenant.à.l'état.0}}{\text{Nombre.total.d'éléments}}$$

$$P_1 = \frac{\text{Nombre.d'éléments.appartenant.à.l'état.1}}{\text{Nombre.total.d'éléments}}$$

### III.3.2.2 Génération par les chaînes de Markov

Après avoir construit la matrice de passage d'ordre zéro (0), on détermine les fonctions de répartition des valeurs appartenant aux états 0 et 1.

La simulation se fait en générant uniformément un nombre aléatoire  $y$  dans l'intervalle  $[0,1]$ , qu'on compare aux probabilités inconditionnelles.

Sachant que :  $P_0 + P_1 = 1$

Alors :  $P_1 = 1 - P_0$

Donc si  $y$  est inférieur à  $P_0$ , la projection de  $y$  se fait sur la fonction de répartition correspondant à l'état 0, sinon elle se fera dans celle de l'état 1.

### III.3.3 Simulation par les lois d'ajustement

Les valeurs d'une série de données peuvent avoir une distribution bien spécifique qui est déterminée par l'ajustement des données historiques et vérifiées par les tests d'adéquation.

La génération dans la loi uniforme est la base de la simulation dans la plupart des lois d'ajustements statistiques.

Un tirage aléatoire, dans une loi Laplace - Gauss centrée réduite, peut être effectué selon la méthode de Box et Muller, on tire deux nombres  $U_1$  et  $U_2$  grâce à la fonction RANDOM qui effectue un tirage aléatoire dans la loi uniforme sur  $[0,1]$ .

Et on calcule :

$$Z_1 = \sqrt{-2\text{Ln}U_1} \cdot \cos(2\pi U_2)$$

$$Z_2 = \sqrt{-2\text{Ln}U_1} \cdot \sin(2\pi U_2)$$

Ainsi définies,  $Z_1$  et  $Z_2$  sont des variables aléatoires, de Laplace – Gauss, centrées réduites et indépendantes.

Si la loi d'ajustement est une normale ( $m, \sigma$ ), on peut toujours se ramener à une normale de moyenne nulle et d'écart type unité  $N(0,1)$ , et faire une transformation inverse par la suite.

$$X = m + \sigma.Z$$

Z : Variable de GAUSS

X : Variable initiale

$m, \sigma$  : Moyenne et écart type de la variable X.

### III.4. ORGANIGRAMME

Nous avons dressé l'organigramme, donné en figure III-3, afin d'illustrer les différentes étapes du procédé de simulation, dont les plus importantes sont :

ACP : Calcul des composantes principales.

Reconstitution : Calcul des données reconstituées et des résidus.

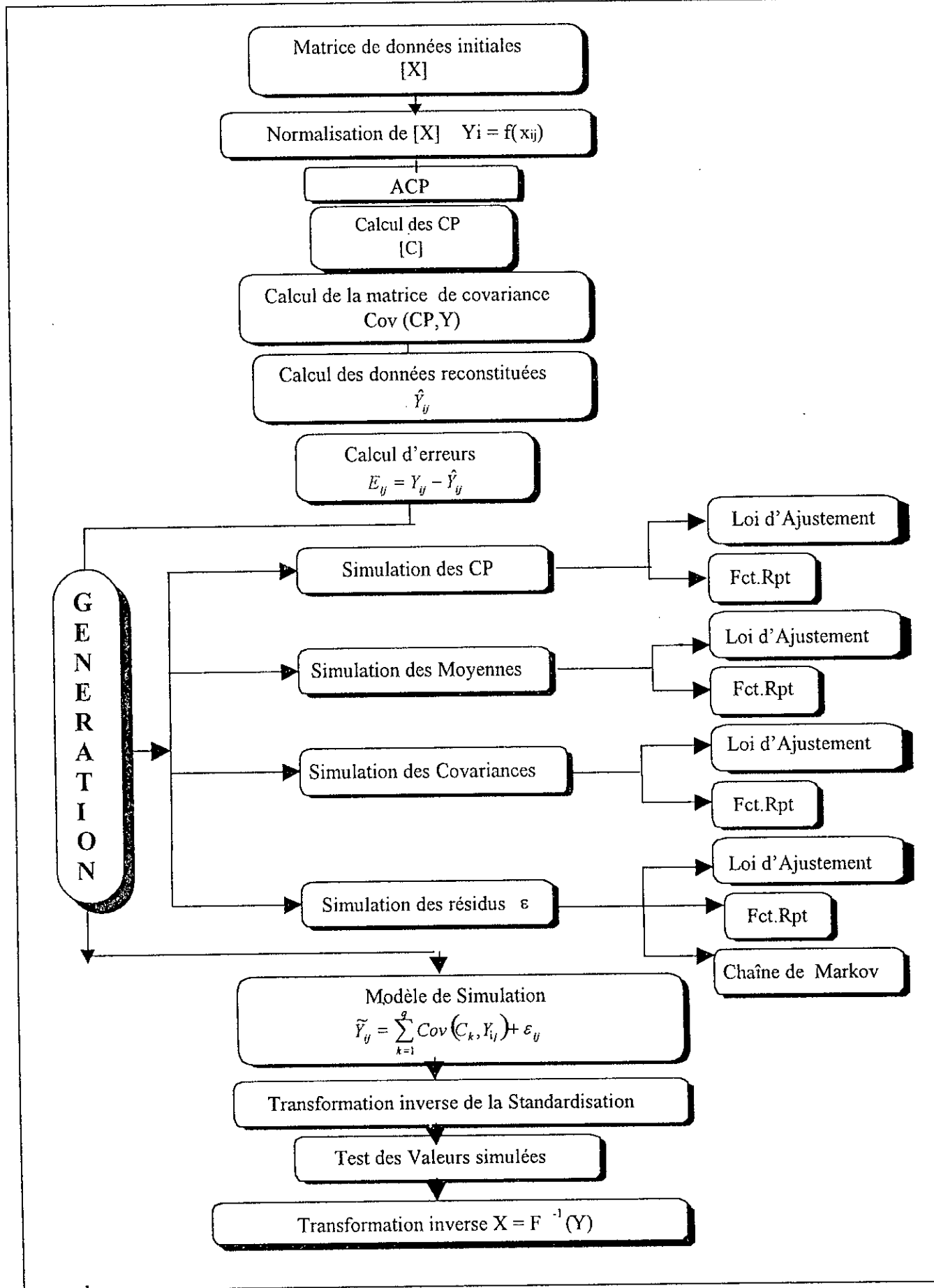
Génération : Génération des différents paramètres du modèle.

Calcul des données simulées et transformations inverses.

La seconde partie de programmation dans cette thèse est donc de transcrire l'organigramme en langage machine, pour automatiser le procédé de simulation.

Fct-Rpt : signifie une génération par la fonction de répartition.

Figure III-3: Organigramme de simulation par le modèle GESTOP



## CHAPITRE IV APPLICATION DU MODELE GESTOP AU PARAMETRES HYDROMETEOROLOGIQUES

### IV.1. CARACTERISTIQUES DE LA REGION

L'aire d'étude est située sur la côte nord-est de l'Algérie, elle englobe la région Constantinoise Seybouse Mellegue couvrant une superficie de 44 719 km<sup>2</sup>. Physiographiquement, on peut identifier trois grandes sous régions qui s'étendent approximativement d'Ouest en Est.

Le climat est de type méditerranéen, caractérisé par des hivers relativement doux et humides (de novembre à février), et par des étés chauds et secs (de mai à août).

En hivers il y a des précipitations relativement importantes, au printemps la région continue d'être affectée par des dépressions liées au mouvement des masses d'airs à travers l'Europe, cela tend à être la période des vents forts le long de la cote mais les précipitations sont moindres, l'été est caractérisé par des vents légers venant du nord et généralement un temps chaud et sec.

En octobre, des précipitations appréciables sont à noter. En hiver, les précipitations maximales coïncident avec la période des températures et d'évaporations les plus basses. Parmi l'ensemble des paramètres hydrométéorologiques, on s'est intéressé aux trois variables prépondérantes à savoir l'évapotranspiration, les précipitations et les apports.

#### IV.1.1 L'Evapotranspiration

La méthode de Penman modifiée (Doorenbos et Pruitt, 1977) a été utilisée pour calculer l'ETP ; pour chaque mois au pas de temps journalier en utilisant la température, l'humidité relative, l'insolation et la vitesse du vent en plus d'autres paramètres caractérisant le milieu physique.

#### IV.1.2 Les Précipitations

La station pluviométrique se trouve à l'intérieur de l'étroite plaine côtière, les terres s'élèvent jusqu'à 2 000 m d'altitude dans les montagnes de la petite Kabylie, et arrive à 1200 m dans les monts de la Medjerdah à l'Est.

La répartition spatiale et saisonnière des précipitations peut être attribuée à trois causes principales ; les dépressions cycliques d'hiver, la topographie et la convection, ces causes combinées avec la grande variété des caractéristiques du relief, provoquent une répartition complexe des saisons pluvieuses. Le régime pluviométrique est assez bien connu grâce à l'existence d'environ 219 postes de mesures, dont la plupart ont plus de 20 années d'observations.

### IV.1.3 Les Apports

La répartition saisonnière des apports est étroitement liée à celle des précipitations, avec un déphasage dû à la nature du sol, la géologie et la topographie, le régime des oueds dans cette région est très variable allant de torrentiel en période de crue à fluviale après une période de pluies abondantes, en été la majorité des oueds est quasiment sèche.



## IV.2. METHODOLOGIE

### IV.2.1 Fichier de données

Les données, concernant les paramètres à étudier, sont disponibles au pas de temps journalier, étalées sur un certain nombre d'années. En vue de construire la matrice d'entrée du modèle, on range les données dans un tableau vertical de P colonnes et N lignes.

### IV.2.2 Calcul des C.P.

#### IV.2.2.1 Calcul des paramètres statistiques

Considérons la matrice [X] et calculons :

- La moyenne et l'écart type des P variables

$$\bar{X}_j = \frac{1}{N} \sum_{i=1}^N X_{ij}$$

$$\sigma_j^2 = \frac{1}{N} \sum (X_{ij} - \bar{X}_j)^2$$

- La covariance inter variables

$$\text{COV}(X_j, X_k) = \frac{1}{N} \sum_{i=1}^N (X_{ij} - \bar{X}_j) \cdot (X_{ik} - \bar{X}_k)$$

#### IV.2.2.2 Application de l'ACP

l'Analyse en composantes principales possède trois variantes de base qui sont:

ACP brute  
 ACP centrée  
 ACP centrée réduite

L'ACP peut être effectuée à partir d'une matrice des covariances ou des corrélations, si la matrice des données utilisées est constituée de valeurs de différentes unités avec la même échelle de mesure l'utilisation de la matrice de corrélation (autrement dit une ACP normée) s'avère avantageuse.

### IV.2.3 Reconstitution

#### IV.2.3.1 Choix du nombre de C.P.

Le choix du nombre de CP à prendre en compte pour la simulation ne peut se faire d'une manière automatique, en se basant uniquement sur le taux de variance, pour cela on procédera à la reconstitution avec différents seuils d'informations en commençant par 100% ce qui correspond au nombre total de CP jusqu'à obtenir une reconstitution satisfaisante avec un nombre minime de CP qu'on notera Q.

#### IV.2.3.2 Calcul des covariances (Variable, C.P.)

Soit [Y] la matrice des données historiques transformées suivant le type d'ACP, la covariance est calculée de la manière suivante:

$$COV(Y_j, C'_k) = \frac{1}{N} \sum_{i=1}^N (Y_{ij} - \bar{Y}_j) \cdot (C'_{ik} - \bar{C}'_k)$$

Comme  $Y_j$  est une variable centrée :  $\bar{Y}_j = 0$   
 et comme :  $\bar{C}'_k = 0$  (par construction)

alors :

$$COV(Y_j, C'_k) = \sum_{i=1}^N Y_{ij} \cdot C'_{ik}$$

#### IV.2.3.3 Calcul des données reconstituées

Le modèle utilisé pour la reconstitution des données observées est donné par :

$$\hat{Y}_{ij} = \sum_{k=1}^Q COV(Y_{jk}, C'_{ik}) C'_{ij}$$

$Y_{jk}$  : Variable historique transformée suivant le type d'A.C.P.

$C'_{ik}$  : Composante Principale normée

Pour juger la qualité de la reconstitution, on est dans l'obligation de superposer les données historiques et les données reconstituées, donc il faut faire subir à la matrice  $[\hat{Y}]$  la transformation inverse ; il arrive qu'après reconstitution on obtienne des valeurs négatives dans ce cas on se doit de transformer les données brutes.

#### IV.2.3.4 Choix du type de transformations

Le modèle utilisé sur des données brutes peut donner lieu à des valeurs négatives, pour cela il faut faire subir une transformation de normalisation appropriée, trois types de transformations sont étudiées:

Log.  
Log translaté.  
Racine carrée.

Le choix de la transformation se fait en confrontant les données reconstituées avec les trois variantes aux données historiques après transformation inverse.  
Une fois les données transformées, il est impératif de reprendre le calcul des C.P. avec la nouvelle matrice des données.

#### IV.2.3.5 Calcul des résidus

Les résidus obtenus lors de la reconstitution sont quantifiés, en déterminant l'écart entre valeur initiale et valeur reconstituée, le résultat sera une matrice de résidus [E] dont les éléments  $\varepsilon_{ij}$  sont calculés comme suit :

$$\varepsilon_{ij} = X_{ij} - \hat{X}_{ij}$$

### IV.2.4 Simulation

#### IV.2.4.1 Ajustement des paramètres

Le modèle de simulation s'articule autour de trois principaux paramètres qui sont :

Les C.P.  
Les résidus  
Les covariances (variable, C.P.)

Pour cela il est utile de savoir à quelle loi statistique s'ajuste chacun de ces paramètres. Les C.P. par construction, (vu qu'ils représentent la projection d'un nuage de points dans un repère orthonormé), suivent une loi normale de moyenne nulle et d'écart type unité, les résidus de part leur origine, suivent une loi normale de moyenne nulle et d'écart type  $\sigma_\varepsilon$ , l'ajustement des covariances se fera cas par cas au niveau de chaque variable hydrométéorologique étudiée.

#### IV.2.4.2 Génération des CP, Résidus et Covariances

Il est plus simple de générer une variable dans une loi statistique, si cette dernière s'ajuste à la variable en question, pour cela les CP et les résidus seront simulés dans la loi normale, si l'ajustement est jugé acceptable, sinon on optera pour la génération à partir de la fonction de répartition.

#### IV.2.4.3 Calcul des données simulées $\tilde{Y}_{ij}$

Les paramètres générés servent d'entrée au modèle de simulation

$$\tilde{Y}_{ij} = \sum_{k=1}^Q COV(Y_{kj}, C'_{ik}) C'_{ij} + \varepsilon_{ij}$$

#### IV.2.4.4 Transformation inverse

Une fois la matrice  $[\tilde{Y}]$  obtenue, on se doit d'opérer les transformations inverses afin de retrouver des données comparables à celles observées.

La première transformation est celle due au type d'A.C.P. :

$$\begin{aligned} \text{si A.C.P. centrée} & : & \tilde{Y}_{ij} = \tilde{X}_{ij} - \bar{X}_j \Rightarrow \tilde{X}_{ij} = \tilde{Y}_{ij} + \bar{X}_j \\ \text{si A.C.P. centrée réduite} & : & \tilde{Y}_{ij} = \frac{(\tilde{X}_{ij} - \bar{X}_j)}{\sigma_j} \Rightarrow \tilde{X}_{ij} = \sigma_j + \bar{X}_j \cdot \tilde{Y}_{ij} \end{aligned}$$

En vue d'introduire le maximum de variabilité dans les séries simulées, on génère, (en plus des CP et des covariances,) la moyenne et l'écart type.

La transformation inverse s'écrira donc comme suit:

$$\begin{aligned} \text{si ACP centrée:} & & \tilde{Y}_{ij} = \tilde{X}_{ij} - \bar{\tilde{X}}_j \Rightarrow \tilde{X}_{ij} = \tilde{Y}_{ij} + \bar{\tilde{X}}_j \\ \text{si ACP centrée réduite:} & & \tilde{Y}_{ij} = \frac{(\tilde{X}_{ij} - \bar{\tilde{X}}_j)}{\tilde{\sigma}_j} \Rightarrow \tilde{X}_{ij} = \tilde{\sigma}_j \tilde{Y}_{ij} + \bar{\tilde{X}}_j \end{aligned}$$

Où  $\bar{\tilde{X}}_j$  et  $\tilde{\sigma}_j$  : Sont respectivement la moyenne et l'écart type simulés d'ordre j

La seconde transformation est celle due à la normalisation des données historiques:

$$\begin{aligned} \tilde{X}_{ij} = \tilde{X}_{ij}^2 & \quad \text{si :} & f(X_{ij}) = \sqrt{X_{ij}} \\ \tilde{X}_{ij} = e^{\tilde{X}_{ij}} + \alpha & \quad \text{si :} & f(X_{ij}) = \log(X_{ij} - \alpha) \\ \tilde{X}_{ij} = e^{\tilde{X}_{ij}} & \quad \text{si :} & f(X_{ij}) = \log(X_{ij}) \end{aligned}$$

#### IV.2.4.5 Traitement des résultats de Simulations

On fait subir aux séries synthétiques différents tests afin de décider de l'aptitude du modèle à simuler le phénomène hydrométéorologique en question.

Les tests sur lesquels on se base pour examiner les résultats sont :

Simulation au pas de temps mensuel :

- Moyennes et écart types des totaux mensuels interannuels.
- Moyennes et écart types des totaux annuels.
- Moyennes et écart types mensuels.

Simulation au pas de temps journalier :

- Moyennes et écart types des totaux mensuels.
- Moyennes et écart types des totaux journaliers interannuels.
- Moyennes et écart types mensuels interannuels.
- Ajustement des valeurs annuelles.
- Ajustement des maxima journaliers annuels.

Simulation au pas de temps horaire :

- Moyennes et écart types des valeurs journalières.
- Reproduction des valeurs maximales.

### IV.3. SIMULATION DES PARAMETRES HYDROMETEOROLOGIQUES AU PAS DE TEMPS MENSUEL

#### IV.3.1 Simulation des ETP

##### IV.3.1.1 Données utilisées

La station hydrométéorologique d'Annaba dispose des données mensuelles d'ETP, étalées sur 22 ans, ce qui nous amène à considérer une matrice de dimension [22x12].

L'évolution chronologique mensuelle des ETP au cours des 22 années relevées, (Figure IV-1) montre nettement l'aspect cyclique du phénomène ETP.

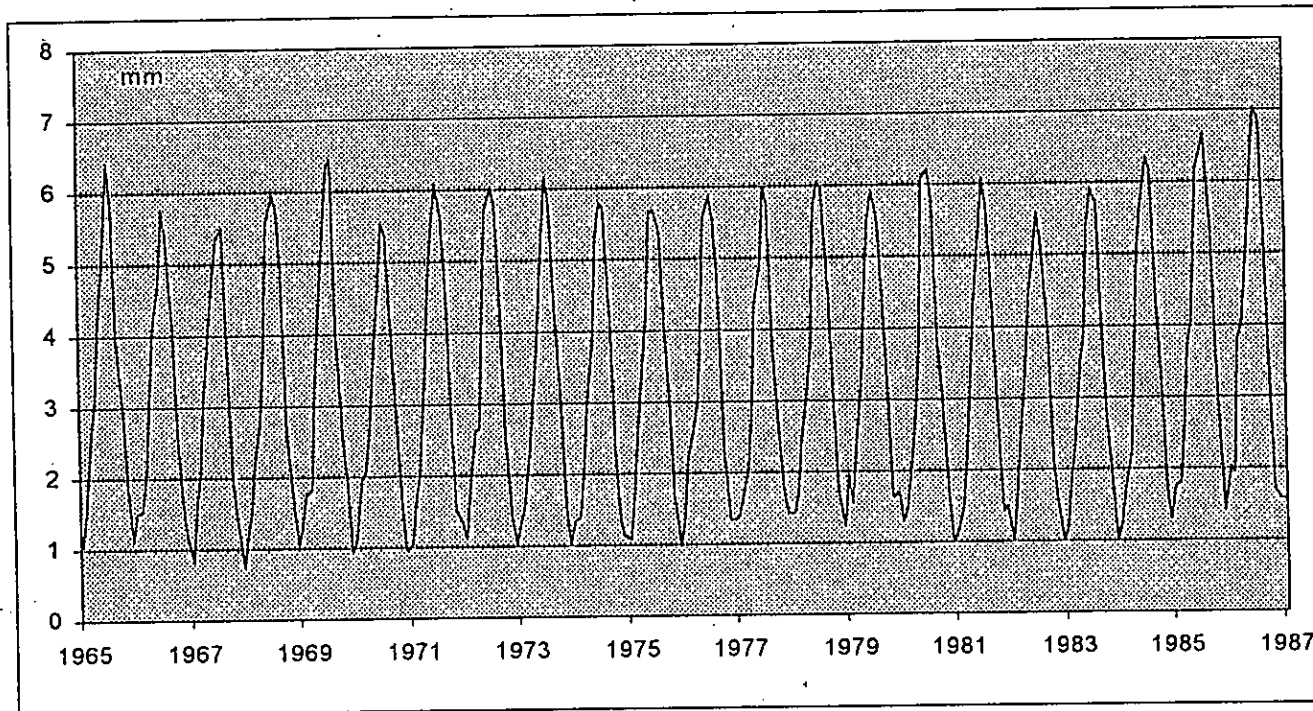


Figure IV-1: Evolution mensuelle des ETP (1965 - 1985)

##### IV.3.1.2 A.C.P.

###### IV.3.1.2.1 Calcul des paramètres statistiques

Les paramètres statistiques calculés sur la série d'ETP sont résumés dans le tableau ci-après.

Tableau IV.3-1: Paramètres statistiques de la série d'ETP

	JAN	FEV	MAR	AVR	MAI	JUN	JUL	AOÛ	SEP	OCT	NOV	DEC
Moyenne (mm)	1,3	1,7	2,3	3,2	4,2	5,4	6,0	5,8	4,3	2,7	1,6	1,2
Ecart type (mm)	0,3	0,2	0,4	0,4	0,4	0,4	0,4	0,4	0,5	0,4	0,2	0,2

On note que les mois, juin juillet et août, présentent les plus fortes moyennes d'évapotranspiration, en raison de l'étroite corrélation qui existe entre l'ETP et la température, cependant les écarts types sont relativement faibles, ceci est du à la nature cyclique du phénomène étudié a sa faible variabilité.

#### IV.3.1.2.2 Calcul des CP

On effectue une ACP normée sur les données brutes. Les CP obtenues sont représentées dans le tableau IV.3.1-2. On remarque que la première CP approche les 50% de variance expliquée, alors que la seconde est nettement inférieure avec environ 12% seulement.

Tableau IV.3-2: CP calculées sur la série d'ETP (ACP normée)

N°	CP	
	Valeurs (%)	Cumul (%)
1	48,2	48,2
2	11,8	59,9
3	9,0	68,9
4	7,9	76,8
5	6,4	83,2
6	4,7	88,0
7	4,4	92,4
8	2,7	95,2
9	2,3	97,4
10	1,3	98,7
11	0,9	99,6
12	0,4	100

### IV.3.1.3 Reconstitution

#### IV.3.1.3.1 Choix du nombre de CP

Avec une CP, on a plus de 48% de la variance expliquée totale. On a considéré 6 CP pour atteindre le seuil de 88% d'informations.

#### IV.3.1.3.2 Choix du type de transformation

La superposition des résultats de la reconstitution (Figure IV.2), n'avantage aucune des deux variantes, elles reconstituent toutes aussi bien l'historique. Donc l'ACP normée peut être effectuée sur des données brutes sans influencer la qualité de la reconstitution, et sans avoir des valeurs reconstituées négatives.

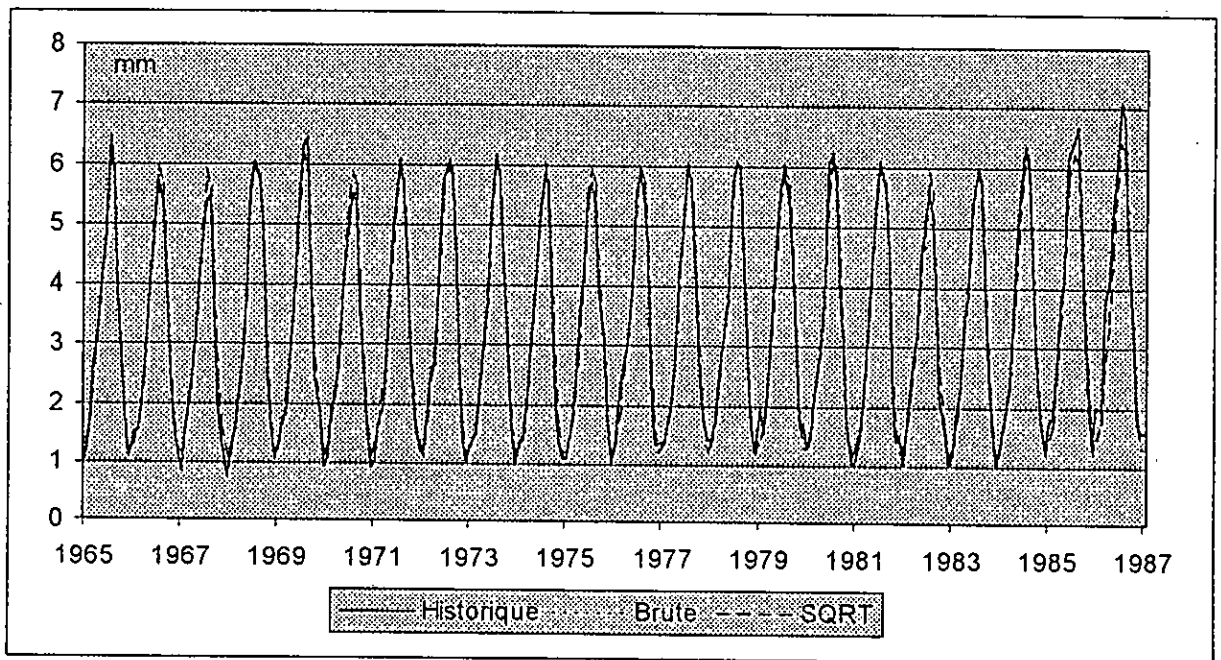


Figure IV-2: Reconstitution de l'ETP (ACP normée)

### IV.3.1.4 Simulation

#### IV.3.1.4.1 Ajustement des paramètres

L'ajustement des différents paramètres à la loi normale est donné en Annexe A.1 (1).

Les CP: Pour les CP 2, 3 et 4 la loi normale donne un bon ajustement avec un seuil de confiance de 95%, alors que pour la première CP, on remarque que l'allure des points calculés n'est pas linéaire, elle présente une zone concave qui sort légèrement de l'intervalle et une extrémité au-dessus de la limite supérieure.



Les Covariances: L'ajustement à la loi normale donne de bons résultats, les points calculés appartiennent à l'intervalle de confiance. Cependant on observe deux points (le premier appartenant à COV1 le second à COV2) situés à la limite de l'intervalle.

Les résidus: l'ajustement de la loi normale aux résidus est assez satisfaisant, on observe néanmoins, quelques points qui sortent de l'intervalle de confiance.

#### *IV.3.1.4.2 Génération des paramètres*

On maintient la génération dans la loi normale et ce pour l'ensemble des paramètres. Les paramètres générés sont classés puis portés sur leurs intervalles de confiance respectifs. Les résultats sont donnés en annexe A.1(2).

Les CP: On obtient des séries qui appartiennent à l'intervalle de confiance, on observe pour la première CP une allure plus ou moins linéaire par rapport à la série calculée, pour les trois autres CP, la génération est bonne. On note l'apparition de quelques valeurs générées à l'extérieure de l'intervalle, ceci permet de rattraper les erreurs d'ajustement des CP calculées.

Les Covariances: On remarque que la génération dans la loi normale d'écart type  $\sigma$ , donne des séries qui appartiennent à l'intervalle de confiance, les séries générées sont jugées bonnes, car avec un seuil inférieure à 95%, on peut avoir toutes les séries à l'intérieur de l'intervalle de confiance.

Les résidus: On remarque que la génération dans la loi normale de moyenne  $\mu$  et d'écart type  $\sigma$ , offre une gamme de séries générées un peu plus large que le domaine de confiance. Cependant la génération est jugés satisfaisante.

#### *IV.3.1.4.3 Résultats des simulations*

Les résultats des tests (Annexe A.2) montrent une bonne approche du phénomène en terme de moyenne (N° 1, 2 et 5) mais les écarts types des totaux mensuels interannuels ainsi que les écarts types mensuels sont au-dessous de la limite inférieure de l'intervalle de confiance (N° 4 et 6).

On se propose de transformer les données historiques par passage à la racine carré. Les nouvelles CP sont données par le tableau suivant.

Tableau IV.3-3: CP calculées sur la série d'ETP (ACP normée) données transformées (SQRT)

N°	CP	
	Valeurs (%)	Cumul (%)
1	47,4	47,4
2	11,6	59,0
3	9,0	68,0
4	8,2	76,2
5	6,6	82,9
6	4,8	87,7
7	4,5	92,2
8	2,8	95,0
9	2,3	97,3
10	1,3	98,7
11	0,9	99,6
12	0,4	100

Les résultats des tests, comme le montre l'Annexe A.3 se sont nettement améliorés, les écarts types se rangent dans l'intervalle de confiance (N° 4 et 6).

Donc la simulation de l'ETP requiert, elle aussi, la transformation des données initiales, même si la reconstitution avec les données brutes est parfaite.

### IV.3.2 Simulation des apports

#### IV.3.2.1 Données utilisées

Les données recueillies à la station Cheffia sont des mesures journalières de débits à l'amont du barrage, à partir desquelles on a calculé les moyennes mensuelles ( $\text{hm}^3/\text{j}$ ) sur un historique de 12 ans. La matrice de données qui servira d'entrée au modèle de simulation sera donc de dimensions  $[12 \times 12]$ .

La Figure IV-3 montre la nature du phénomène étudié, en examinant l'évolution mensuelle au cours des 12 années relevées (1978 - 1990) on note une irrégularité dans la répartition temporelle des apports, alors que les années 78 à 82 semblent des années pluvieuses les années allant de 82 à 88 représentent six années de sécheresse.

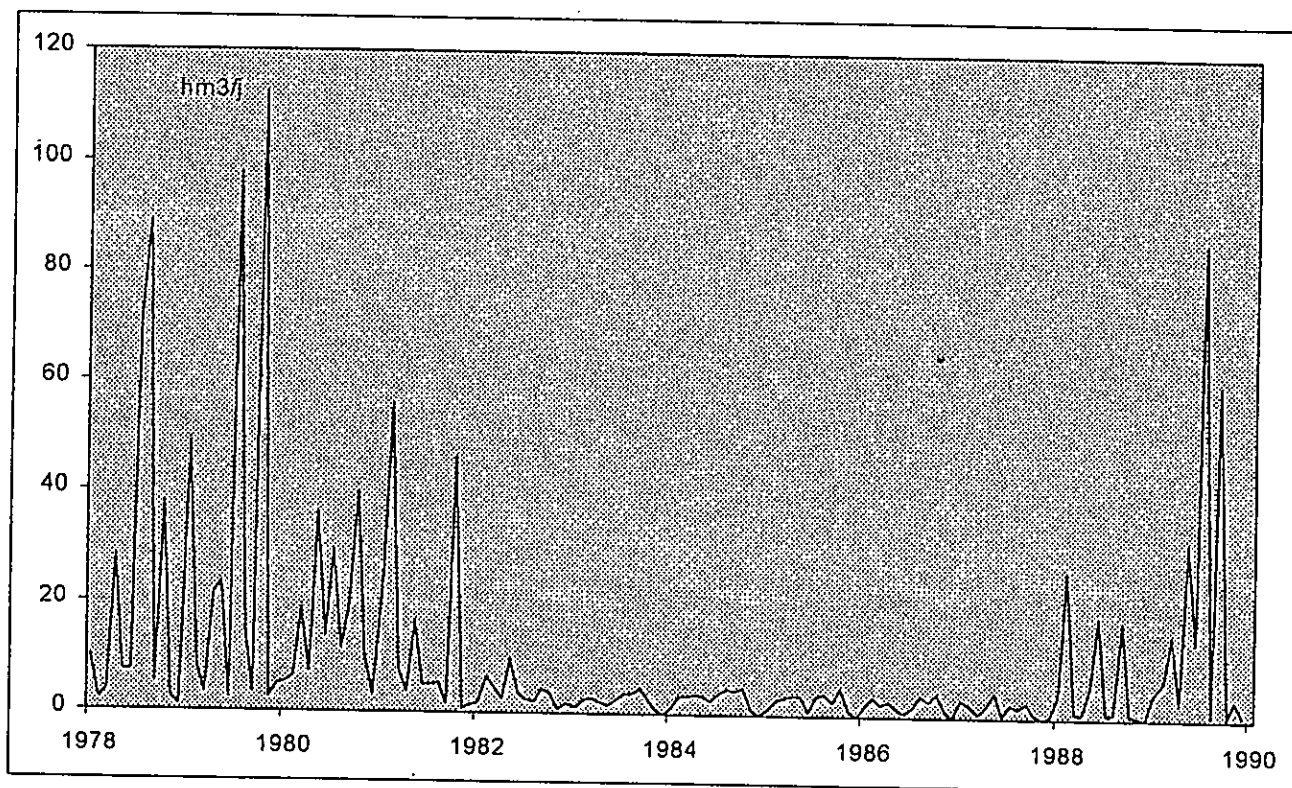


Figure IV-3: Evolution mensuelle des Débits - Station Cheffia –

## IV.3.2.2 A.C.P.

## IV.3.2.2.1 Calcul des paramètres statistiques

Les paramètres statistiques calculés sur la série de débits sont résumés dans le tableau ci-après.

Tableau IV.3-4 : Paramètres statistiques des Débits mensuels

	JAN	FEV	MAR	AVR	MAI	JUN	JUL	AOU	SEP	OCT	NOV	DEC
Moyenne (m <sup>3</sup> /j)	22,2	28,7	16,8	14,6	3,7	2,4	3,1	2,6	2,3	1,8	6,5	19,0
Ecart type (m <sup>3</sup> /j)	29,7	38,3	12,4	18,6	2,5	1,4	1,5	1,6	1,3	1,4	9,2	27,6

## IV.3.2.2.2 Calcul des CP

Une ACP normée est effectuée sur la matrice [X], les CP résultant sont illustrées dans le tableau suivant:

Tableau IV.3-5: CP calculées sur la série d'apports de la Cheffia - ACP normée –  
- Données brutes –

N°	CP	
	Valeurs (%)	Cumul (%)
1	40,54	40,5
2	18,55	59,1
3	14,29	73,4
4	11,56	84,9
5	6,20	91,1
6	5,29	96,4
7	1,97	98,4
8	0,90	99,3
9	0,45	99,7
10	0,24	100
11	0,02	100
12	0,00	100

## IV.3.2.3 Reconstitution

## IV.3.2.3.1 Choix du nombre de CP

L'examen du Tableau IV.3.2-5 révèle que la première CP s'octroie à elle seule plus de 40% de l'information. En prenant 4 CP on totalise environ 85% de la variance expliquée globale.

#### IV.3.2.3.2 Choix du type de transformation

La reconstitution des valeurs historiques brutes donne lieu à des débits reconstitués négatifs erronés, donc une transformation à priori s'impose, la Figure IV.4 montre que la transformation racine carrée permet une bonne reconstitution avec environ 64% de variance expliquée, on remarque que la période sèche est reconstituée invariablement par les données brutes que par les données transformées. On choisit donc la transformation racine carrée pour la simulation des apports.

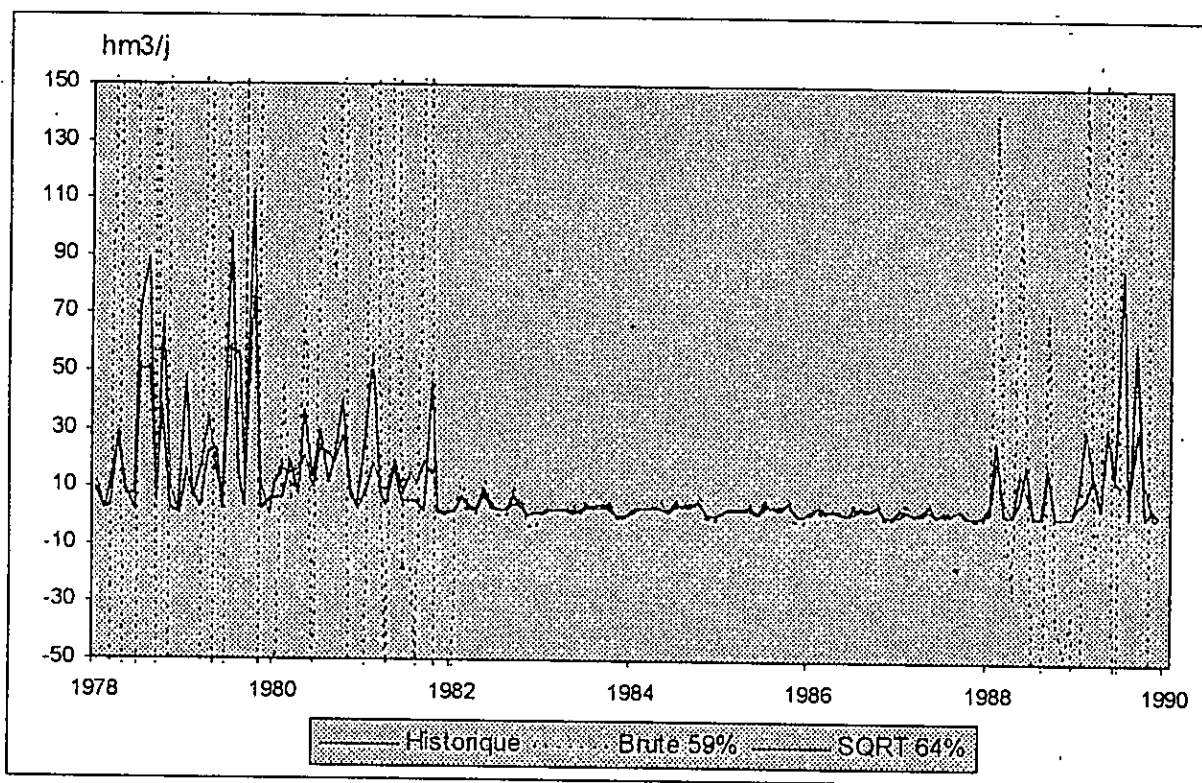


Figure IV-4: Reconstitution des Débits – ACP normée -

#### IV.3.2.3.3 Calcul des nouvelles CP

Une fois les données transformées, on recalcule les nouvelles CP, dont les valeurs sont reportées sur le tableau IV.3-6. Le nombre de CP à prendre en considération ne change pas (on maintient 4 CP), avec une augmentation de moins de 3%, par rapport au cas précédent.

Tableau IV.3-6: CP calculées sur les données de débits transformées (ACP normée)

N°	CP	
	Valeur (%)	Cumul (%)
1	45,33	45,33
2	18,72	64,06
3	12,36	76,41
4	9,72	86,14
5	6,33	92,47
6	4,02	96,49
7	2,07	98,56
8	0,78	99,34
9	0,38	99,72
10	0,21	99,94
11	0,06	100
12	0,00	100

#### IV.3.2.4 Simulation

##### IV.3.2.4.1 Ajustement des paramètres

L'ajustement des différents paramètres, Annexe A.4 (1), a mené aux constatations suivantes :

Les CP: On observe que les points calculés ne présentent pas une allure linéaire, mais restent inscrits dans le domaine de confiance. On note pour la CP une, que la valeur extrême est en dehors de l'intervalle de confiance, même remarque pour la CP 4.

Les Covariances: La loi normale s'ajuste bien aux covariances calculées, les quatre vecteurs de covariances ne présentent pas de points en dehors de l'intervalle de confiance.

Les résidus: La loi normale donne de bon résultats avec les résidus calculés, On ne note pas de points à l'extérieur de l'intervalle de confiance.

##### IV.3.2.4.2 Génération des paramètres

L'ensemble des paramètres est généré dans la loi normale. Les résultats obtenus sont donnés en Annexe A.4 (2).

Les CP: Les séries générées dans la loi normale  $N(0,1)$ , appartiennent toutes à l'intervalle de confiance.

Les Covariances: Les séries générées dans la loi normale  $N(\mu, \sigma)$ , sont en générale, à l'intérieure de l'intervalle de confiance.

Les résidus : Les séries générées sont acceptables. Leurs aspect est dû au fait qu'il s'agit de la génération de valeurs très faibles ( $\mu \approx 0, \sigma \approx 0.2$ ).

#### IV.3.2.4.3 Résultats des simulations

L'ensemble des tests est donné en Annexe A.5. On remarque que tous les tests sont vérifiés aussi bien en terme de moyenne (N° 1, 2 et 5) que d'écart types (N° 3, 4 et 6). Les différents paramètres sont inscrits dans leurs intervalles de confiance respectifs.

### IV.3.3 Simulation des pluies

#### IV.3.3.1 Données utilisées

Nous disposons de données pluviométriques journalières de la station d'Annaba sur un historique de 36 années, à partir duquel nous déterminons le cumul mensuel pour obtenir une matrice  $[X]$  de dimension  $[36 \times 12]$ .

L'évolution mensuelle des pluies au cours des 36 années relevées (Figure IV-5), révèle la nature purement aléatoire du phénomène étudié.

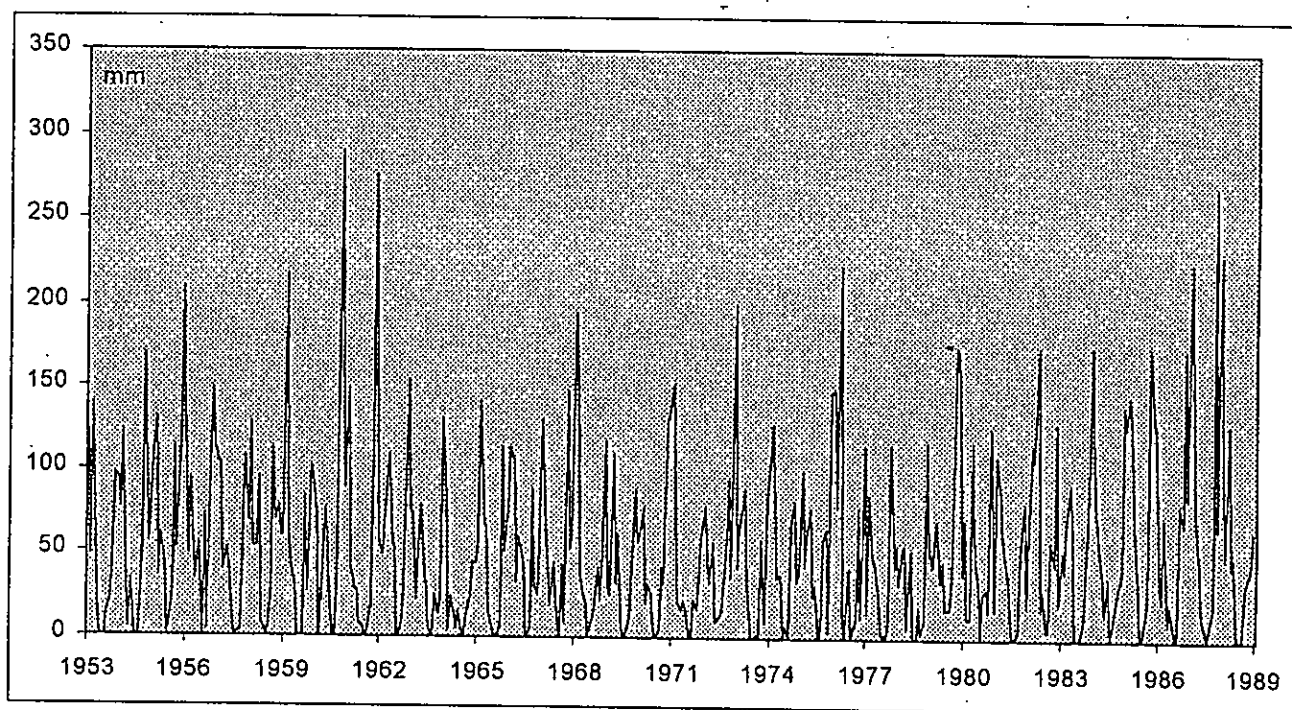


Figure IV-5: Evolution mensuelle de la pluviométrie (1953 - 1988)

#### IV.3.3.2 A.C.P.

##### IV.3.3.2.1 Calcul des paramètres statistiques

Les paramètres statistiques calculés sur la série de pluies sont résumés dans le tableau suivant.

Tableau IV.3-7 : Paramètres statistiques de la série de Pluies d'Annaba (1953 - 1986 )

	JAN	FEV	MAR	AVR	MAI	JUN	JUL	AOU	SEP	OCT	NOV	DEC
Moyenne (mm)	93,5	76,2	65,8	52,8	31,8	15,2	2,1	10,2	35,4	80,6	86,2	92,2
Ecart type (mm)	47,1	49,4	42,8	36,1	17,9	16,1	4,3	15,0	31,4	67,5	59,5	53,1

#### IV.3.3.2 Calcul des CP

On effectue sur la matrice [X] une ACP normée. On peut voir la répartition des CP dans le Tableau IV.3.3-8. La différence observée entre les quatre premières CP n'est pas importantes, la première CP ne représente que 17.4% de l'information totale, tandis que les trois CP suivantes totalise près de 37%.

Tableau IV.3-8: Les CP calculées avec ACP normée sur la série de pluies

N°	CP	
	Valeurs (%)	Cumul (%)
1	17,42	17,42
2	13,69	31,11
3	12,93	44,05
4	10,60	54,64
5	8,80	63,44
6	7,44	70,88
7	7,11	77,99
8	6,29	84,28
9	5,10	89,39
10	4,37	93,76
11	3,26	97,03
12	2,97	100

#### IV.3.3.3 Reconstitution

##### IV.3.3.3.1 Choix du nombre de CP

La répartition de la variance expliquée est spécifique au phénomène pluie. On observe une diminution graduelle du pourcentage d'information. Avec une CP on ne peut avoir que 17% de l'information total, on a fixé le seuil à 44% ce qui revient à considérer 3 CP.

##### IV.3.3.3.2 Choix du type de transformation

Le modèle de reconstitution donne lieu à des valeurs de pluies négatives, il est donc impératif de procéder à la transformation des données initiales. On remarque, à partir de la figure IV-6, que l'utilisation des données brutes, fournit une reconstitution erronée, puisqu'elle surestime certaines valeurs, alors que la transformation racine carré permet de donner une bonne reconstitution.



On a choisi donc la transformation racine carré.

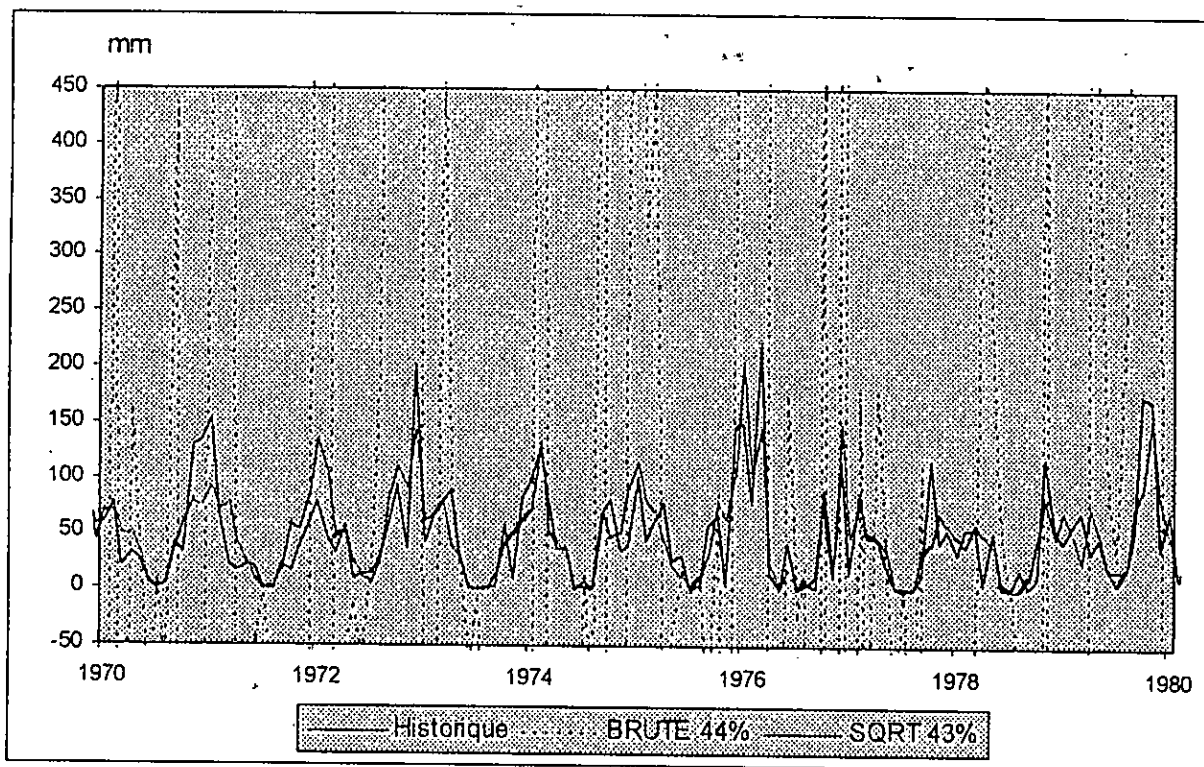


Figure IV-6: Reconstitution de la série de pluies – ACP normée –

#### IV.3.3.3.3 Calcul des nouvelles CP

Les nouvelles CP calculées après transformation sont reportées dans le tableau IV.3.3-9. Le nombre de CP à prendre en considération est de 4, ce qui signifie une augmentation de 9%. Par rapport au cas précédent.

Tableau IV.3-9: CP calculées sur la série de pluie - ACP normée - Données transformées-

N°	CP	
	Valeurs (%)	Cumul (%)
1	16,90	16,90
2	13,45	30,35
3	12,16	42,51
4	10,39	52,90
5	9,4	62,34
6	7,69	70,03
7	7,39	77,42
8	6,98	84,39
9	5,35	89,74
10	4,28	94,02
11	3,10	97,12
12	2,88	100

#### IV.3.3.4 Simulation

##### IV.3.3.4.1 Ajustement des paramètres

L'ajustement des CP, résidus ainsi que des covariances, est donné en Annexe A.6 (1).

Les CP: Les quatre composantes présentent un bon ajustement à la loi normale.

Les covariances: Les covariances s'ajustent à la loi normale avec un seuil de 95% de confiance.

Les résidus: Les résidus s'ajustent bien à la loi normale

##### IV.3.3.4.2 Génération des paramètres

On maintient la génération dans la loi normale et ceci pour l'ensemble des paramètres.

Les résultats de la génération sont donnés en Annexe A.6 (2).

Les CP: Les séries générées appartiennent à l'intervalle de confiance

Les covariances: la génération est acceptable dans l'ensemble, on remarque pour la troisième covariance une sous estimation des valeurs extrêmes, si on revient à l'ajustement de la troisième covariance calculé, on remarque que la dernière valeur est trop éloignée par rapport à la valeur antécédente, est qu'elle est à la limite de l'intervalle, le processus de génération n'a pas pu reproduire cette valeur extrême.

Les résidus: dans ce cas ils sont très faibles ( $\mu \approx 0, \sigma \approx 0.1$ ).

##### IV.3.3.4.3 Résultats des simulations

Les tests sont vérifiés, tel que le montre l'Annexe A.7 on voit bien que l'ensemble des paramètres sont inscrits dans leurs intervalles de confiance respectifs.

## IV.4. SIMULATION DES PARAMETRES HYDROMETEOROLOGIQUES AU PAS DE TEMPS JOURNALIER

Le passage du pas de temps mensuel (lors de la simulation des variables hydrométéorologiques) au pas de temps journalier ne se fait pas sans difficulté, du fait de la forte corrélation existante entre les variables successives, et surtout la non stationnarité des variables aléatoires prises en considération, ajoutez y à cela la masse de données et de calculs nécessaires pour la simulation de ces derniers.

### IV.4.1 Simulation des variables cycliques (ETP)

#### IV.4.1.1 Données utilisées

Les mêmes données utilisées dans le cas mensuel, sont utilisées dans le cas journalier, rappelons que l'on dispose d'un historique de 22 années. L'évolution chronologique des ETP, de 1965 à 1986 (Annexe B.1) montre que le phénomène ETP résulte de la combinaison de deux composantes bien distinctes une cyclique (de forme sinusoïdale de période une année) et l'autre aléatoire.

#### IV.4.1.2 A.C.P.

##### IV.4.1.2.1 Calcul des paramètres statistiques

Les paramètres statistiques calculés sur la série d'ETP sont résumés dans le tableau suivant:

Tableau IV.4-1 : Paramètres statistiques de la série d'ETP.

	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975
Moyennes (mm)	3,3	3,2	2,9	3,3	3,3	3,0	3,4	3,3	3,3	3,1	3,3
Ecart types (mm)	1,9	1,8	1,8	2,0	2,1	1,7	2,0	2,0	2,0	1,9	1,9

	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
Moyennes (mm)	3,3	3,2	3,4	3,5	3,3	3,3	3,2	3,3	3,6	3,8	4,1
Ecart types (mm)	1,9	1,9	1,9	1,8	2,1	2,0	1,8	1,8	2,1	2,1	2,2

On remarque une stabilité dans les moyennes et les écarts types, qui est liée à la nature cyclique du phénomène évapotranspiration.

##### IV.4.1.2.2 Calcul des CP

On a effectué une ACP normée sur les données brutes, puis sur les données transformées par racine carrée, la variation des composantes principales est donnée par la figure IV.-7.

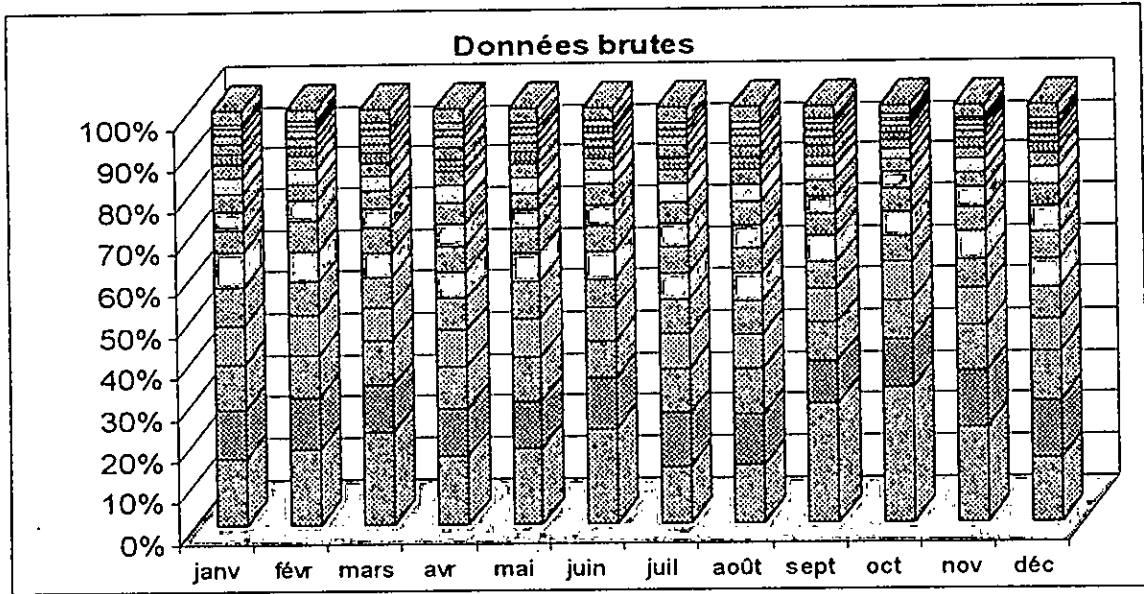


Figure IV-7: Variation des CP (ETP journalières)

On remarque une répartition saisonnière de la première CP, elle augmente de janvier à mars, diminue en avril puis augmente au mois de juin pour atteindre son minimum au mois d'août, et remonte par la suite pour atteindre son maximum au mois d'octobre, et diminue jusqu'à décembre. De la seconde à environ la cinquième composante l'apport est quasiment constant, cependant on observe des fluctuations dans l'apport des dernières CP.

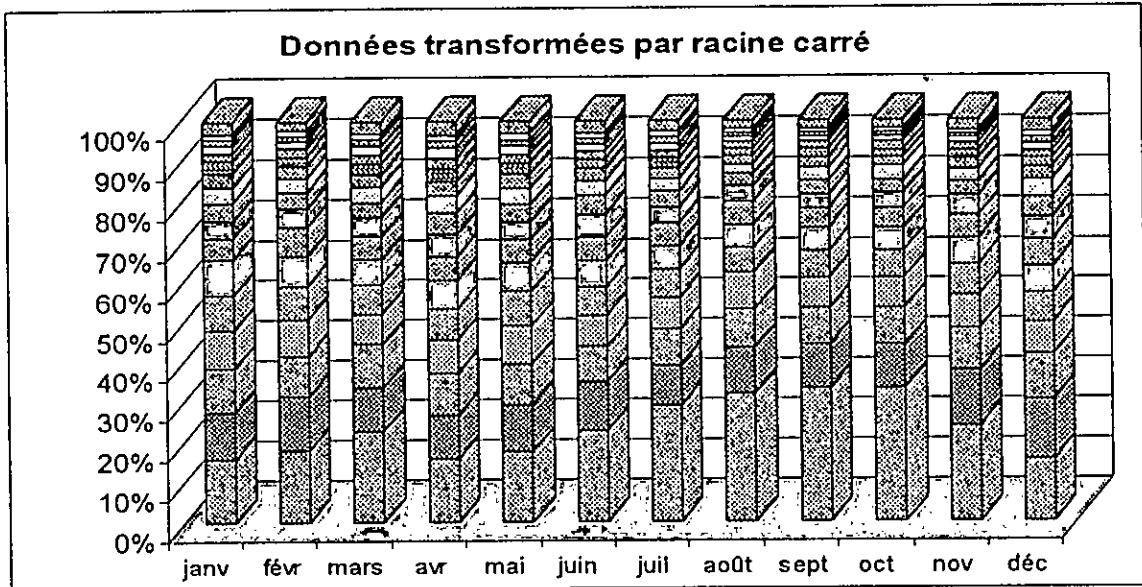


Figure IV-8: Variation des CP (ETP journalières transformées)

Après transformation des données par la fonction racine carré, la répartition de la première CP se trouve modifiée (Figure IV-8), on observe deux saisons: De janvier à mai et de mai à décembre, tout en restant maximale pour les mois de septembre et octobre et minimale pour le mois d'avril.

### IV.4.1.3 Reconstitution

#### IV.4.1.3.1 Choix du nombre de CP

La première CP explique en moyenne 20% de la variance globale. Nous avons pris en compte 6 CP sur un total de 22, qui expliquent à elle seules 60% de la variance totale du phénomène étudié.

#### IV.4.1.3.2 Choix du type de transformation

Le choix de la transformation à effectuer sur les données journalières initiales, se fait par superposition des données reconstituées, et des données historiques. Après examen de la reconstitution, on a constaté que la transformation racine carrée donne une reconstitution très proche de l'historique.

### IV.4.1.4 Simulation

#### IV.4.1.4.1 Ajustement des paramètres

On maintient l'ajustement des différents paramètres dans la loi normale. Les résultats sont donnés en Annexe B.2 (1), on a donné à titre indicatif les ajustements concernant deux mois le mois de juillet représentant la saison sèche et le mois de décembre représentant la saison humide.

Les CP: On remarque qu'en général les CP s'apprêtent bien à l'ajustement. Cependant Les deux premières CP concernant le mois de juillet présente une extrémité au dessous de la limite inférieure de l'intervalle de confiance. On peut faire la remarque inverse pour les CP du mois de décembre.

Les Covariances: les covariances calculées présentent un bon ajustement .

Les résidus: On considère pour l'ajustement le vecteur résidus de la matrice entière, on remarque que pour le mois de juillet les résidus les plus faibles sont en dehors de l'intervalle, alors que pour le mois de décembre se sont les résidus les plus élevés qui sortent de l'intervalle.

#### IV.4.1.4.2 Génération des paramètres

Les résultats de la génération dans la loi normale, sont donnés en Annexe B.2 (2).

La génération des CP et des covariances donne de bon résultats, toutes les séries générées appartiennent à l'intervalle, avec un seuil de confiance 95% . Le terme résiduel est jugé acceptable en moyenne.

IV.4.1.4.3 Résultats des simulations

On calcule les valeurs annuelles des séries simulées pour lesquelles on détermine la moyenne et l'écart type, qu'on compare par la suite aux valeurs calculées sur la série historique (Figure IV.-9)

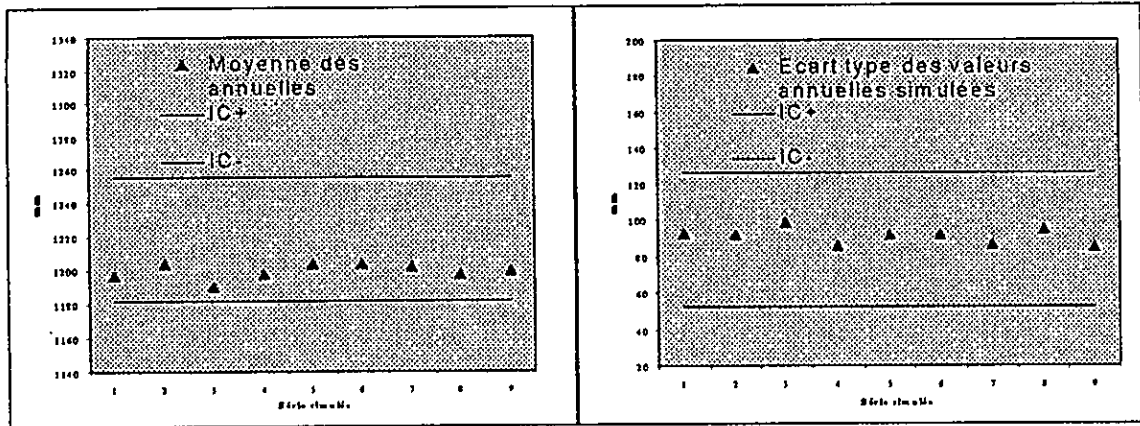


Figure IV-9: Tests des moyennes et écarts types des valeurs d'ETP annuelles.

On remarque que pour l'ensemble des séries simulées, la moyenne et l'écart type des valeurs annuelles appartiennent à l'intervalle de confiance à un seuil de 95%.

On a testé également l'ajustement des valeurs annuelles, on remarque (figure IV.-10) que ces dernières suivent la loi normale, et on arrive à obtenir des valeurs annuelles simulées appartenant à l'intervalle de confiance et reproduisant la même allure que les valeurs annuelles historiques.

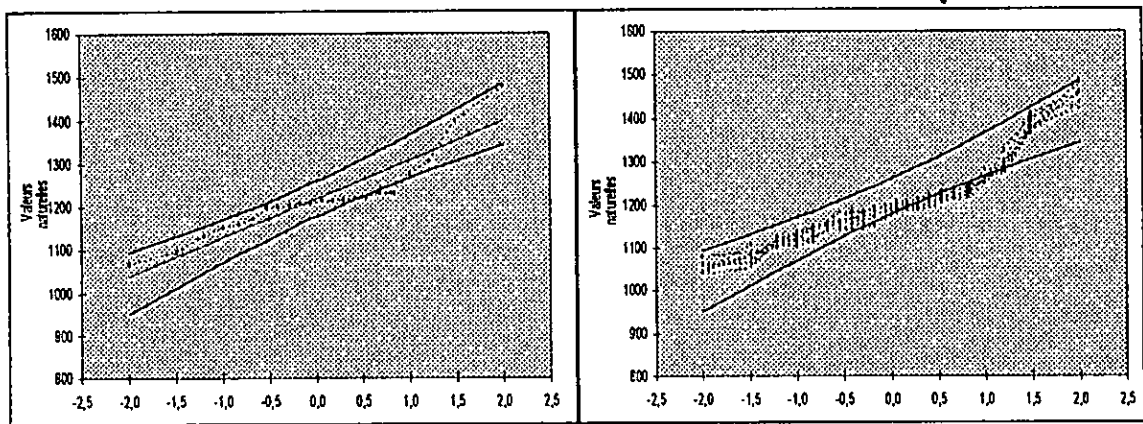


Figure IV-10: Ajustement des valeurs d'ETP annuelles observées et simulées

On a déterminé les moyennes mensuelles historiques, dont on a calculé la moyenne et l'écart type ainsi que leurs intervalles de confiances ( $\alpha = 5\%$ ).

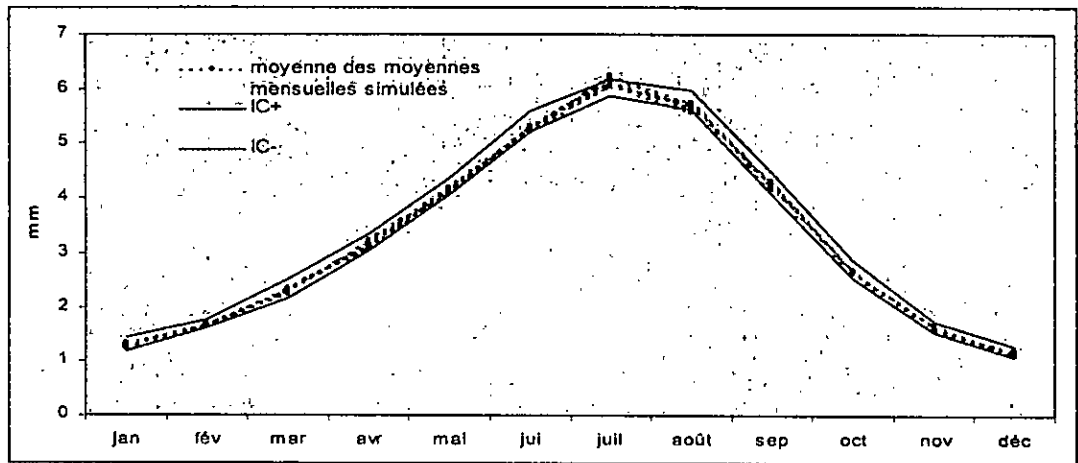


Figure IV-11: Test des moyennes des moyennes mensuelles d'ETP

La courbe donnée en Figure IV.-11 reproduit l'aspect physique du phénomène, elle a la forme d'une cloche qui atteint son maximum durant les mois de juillet et août, et son minimum durant l'hiver, ceci et du à la forte intercorrélation des ETP, et surtout à leurs très faible variabilité mensuelle. Les moyennes mensuelles des séries simulées sont toutes inscrites dans l'intervalle de confiance.

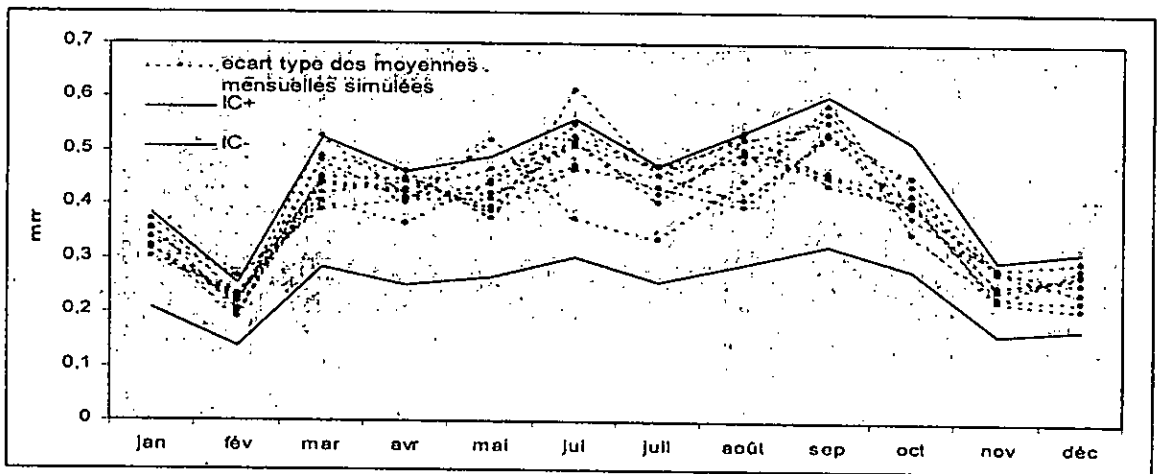


Figure IV-12: Test des écart types des moyennes mensuelles d'ETP

Les écarts types des séries simulées (figure IV.-12) appartiennent à l'intervalle de confiance ( $\alpha = 5\%$ ).

Les valeurs maximales d'ETP de la série historique, suivent la loi de Gumbel (Figure IV.-13). Avec le modèle GESTOP les valeurs simulées sont à l'intérieur de l'intervalle de confiance ( $\alpha = 5\%$ ), cela dénote la qualité de restitution de ce paramètre qui est très difficile à appréhender.

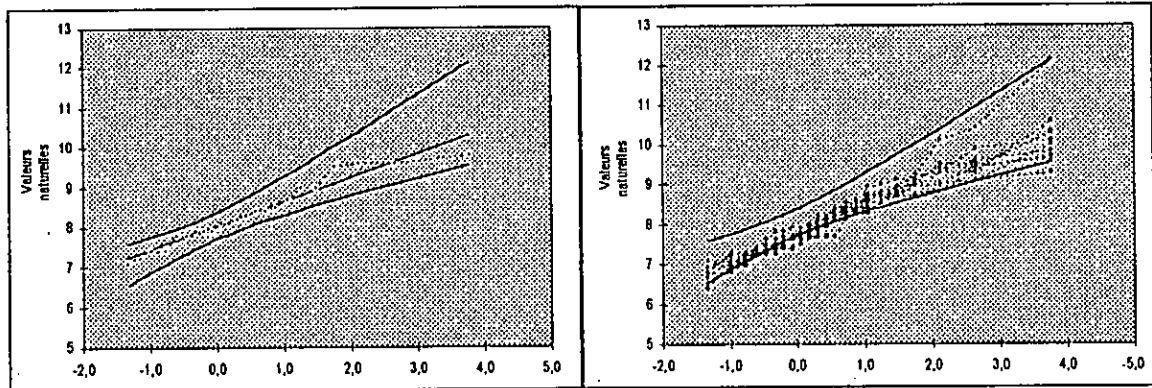


Figure IV-13: Ajustement des valeurs d'ETP maximales (Loi de Gumbel)



## IV.4.2 Simulation des apports

### IV.4.2.1 Données utilisées

Les données disponibles concernant les apports journaliers de la station de Cheffia s'étalent sur un historique de 12 années.

L'évolution journalière des apports (Annexe B.3) montre d'une part, l'aspect aléatoire du phénomène débit, et fait ressortir d'autre part la faiblesse de ces apports pour les années 1988 et 1989.

#### IV.4.2.2 A.C.P.

##### IV.4.2.2.1 Calcul des paramètres statistiques

Les paramètres statistiques calculés sur la série de débits sont résumés dans le Tableau IV.4-2.

Tableau IV.4-2: Paramètres statistiques de la série des débits historiques

	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Moyenne hm3/j	0,15	0,22	0,49	0,1	1,03	0,46	2,79	0,02	1,96	0,01	0,11	0,02	0,09	0,12
Ecart type hm3/j	0,34	0,32	0,79	0,14	1,68	0,46	8,89	0,04	3,77	0,02	0,27	0,03	0,14	0,16

##### IV.4.2.2.2 Calcul des CP

On a effectué sur la matrice [X] une ACP normée, avec deux variantes, les données brutes et les données transformées par la racine carrée. On a observé la répartition des CP présentées dans les figures IV.-14 et IV-15.

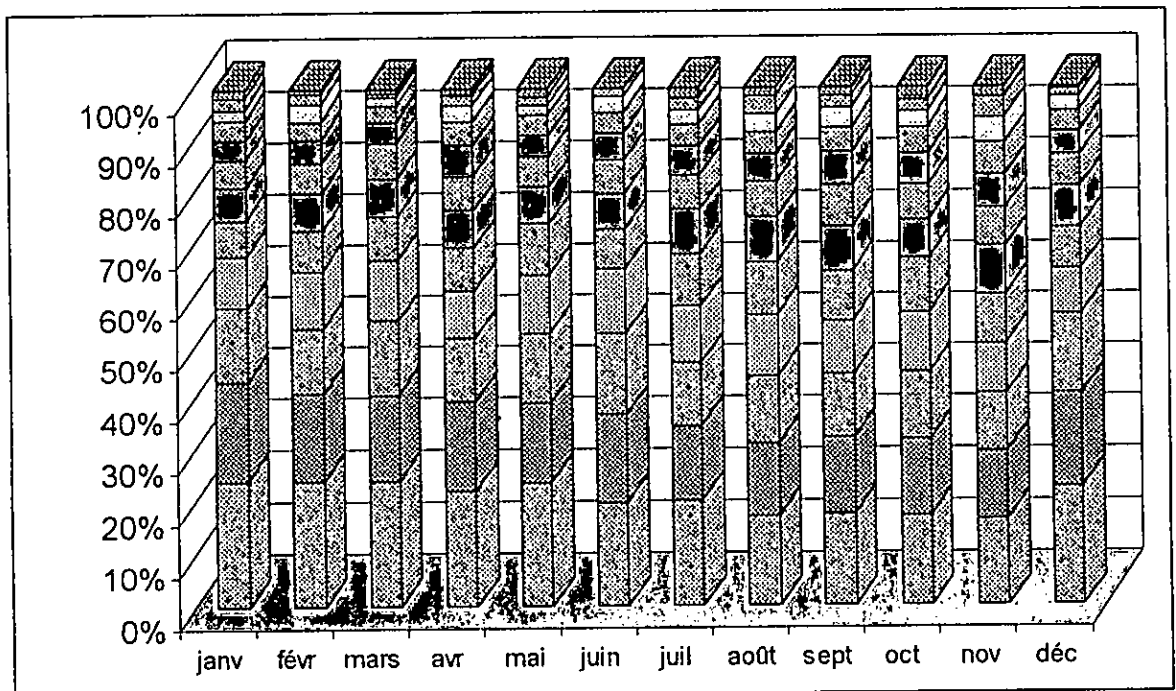


Figure IV-14: Variation des CP (données de débits brutes)

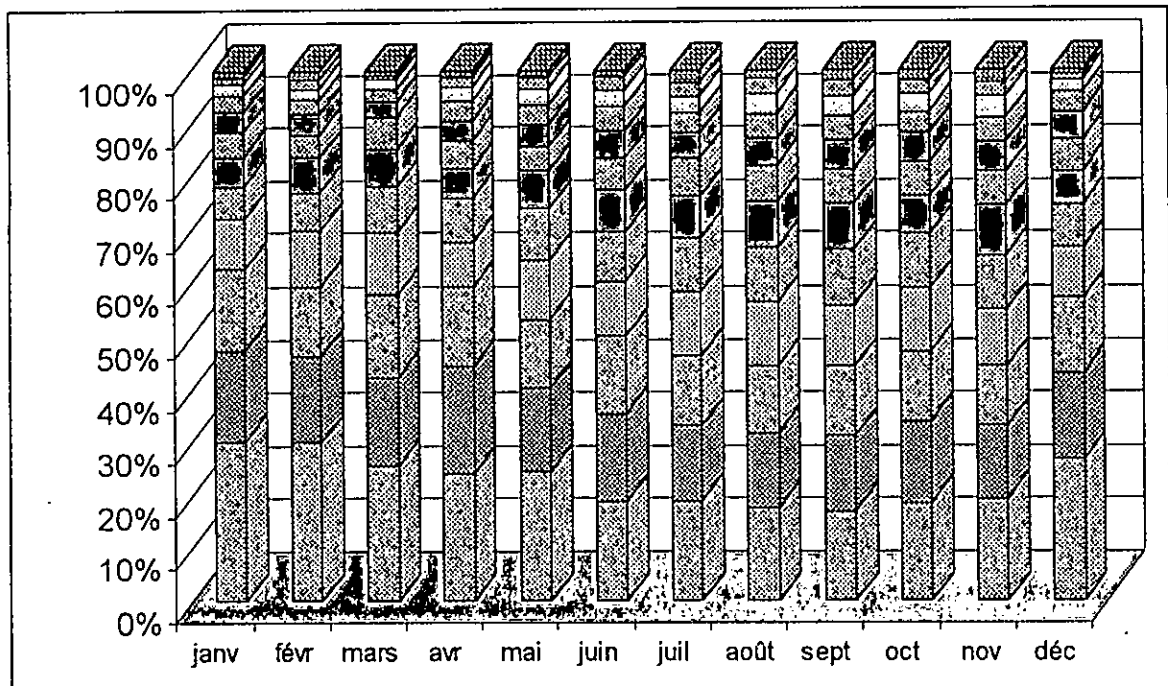


Figure IV-15: Variation des CP (données de débits transformées)

#### IV.4.2.3 Reconstitution

##### IV.4.2.3.1 Choix du nombre de CP

On a fixé le seuil d'information à 60%, ce qui revient à considérer 3 ou 4 CP selon le cas.

##### IV.4.2.3.2 Choix du type de transformation

Les données brutes reconstituées sont surestimées, le maximum historique qui est égale à 42,8 hm<sup>3</sup>/j, est reconstitué avec 129 hm<sup>3</sup>/j, on a retrouvé également, des valeurs négatives de l'ordre de -68 hm<sup>3</sup>/j. Donc pour la simulation des apports, on a été contraint de transformer les données initiales, pour pallier aux aberrations constatées lors de la reconstitution.

#### IV.4.2.4 Simulation

##### IV.4.2.4.1 Ajustement des paramètres

On a choisit d'ajuster les paramètres concernant un mois sec (juillet), et un mois humide (décembre). L'ajustement des paramètres calculés avec une ACP normée sur les données transformées par racine carrée, donné en Annexe B.4 (1), et Annexe B.5 (1), a mené aux constatations suivantes :

Les CP: les composantes principales des deux mois considérés s'ajustent bien à la loi normale avec un seuil de 95%.

Les covariances: elles présentent un bon ajustement à la loi normale; pour le mois de juillet, les covariances calculées sont très proches de la courbe théorique.

Les résidus: la loi normale donne les ajustements attendus pour les résidus, qui par construction suivent une loi normale  $N(0, \sigma\varepsilon)$ .

#### IV.4.2.4.2 Génération des paramètres

L'ajustement des paramètres générés par la loi normale est donné en Annexe B.4 (2) pour le mois de décembre et en Annexe B.5 (2) pour le mois de juillet.

#### IV.4.2.4.3 Résultats des simulations

L'approche adoptée pour la simulation des débits journaliers a été la même que celle des ETP, mais a un tel pas de temps il y a la stationnarité du phénomène qui entre en jeu, en effet dans ce cas le phénomène est intermittent, après une période relativement longue de valeurs faibles de débits, on se retrouve avec un maximum difficile à appréhender et à modéliser, c'est pour cela que plusieurs variantes (Tableau IV.4-2) ont été développées et testées pour essayer d'arriver à un modèle de simulation restituant correctement les caractéristiques statistiques de l'historique et surtout les valeurs maximales journalières annuelles des débits. Plusieurs lois d'ajustements ont été testées et les meilleurs résultats ont été obtenus pour la loi de Goodrich.

La fonction de densité de probabilité (FDP) de Goodrich s'écrit :

$$F(x) = 1 - e^{-a(x-b)^{1/n}} \quad \text{pour } b \leq x \leq +\infty$$

Où a, b et n sont les paramètres à calculer.

Comme la FDP de Goodrich contient trois paramètres, ce type d'approximation est assez flexible, elle a été utilisée pour des valeurs moyennes, pour lesquelles elle a donné de bon résultats.

Les paramètres de Goodrich sont déterminés comme suit :

$$\begin{aligned} m_1 &= \bar{X} \\ m_2 &= \frac{1}{N} \sum X^2 \\ m_3 &= \frac{1}{N} \sum X^3 \\ S_2 &= \frac{1}{N} \sum (X - \bar{X})^2 \\ S_3 &= \frac{1}{N} \sum (X - \bar{X})^3 \\ m'_3 &= m_3 - (3\bar{X}S_2 + \bar{X}^3) \\ C_s &= \frac{m'_3}{S_3} \end{aligned}$$

On tire à partir d'un abaque la valeur de n correspondant à  $C_s$ , et on détermine les coefficients a et b :

$$\begin{aligned} a &= \frac{1}{S_2} \left( \Gamma_{(2n+1)} - \Gamma_{(n+1)}^2 \right)^{1/2n} \\ b &= \bar{X} - \frac{\Gamma_{(n+1)}}{a^n} \end{aligned}$$

Pour générer dans la loi de Goodrich, on tire aléatoirement une fréquence notée  $y$ , qu'on injecte dans la fonction inverse de la FDP.

$$x = b + \frac{[\log(1-y)]^n}{a^n}$$

La simulation avec la première variante est basée sur la qualité des ajustements, la génération dans la fonction de répartition a été faite dans le but d'englober l'information perdue avec l'approximation des lois statistiques, la variante trois permet l'introduction d'une information nouvelle, la variante quatre suit la même logique que la variante trois sauf que son domaine de définition est plus large que celui de la fonction de répartition.

Tableau IV.4-3: Les différentes variantes de simulation

	Covariances	CP		Résidus
		CP1	2 Q/CP	
Variante 1	N	N	N	N
Variante 2	FR	FR	FR	FR
Variante 3	FR étendue	FR étendue	FR étendue	FR étendue
Variante 4	G	G	G	G
Variante 5	G	N	G	G

N : Simulation dans la loi normale

FR : Simulation dans la fonction de répartition

FR étendue : Simulation dans la fonction de répartition étendue

G: Simulation dans la loi de Goodrich

L'acceptation ou le rejet d'une variante se fait sur la base des critères pris en considération pour la variable ETP à savoir: les moyennes et écarts types annuelles et mensuelles ainsi que les ajustements des valeurs annuelles et maximales journalières annuelles

Tableau IV.4-4: Résultats des simulation des débits

	Valeurs annuelles			Valeurs mensuelles		Ajustement Val. Max. journalières
	$\bar{X}$	$\sigma$	Ajust.	$\bar{X}$	$\sigma$	
Variante 1	+	-	+	+	-	-
Variante 2	-	-	+	+	-	+
Variante 3	-	+	+	+	+	+
Variante 4	+	-	+	+	-	+
Variante 5	+	+	+	+	+	+

les valeurs annuelles des séries de débits simulées (Figure IV-16) appartiennent en moyenne à l'intervalle de confiance, mais les écarts types sont au dessus de la limite supérieure.

Les valeurs annuelles des débits historiques suivent une loi normale (Figure IV-17), et on arrive à obtenir des valeurs annuelles simulées appartenant à l'intervalle et reproduisant la même allure que les valeurs annuelles historiques.

La Figure IV.-18 permet de discerner une période sèche allant de juin à septembre et une saison humide allant du mois d'octobre au mois de mai, la figure montre également que les moyennés des valeurs mensuelles des séries simulées appartiennent à l'intervalle de confiance aussi bien pour la saison sèche que pour la saison humide.

Pour ce qui est des écarts types des valeurs mensuelles on observe une surestimation, principalement pour les mois humides (Figure IV-19).

Les valeurs maximales des débits historiques suivent la loi de Gumbel (Figure IV.-20). Le modèle GESTOP permet d'obtenir des valeurs maximales simulées qui appartiennent à l'intervalle de confiance de la série historique.

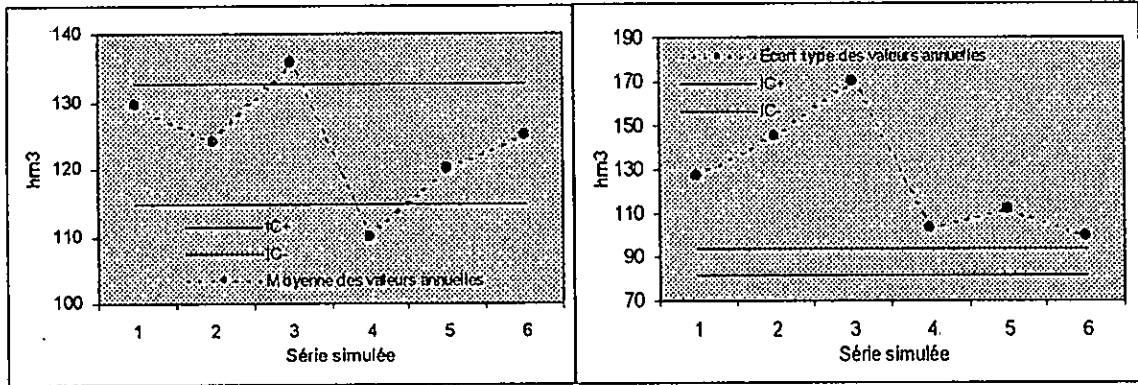


Figure IV-16: Tests des moyennes et écarts types des valeurs de débits annuelles (V1).

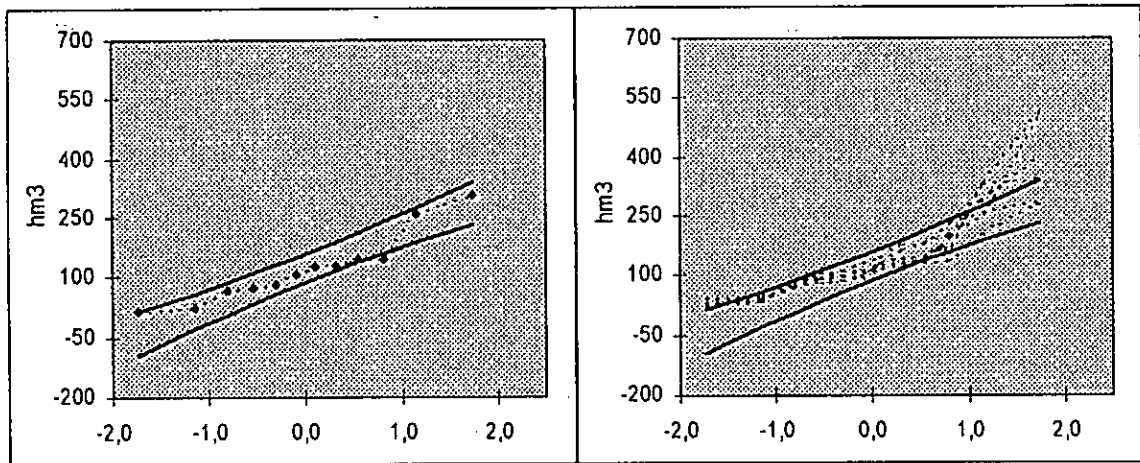


Figure IV-17: Ajustement des valeurs de débits annuelles observées et simulées (V1)

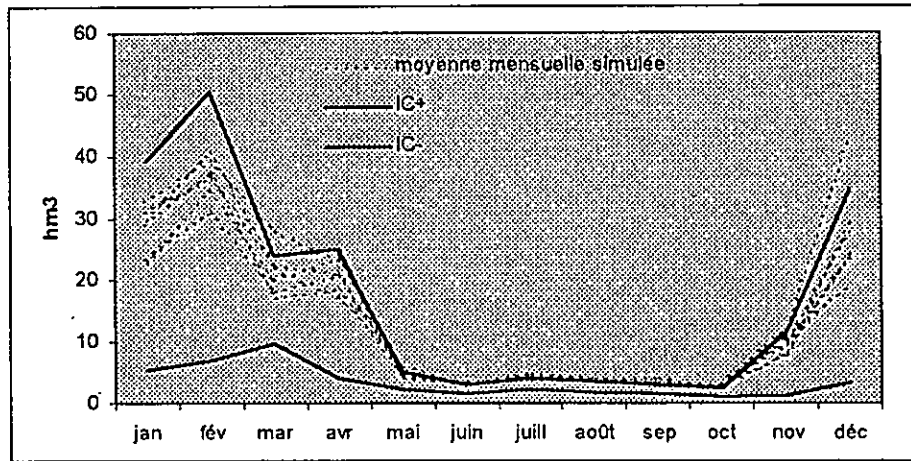


Figure IV-18: Test des moyennes des valeurs mensuelles de Débit (V1)

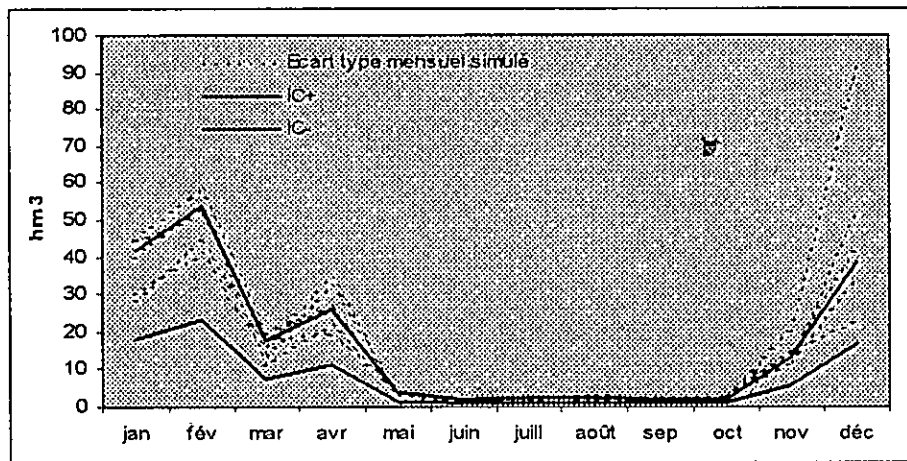


Figure IV-19: Test des écart types des valeurs mensuelles de Débits (V1)

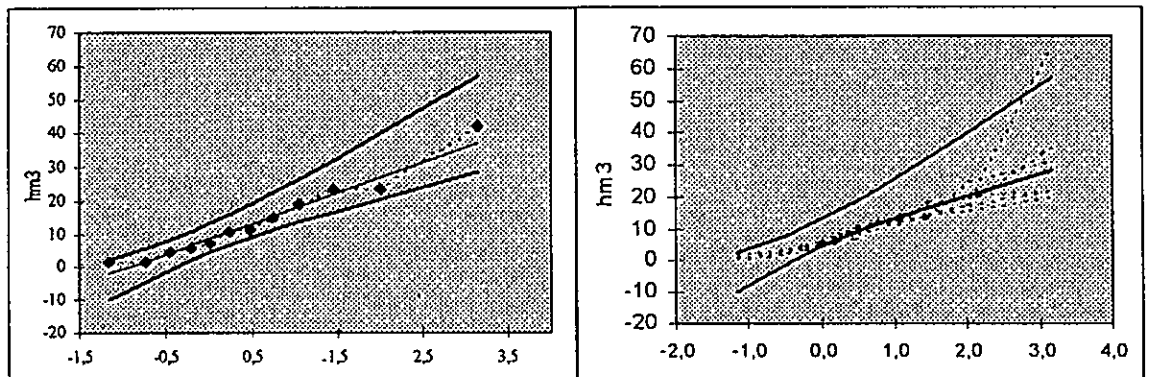


Figure IV-20: Ajustement des valeurs de Débits maximales (Loi de Gumbel) (V1)

La deuxième variante maintient la surestimation des moyennes et des écarts types lors du test des valeurs annuelles (Figure IV-21).

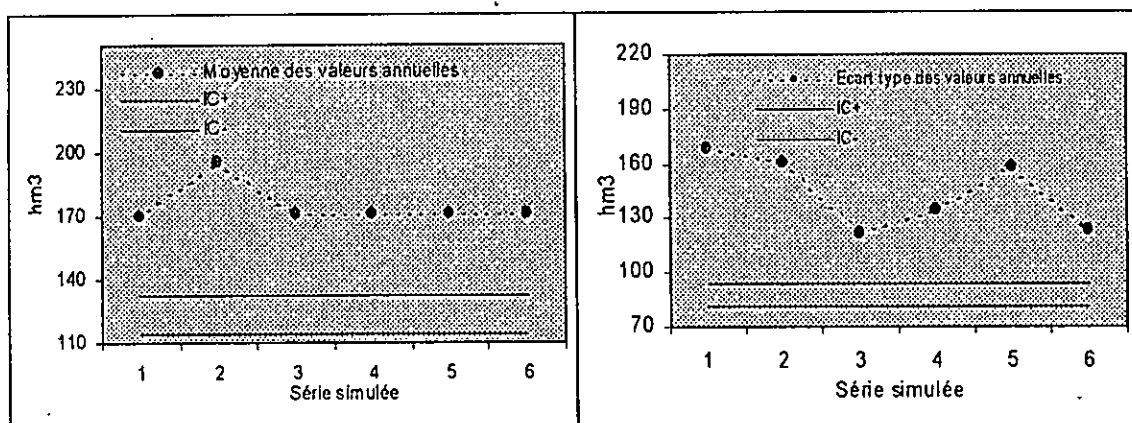


Figure IV-21 : Tests des moyennes et écarts types des valeurs de débits annuelles (V2).

Les valeurs annuelles simulées (Figure IV-22) présentent un bon ajustement à la loi normale, mais on observe quelques valeurs extrêmes largement supérieures à la valeur annuelle maximale historique.

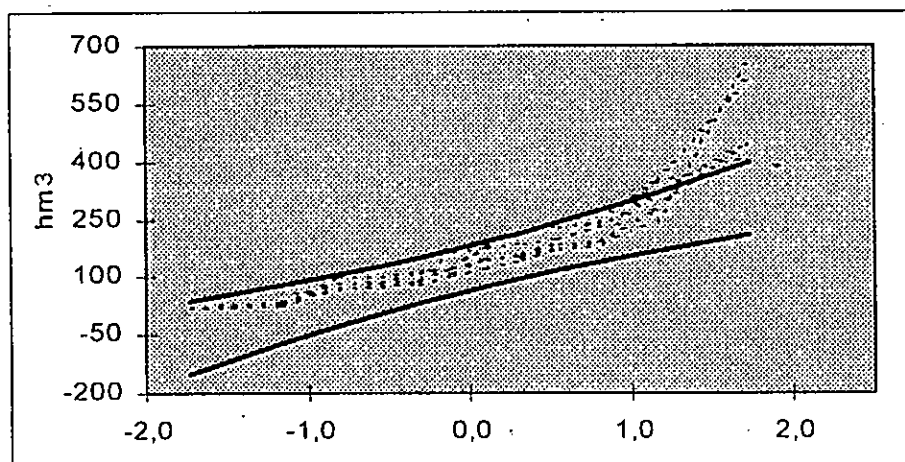


Figure IV-22: Ajustement des valeurs de débits annuelles simulées (V2)

Les moyennes des valeurs mensuelles sont à l'intérieur de l'intervalle de confiance (Figure IV-23).

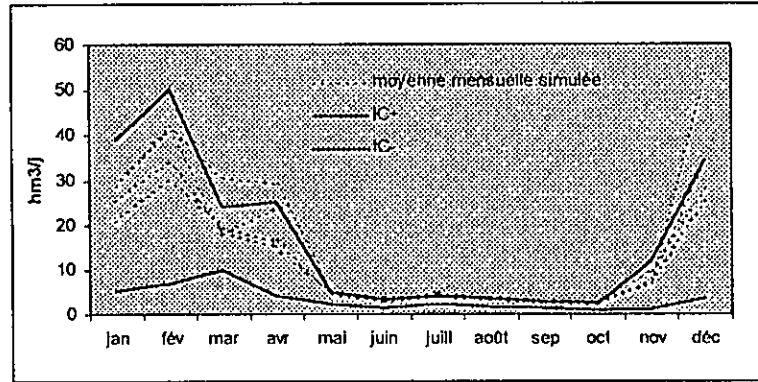


Figure IV-23: Test des moyennes des valeurs mensuelles de Débit (V2)

Les écarts types des valeurs mensuelles (Figure IV-24) sont satisfaisant sauf pour le mois de décembre où on observe des écarts types supérieures à l'extrémité supérieure de l'intervalle de confiance.

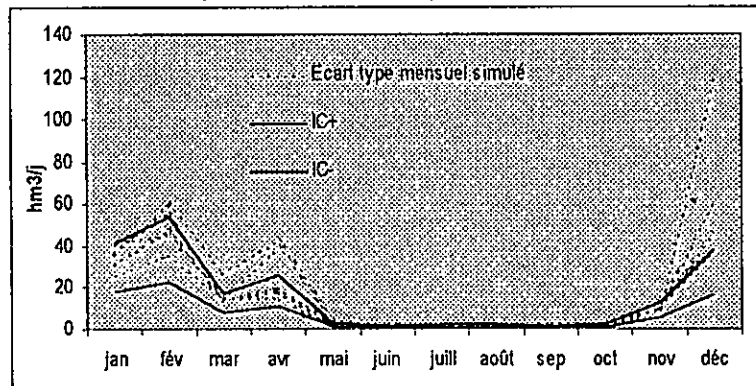


Figure IV-24: Test des écart types des valeurs mensuelles de Débits (V2)



Pour ce qui est des valeurs maximales journalières (Figure IV-25), on a obtenu des valeurs simulées qui appartiennent à l'intervalle de confiance.

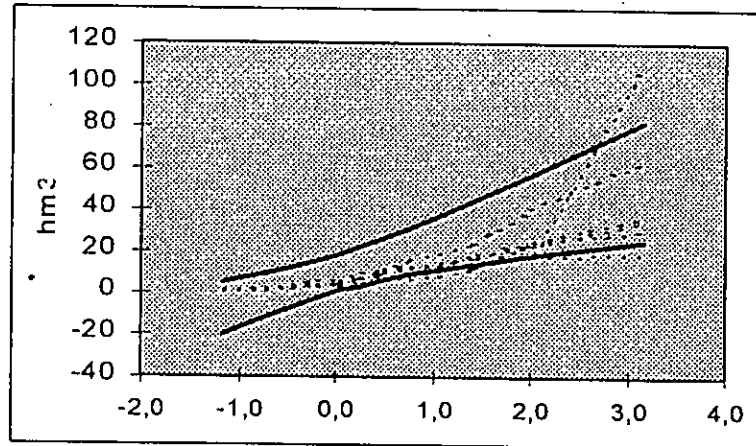


Figure IV-25: Ajustement des valeurs de Débits maximales (Loi de Gumbel) (V2)

La troisième variante permet d'améliorer les écarts types des moyennes des valeurs annuelles (Figure IV-26)

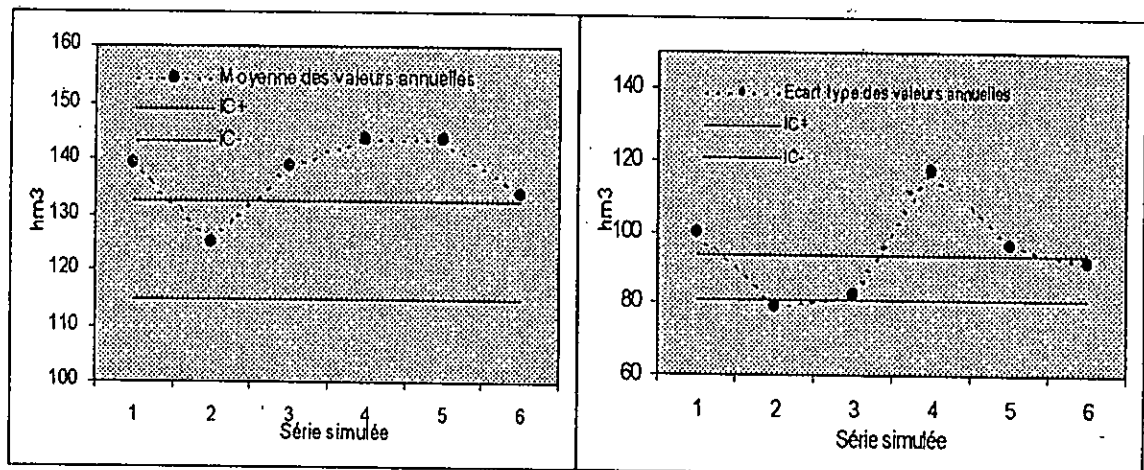


Figure IV-26: Tests des moyennes et écarts types des valeurs de débits annuelles (V3).

Les valeurs annuelles simulées appartiennent en totalité à l'intervalle de confiance (Figure IV-27).

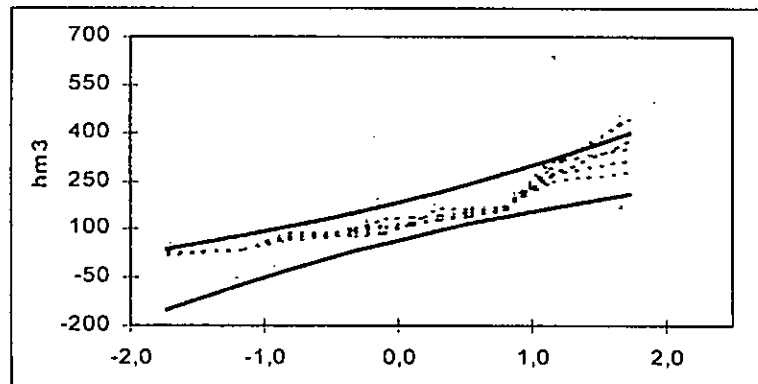


Figure IV-27: Ajustement des valeurs de débits annuelles observées et simulées (V3).

Les moyennes et les écarts types des valeurs mensuelles (Figure IV-28 et Figure IV-29) sont très bien restituées, toutes les valeurs sont à l'intérieure de l'intervalle de confiance.

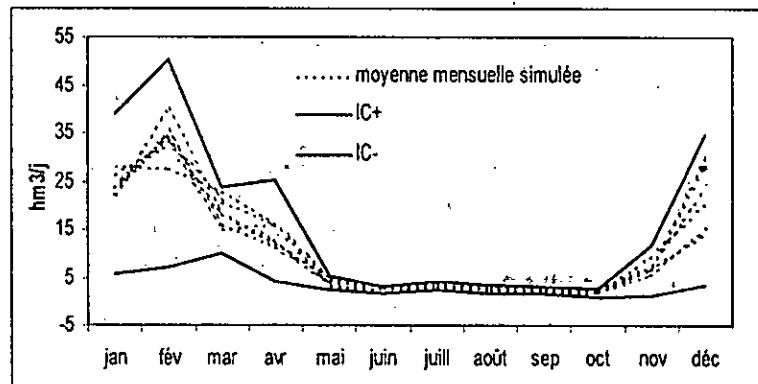


Figure IV-28: Test des moyennes des valeurs mensuelles de Débit (V3).

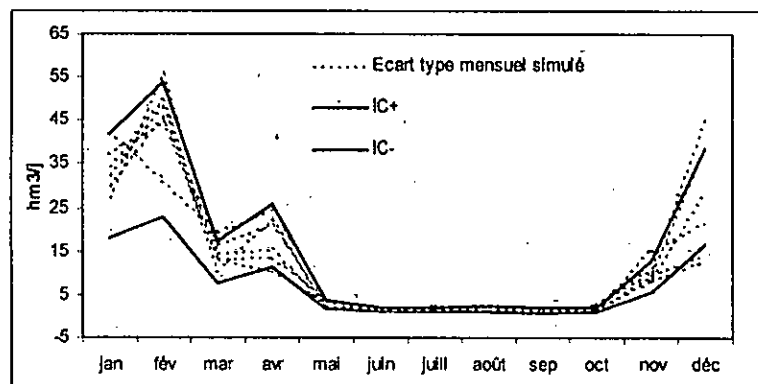


Figure IV-29: Test des écart types des valeurs mensuelles de Débits (V3).

Les valeurs journalières maximales obtenues en prolongeant la fonction de répartition, sont donc l'intervalle de confiance, mais ne présentent pas de maxima notables (Figure IV-30).

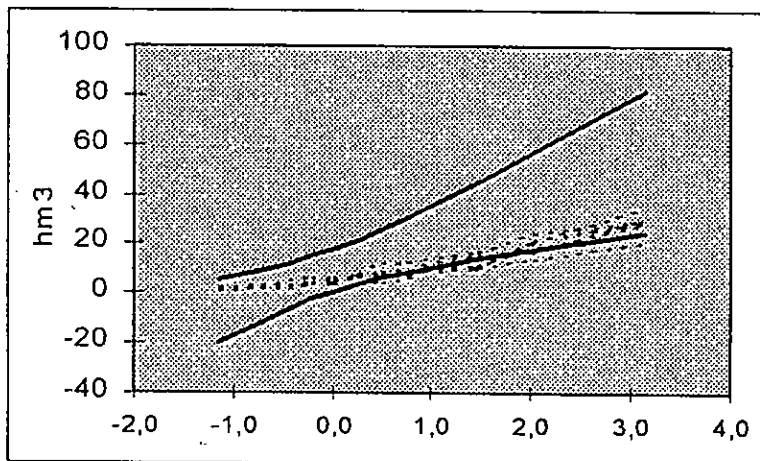


Figure IV-30: Ajustement des valeurs de Débits maximales (loi de Gumbel) (V3).

Les résultats obtenus avec la quatrième Variante, basée sur la génération dans la loi de Goodrich, donnent des valeurs annuelles dont la moyenne et l'écart type (Figure IV-31) sont à l'intérieur de l'intervalle de confiance.

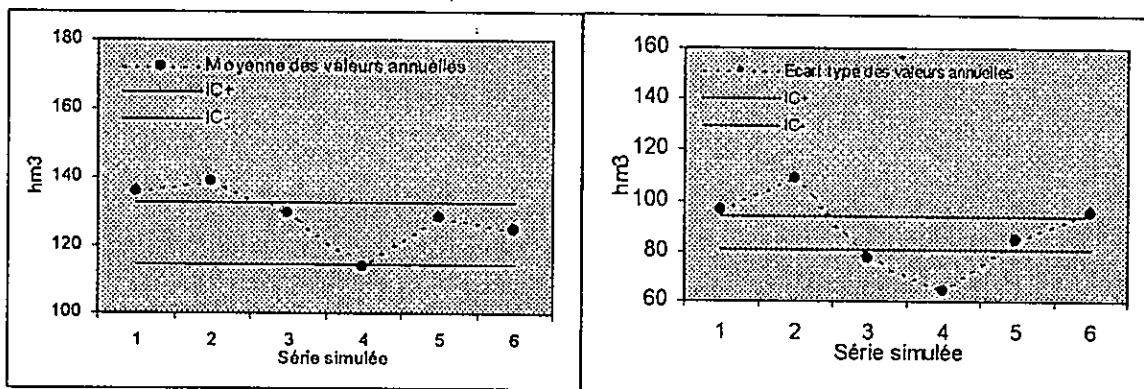


Figure IV-31: Tests des moyennes et écarts types des valeurs de débits annuelles (V4).

Le test de l'ajustement des valeurs annuelles simulées est vérifiés à un degré très satisfaisant (Figure IV-32).

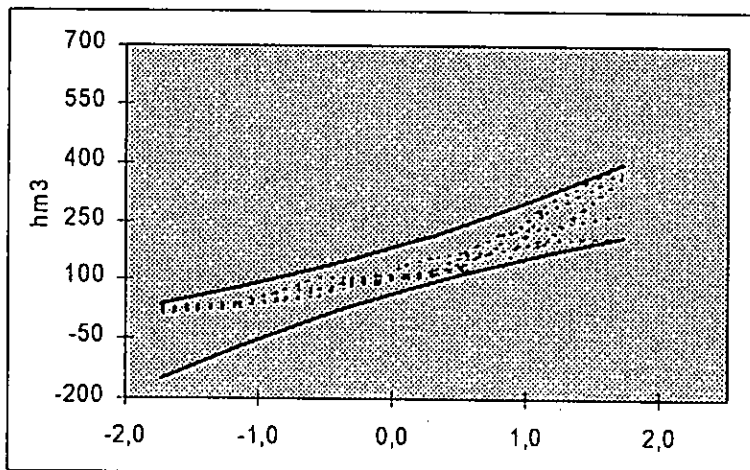


Figure IV-32: Ajustement des valeurs de débits annuelles simulées (V4).

Les moyennes des valeurs mensuelles (Figure IV-33) appartiennent à l'intervalle de confiance, aussi bien pour les mois secs que pour les mois humides.

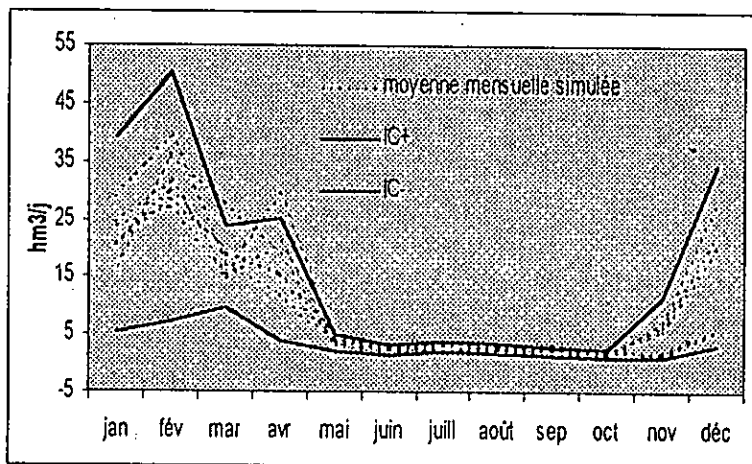


Figure IV-33: Test des moyennes des valeurs mensuelles de Débit (V4).

Les écarts types des valeurs mensuelles (Figure IV-34) sont en majorité dans l'intervalle de confiance, on observe cependant une forte variabilité au mois d'avril et au mois de décembre.

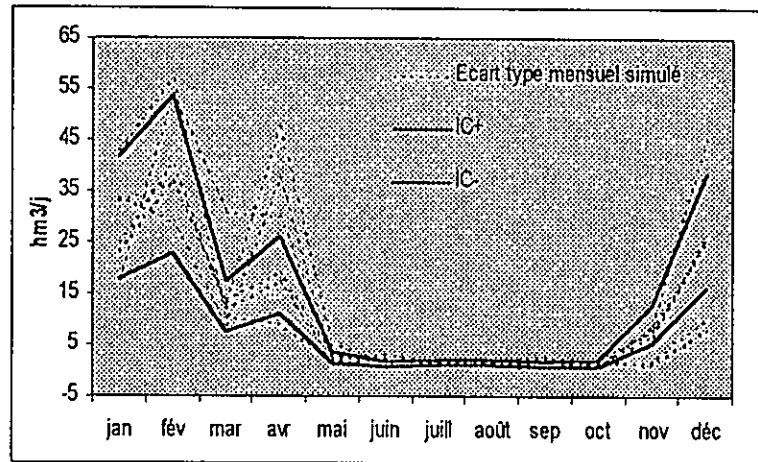


Figure IV-34: Test des écarts types des valeurs mensuelles de Débits (V4).

Les valeurs maximales journalières simulées (Figure IV-35) sont semblables à ceux de la variante trois. Le test est vérifié mais on n'a toujours pas des valeurs extrêmes appréciables.

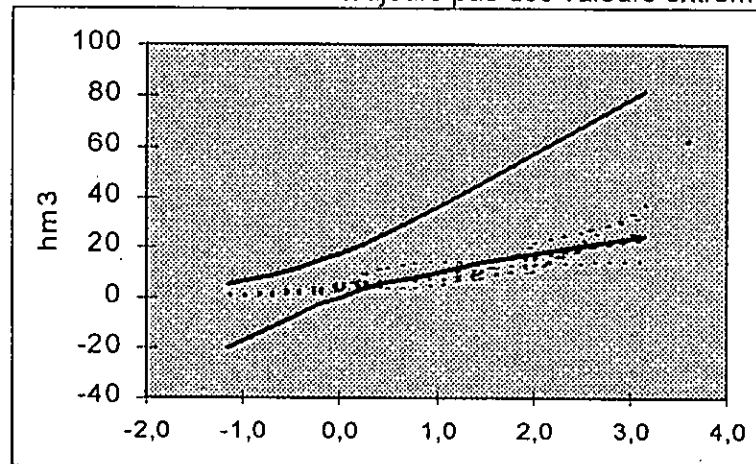


Figure IV-35: Ajustement des valeurs de Débits maximales (Loi de Gumbel) (V4).

Les séries simulées obtenues avec la dernière variante, possèdent des valeurs annuelles (Figure IV-36) dont la moyenne est légèrement au-dessus de l'intervalle et l'écart type oscille autour de l'intervalle. Le test est jugé acceptable.

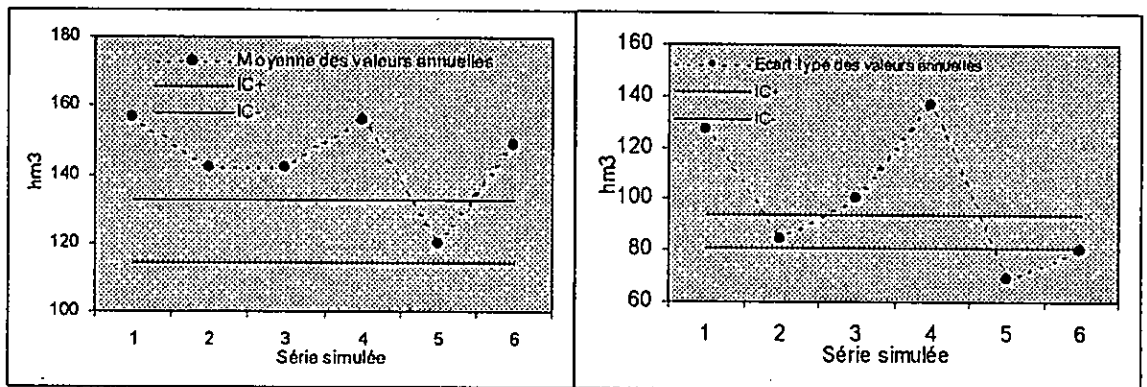


Figure IV-36: Tests des moyennes et écarts types des valeurs annuelles (V5).

L'ajustement des valeurs annuelles est vérifié (Figure IV-37).

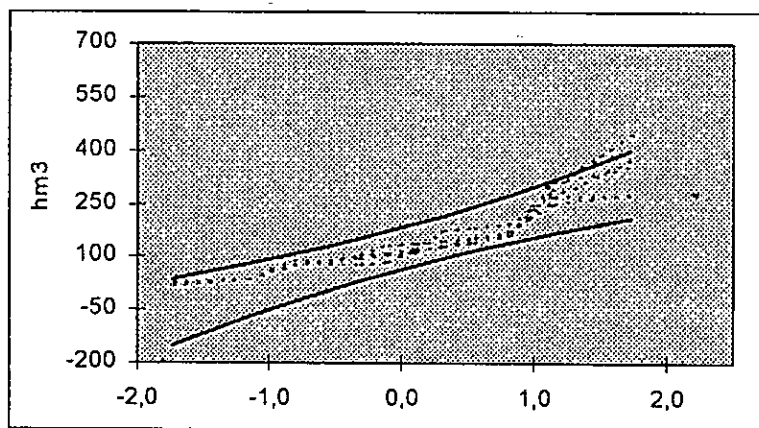


Figure IV-37: Ajustement des valeurs de débits annuelles simulées (Loi normale) (V5).

Les valeurs mensuelles calculées sur les séries simulées, possèdent des moyennes (Figure IV-38) appartenant à l'intervalle de confiance.

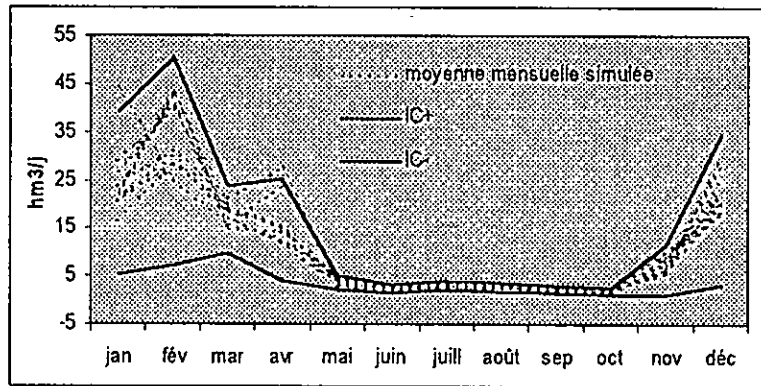


Figure IV-38: Test des moyennes des valeurs mensuelles de Débit (V5).

Le test des écarts types des valeurs mensuelles est également vérifié (Figure IV-39).

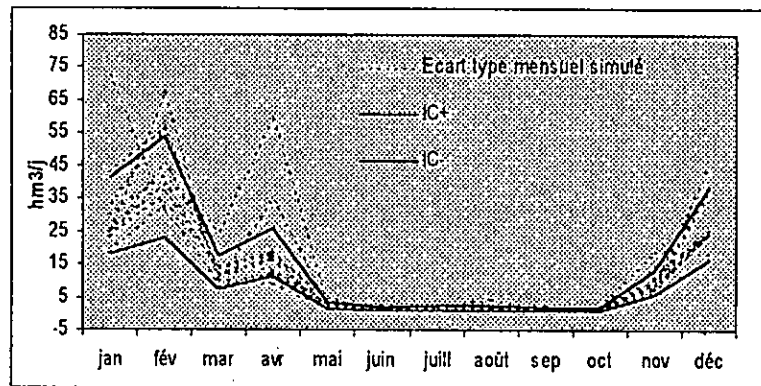


Figure IV-39: Test des écart types des valeurs mensuelles de Débits (V5).

Les séries simulés avec cette variante vérifient le test de l'ajustement des valeurs maximales (figure IV-40).

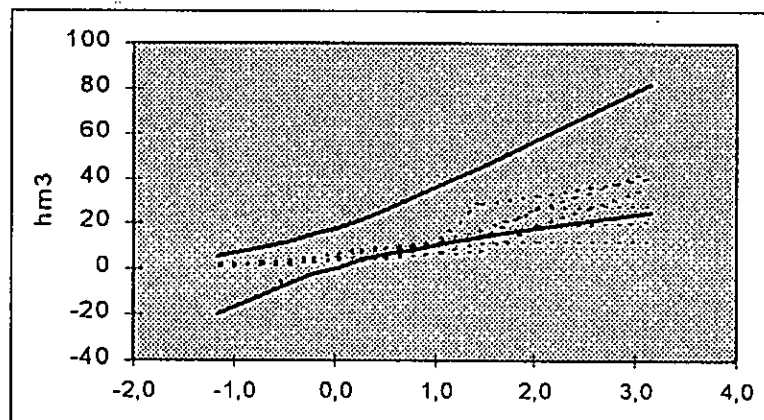


Figure IV-40: Ajustement des valeurs de Débits maximales (Loi de Gumbel) (V5).

Les phénomènes intermittant tel que la variable débit sont plus complexes a simuler que les phénomènes cycliques (ETP) du fait de la très grande variabilité au pas de temps journalier

### IV.4.3 Simulation des pluies

#### IV.4.3.1 Données utilisées

On utilise pour la simulation de la pluviométrie au pas de temps journalier, les données de la station d'Annaba de 1951 à 1986, dont l'évolution est donnée en Annexe B.6.

#### IV.4.3.2 A.C.P.

##### IV.4.3.2.1 Calcul des paramètres statistiques

Les paramètres statistiques calculés sur la série de pluies sont résumés dans le tableau suivant. On relève la plus forte moyenne en 1985 et la plus faible en 1962. Alors le plus haut écart types est relevé durant l'année 1983 coïncidant avec une pluie maximale de 108 mm.

Tableau IV.4-5: Paramètres statistiques de la série de Pluies d'Annaba (1951 - 1986)

	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962
Moyennes (mm)	1,7	1,7	2,5	2,4	1,7	1,9	2,0	2,3	2,0	2,0	1,4	0,8
Ecart types (mm)	4,8	5,4	7,3	6,1	4,3	5,3	6,1	7,8	7,9	5,5	4,1	2,6

	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974
Moyennes (mm)	1,7	1,7	1,7	1,8	1,3	1,6	1,2	1,8	1,3	1,7	1,7	2,0
Ecart types (mm)	5,1	4,5	5,1	5,2	3,7	5,9	4,5	4,8	4,5	4,4	4,8	7,2

	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
Moyennes (mm)	1,2	1,1	2,0	1,4	1,4	2,1	1,7	1,5	2,7	1,6	3,1	1,5
Ecart types (mm)	3,2	3,1	5,5	4,6	4,0	6,0	4,5	4,2	8,4	5,8	7,9	3,6

##### IV.4.3.2.2 Calcul des CP

On effectue une ACP normée sur la série de pluies brutes, les CP obtenues sont illustrées par la Figure IV-41 ci-dessous. On remarque la différence de l'apport de la première CP d'un mois à un autre, les mois secs trouvent le maximum de l'information concentré dans la première CP, on donne l'exemple de juillet qui avec 2CP totalise 80% de l'information.

Les CP obtenues après transformation des données est données dans la Figure IV-42; On n'observe pas de variation des apports par rapport au cas des données non transformées, la première CP reste maximale pour les mois secs, mais avec une nette diminution de l'amplitude.



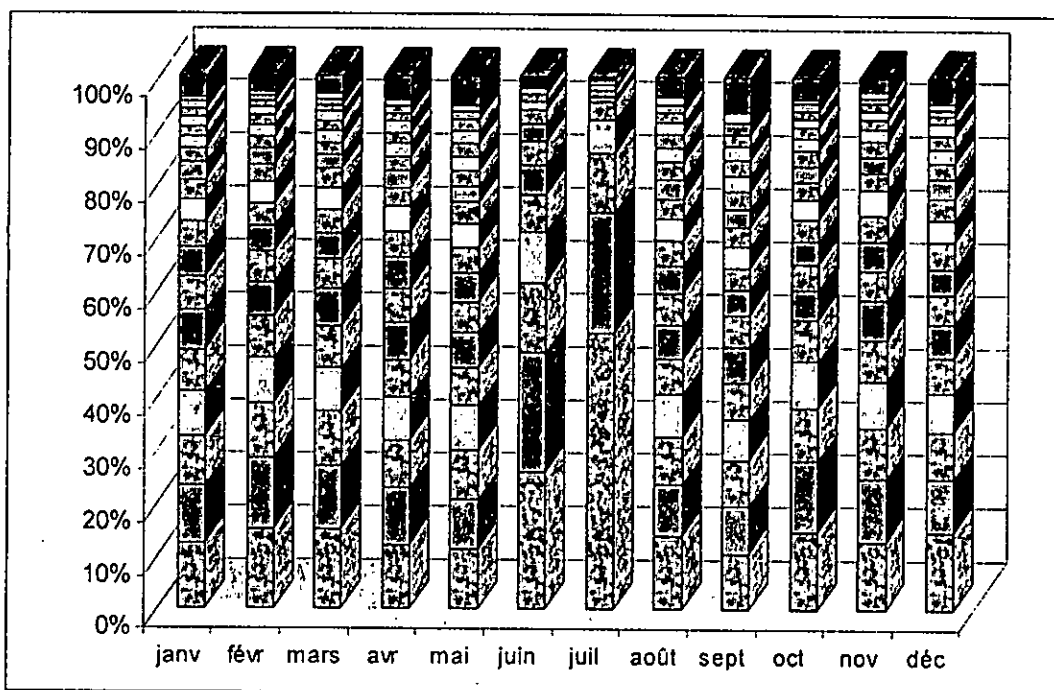


Figure IV-41: Les CP calculées sur la série de pluies brutes

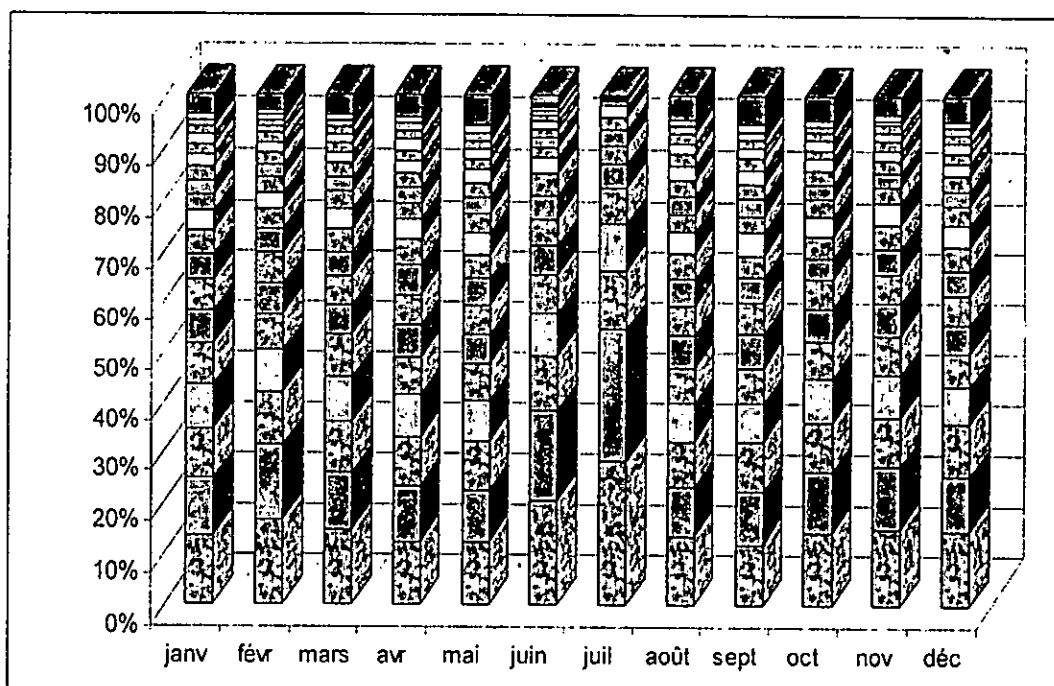


Figure IV-42: CP calculées sur la série de pluies transformée

### IV.4.3.3 Reconstitution

#### IV.4.3.3.1 Choix du nombre de CP

Vu la complexité du phénomène pluies on fixe un seuil d'information à environ 80% ce qui revient à considérer entre 6 et 7 CP avec exceptionnellement 3 CP pour le mois de juillet.

#### IV.4.3.3.2 Choix du type de transformation

Le modèle de reconstitution donne lieu à des valeurs de pluies négatives, il est donc impératif de procéder à la transformation des données initiales.

### IV.4.3.4 Simulation

#### IV.4.3.4.1 Ajustement des paramètres

Afin de mieux étudier le comportement des différents paramètres du modèle on choisit un mois sec le mois de juillet et un mois humide le mois d'Octobre et un mois sec le mois de juillet.

L'ajustement des CP, des résidus ainsi que des covariances, est donné en Annexe B.7. (1) pour le mois de décembre et en Annexe B.8 (1) pour le mois de juillet.

Les CP: Les quatre composantes présentent un bon ajustement à la loi normale.

Les covariances: Les covariances s'ajustent à la loi normale avec un seuil de 95% de confiance.

Les résidus: Les résidus s'ajustent bien à la loi normale

#### IV.4.3.4.2 Génération des paramètres

On maintient la génération dans la loi normale et ceci pour l'ensemble des paramètres.

Les résultats de la génération sont donnés en Annexe B.7. (2) pour le mois de décembre et en Annexe B.8 (2) pour le mois de juillet

Les CP: Les séries générées appartiennent à l'intervalle de confiance

Les covariances: la génération est acceptable dans l'ensemble, on remarque pour la troisième covariance une sous estimation des valeurs extrêmes, si on revient à l'ajustement de la troisième covariance calculé, on remarque que la dernière valeur est trop éloignée par rapport à la valeur précédente, est qu'elle est à la limite de l'intervalle, le processus de génération n'a pas pu reproduire cette valeur extrême.

Les résidus: Les résidus dans ce cas sont très faibles ( $\mu \approx 0, \sigma \approx 0.1$ ), le générateur de nombre aléatoire ne peut générer des valeurs aussi faibles avec un faible écart type.

IV.4.3.4.3 Résultats des simulations

L'approche adoptée pour la simulation des pluies journalières a été la même que celle des débits, le phénomène est également intermittent et difficile à modéliser. On maintient les même variantes utilisées pour la simulation des débits (Tableau IV.4-3).

Les résultats de la simulation des pluies avec les différentes variantes est données dans le Tableau IV.4-6.

Tableau IV.4-6: Résultats des simulation des pluies

	Valeurs annuelles			Valeurs mensuelles		Ajustement Val Max journalières
	$\bar{X}$	$\sigma$	Ajust	$\bar{X}$	$\sigma$	
Variante 1	+			+	-	
Variante 2	+			+	+	
Variante 3	+			+	-	
Variante 4	+	+	+	+	+	
Variante 5	+	+	-	+	+	+

La première variante permet d'avoir des séries simulées dont la moyenne valeurs annuelles appartient à l'intervalle de confiance mais l'écart type de ces dernières sont sous estimés (Figure IV.43), pour ce qui est des valeurs mensuelles, le test de la moyenne est vérifié (Figure IV.45), mais pas celui des écarts types, on observe une surestimation de la variabilité des mois humides (Figure IV.46). Les valeurs maximales de pluies journalières sont nettement au-dessous de l'intervalle de confiance (Figure IV.47), le maximum simulé ne dépasse guère 40 mm alors que le maximum historique est de 118 mm.

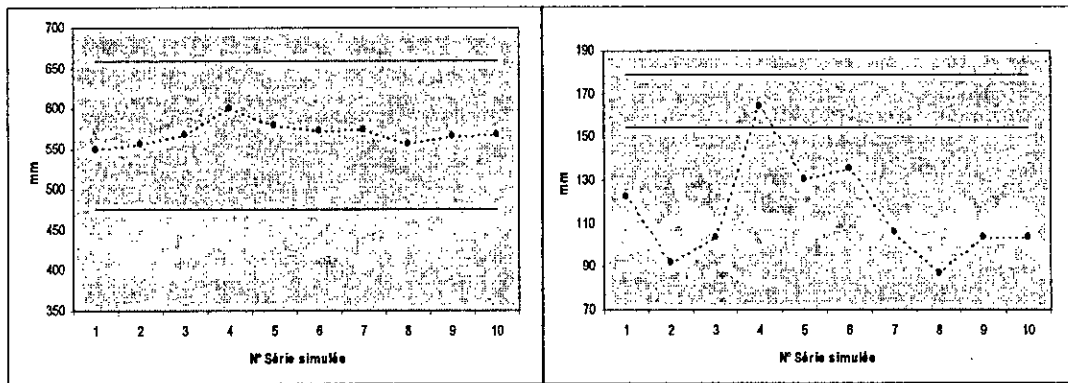


Figure IV-43: Tests des moyennes et écarts types des valeurs de pluies annuelles (V1).

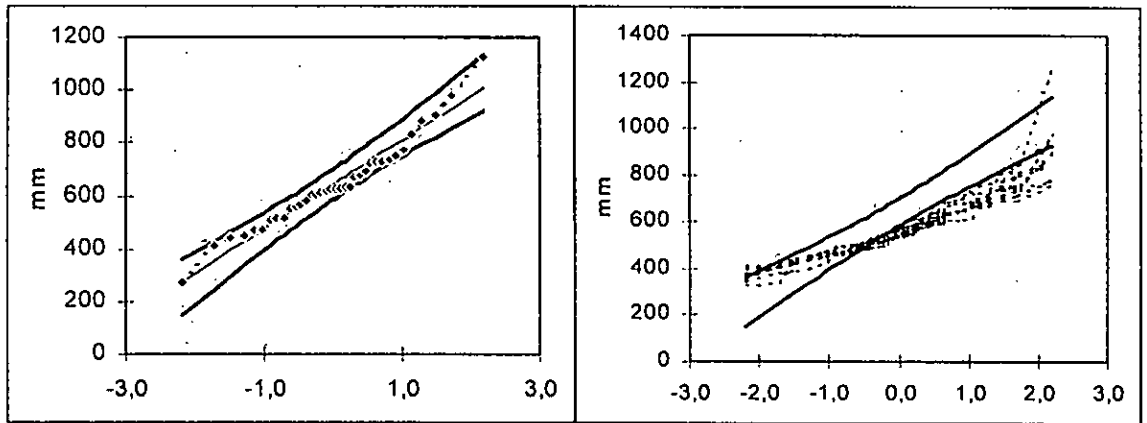


Figure IV-44: Ajustement des valeurs annuelles de pluies observées et simulées (V1).

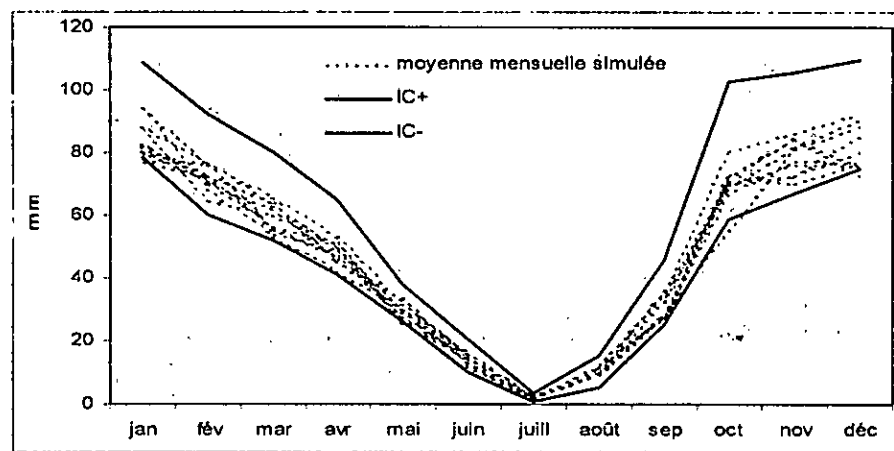


Figure IV-45: Test des moyennes des valeurs de Pluies mensuelles (V1).

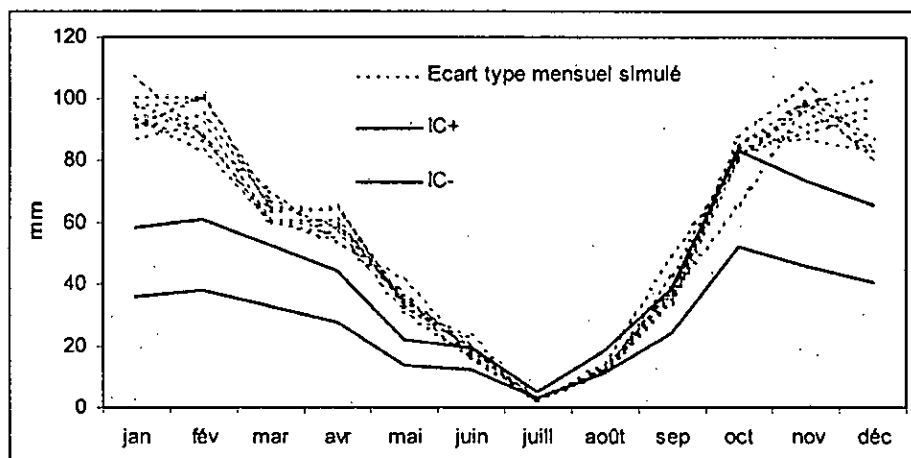


Figure IV-46: Test des écart types des valeurs de Pluies mensuelles (V1).

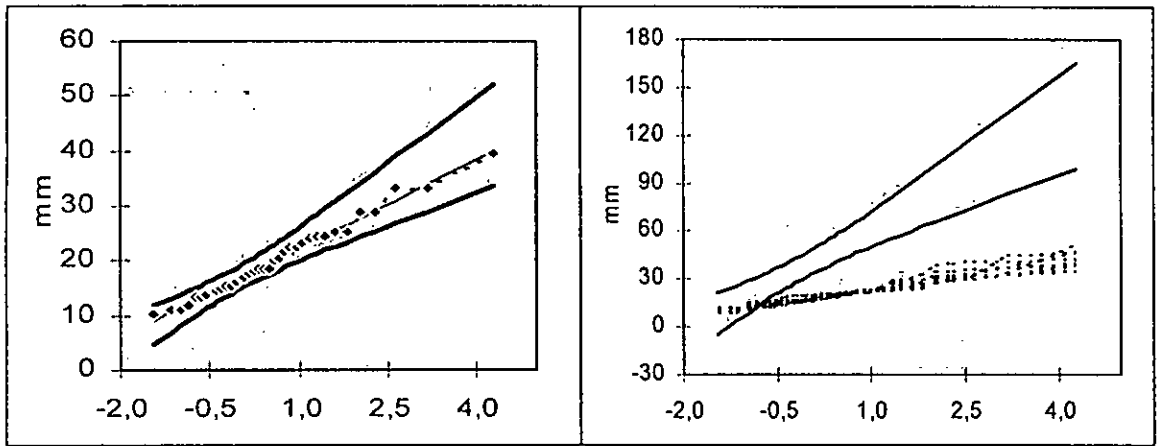


Figure IV-47: Ajustement des pluies maximales observées et simulées (Loi de Gumbel) (V1).

La seconde variante présente une amélioration de la variabilité aussi bien pour la moyenne des valeurs annuelles (Figure IV-48) que pour les valeurs mensuelles (Figure IV-51). Les pluies maximales journalières s'améliorent en présentant une moyenne des valeurs extrêmes de 60 mm avec un maximum de 180 mm (Figure IV-52), mais ceci reste au-dessous des résultats escomptés.

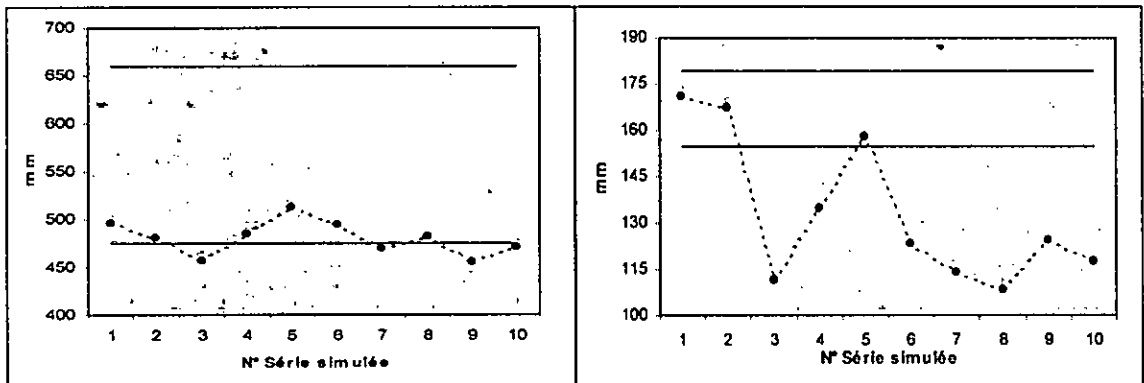


Figure IV-48: Tests des moyennes et écarts types des valeurs de pluies annuelles (V2).

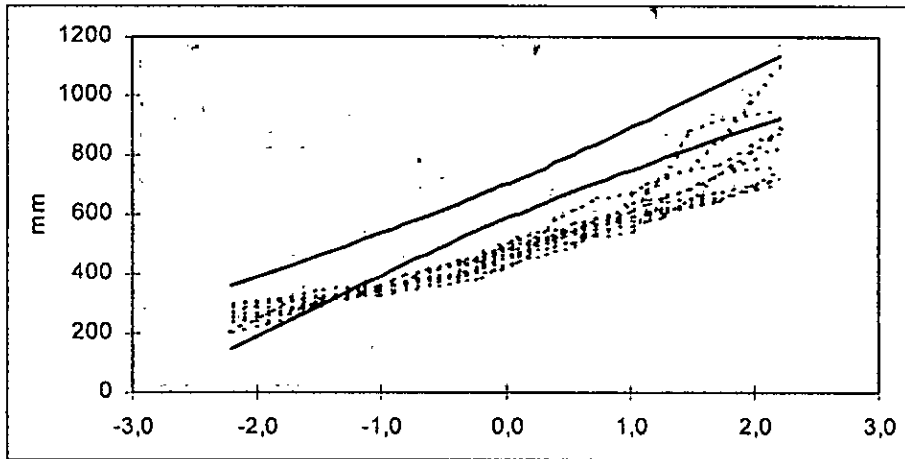


Figure IV-49: Ajustement des valeurs annuelles de Pluies simulées (V2).

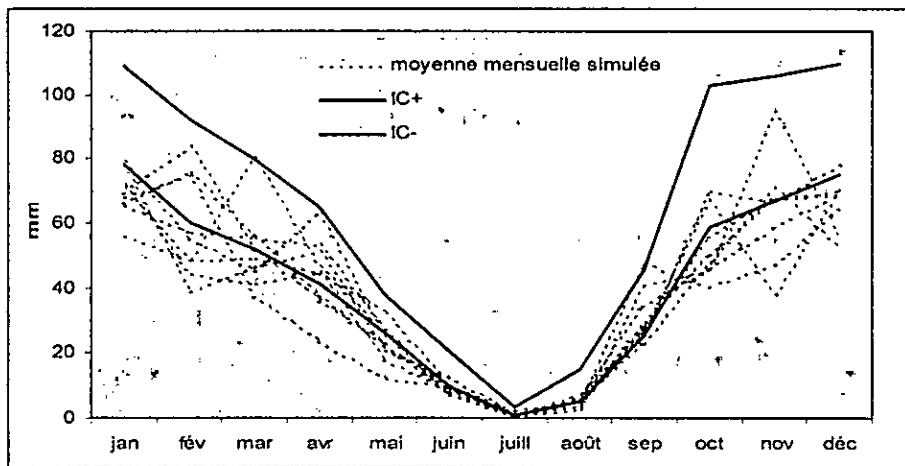


Figure IV-50: Test des moyennes des valeurs de Pluies mensuelles (V2).

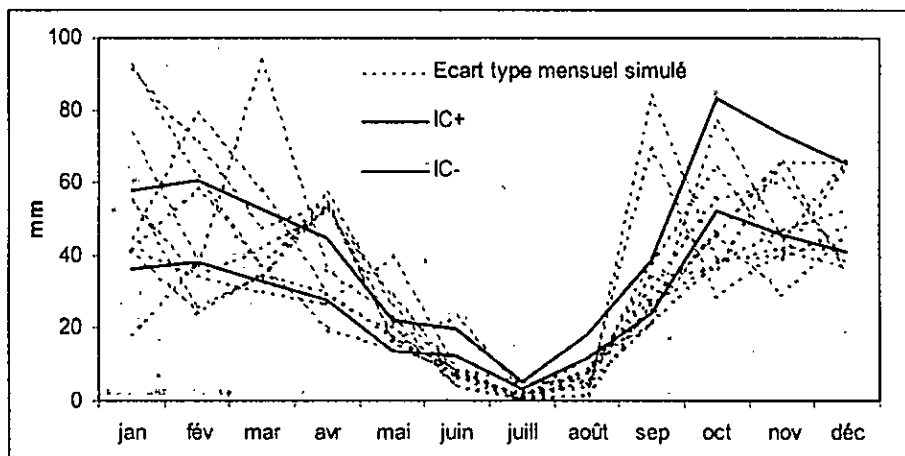


Figure IV-51: Test des écart types des valeurs de Pluies mensuelles (V2).

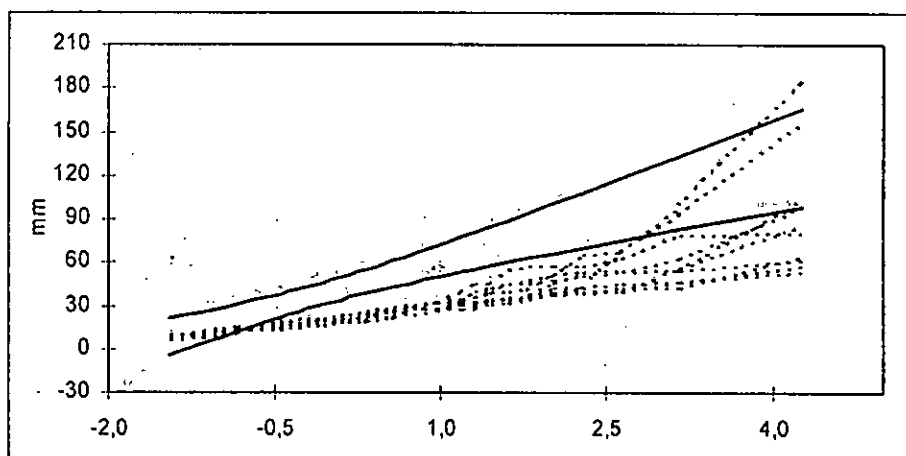


Figure IV-52: Ajustement des pluies maximales observées et simulées (Loi de Gumbel) (V2).

Avec la variante trois on passe a une surestimation des écarts types des moyennes des valeurs annuelles (Figure IV-53) et des valeurs mensuelles. On a obtenue de meilleures résultats pour les valeurs journalières maximales (Figure IV-57).

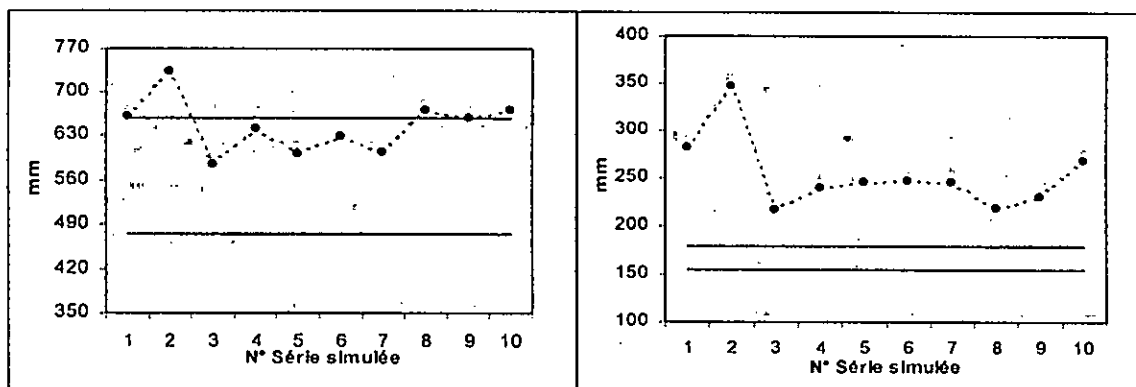


Figure IV-53: Tests des moyennes et écarts types des valeurs annuelles (V3).

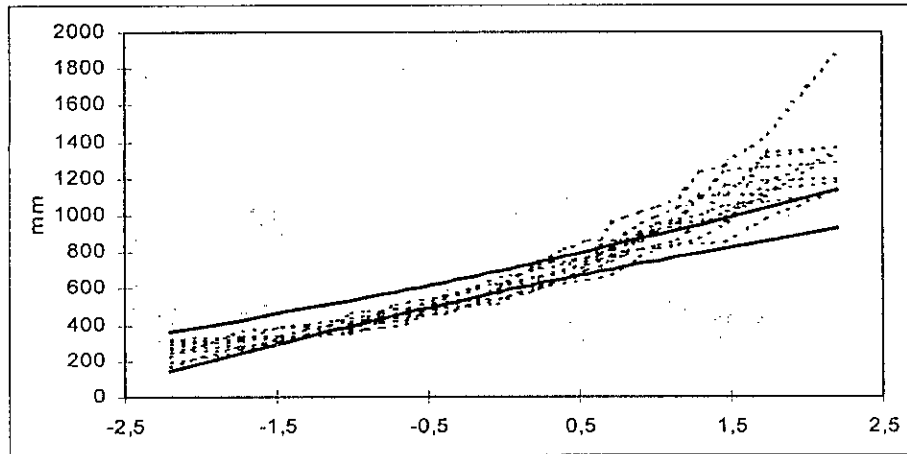


Figure IV-54: Ajustement des valeurs annuelles de Pluies observées et simulées (V3).

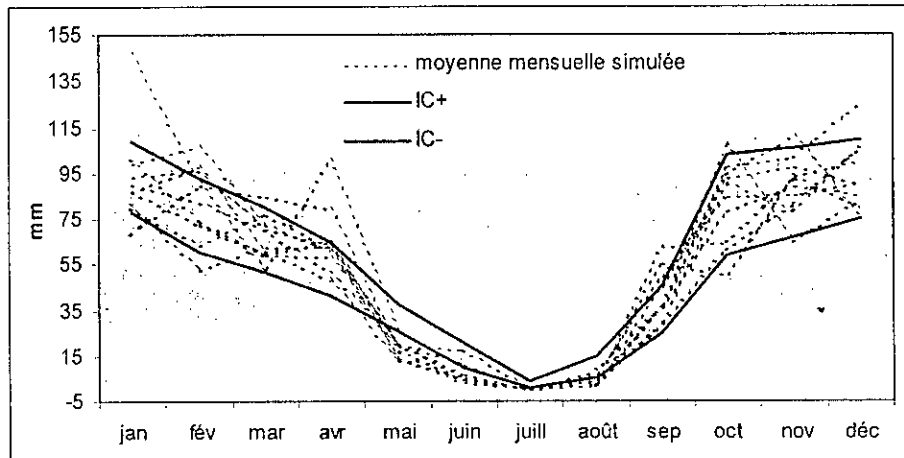


Figure IV-55: Test des moyennes des valeurs de Pluies mensuelles (V3).

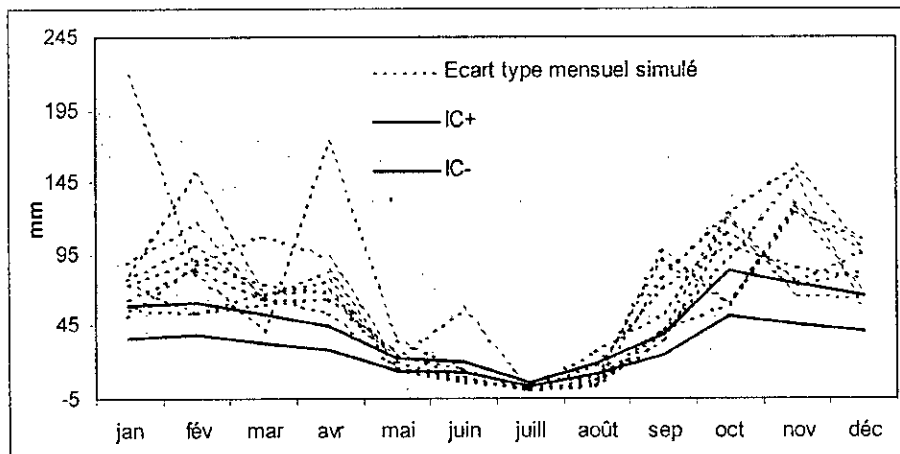


Figure IV-56: Test des écart types des valeurs de Pluies mensuelles (V3).



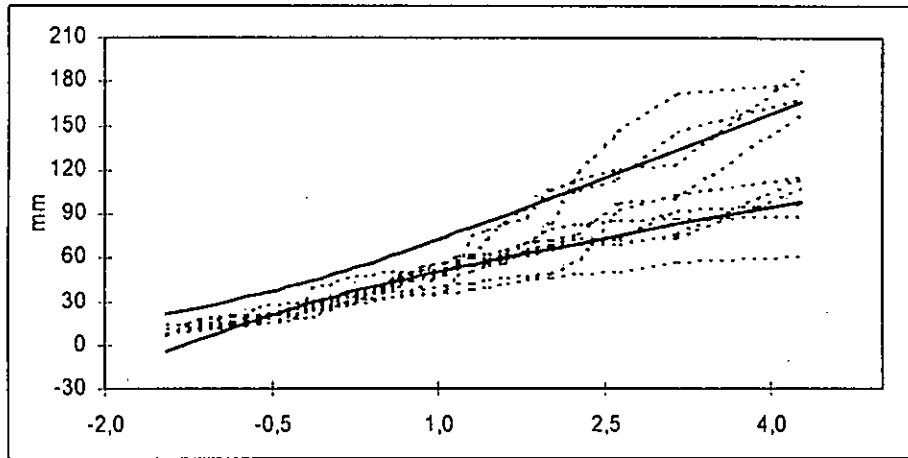


Figure IV-57: Ajustement des pluies maximales observées et simulées (Loi de Gumbel) (V3).

La variante quatre donne de bon résultats en terme de valeurs annuelles (Figure IV-58) et de valeurs mensuelles (Figure IV-60 et Figure IV-61) mais les maximums journaliers sont au dessus de l'intervalle de confiance (Figure IV-62), un maximum de 430 mm est observé.

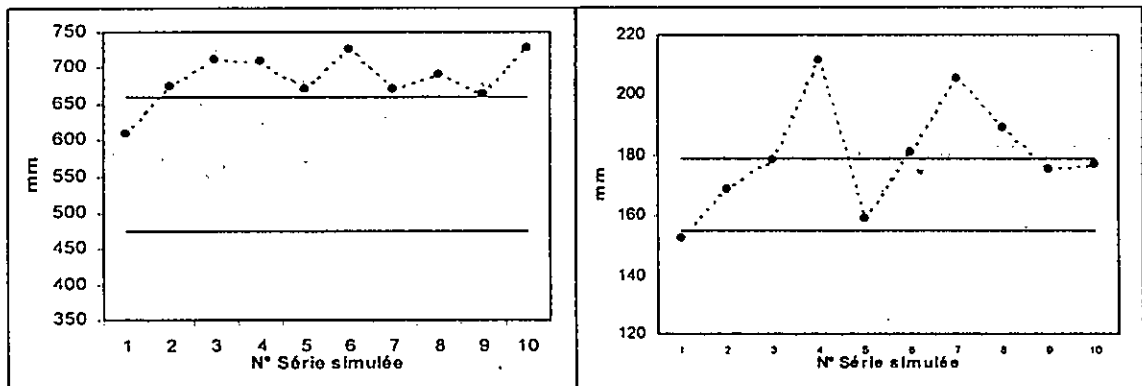


Figure IV-58: Tests des moyennes et écarts types des valeurs de pluies annuelles (V4).

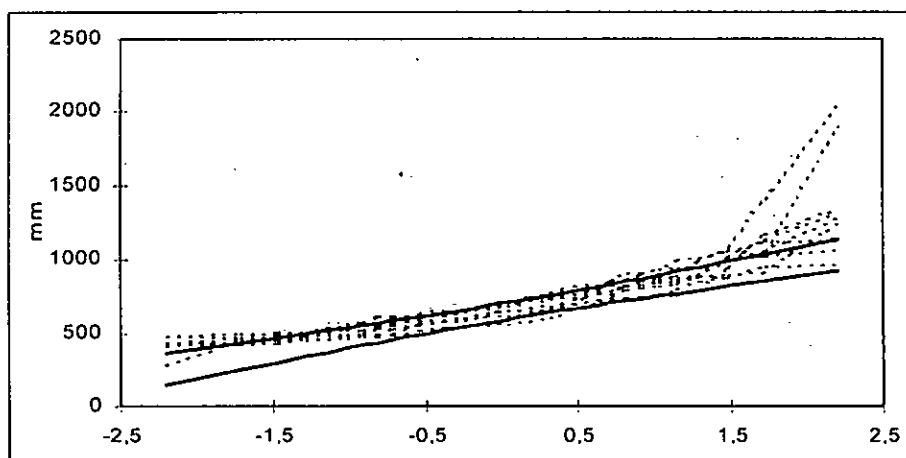


Figure IV-59: Ajustement des valeurs annuelles de Pluies simulées (V4).

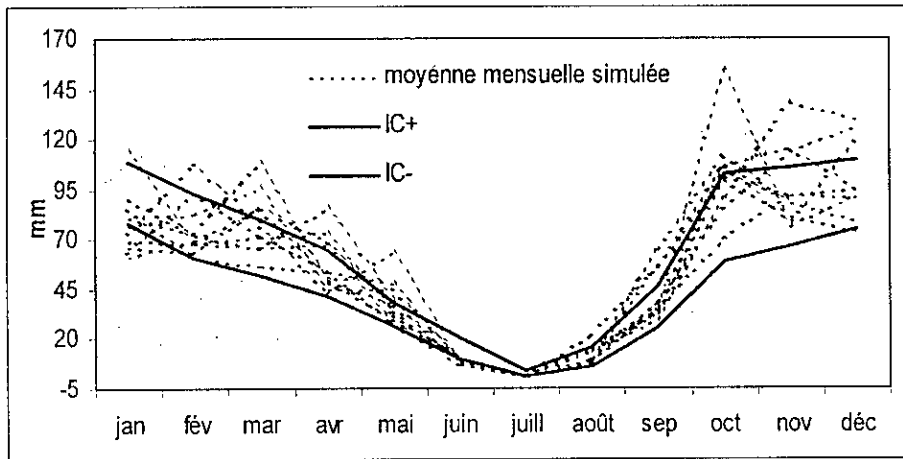


Figure IV-60: Test des moyennes des valeurs de Pluies mensuelles (V4).

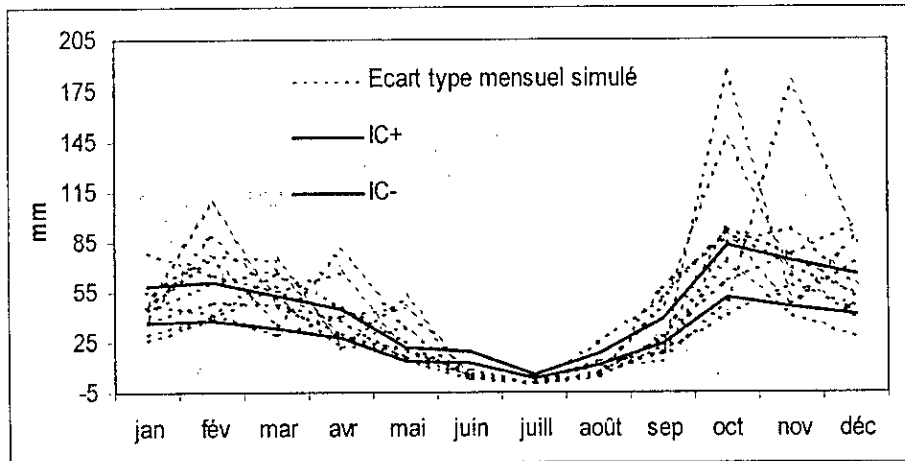


Figure IV-61: Test des écart types des valeurs de Pluies mensuelles (V4).

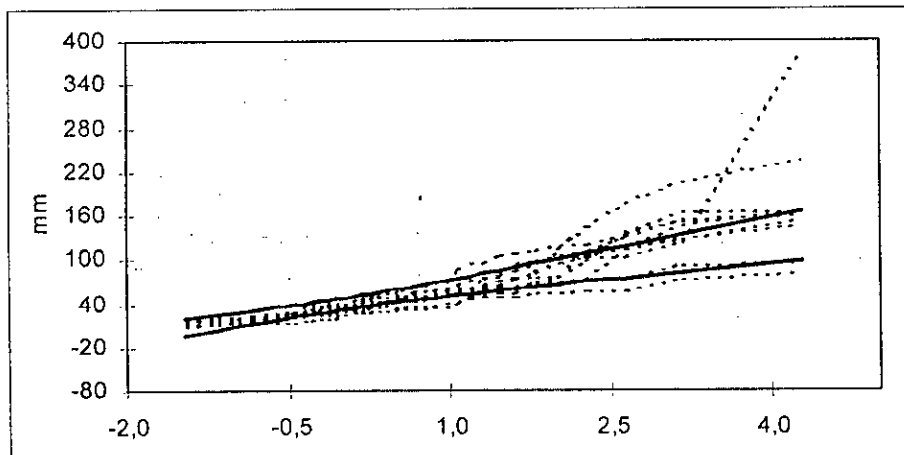


Figure IV-62: Ajustement des pluies maximales observées et simulées (Loi de Gumbel) (V4).

La variante finale qui combine la génération dans la loi normale et la loi de Goodrich permet d'avoir des résultats acceptables pour les différents tests (de Figure IV-63 à Figure IV-67).

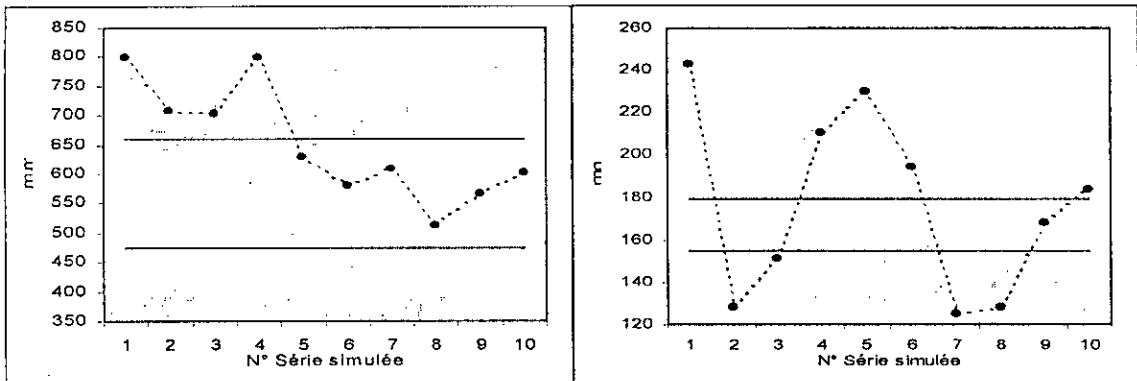


Figure IV-63: Tests des moyennes et écarts types des valeurs de pluies annuelles (V5).

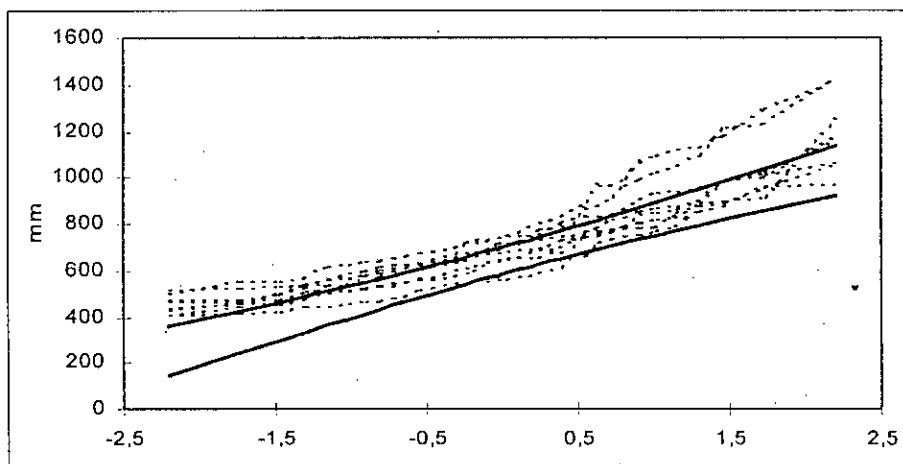


Figure IV-64: Ajustement des valeurs annuelles de Pluies observées et simulées (V5).

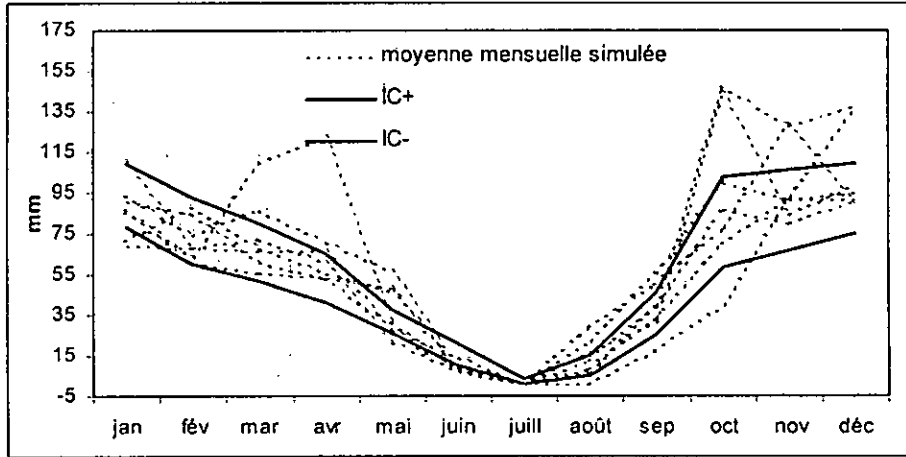


Figure IV-65: Test des moyennes des valeurs de Pluies mensuelles (V5).

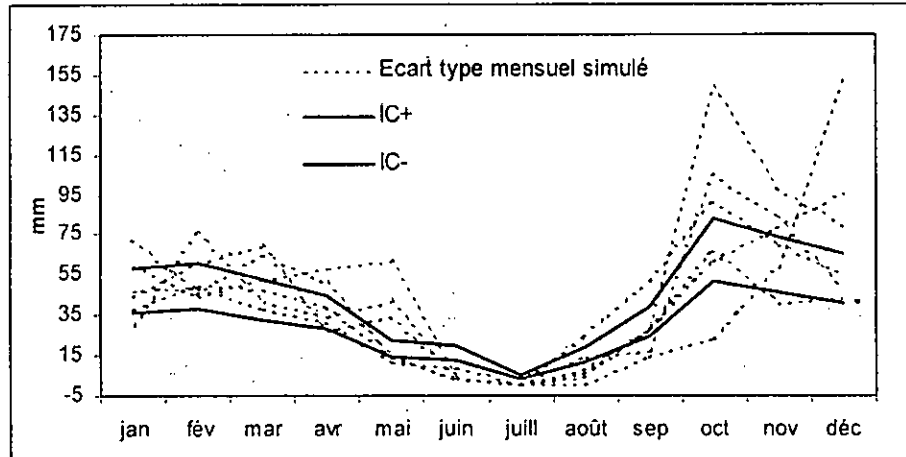


Figure IV-66: Test des écart types des valeurs de Pluies mensuelles (V5).

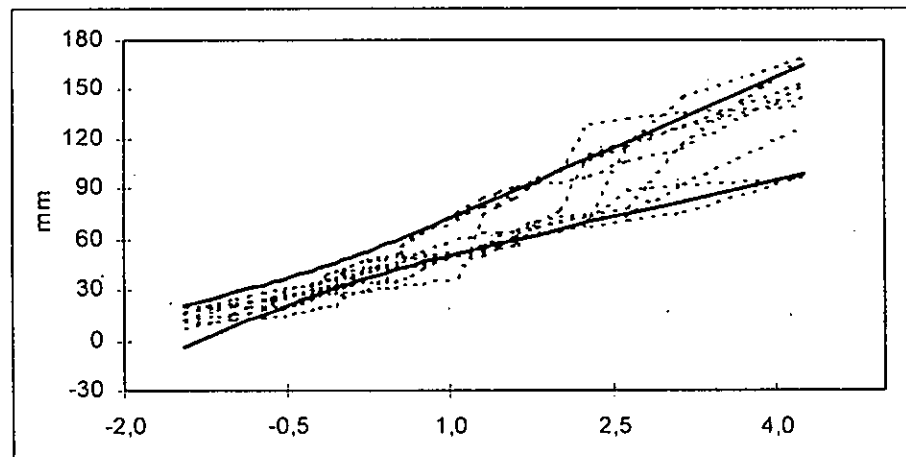
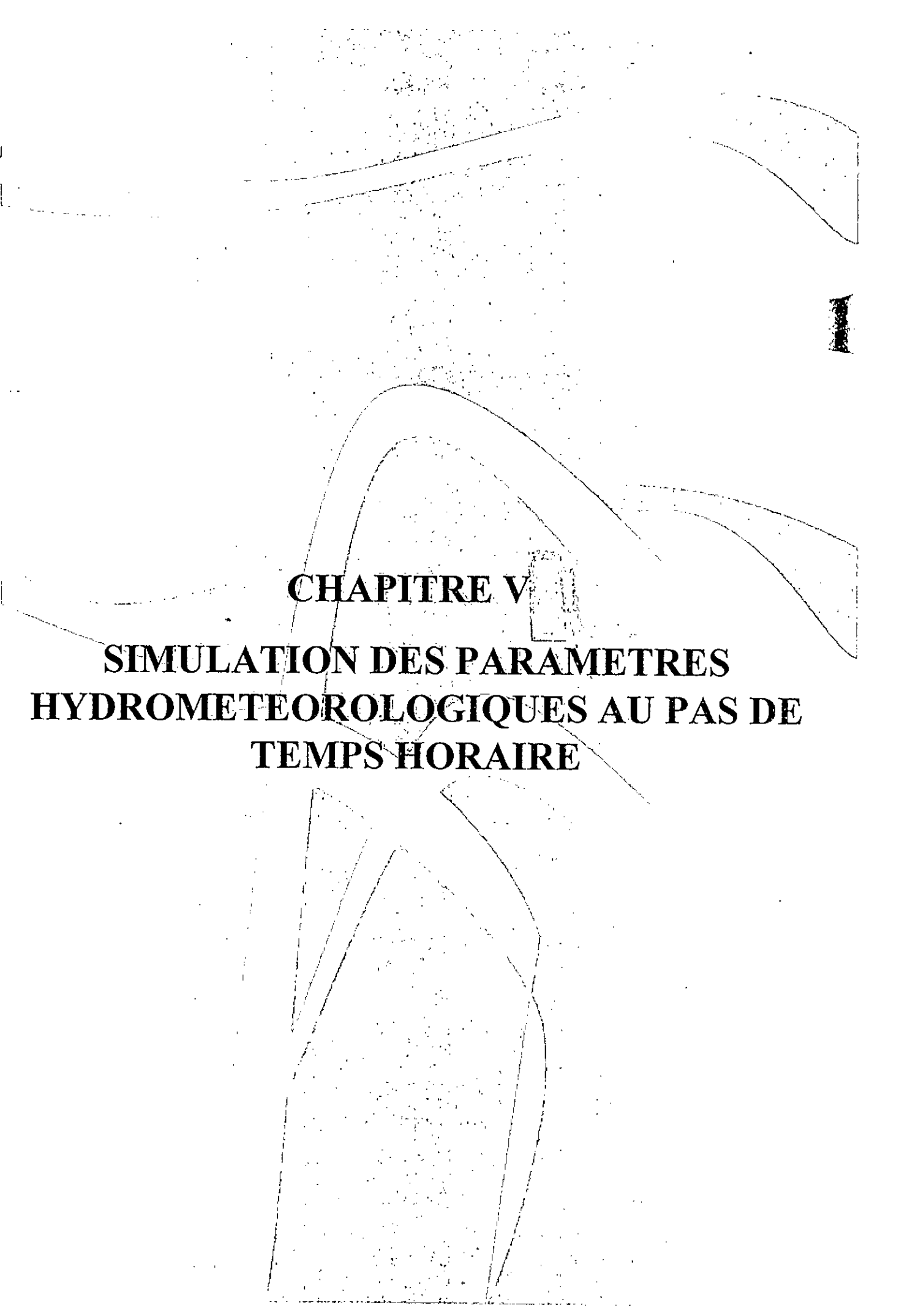


Figure IV-67: Ajustement des pluies maximales simulées (Loi de Gumbel) (V5).



**CHAPITRE V**  
**SIMULATION DES PARAMETRES**  
**HYDROMETEOROLOGIQUES AU PAS DE**  
**TEMPS HORAIRE**

## CHAPITRE V SIMULATION DES PARAMETRES HYDROMETEOROLOGIQUES AU PAS DE TEMPS HORAIRE

### INTRODUCTION

Le passage du pas de temps journalier (lors de la simulation des variables hydrométéorologiques) au pas de temps horaire, rend les corrélations entre variables successives plus importantes, et introduit de nouveaux aspects des phénomènes hydrométéorologiques, non visibles à l'échelle mensuelle et journalière.

### V.1. SIMULATION DES ETP

#### V.1.1 Données utilisées

On dispose d'une campagne de mesure d'ETP à l'échelle horaire allant de la 41<sup>ème</sup> semaine de l'année 1991 à la 13<sup>ème</sup> semaine de l'année 1992, on omettant les jours lacunaires, la série d'ETP compte 76 jours d'observations non lacunaires. L'examen de l'évolution chronologique des ETP au pas de temps horaire (Annexe C.1) montre l'aspect cyclique des ETP de période 24 heures mais avec des amplitudes différentes, ce qui est dû à l'effet saisonnier vu que la série va de la mi-octobre à la fin du mois de mars.

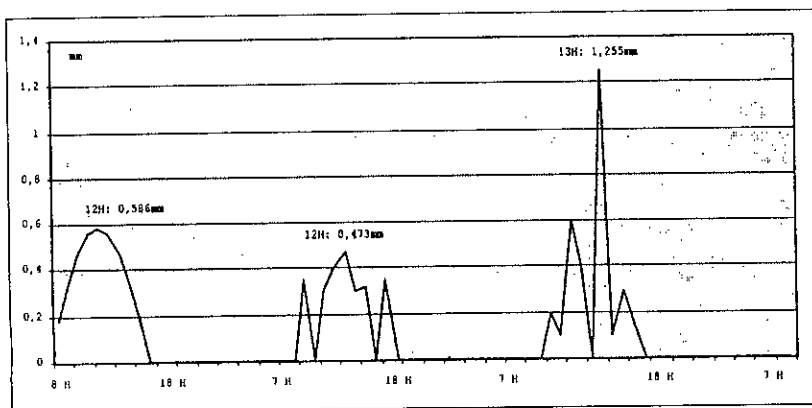
La matrice de données est de dimensions, équivalent à 76 jours d'observations fois 24 heures, les caractéristiques statistiques de la série d'ETP sont données ci-après.

Tableau V.1-1 : Caractéristiques statistiques de la série d'ETP horaires

	8 H	9 H	10 H	11 H	12 H	13 H	14 H	15 H	16 H	17 H	18 H	19 H
Moyenne (mm)	0,0	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0
Ecart type (mm)	0,1	0,1	0,1	0,2	0,2	0,2	0,1	0,1	0,1	0,0	0,0	0,0

	20 H	21 H	22 H	23 H	24 H	1 H	2 H	3 H	4 H	5 H	6 H	7 H
Moyenne (mm)	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Ecart type (mm)	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Le tableau V.1.1 permet d'observer la différence entre le jour et la nuit, entre 18 heures et 7 heures du matin, les ETP sont en moyennes nulles, pour mieux approcher ce phénomène ont reporté sur un graphe l'évolution de l'ETP sur trois jours (Figure V.1), on remarque une variation en cloche allant de 8H du matin à environ 18H le soir, marquant un maximum vers midi, pour quasiment s'annuler entre 18H et 7H du matin. Lorsque les conditions climatiques changent par exemple une pluie est enregistrée, l'ETP chute directement à zéro pour remonter juste après la fin de l'averse.



FigureV.1 : Evolution de l'ETP horaire sur 72 heures

**V.1.2 Résultats de l'ACP**

On a effectué une ACP centrée sur les données transformées par racine carré, la première CP monopolise toujours le maximum de variance expliquée près de 77% (Tableau V.1.2), suivi de la seconde CP avec 9%. Cette concentration de l'information dans les premières CP est une caractéristique des phénomènes cycliques. On a procédé à la simulation avec 2 CP ce qui implique un cumul d'information de 85%.

Tableau V.1-2 : Les CP calculées sur la série d'ETP (ACP centrée, données transformées)

N° CP	CP		N° CP	CP	
	Valeurs (%)	Cumul (%)		Valeurs (%)	Cumul (%)
1	76,8	76,8	13	0,0	100
2	8,4	85,3	14	0,0	100
3	5,4	90,7	15	0,0	100
4	3,5	94,2	16	0,0	100
5	2,3	96,5	17	0,0	100
6	1,6	98,1	18	0,0	100
7	1,5	99,6	19	0,0	100
8	0,2	99,8	20	0,0	100
9	0,1	100	21	0,0	100
10	0,0	100	22	0,0	100
11	0,0	100	23	0,0	100
12	0,0	100	24	0,0	100

**V.1.3 Résultats de la simulation des ETP**

Les résultats de la simulation des données horaires (Figure V-2) montre une légère surestimation des moyennes, les écarts type sont acceptables. On examine l'ajustement des différents paramètres (Annexe C.2), on a observé que la première CP s'ajuste bien à la loi normale, mais l'ajustement de la seconde CP est juste moyen, cependant les covariances est les résidus sont bien ajustés.

On remarque également que les paramètres générés suivent la courbe théorique de l'ajustement (Annexe C.3) et néglige les points qui n'appartiennent pas à l'intervalle de confiance.

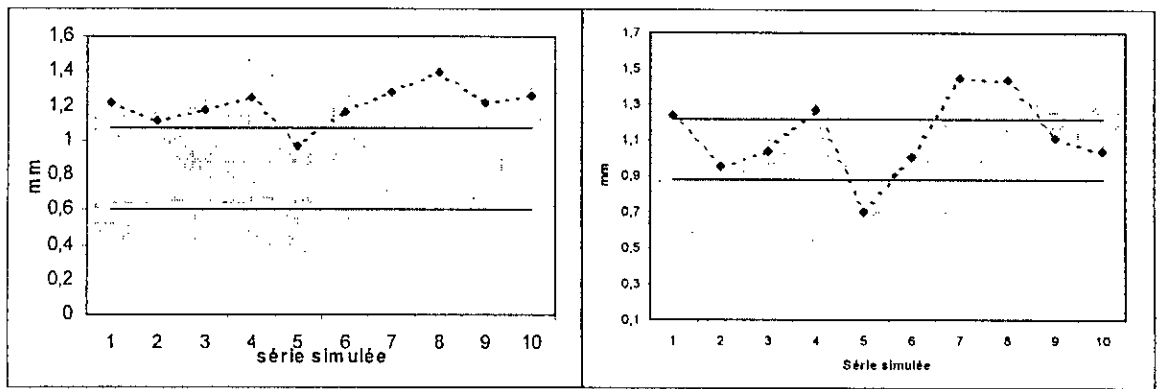


Figure V.2 : Test de la moyenne et l'écart type des valeurs journalières (ETP)

On essaye comme seconde variante la génération avec la fonction de répartition, les moyennes et les écarts types se rangent dans l'intervalle de confiance (Figure V-3).

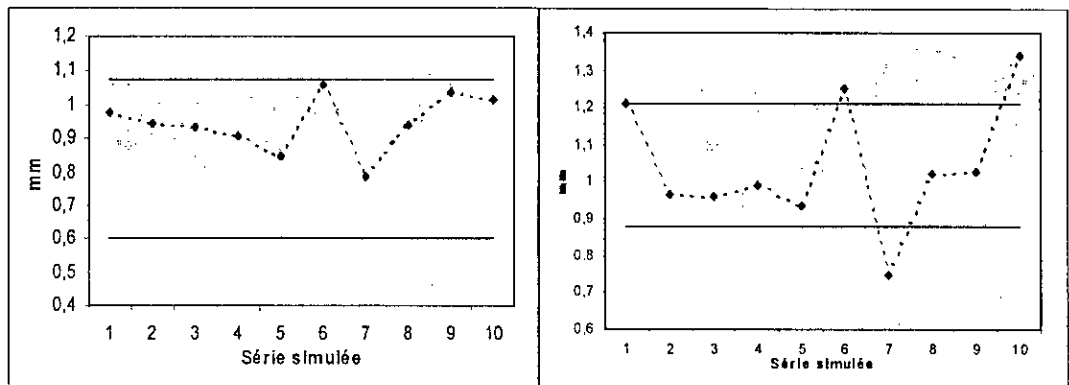


Figure V.3 : Test de la moyenne et l'écart type des valeurs journalières (ETP)

La génération de la seconde CP par la fonction de répartition permet de mieux ressortir les particularités de cette dernière (Figure V.4).

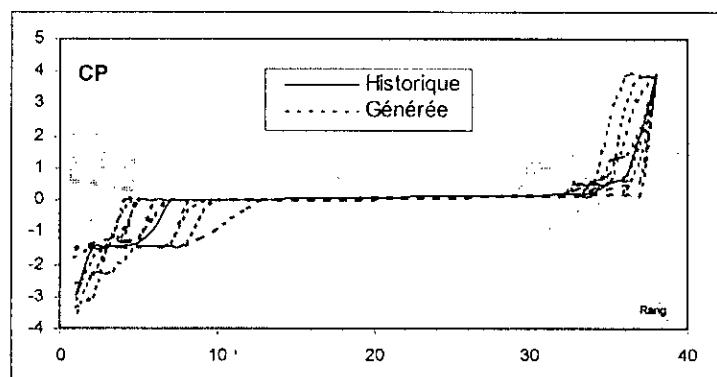


Figure V.4 : génération de la seconde CP par la fonction de répartition



Le phénomène ETP englobe deux types d'évaporation : l'évaporation de la l'eau qui imprègne les couches superficielles du terrain (humidité du sol), soit qu'elle provienne des pluies récentes infiltrées à faibles profondeurs, soit qu'elle remonte par capillarité de la nappe phréatique, l'eau s'évapore aussi des surfaces d'eau stagnantes ou courantes (mares, étangs, rivières, lacs...) ; le second type d'évaporation est la transpiration de la couverture végétale à travers les pores que présentent les plantes en général sur leur feuillage. L'ETP présente des variations diurnes en plus des variations saisonnières

**Variations saisonnières** : elles sont à la fois liées à l'activité végétative de la plante et aux variations du pouvoir évaporant de l'atmosphère : comparons la transpiration d'un sol couvert de végétation et l'évaporation d'une surface libre ; pendant la période de croissance des plantes (saison chaude) la transpiration varie à peu près comme la l'évaporation par contre, en dehors de la période de végétation (après la mort des plantes annuelles ou la chute des feuilles), la transpiration cesse alors que l'évaporation de l'eau continue à une allure réduite par la diminution du pouvoir évaporant de l'atmosphère.

**Variations diurnes** : elles sont liées aux variations de la température, de l'humidité et plus particulièrement de l'intensité de la lumière. Pendant la journée la transpiration varie à peu près comme l'évaporation jusqu'à la tombée du jour ; à ce moment, elle devient pratiquement nulle en raison de la fermeture des stomates des feuilles et ne reprendra qu'au lever du soleil.

Cet aspect du phénomène ETP, qui ne peut être appréhender, qu'a l'échelle horaire, explique pourquoi la méthode de simulation selon qu'on travail en journalier ou en horaire. On observe en Annexe C.4 l'évolution d'une série d'ETP simulées, on note la reproduction de l'effet saisonnier, tout en maintenant la cyclicité du phénomène de façon moins lissée que la série d'origine.

## V.2. SIMULATION DES PLUIES

### V.2.1 Les données utilisées

Les données de pluies horaires disponibles vont de la 47<sup>ième</sup> semaine de l'année 1989 à la 24<sup>ième</sup> semaine de l'année 1990 (Annexe C.4), on enlève les jours lacunaires pour obtenir 162 jours complets sur un total de 210 jours.

Les caractéristiques statistiques de la série de pluies sont données ci-après.

Tableau V.2.1 : Caractéristiques statistiques de la série de pluies horaires

	1 H	2 H	3 H	4 H	5 H	6 H	7 H	8 H	9 H	10 H	11 H	12 H
Moyenne (mm)	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,1	0,1
Ecart type (mm)	0,2	0,3	0,3	0,3	0,3	0,4	0,7	0,4	0,4	0,7	0,4	0,7
Maximum (mm)	1,1	1,9	2,0	2,1	1,8	2,3	5,6	2,8	4,4	6,4	2,9	6,5

	13 H	14 H	15 H	16 H	17 H	18 H	19 H	20 H	21 H	22 H	23 H	24 H
Moyenne (mm)	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,1	0,1	0,1	0,1
Ecart type (mm)	0,5	0,2	0,4	0,3	0,5	0,7	0,7	0,6	0,5	0,4	0,4	0,4
Maximum (mm)	3,1	1,9	2,5	2,1	4,4	4,9	6,4	3,8	3,6	3,1	2,4	3,8

De part la nature aléatoire du phénomène, aucune conclusion ne peut être faite, afin d'expliquer la variabilité observée.

### V.2.2 Résultats de l'ACP

On a effectué une ACP centrée sur les données pluviométriques transformées par la fonction racine carré, les résultats sont donnés dans le tableau V.2.2, la première CP apporte environ 47%, on a choisi quatre CP ce qui signifie 82% d'informations.

Tableau V.2.2 Les CP calculées sur la série de pluies (ACP centrée, données transformées)

N° CP	CP		N° CP	CP	
	Valeurs (%)	Cumul (%)		Valeurs (%)	Cumul (%)
1	46,6	46,6	13	0,4	99,0
2	15,9	62,5	14	0,4	99,3
3	12,1	74,7	15	0,2	99,6
4	6,8	81,5	16	0,2	99,8
5	4,7	86,1	17	0,1	99,9
6	3,5	89,7	18	0,1	100
7	2,5	92,2	19	0,0	100
8	2,0	94,2	20	0,0	100
9	1,7	95,9	21	0,0	100
10	1,1	97,1	22	0,0	100
11	0,9	97,9	23	0,0	100
12	0,6	98,6	24	0,0	100

### V.2.3 Résultats de la simulation

L'ajustement des paramètres (Annexe C.5) a mené aux constatations suivantes, la première CP présente des extrémités qui sortent de l'intervalle de confiance, pour le reste des paramètres l'ajustement est bon, on voit en Annexe C.6 l'ajustement des paramètres générés.

On maintient la variante qui a donné les meilleurs résultats dans le chapitre IV qui consiste à générer la première CP dans la loi normale, les autres CP avec la fonction de Goodrich, les covariances dans la loi normale et les résidus par la fonction de Goodrich.

Les moyennes appartiennent à l'intervalle de confiance, alors que les écarts types sont au-dessous (Figure V-5). Ceci est du à la qualité de l'ajustement de la première CP, la figure montre que cette CP présente des valeurs extrêmes n'appartenant pas à l'intervalle de confiance. Donc la génération par la loi normale sous-estime cette composante.

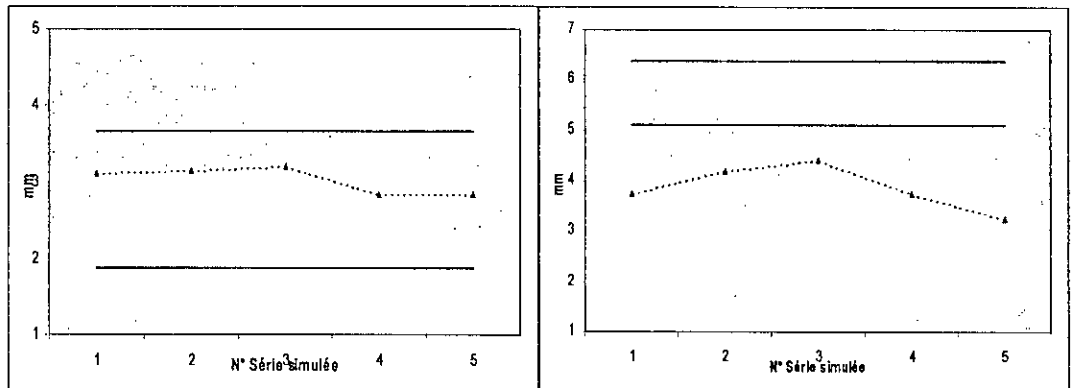


Figure V-5 : Test de la moyenne et l'écart type des valeurs journalières (Pluies)

On propose comme seconde variante, de générer les CP et les résidus par la fonction de Goodrich, les covariances par la loi normale. On observe une amélioration des écarts types (Figure V-6).

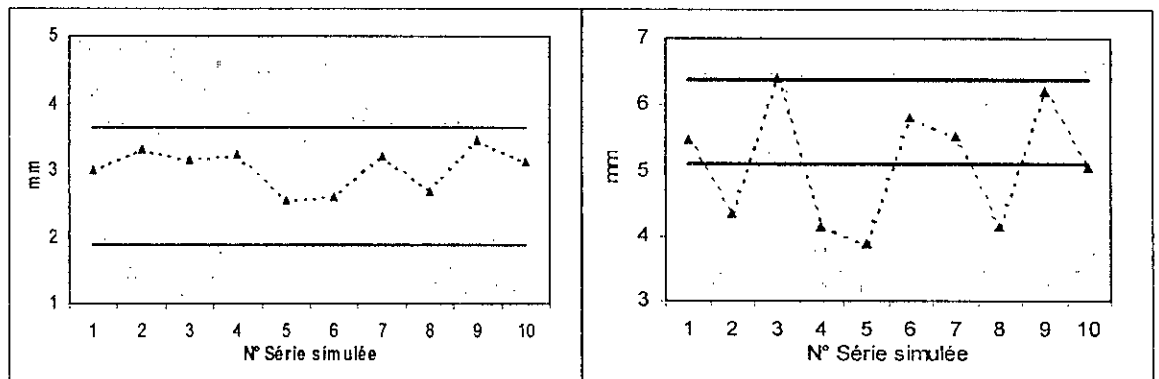


Figure V-6 : Test de la moyenne et l'écart type des valeurs journalières (Pluies)

Effectivement en générant la première CP dans la loi de Goodrich on arrive à restituer les valeurs extrêmes de cette CP négligé lors de la génération dans la loi d'ajustement (Figure V-7).

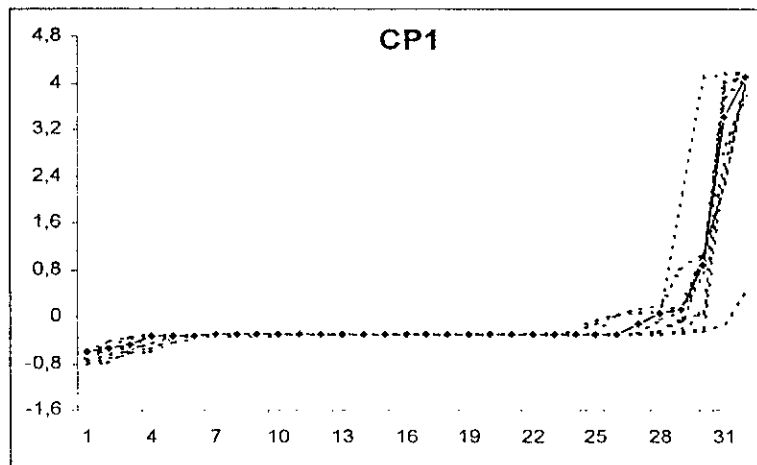


Figure V-7: Génération de la première CP par la loi de Goodrich

## **CHAPITRE VI**

# **ETUDE COMPARATIVE DES MODELES DE SIMULATION EN HYDROLOGIE**

## CHAPITRE VI ETUDE COMPARATIVE DES MODELES DE SIMULATION EN HYDROLOGIE

### INTRODUCTION

Quatre modèles de simulation ont été pris en considération dans ce chapitre à savoir, les modèles Markoviens (binaires et multiclassés), le modèle de la décomposition de la série historique, le modèle SVD (Singular Value Décomposition) et le modèle GESTOP. Cette étude comparative a pour but de ressortir les avantages et les inconvénients, ainsi que les contraintes d'utilisation de chaque modèle.

Les mêmes critères adoptés pour l'acceptation ou le rejet des séries simulées par le modèle GESTOP (tests sur les valeurs annuelles, mensuelles et maximales journalières annuelles) ont été utilisés dans le cadre de cette étude comparative.

### VI.1. SIMULATION DES ETP

Les ETP journalières suivent un phénomène cyclique, c'est pour cela que les modèles Markoviens ne peuvent pas être utilisés dans ce cas (ces derniers exigent l'intermittence du phénomène étudié). Les modèles GESTOP, SVD et décomposition de la série historique (en tenant compte des variantes à 2, 4 et 6 harmoniques) ont été appliqués aux ETP. Il en ressort que tous ces modèles donnent des résultats très satisfaisant (Figure VI-1, Figure VI-2, Figure VI-3, Figure VI-4 et Figure VI-5), en effet toutes les moyennes et les écarts types annuels et mensuels des séries simulées sont à l'intérieur de l'intervalle de confiance ( $\alpha = 5\%$ ) de la série historique, ainsi que les ajustements des ETP annuelles et maximales journalières annuelles. Toutes les séries simulées ont été acceptées et cela pour l'ensemble des modèles, cela dénote la relative simplicité de la simulation des variables hydrométéorologiques cycliques (ETP, température, insolation, etc....) même au pas de temps journalier du fait de la très faible variabilité de ces phénomènes saisonniers, et de l'inexistence de valeurs maximales brusques.

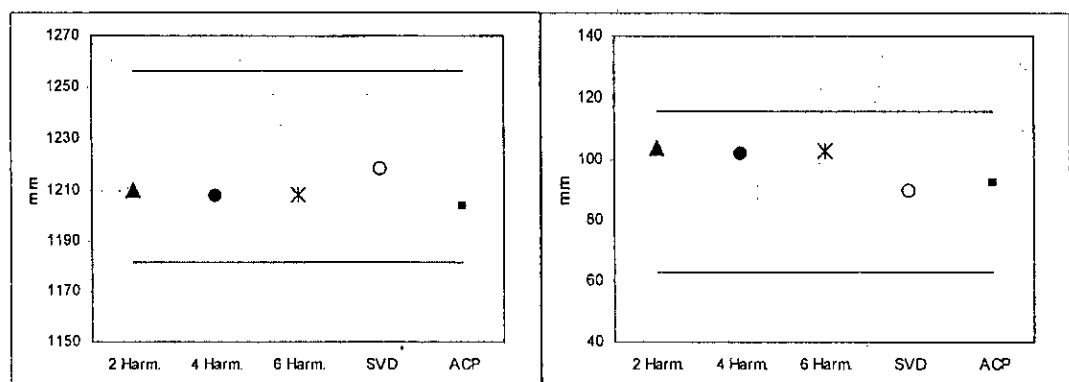


Figure VI-1: Tests sur les moyennes et écarts types des valeurs d'ETP annuelles

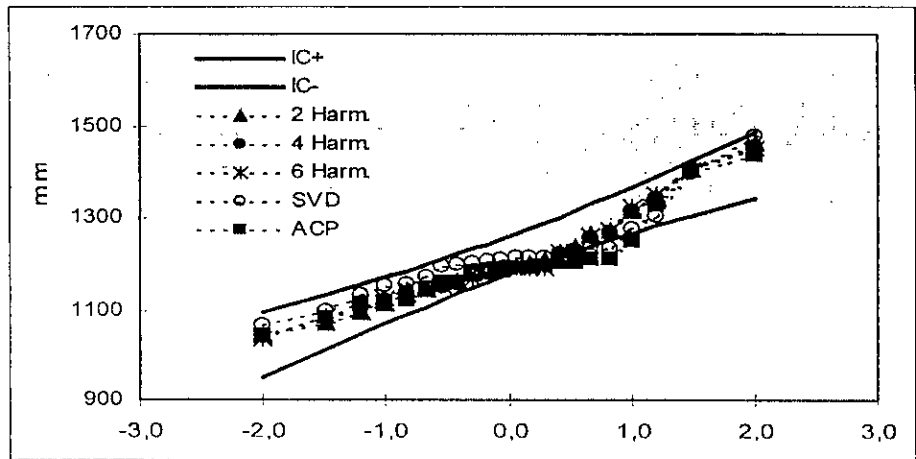


Figure VI-2 : Ajustement des valeurs d'ETP annuelles

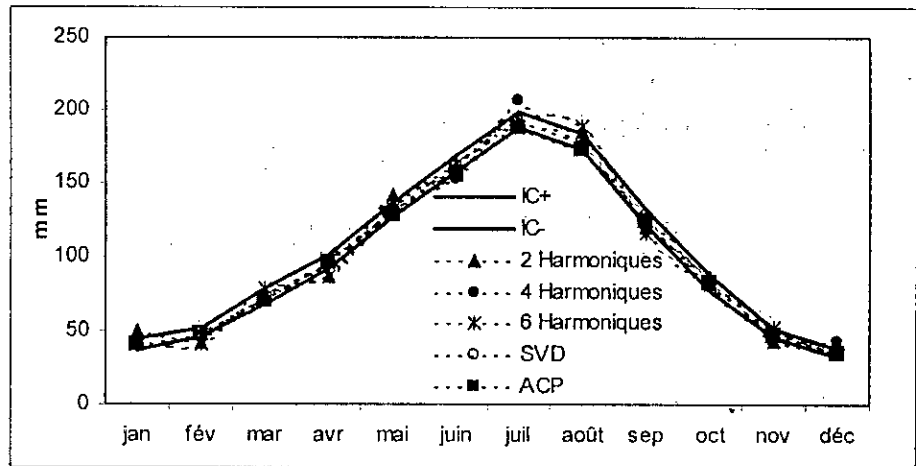


Figure VI-3 : Tests sur les moyennes des valeurs mensuelles

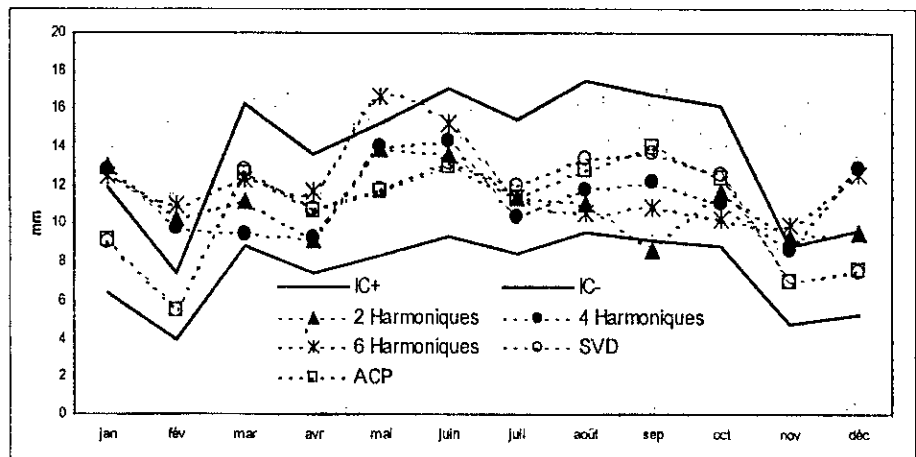


Figure VI-4 : Tests sur les écarts types des valeurs d'ETP mensuelle

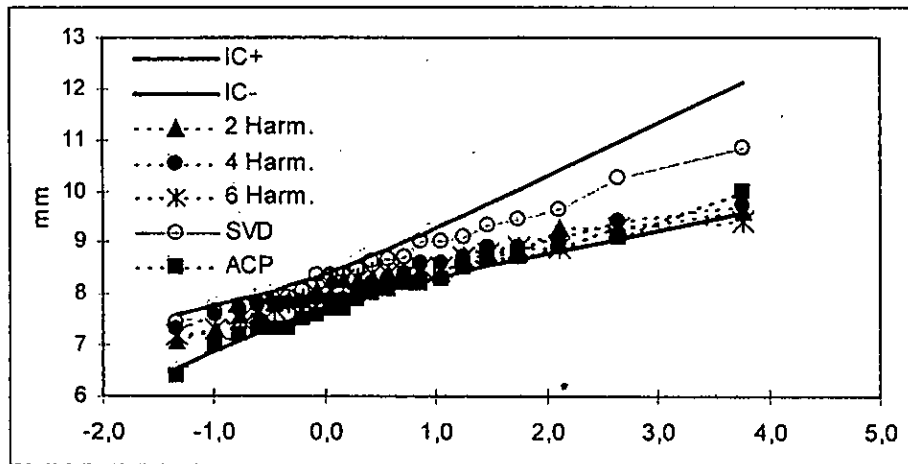


Figure VI-5 : Ajustement des valeurs d'ETP maximales journalières

## VI.2. SIMULATION DES DEBITS

Le modèle de la décomposition de la série historique n'est pas utilisable pour la simulation de débits du fait qu'il exige la cyclicité du phénomène étudié. Dans ce cas on a pris en considération les modèles Markoviens multiclassés, le modèle SVD et le modèle GESTOP. En ce qui concerne le modèle Markovien multiclassés, la subdivision de la série historique c'est faite suivant trois variantes (3 classes).

Aléatoirement (Cas 1)

- C1 : 0 – 0,0001 hm<sup>3</sup>
- C2 : 0,0001 – 9,9999 hm<sup>3</sup>
- C3 : > 9,999 hm<sup>3</sup>

Les classes (Cas 2) ont été subdivisées en fonction de la moyenne et de l'écart type de la série historique, dans ce cas on a subdivisé cette dernière en quatre classes

- Les débits nuls ( C1 < 0,001 hm<sup>3</sup>)
- Les débits entre  $\bar{X} - \sigma$  et  $\bar{X} + \sigma$  (0,0001 < C2 < 1,785)
- Les débits entre  $\bar{X} + \sigma$  et  $\bar{X} + 2\sigma$  (1,785 < C3 < 4,58)
- Les débits supérieurs à  $\bar{X} + 2\sigma$  : (C4 > 4,58)

La troisième variante (Cas 3) a été faite en tenant compte de la fonction de répartition

- Fréquence < 0,7401 (C1 < 0,0001 hm<sup>3</sup>)
- Fréquence entre 0,7401 et 0,8606 (0,0001 < C2 < 0,21 hm<sup>3</sup>)
- Fréquence entre 0,8606 et 0,9049 (0,21 < C3 < 0,4231 hm<sup>3</sup>)
- Fréquence supérieure à 0,9049 (0,21 < C3 < 0,4231 hm<sup>3</sup>)

Tous les modèles permettent une bonne approche de la moyenne interannuelle simulée (Figure VI-6), à part le modèle SVD qui sous-estime cette dernière.

Les modèles Markoviens multiclassés, et le modèle SVD sous-estiment également les écarts types inter annuels simulés .

L'ajustement des valeurs annuelles (Figure VI-7) a donné de bons résultats sauf que le modèle SVD n'arrive pas à reproduire les valeurs maximales annuelles.



L'évolution des moyennes mensuelles interannuelles (Figure VI-8) montre que les valeurs simulées sont en majorité à l'intérieur de l'intervalle de confiance ( $\alpha=5\%$ ) de la série historique. Les résultats sont acceptables pour l'ensemble des modèles de simulation.

En ce qui concerne les écarts types (Figure VI-9) les meilleurs résultats ont été obtenus par le biais du modèle GESTOP, Les autre modèles surestiment certains écarts types mensuels principalement les mois d'Août, Septembre et Octobre.

Les tests d'ajustement des valeurs maximales journalières annuelles (Figure VI-10) permettent de rejeter la simulation des débits par le modèle SVD du fait qu'il n'arrive pas a reproduire les valeurs extrêmes, le modèle Markovien multiclases n'arrive pas a générer des valeurs maximales journalières supérieures à celles de l'historique du fait qu'aucune extension de la fonction de répartition n'a été intégrer dans le modèle.

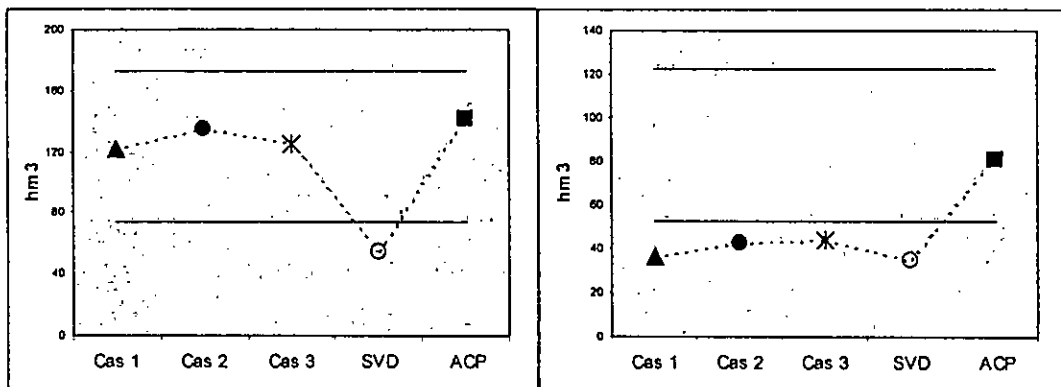


Figure VI-6: Tests sur les moyennes et écarts types des valeurs de débits annuels

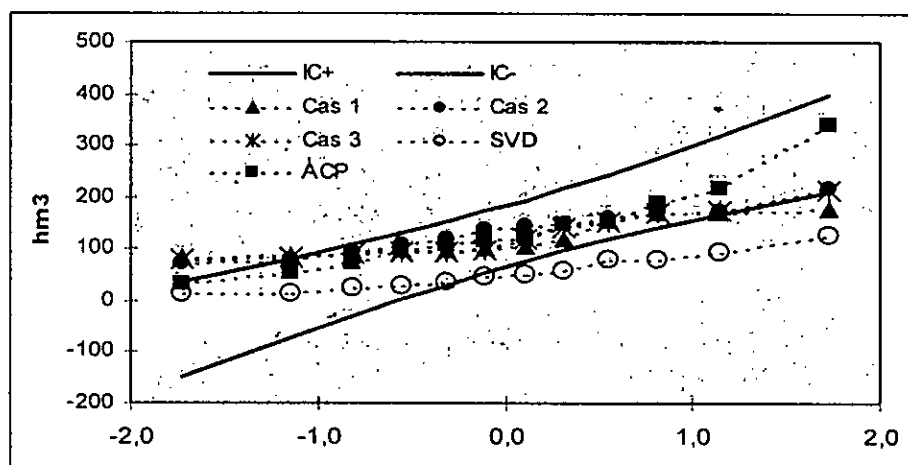


Figure VI-7: Ajustement des valeurs de débits annuels

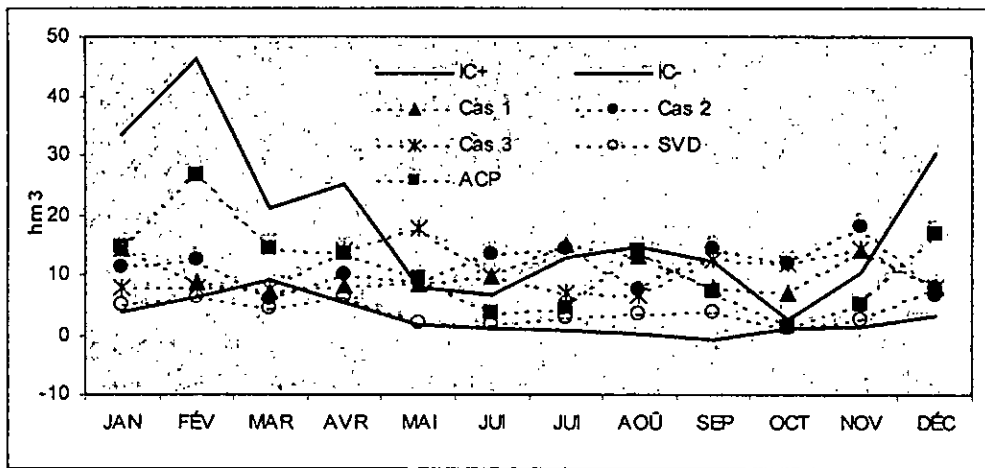


Figure VI-8: Tests sur les moyennes des valeurs de débits mensuels

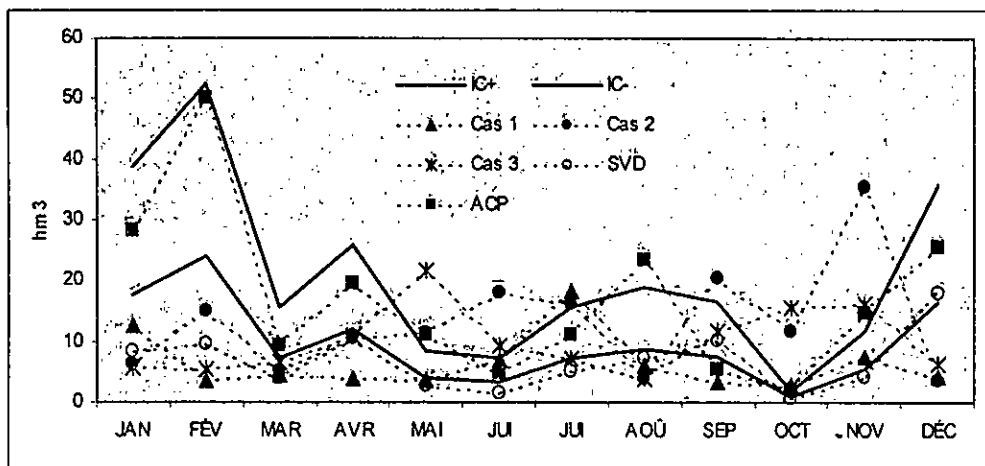


Figure VI-9: Tests sur les écarts types des valeurs de débits mensuels

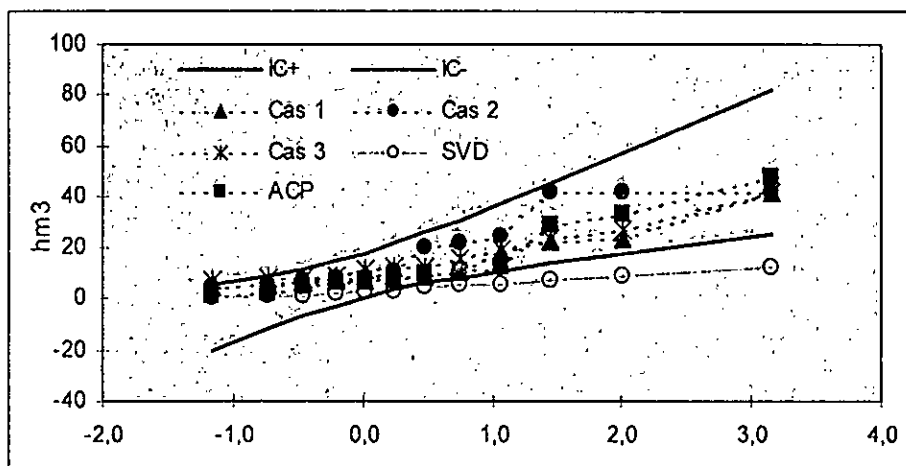


Figure VI-10: Ajustement des valeurs des débits maximums journaliers

### VI.3. SIMULATION DES PLUIES

Les mêmes modèles que ceux utilisés pour la simulation des débits ont été utilisés pour la simulation des pluies, la seule différence c'est que pour ce dernier paramètre, on a travaillé avec des modèles Markoviens binaires (ordre 0, 1 et 2).

En générale les moyennes interannuelles des séries simulées et leurs écarts types (Figure VI- 11) sont assez bien restitués, la même constatation peut être faite pour l'ajustement des valeurs annuelles simulées (Figure VI-12). Ces dernières s'ajustent bien à la loi normale.

Les modèles GESTOP et SVD ont donné les meilleurs résultats pour les tests sur les valeurs mensuelles (moyennes : Figure VI-13 et écarts types :Figure VI-14), en effet ils permettent de ressortir l'effet saisonnier ainsi que la variabilité des mois les plus humides. Les modèles Markoviens ont sous-estimés ces derniers.

Les remarques faites précédemment trouve les explications dans l'ajustement des valeurs journaliers maximales (Figure VI-15), en effet les modèles Markoviens n'arrive pas à reproduire les valeurs historique, les valeurs maximales journalières simulées ne dépasse guerre les 80mm, alors que le maximum historique est de 108 mm, pour les autres modèles GESTOP et SVD, les résultats sont avérés satisfaisant mais le modèle SVD reproduit l'information historique et n'arrive à créer une information nouvelle.

Le tableau 1 récapitule les résultats obtenus lors de cette étude comparative.

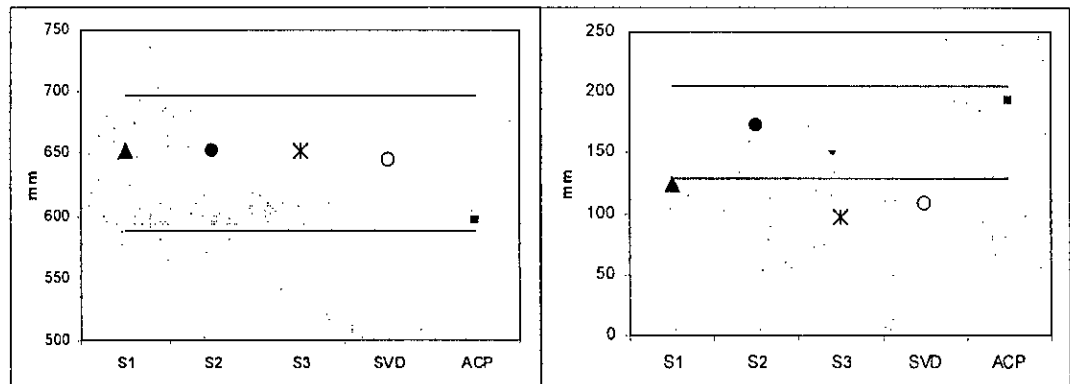


Figure VI-11: Tests des moyennes et écarts types des valeurs de pluies annuelles

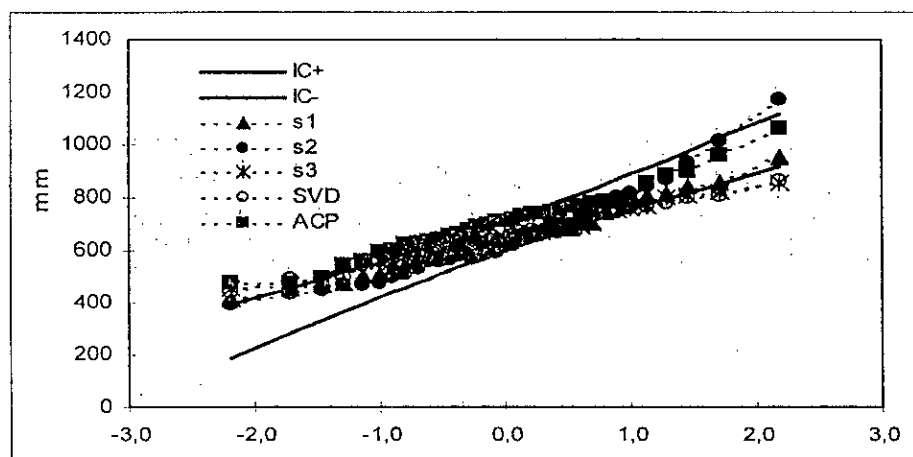


Figure VI-12: Ajustement des valeurs de pluies annuelles

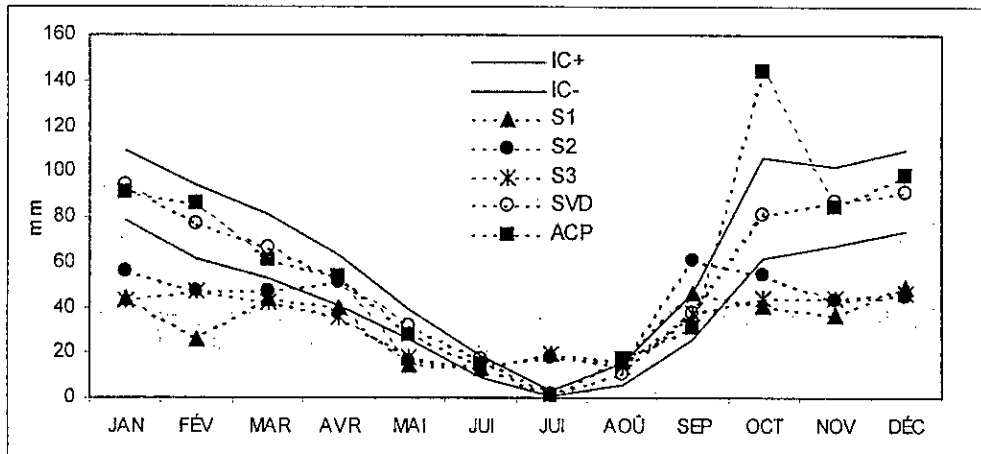


Figure VI-13: Test sur les moyennes des valeurs de pluies mensuelles

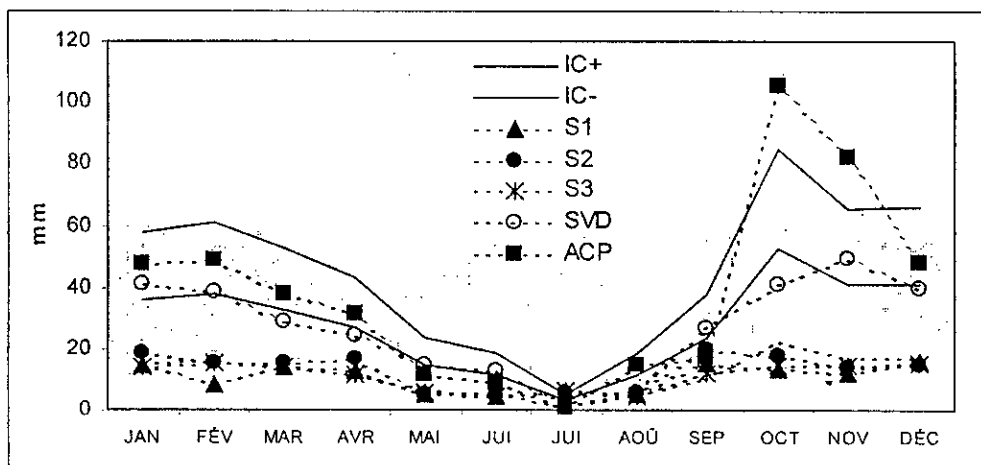


Figure VI-14: Test sur les écarts types des valeurs de pluies mensuelles

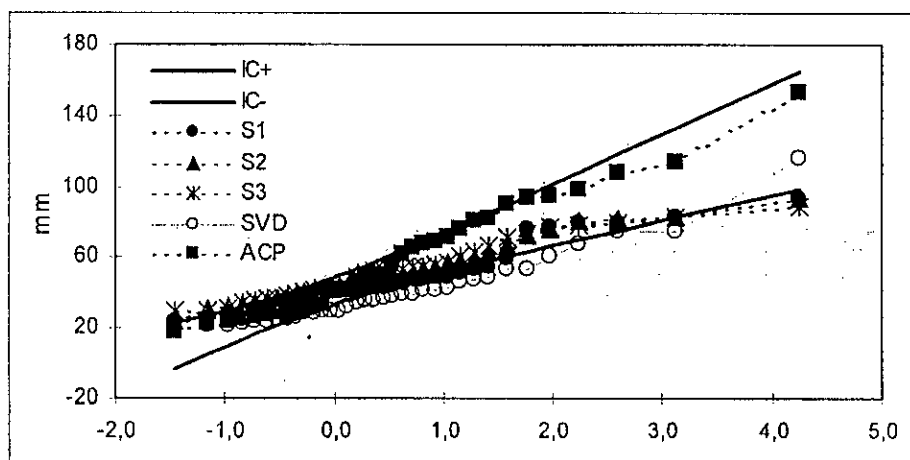


Figure VI-15: Ajustement des valeurs de pluies maximales journalières

Tableau VI-1: Comparaison entre les modèles de simulation

Modèle / Variable	ETP	Débit	Pluie
Markovien	-	+	+
Décomposition	+	Impossible	Impossible
SVD	+	-	+
GESTOR	+	+	+



CONCLUSION

## CONCLUSION

La simulation permet la génération de séries synthétiques des paramètres hydrométéorologiques, qui servent d'entrée aux modèles d'optimisation, de la gestion des ressources hydriques ; les modèles de simulation actuellement utilisés, sont alourdis par diverses contraintes à savoir le type de variable, le pas de temps, la chronologie, la stationnarité, la taille de l'échantillon etc... La démarche suivie, dans le cadre de ce travail de recherche avait pour finalité de trouver une solution unificatrice pour l'obtention de séries simulées pour tout type de variables, et à différents pas de temps et surtout en éliminant les contraintes précitées, qui ne permettent pas une généralisation de la simulation, et qui obligeaient les gestionnaires à faire un choix du modèle à utiliser, qui n'est pas basé sur un jugement scientifique, en fonction de la série prise en considération. Pour cela dans ce travail on s'est basé sur la technique de l'analyse en composantes principales qui est usuellement utilisée dans le domaine spatial et qui a permis dans le domaine temporel d'apporter des solutions à une problématique qui au départ n'avait pas de solution unique.

Les différentes étapes pour la recherche d'une méthodologie unique, de la simulation des variables hydrométéorologiques nous a conduit à faire les remarques suivantes.

La simulation au trois pas de temps exigent la transformation des données initiales. La transformation par la fonction racine carré, représente la seule contrainte pour la simulation au pas de temps mensuel, par la suite la génération des paramètres du modèle dans la loi normale suffit pour obtenir les résultats escomptés.

Lors du passage au pas de temps journaliers une distinction se fait entre les paramètres hydrométéorologiques cyclique (ETP) et les phénomènes aléatoires (débit, pluies). En effet les résultats de la simulation de l'évapotranspiration en empruntant la même procédure que dans le cas mensuel restent très bons.

La simulation au pas de temps journalier des débits et des pluies donne des résultats bons en moyenne, mais pose un problème de variabilité, on a constaté que les écarts types sont surestimés pour la simulation des débits alors qu'ils sont sous-estimés pour la simulation des pluies. Ceci se traduit par des valeurs extrêmes simulées nettement inférieures à celles de l'historique. A quoi sont dues ces différences ? L'examen de l'ajustement des paramètres CP et COV révèle que pour l'ETP, les points sont proches de la courbe théorique, alors que dans le cas de la simulation des débits, les premières CP présentent une extrémité supérieure au-dessous de l'intervalle de confiance. Cette valeur n'étant pas restituée avec la génération dans la loi normale, conduit à une surestimation des débits simulés. On observe le cas contraire pour les pluies, les premières CP ont une extrémité supérieure au-dessus de l'intervalle de confiance, comme les CP générées suivent en général la tendance de la courbe théorique, ceci mène à la sous-estimation des valeurs exceptionnelles des paramètres du modèle, et un examen approfondi à montrer que ces valeurs renferment l'information concernant la variabilité de la série d'origine ainsi que les valeurs maximales qu'elle contient.

A partir de ces constatations, différentes variantes ont été mises au point, basées en un premier temps sur la génération à partir de la fonction de répartition. Cette variante a donné quelques améliorations, qui restent néanmoins insuffisantes, en un second temps on a trouvé que la loi de Goodrich permettait d'obtenir des séries de paramètres générés ayant la même allure que les paramètres calculés initialement, donc elle a la capacité de restituer convenablement les valeurs extrêmes.

La dernière variante avec laquelle on a conclu ce travail consistait à combiner la génération dans la loi normale (lorsque l'ajustement le permet) et la génération avec la loi de Goodrich.

Ce travail de recherche a permis de localiser chaque type d'information dans la structure du modèle, la moyenne du phénomène est directement liée à l'allure moyenne de la première CP, quant à la variabilité elle dépend des valeurs qui s'éloignent de cette allure moyenne pour les paramètres CP et COV. Les phénomènes hydrométéorologiques sont plus complexe à simuler, lors que les valeurs maximales ont des intensités et des fréquences irrégulières

Le modèle GESTOP a permis de simuler les variables hydrométéorologiques, à différents pas de temps (mensuel, journalier, horaire) sans aucune contrainte, malgré la complexité de la génération au pas de temps très court du fait de l'intermittence des phénomènes étudiés.

Ce travail de recherche a comme perspective la simulation spatio-temporelle pour résoudre le problème du manque ou de la mauvaise qualité des données hydrométéorologiques, ce problème se pose avec acuité en Algérie, du fait de la construction de nombreux barrages sur des sites ne disposant pas de données (même très courtes) nécessaire au dimensionnement de ces ouvrages.



## BIBLIOGRAPHIE

- AURAY, J.P* (1990). Analyse des données multidimensionnelles. pp19-75
- BENZECRI, J.P.* (1982). Histoire et préhistoire de l'Analyse des données. DUNOD
- BOUROCHE, J.M.* (1983). L'Analyse des données. 2ème Edition Que Sais-je ? Pp5-45.
- BURDEN, R.L.*, (1981). Numerical Analysis. 2ème Edition BOSTON, Massachusettes, pp 73-107.
- CARLSON, R.E., et AL.* (1970). Application of linear random models to four annual stream flow series. Water Ressource. Res, 6, 1070-8
- CAVADIAS, G.* (1980). An approach to the stochastic simulation of river flows. Rapport N° 80-1, Dept. Of civil Engineering and Applied Mechanics, McGill University, Montréal.
- CAVADIAS, G.* (1985). A multivariate seasonal model for streamflow simulation. Colloque sur la recherche en hydrologie au Quebec. 53ième congrès de l'ACFAS, Chicoutimi, Canada
- CHEVALIER, P.*(1985). Pluie et crue au Sahel. Hydrologie de la marre d'Oursi, Burkina-Faso, Edition l'ORSTOM, pp79-88.
- CLENE, T.B.* (1981). Selecting Seasonal streamflow models Water ressources, Volume17.
- COLSTON et WIGGERT* (1968). The advanced theory of statistics. Fourth edition. Macmillan Oxford University Press.
- COOPER, DM. et WOOD, EF.*, (1982). Edentification of multivariate time series and multivariate INPUT – OUTPUT. Water Ressources Volume 18.
- DECHEMI, N., BERMAD A. et HAMRICHE A.*(1994). Simulation des débits moyens mensuels en zone semi-aride par l'Analyse en composantes principales (ACP). Revue hydrologie continentale.
- DECHEMI, N., BERMAD A. et TACHET K.*(1997). Etude et simulation de l'évapotranspiration par l'Analyse en composantes principales (ACP). Bulletin du comité international de l'irrigation et du drainage . CIID, 46(2) ; Pp. 45-59
- DELAGARDE, J.* (1983). Initiation à l'Analyse des Données ADD. DUNOD, Pp1-91.
- DELLEUR, JW.* (1978). Les processus de type ARIMA pour la prévision et la simulation en hydrometeorologie. La Houille Blanche N°6.

- DIPAY, E.* (1989). *Eléments d'Analyse des Données*. DUNOD, Pp168-221.
- DITLEVSEN, O.* (1971). *Extremes and first passage times with application in civil engineering*, Danmark Ingeniorakademie, Copenhagen
- DUBAND, D.* (1989). *Hydrologie statistique approfondie*. I.N.P.G., PpVI.1-VI.43.
- DUBREUIL, P.* (1974). *Initiation à l'Analyse Hydrologique*. ORSTOM, Paris.
- HACHE, M., P.F. RASMUSSEN et CAVADIAS, G.* (1994). *The multivariate seasonal flow simulation model SVD. Application to the Ottawa river system*. Rapport interne I-133, INRS-EAU, Sainte foy, Québec.
- HARMS et CAMPBELL* (1966). *Analyzing multivariate data*. The Dryden Press.
- HIPEL et AL.* (1977). *Advances in Box-Jenkins modelling, 1, model construction*. Water Ressource. Res., 13, 567-75
- KOTTEGODA*, (1980). *Stochastic Water Ressources Technology*. The Macmillan Press LTD.
- LANE, M.* (1979). *Conditional chance – constrained model for reservoir control*. Water Ressou., 9, 937-48
- LEBART, L.* (1982). *Technique de la description statistique multidimensionnelle*. MASSON, Pp119-148.
- LETTENMAIER, D.P., et BURGESS, S.J.* (1976). *Use of states estimation techniques in water ressource system modelling*. Water Ressource. Bul., 12, 83-99 (1977)
- MORIN, G.* (1974). *Génération de chroniques de débits Conception nouvelle utilisant des fonctions orthogonales*. Thèse de D.E.A. Grenoble.
- MATALAS, N.C.* (1967). *Some comments in the use of factor analysis*. Water Ressources 3 (1).
- MORISSON, D.F.* (1967). *Multivariate statistical methodes*.
- MCGRAW HILL*, (1981). *Physical basis of statistic models of annuel flows*. Water Ressources volume 17 N°2 , Avril) Pp428-430.
- MANDELBROT, BB. et WALLIS, JR.* (1969). *Computer experiments with fractional gaussian noises*. Water ressources Volume, 5).
- MANDELBROT, BB.* (1971). *A fast fractional gaussian noise generator*. Water ressources Volume 7.

- MEJIA, J.M.* (1972). Stream flow simulation 2.  
The Broken line process as a potential model of hydrologic simulation.  
Water resources Volume 8.
- MEJIA, J.M. et ROUSSELLE, J.* (1976). Disaggregation models in hydrology.  
Water resources Volume 12.
- NAKACHE, J.P.* (1970). Analyse en Composantes Principales. PARIS
- O'CONNEL, P.E., Hall, M.G.* (1971). Some applications of stochastic hydrological models.  
Modelling of water resource systems, Vol. 1  
(ed. A.K. Biswas), Harvest house, Montreal pp250-62
- PERI, D. et BRADLEY, E.* (1984). Méthodes de calculs statistiques intensifs  
sur ordinateurs. L'intelligence artificielle, Pp21-33, Bibliothèque pour  
la science Diffusion Berlin
- ROBERT, J.* (1984). Les statistiques: une approche nouvelle.  
Mac Graw Hill, Pp409-418.
- SALAS, J.D.* (1982). ARMA model identification of hydrologic times series.  
Water resources Volume , 18).
- TESSIER, D.* (1983). Analyse Statistique des caractéristiques physiographiques  
de petits bassins versants en relation avec le débits de crue.  
Thèse M.Sc.A. Montreal.
- THOMAS, H.A. et FIERING, M.B.* (1962). Mathematical synthesis of stream  
flow sequences for the analysis of river Basin by simulation.  
Design of Water Resources System.
- TREIBER, B. et PLATE, E.J.* (1972). A Stochastic models for the simulation of  
daily flows. Bulletin des sciences hydrologiques, Volume 22.
- TROUDE, C.* (1992). Analyse des Données. Cycle S.E.A., Pp1-8.
- VALENCIA, R.D. et SCHAAKE, J.C.* (1973). Disaggregation processes  
in stochastic hydrology. Water resources Res. 9, 580-5
- VOLLE, M.* (1985). Analyse des Données. Economica, Pp1-25.
- WALLIS, J.R.* (1965). Multivariate statistical methods in hydrology.  
Water Ressources 1(4).
- WEISS, G.* (1977). Shot Noise Models. Water Ressources Volume 13.

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

# ECOLE NATIONALE POLYTECHNIQUE

DEPARTEMENT D'HYDRAULIQUE

THESE DE MAGISTER

SUJET

*Contribution à la simulation des paramètres hydrométéorologiques  
(aléatoires et cycliques) a différents pas de temps  
(mensuel, journalier, horaire) par le modèle GESTOP.*

## ANNEXES

Proposé par :

M. N. DECHMI

Etudié par :

Mme. A. HAMRICHE

dirigé par :

M. N. DECHMI

PROMOTION

97/98

E.N.P. 10, Avenue Hacén Badi - El Harrach - Alger

**LISTE DES ANNEXES**

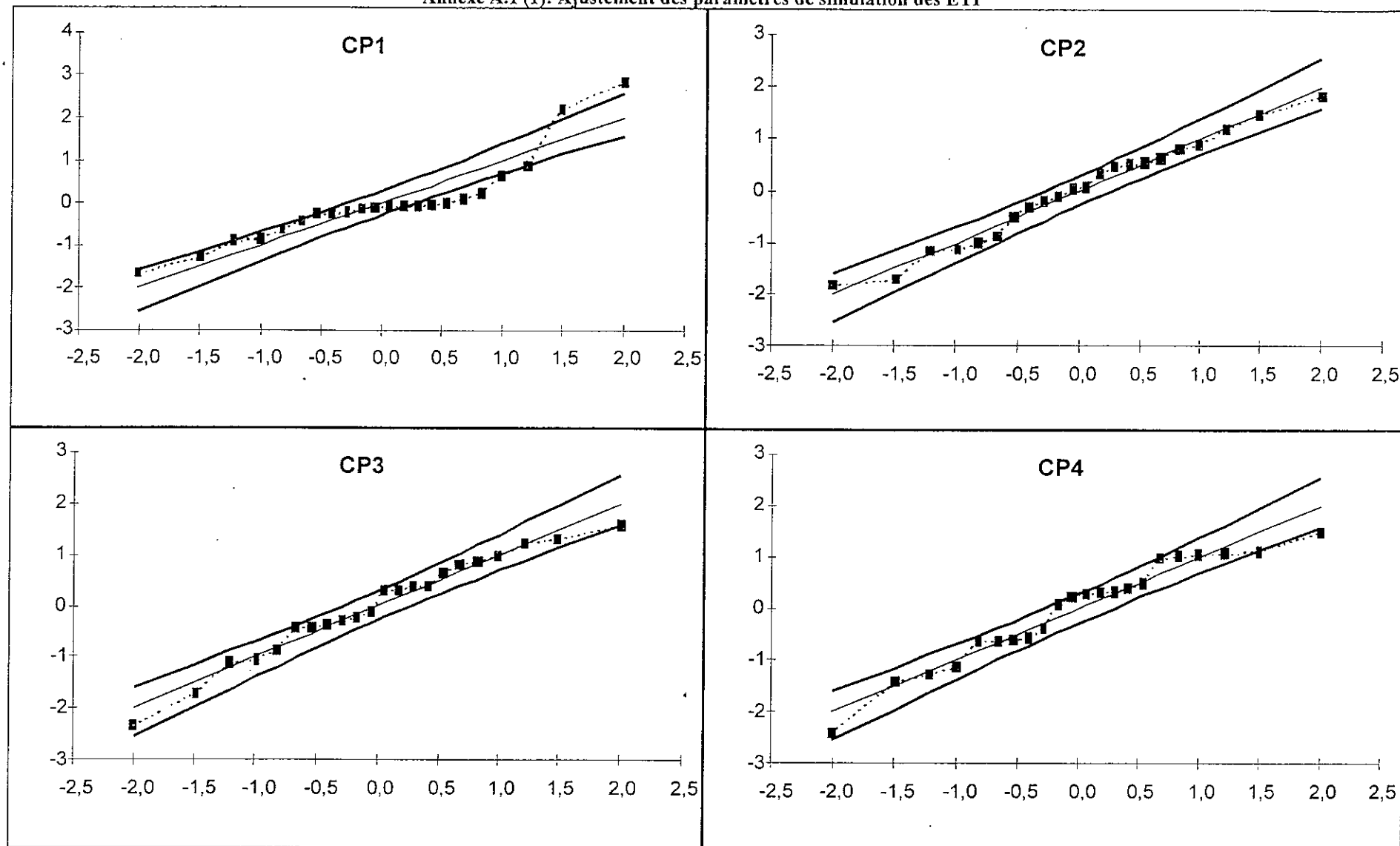
<b>Annexe A</b>		
Annexe A.1	Ajustement des paramètres de simulation des ETP	1
Annexe A.2	Tests des séries d'ETP simulées (Données brutes)	7
Annexe A.3	Tests des séries d'ETP simulées (Données transformées)	8
Annexe A.4	Ajustement des paramètres de simulation des Débits	9
Annexe A.5	Tests des séries de Débits simulées (Données transformées)	12
Annexe A.6	Ajustement des paramètres de simulation des Pluies	16
Annexe A.7	Tests des séries de Pluies simulées (Données transformées)	22
<b>Annexe B</b>		
Annexe B.1	Evolution chronologique des ETP au pas de temps journalier	23
Annexe B.2	Ajustement des paramètres de simulation des ETP	24
Annexe B.3	Evolution chronologique des débits au pas de temps journalier	30
Annexe B.4	Ajustement des paramètres de simulation des Débits (Décembre)	31
Annexe B.5	Ajustement des paramètres de simulation des Débits (Juillet)	37
Annexe B.6	Evolution chronologique des pluies au pas de temps journalier	43
Annexe B.7	Ajustement des paramètres de simulation des Pluies (Décembre)	44
Annexe B.8	Ajustement des paramètres de simulation des pluies (Juillet)	50
<b>Annexe C</b>		
Annexe C.1	Evolution chronologique des ETP au pas de temps horaire	56
Annexe C.2	Ajustement des paramètres calculés pour la simulation des ETP au pas de temps horaire	57
Annexe C.3	Ajustement des paramètres générés pour la simulation des ETP au pas de temps horaire	58
Annexe C.4	Evolution chronologique des ETP simulées au pas de temps horaire	60
Annexe C.5	Evolution chronologique des pluies au pas de temps horaire	61
Annexe C.6	Ajustement des paramètres de simulation des pluies au pas de temps horaire	62
Annexe C.7	Ajustement des paramètres générés pour la simulation des pluies au pas de temps horaire	65
Annexe C.8	Evolution chronologique des pluies simulées au pas de temps horaire	68

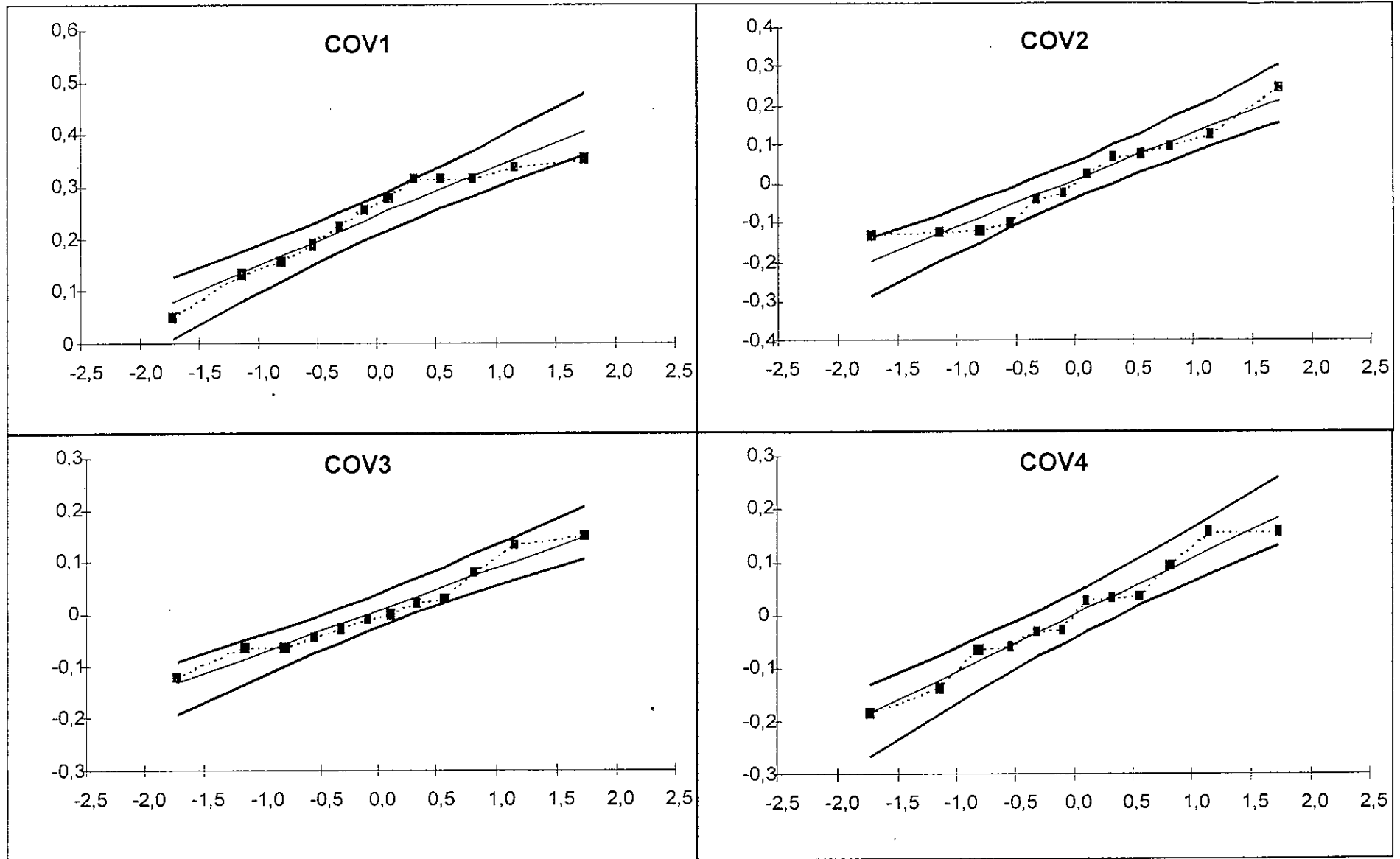
المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

## ANNEXE A

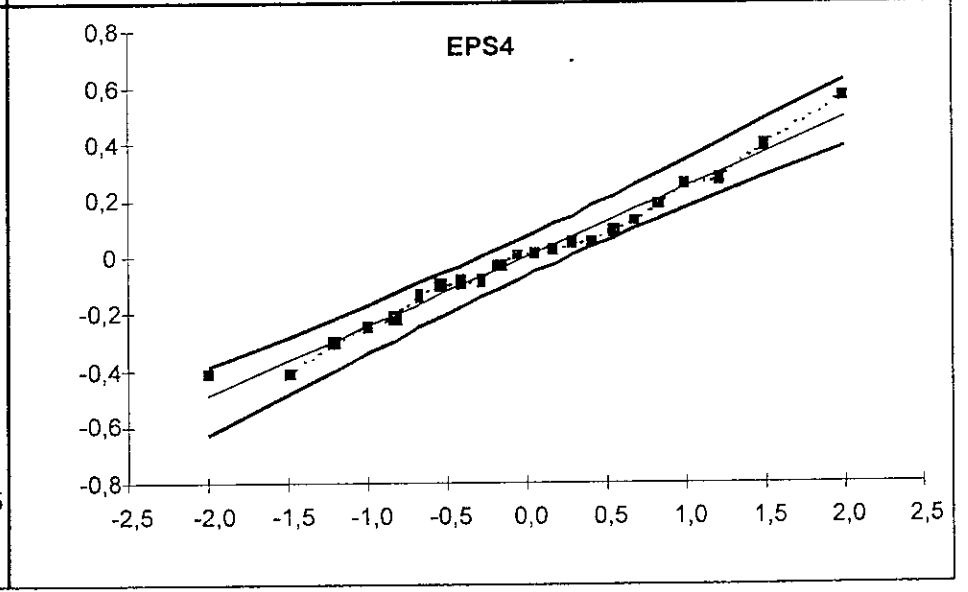
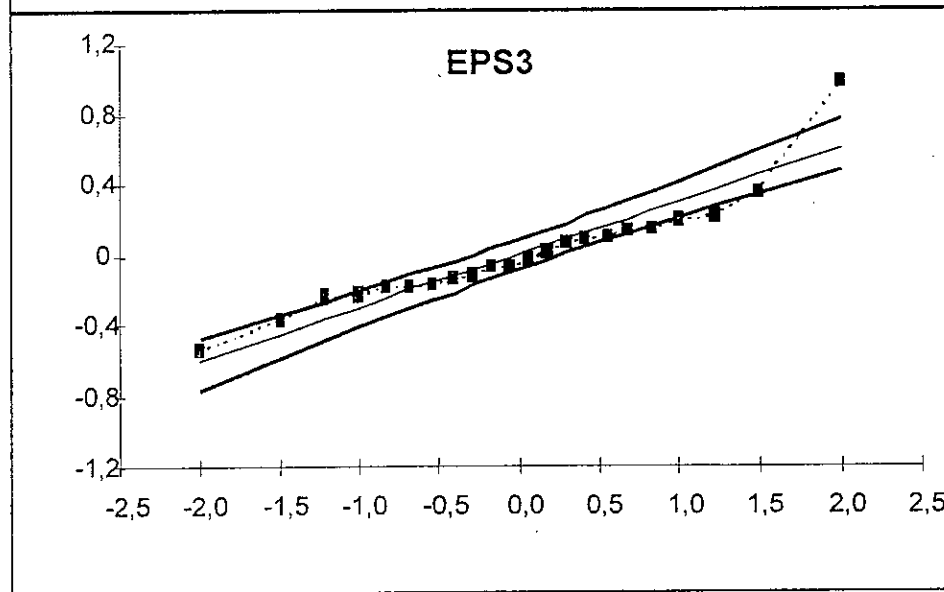
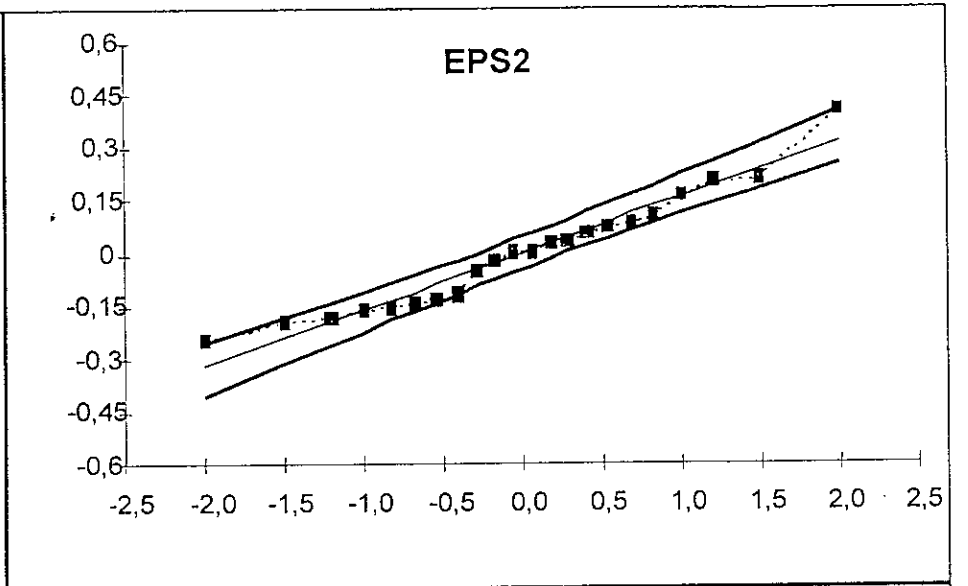
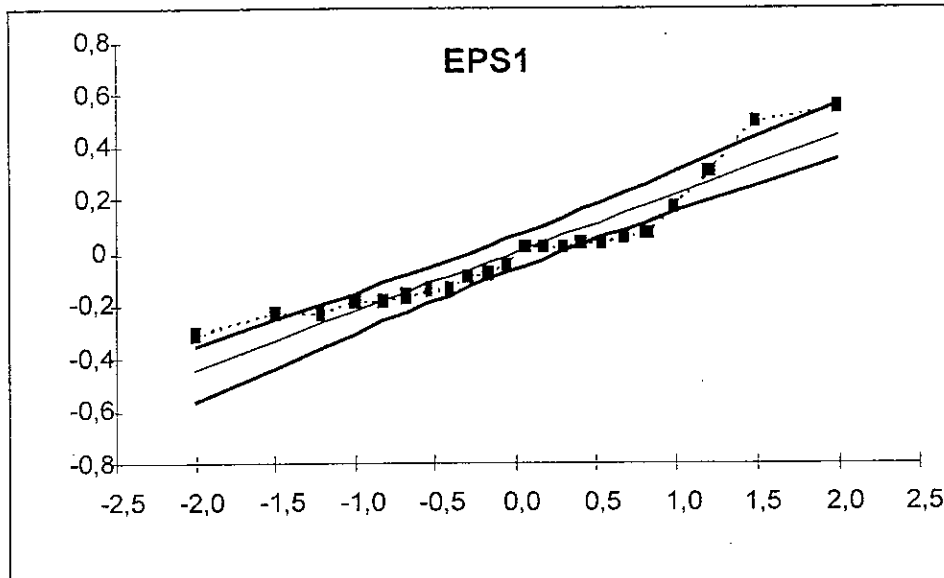
# SIMULATION AU PAS DE TEMPS MENSUEL

Annexe A.1 (1): Ajustement des paramètres de simulation des ETP

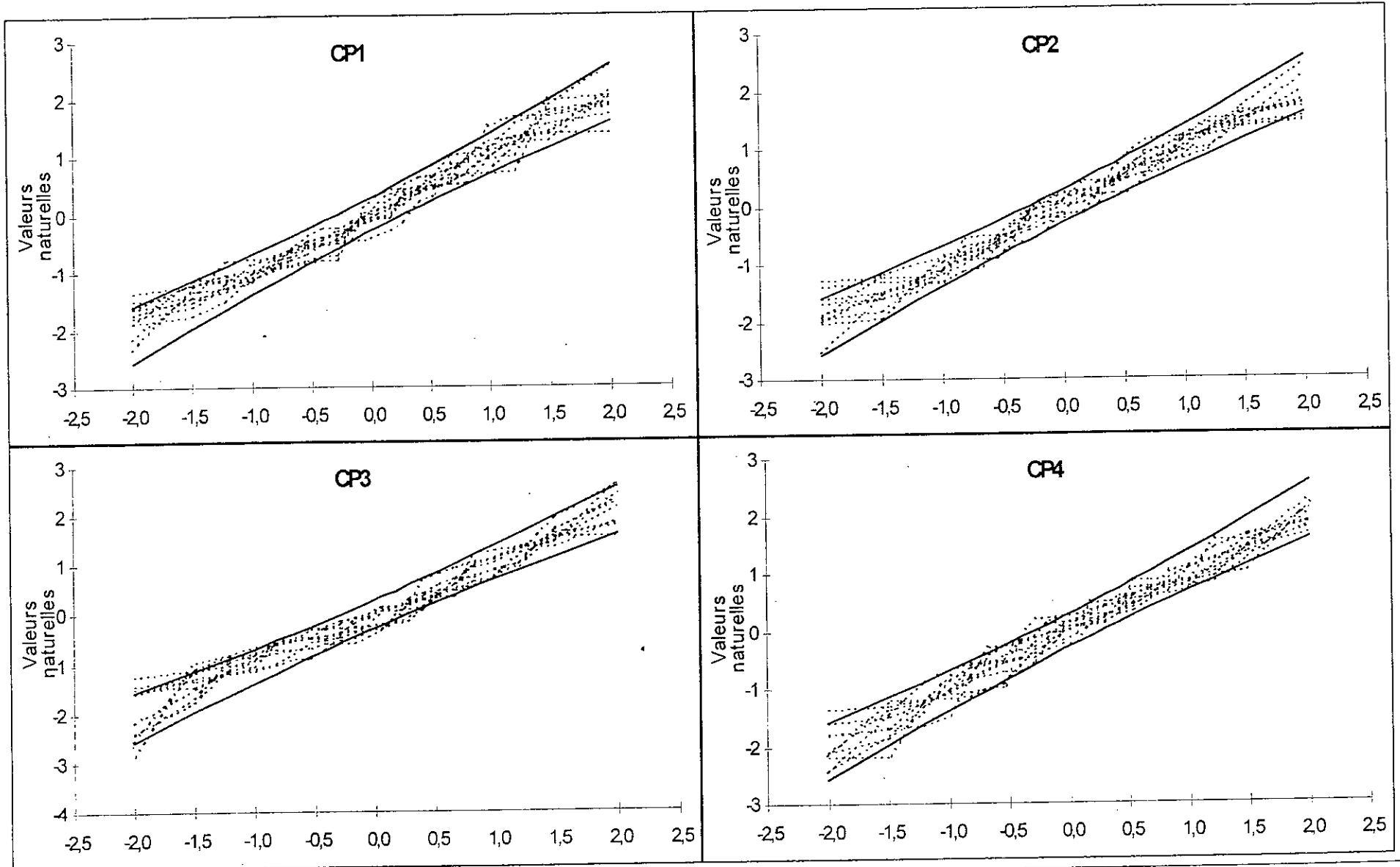


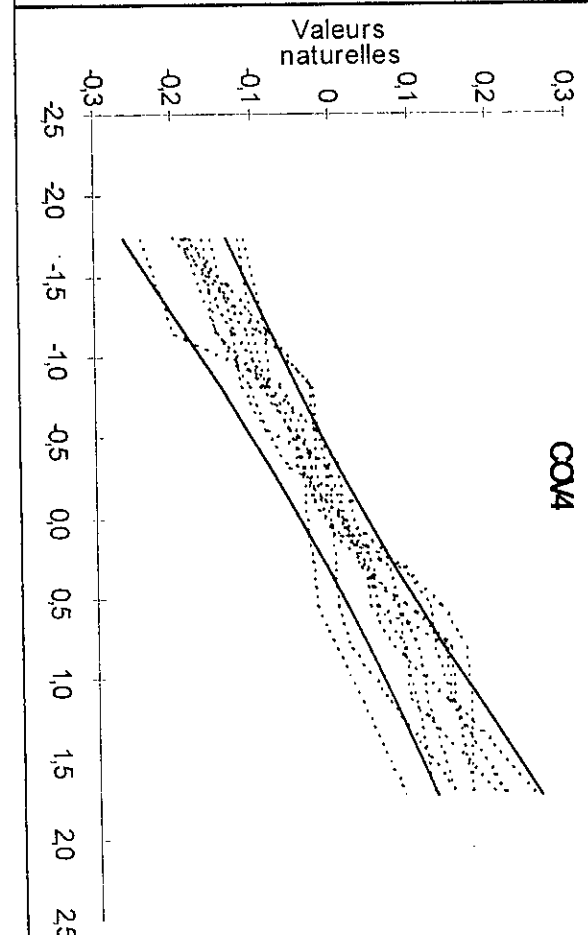
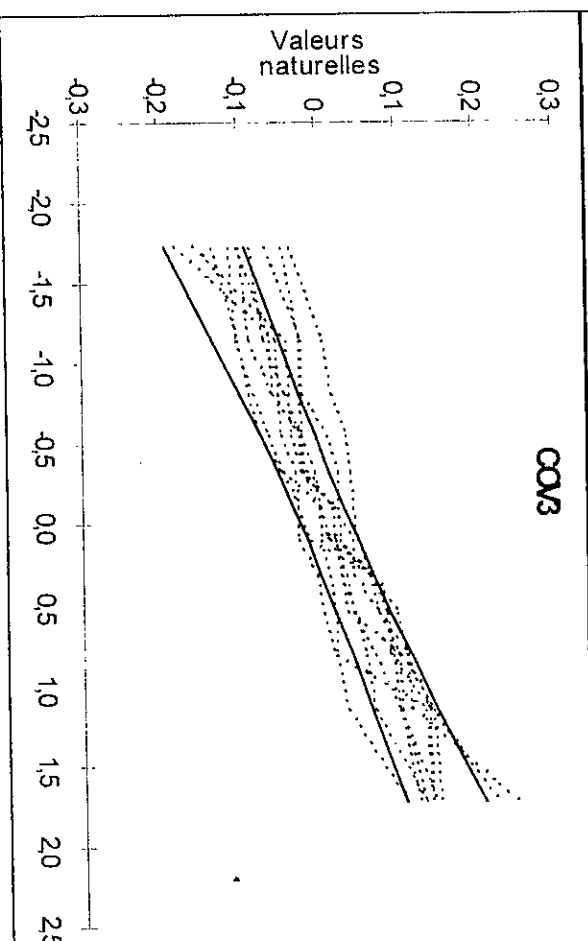
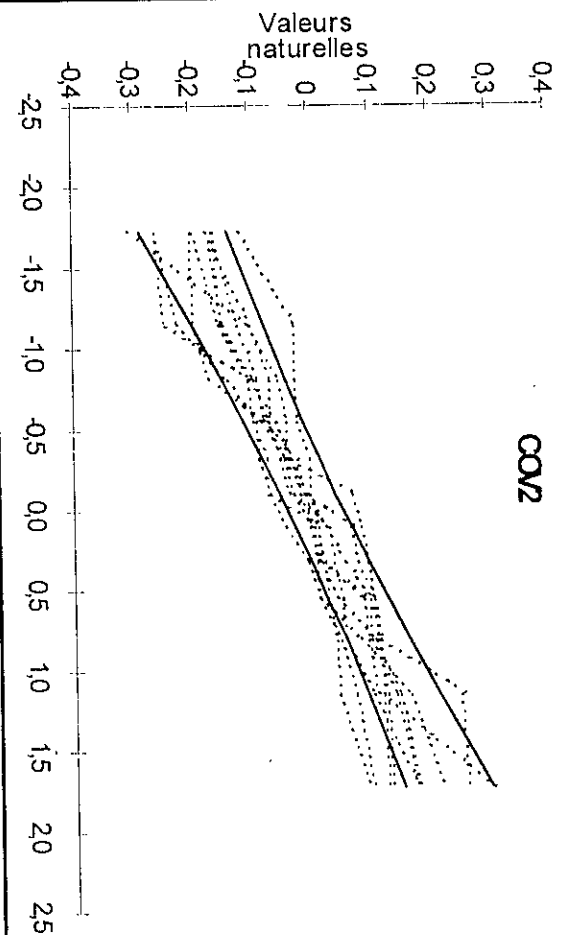
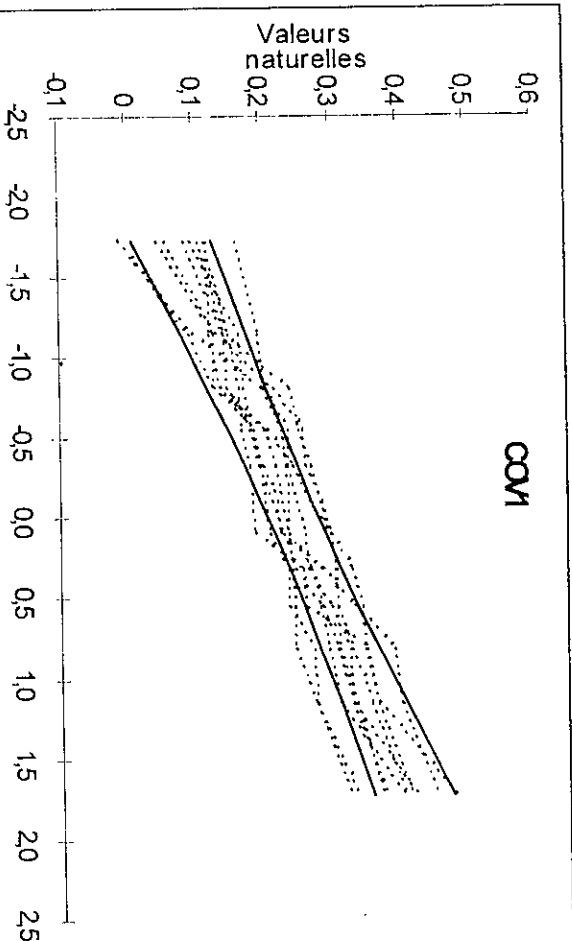


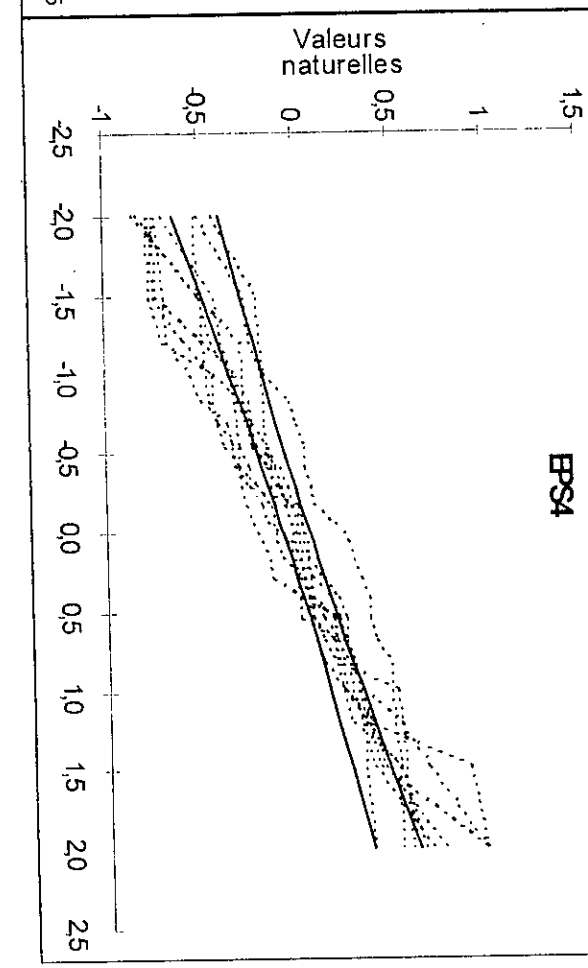
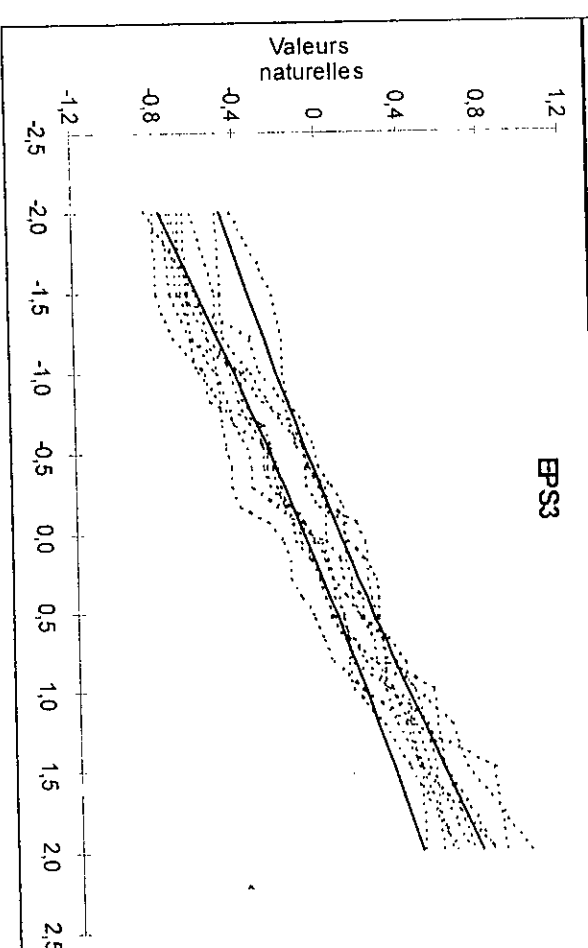
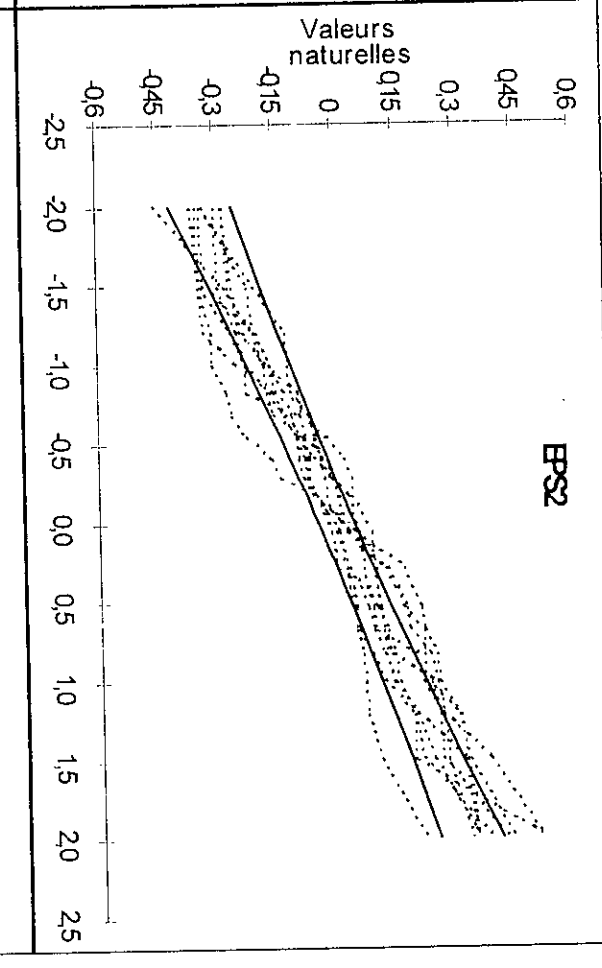
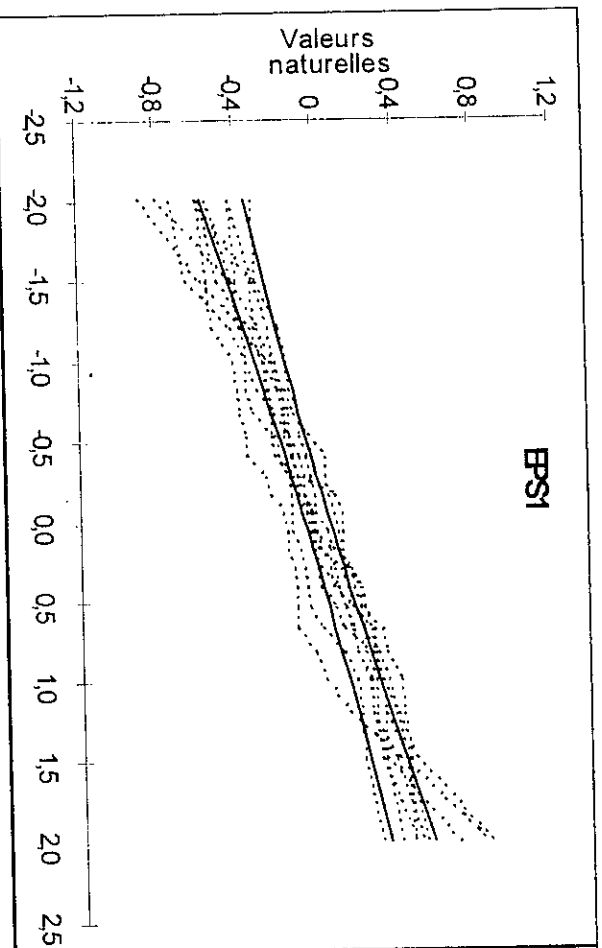




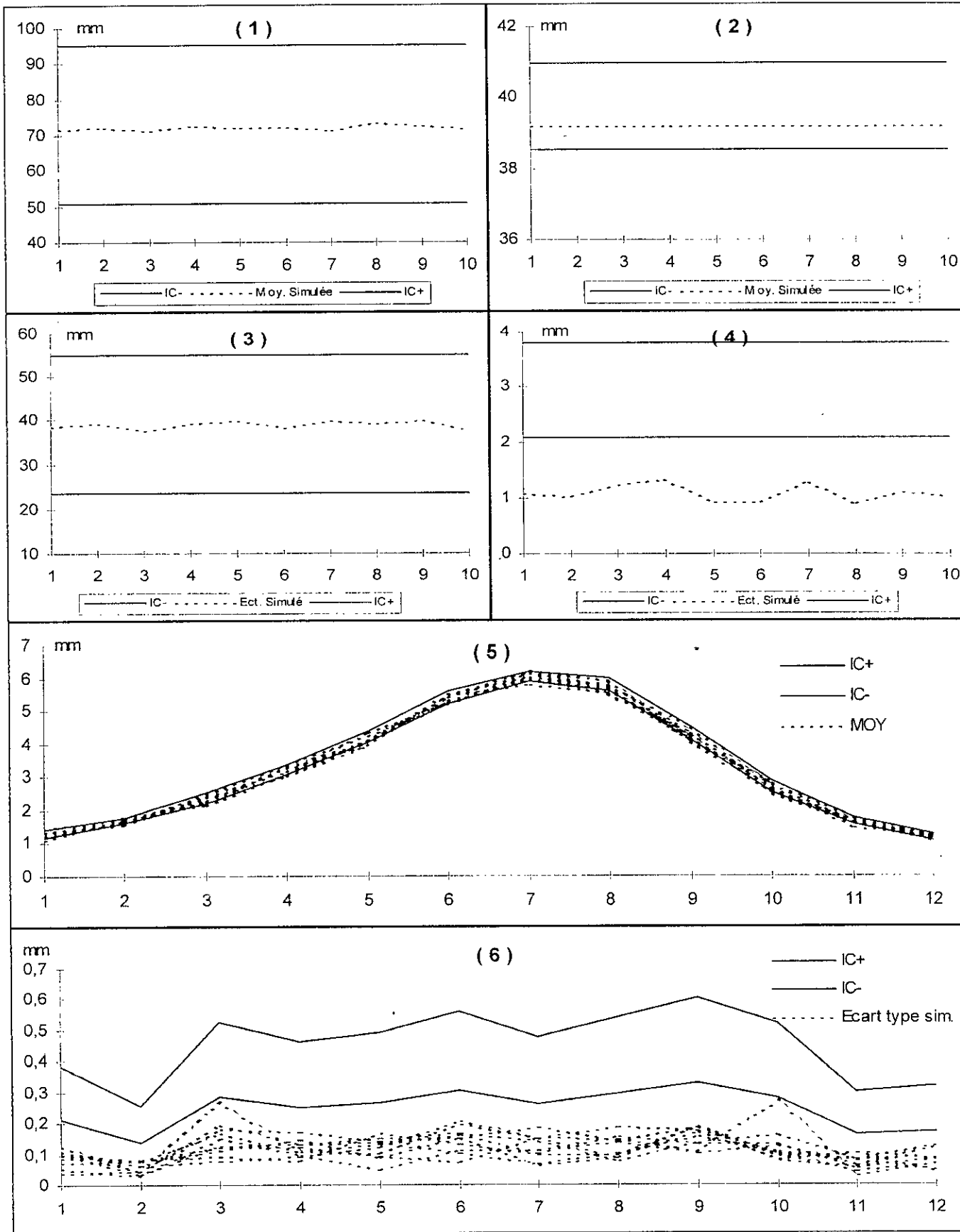
Annexe A.1 (2)



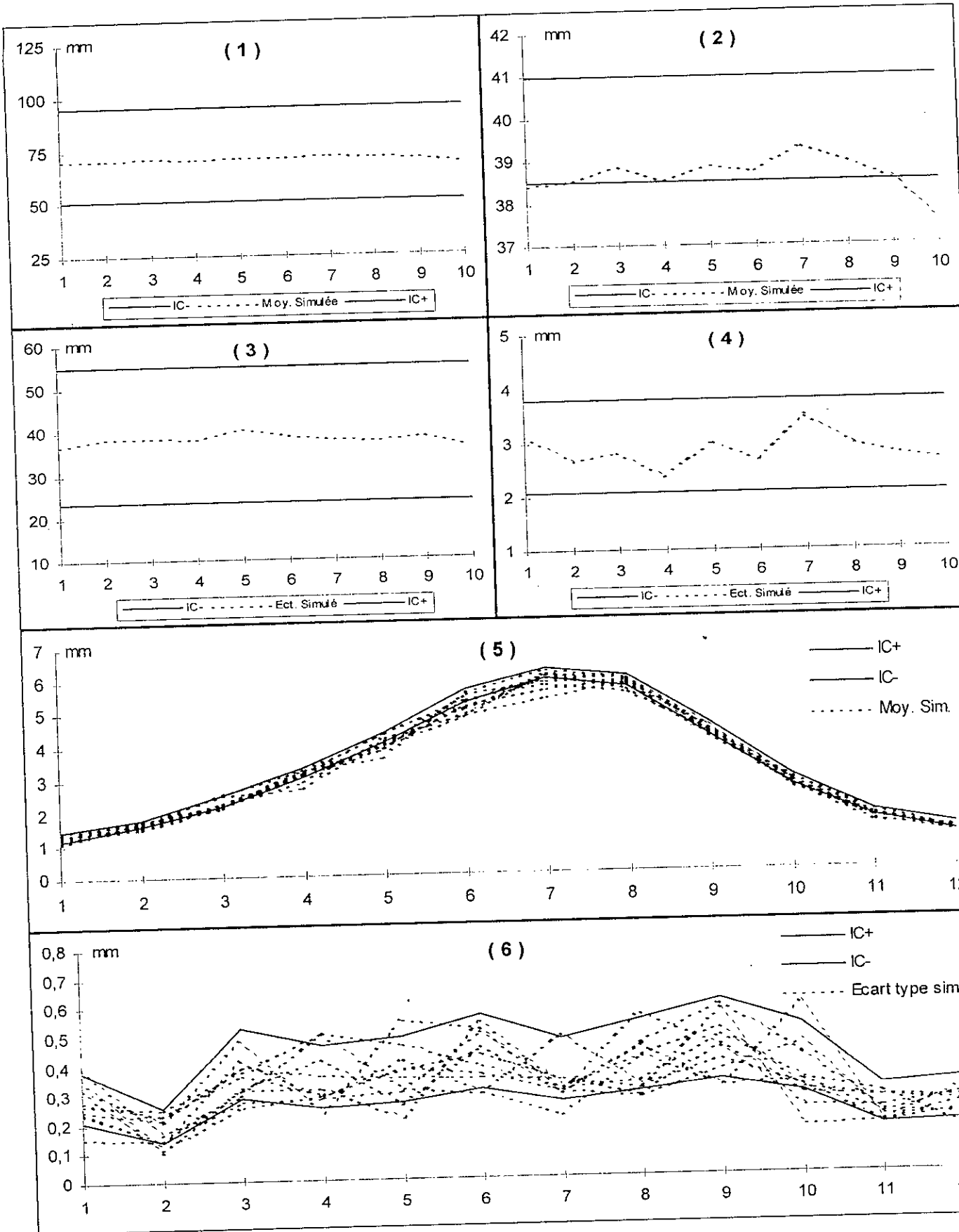




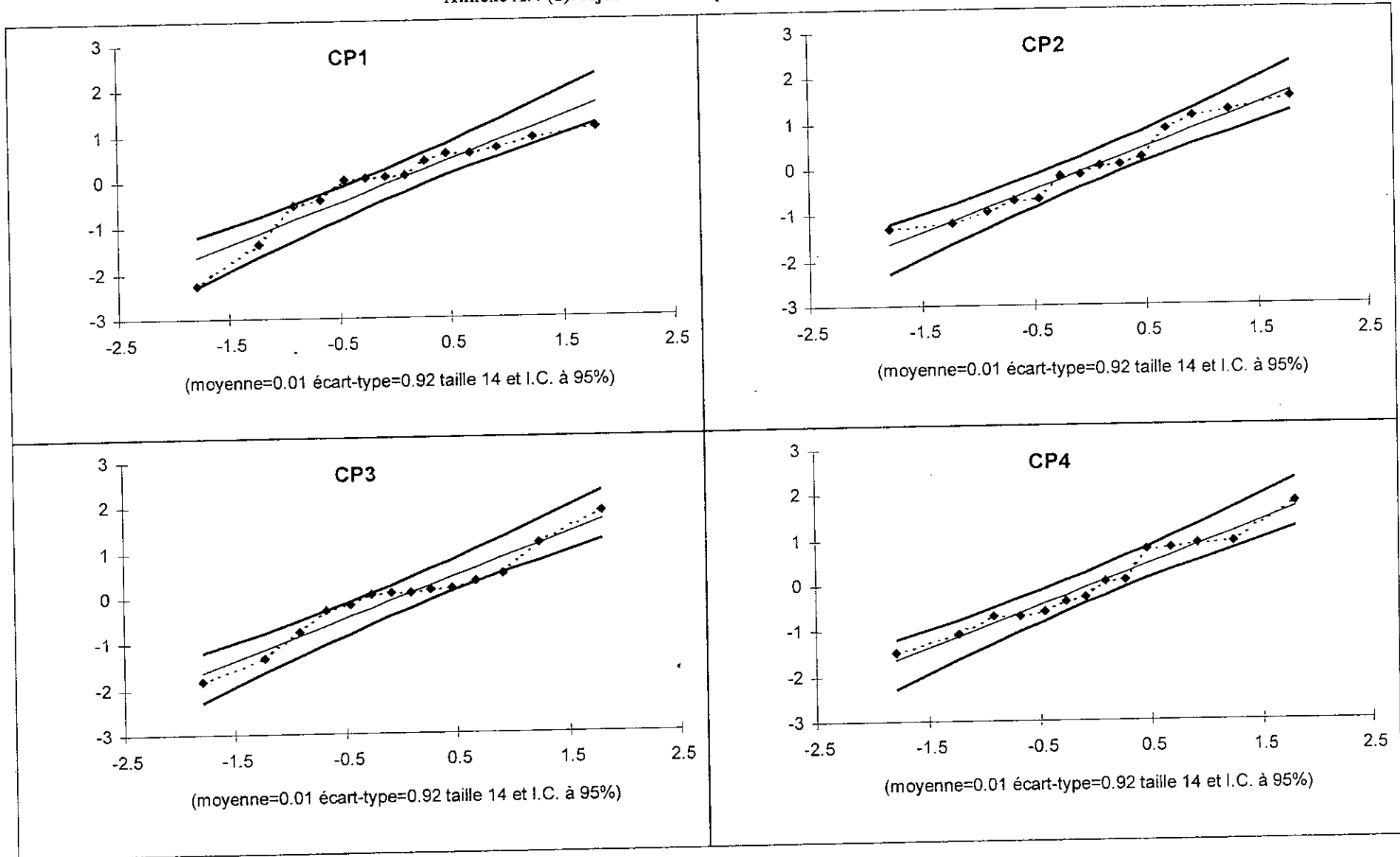
Annexe A.2: Tests des séries d'ETP simulées (Données brutes)

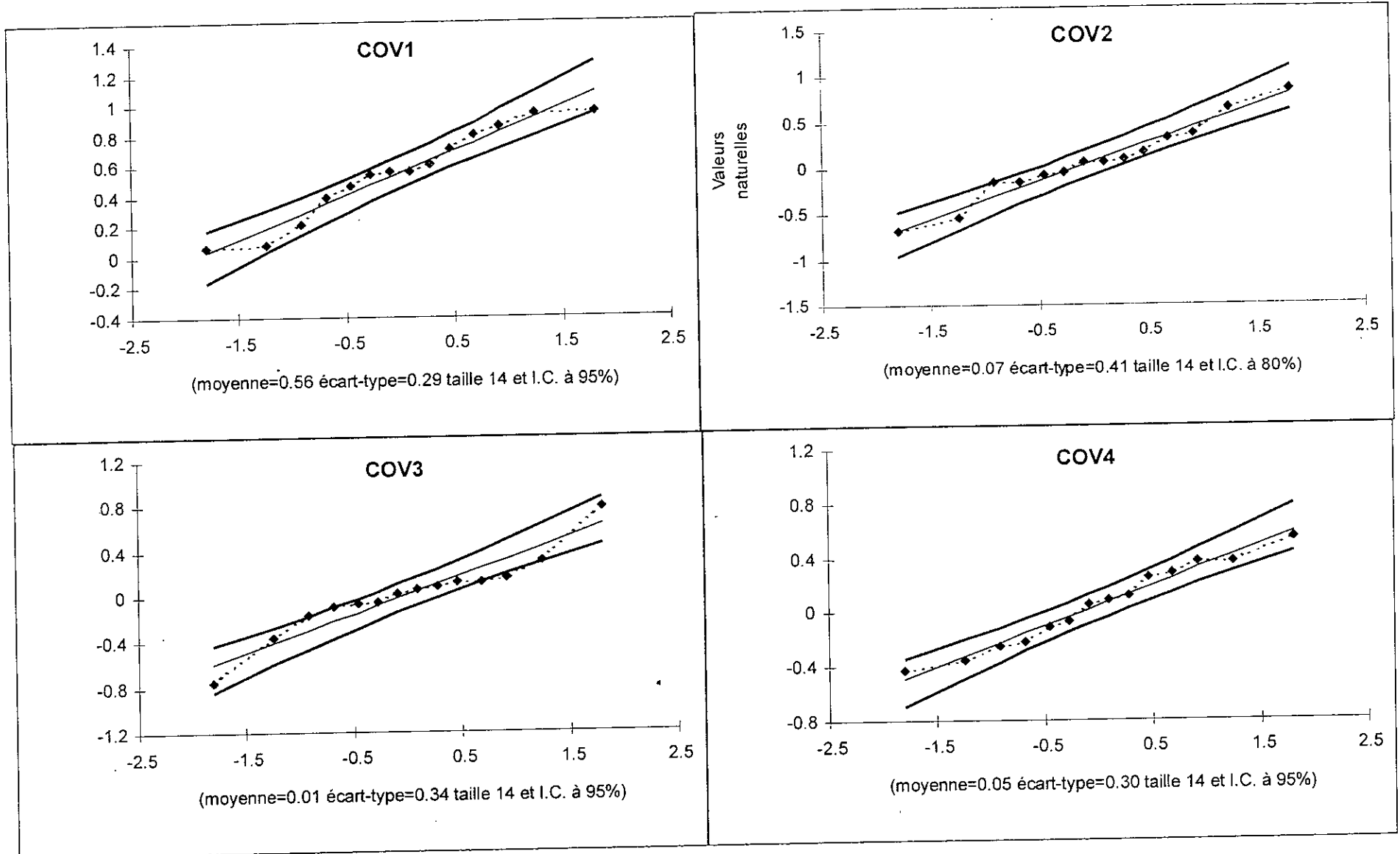


Annexe A.3: Tests des séries d'ETP simulées (Données transformées)

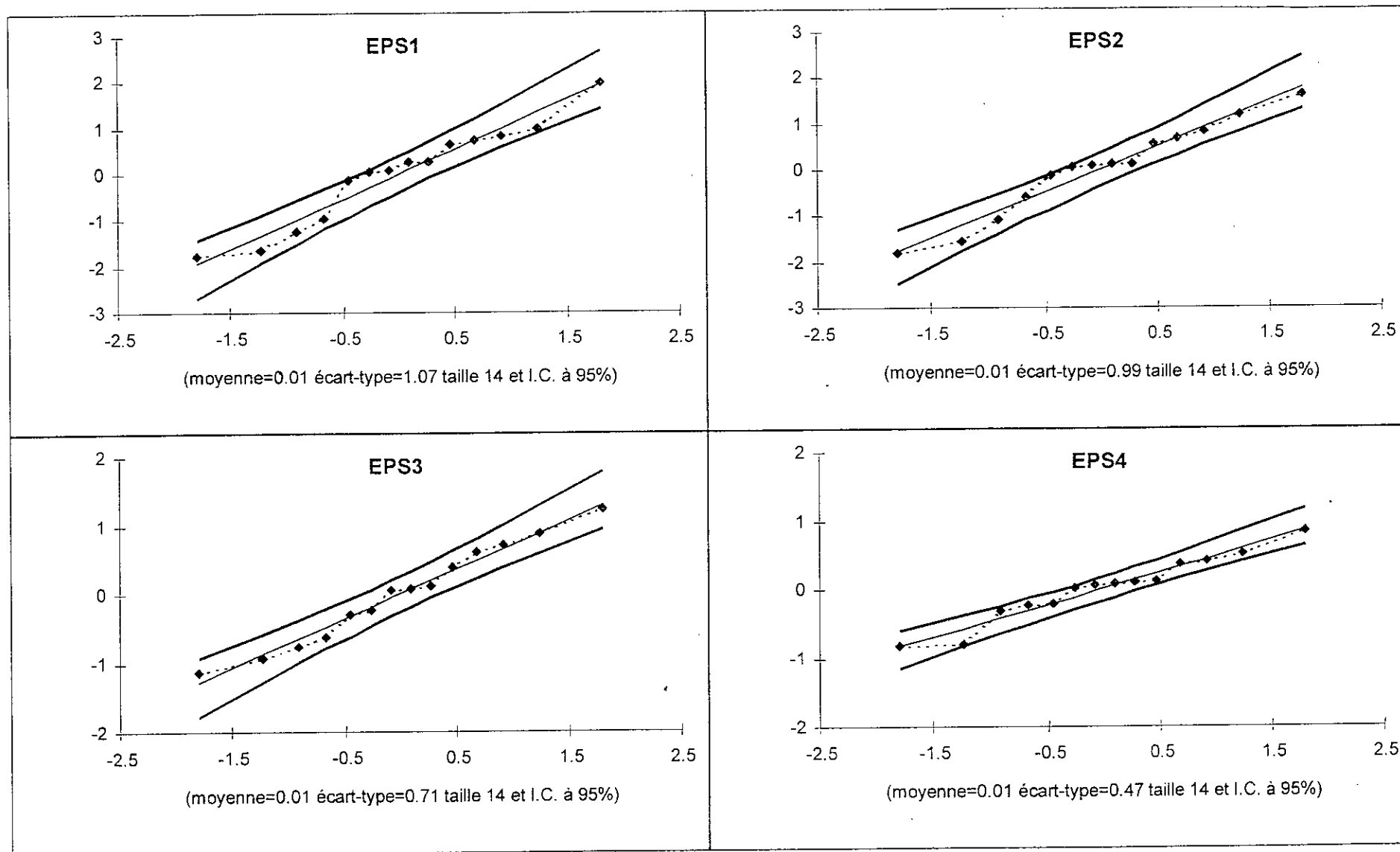


## Annexe A.4 (1): Ajustement des paramètres de simulation des Débits

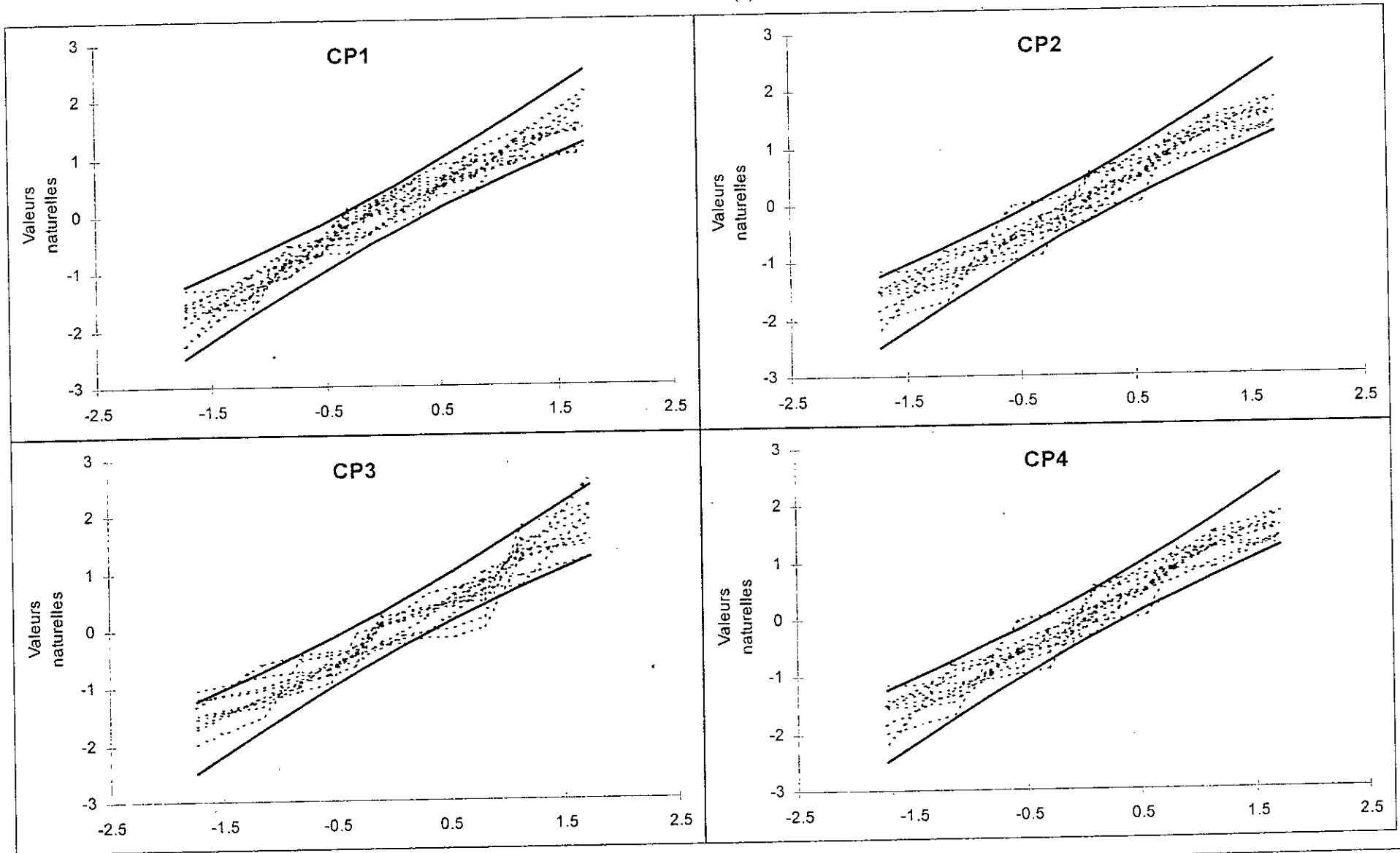


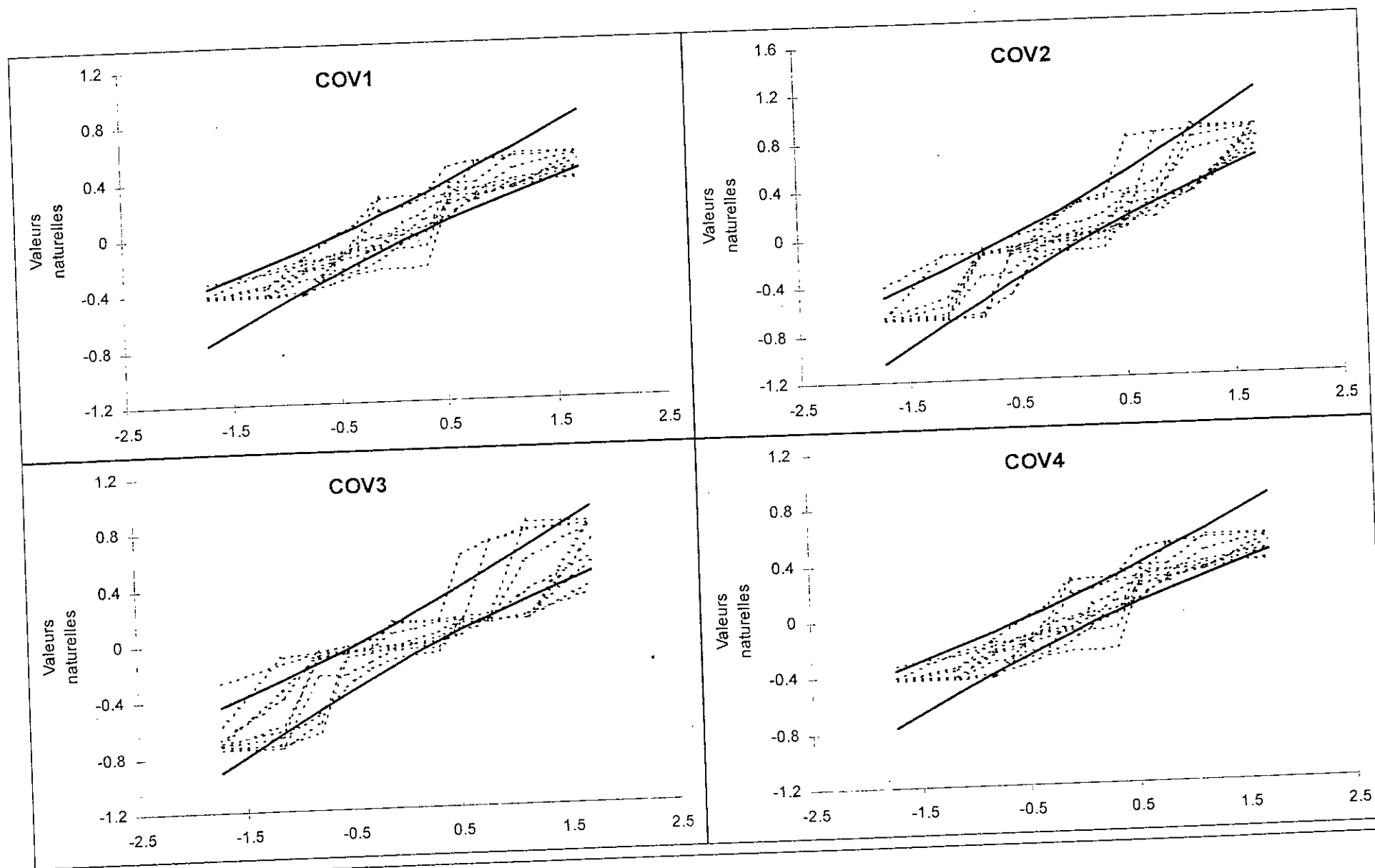


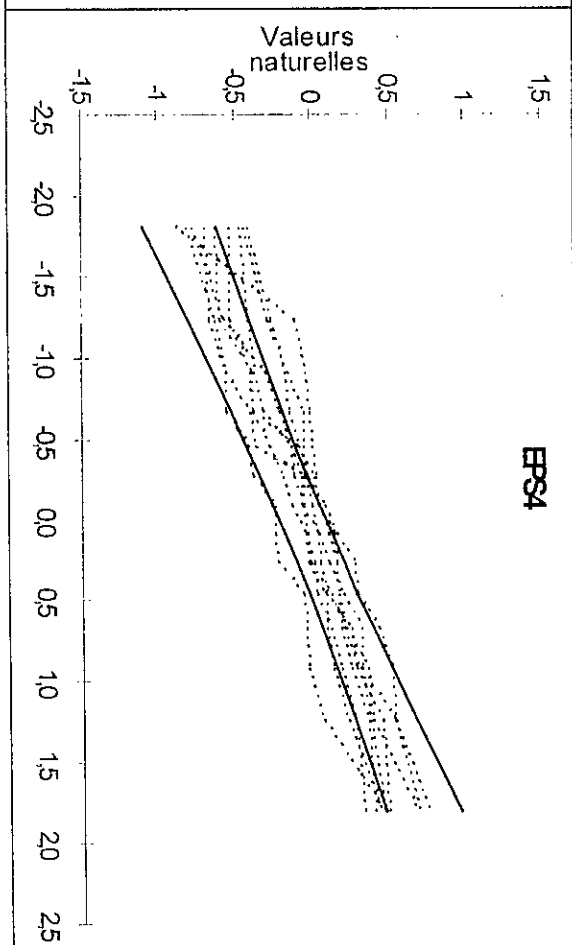
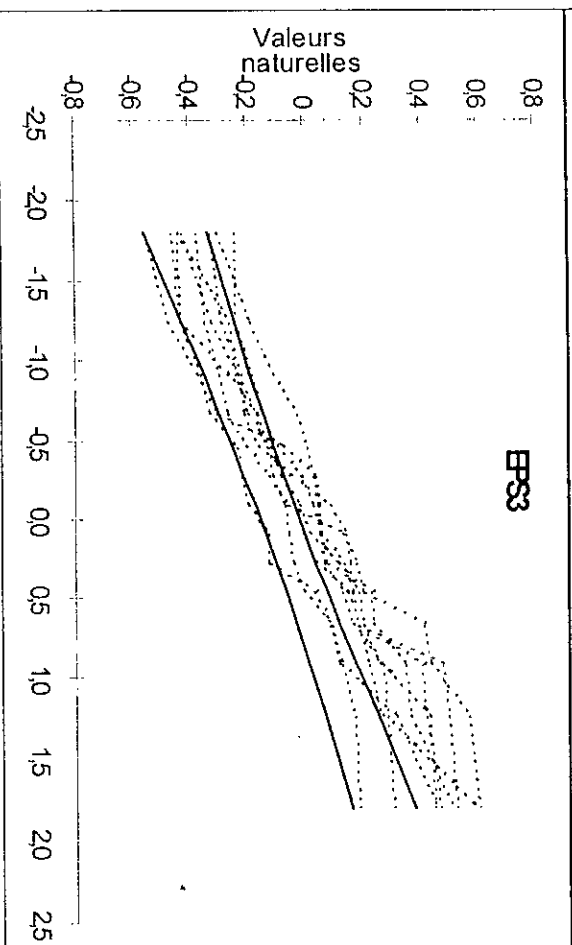
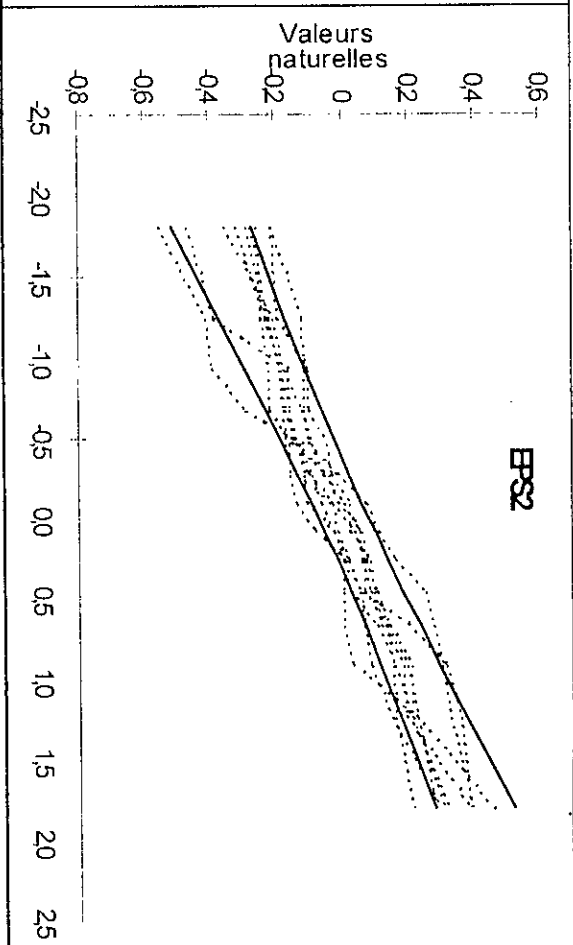
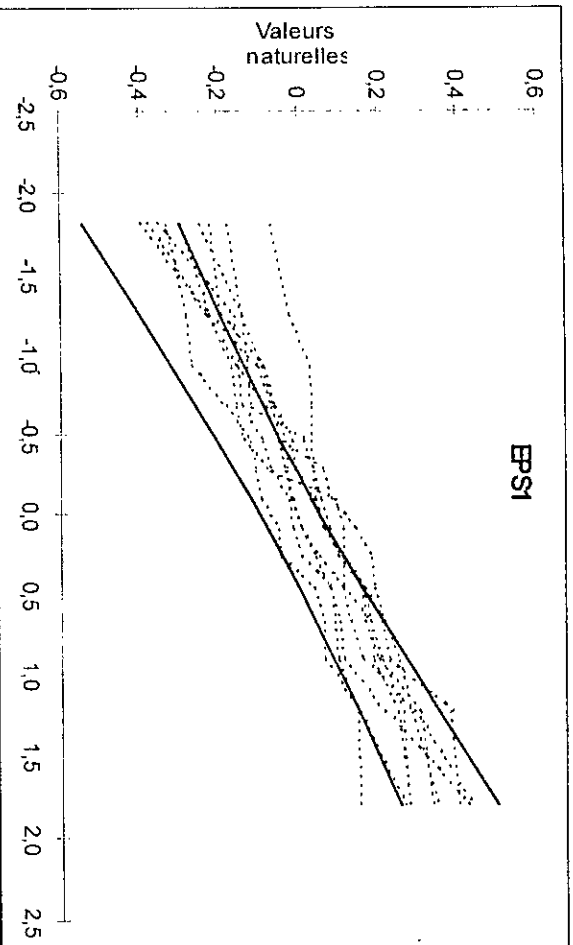




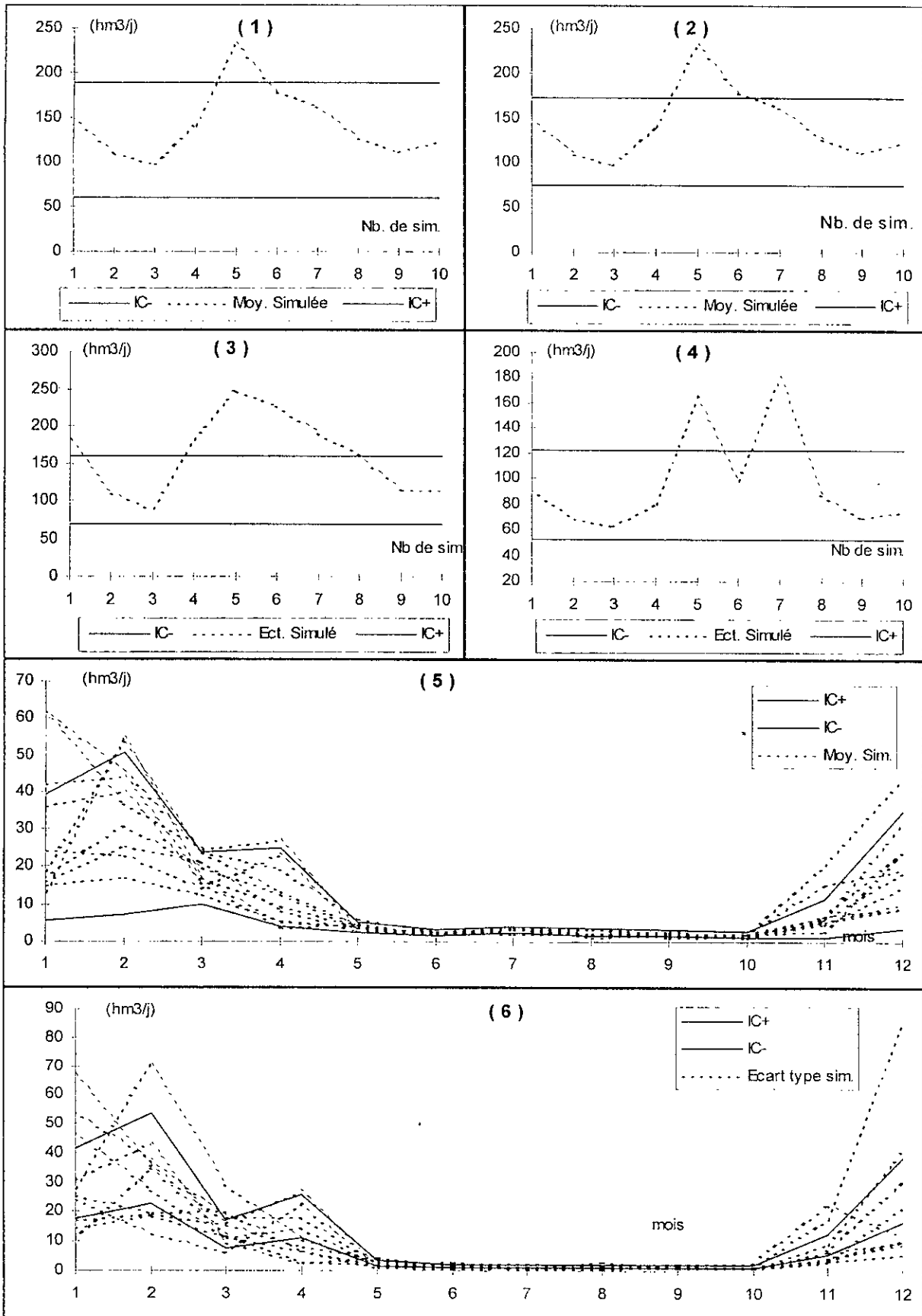
Annexe A.4 (2)



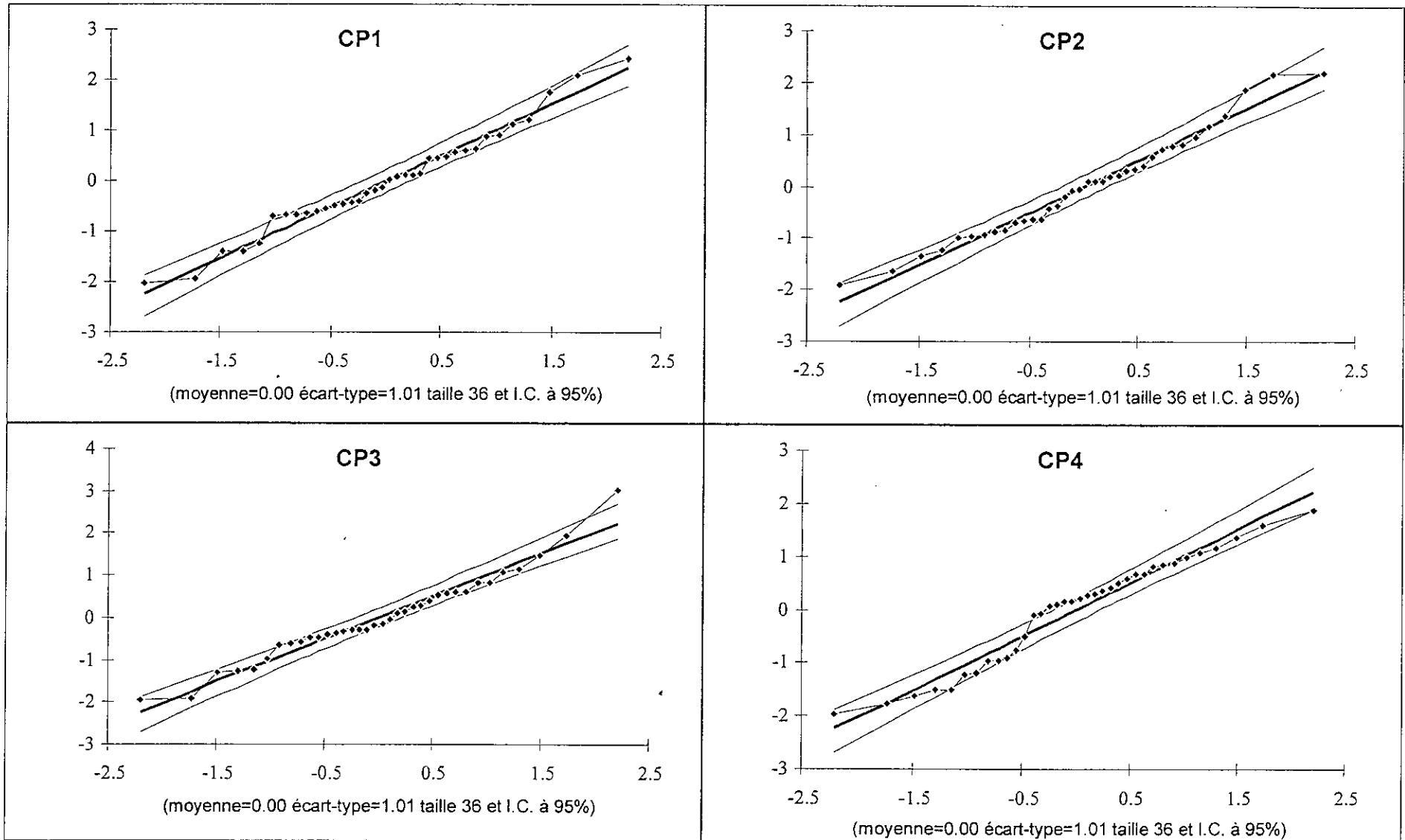


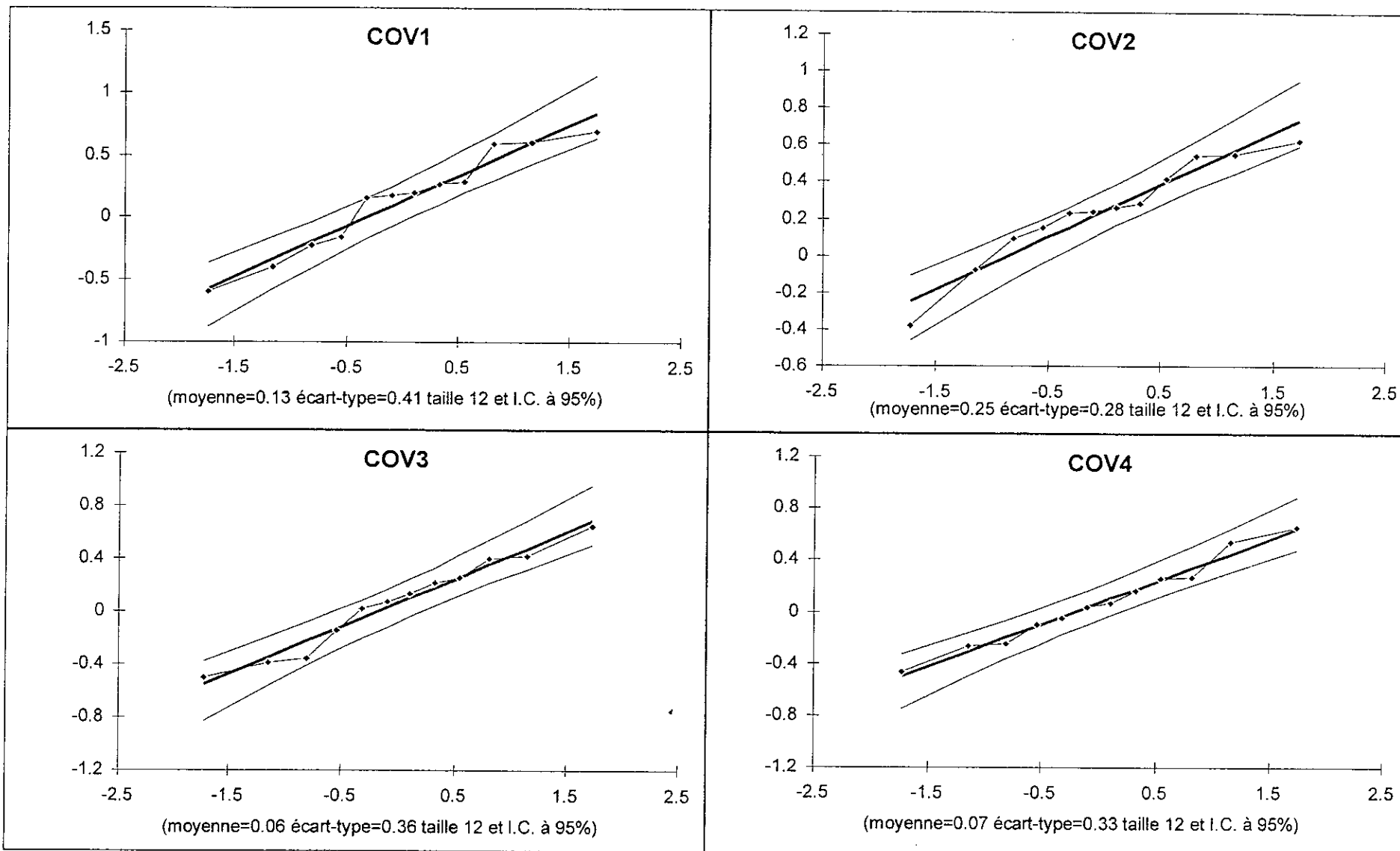


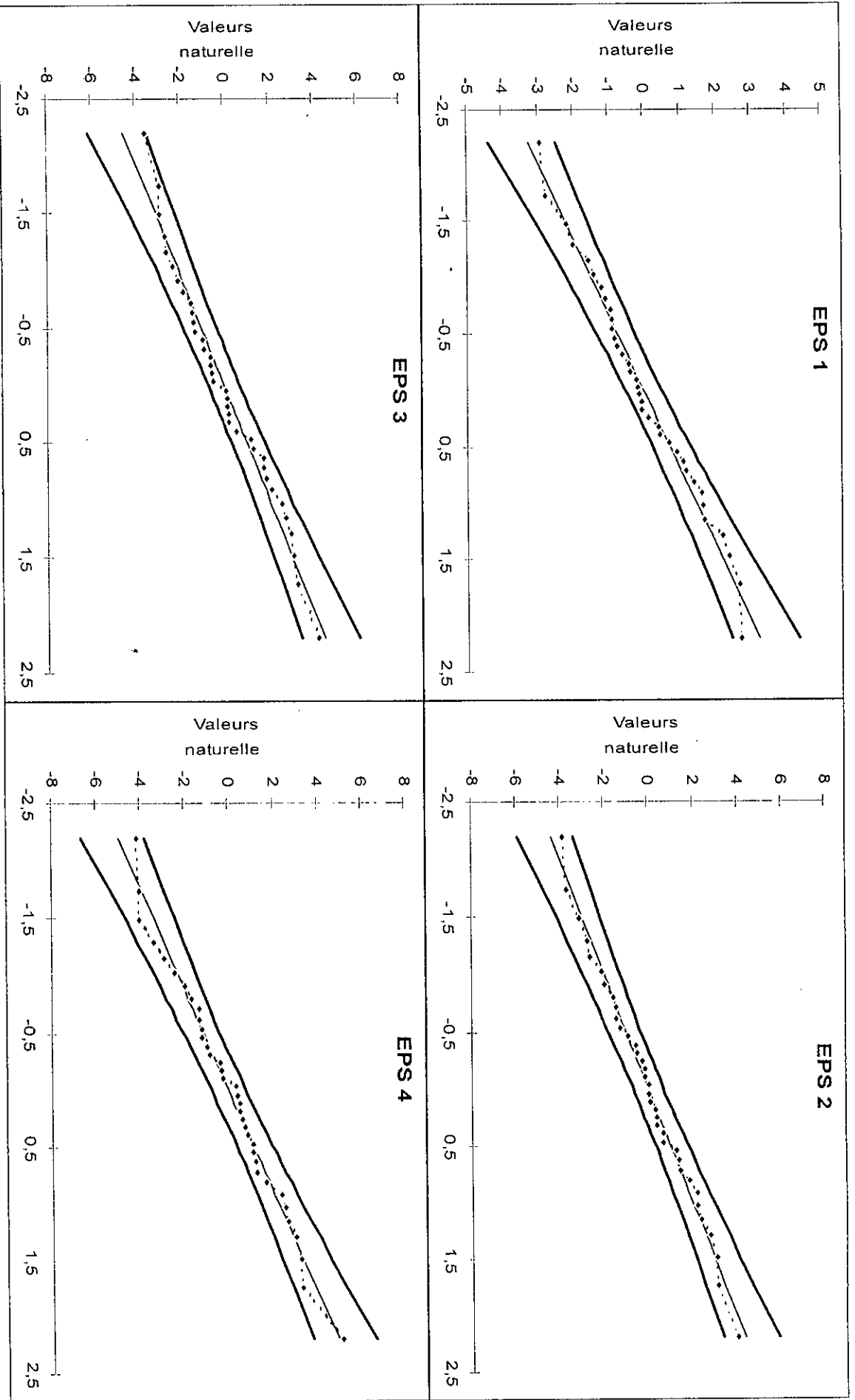
Annexe A.5.: Tests des séries de Débits simulées (Données transformées)



## Annexe A.6 (1): Ajustement des paramètres de simulation des Pluies

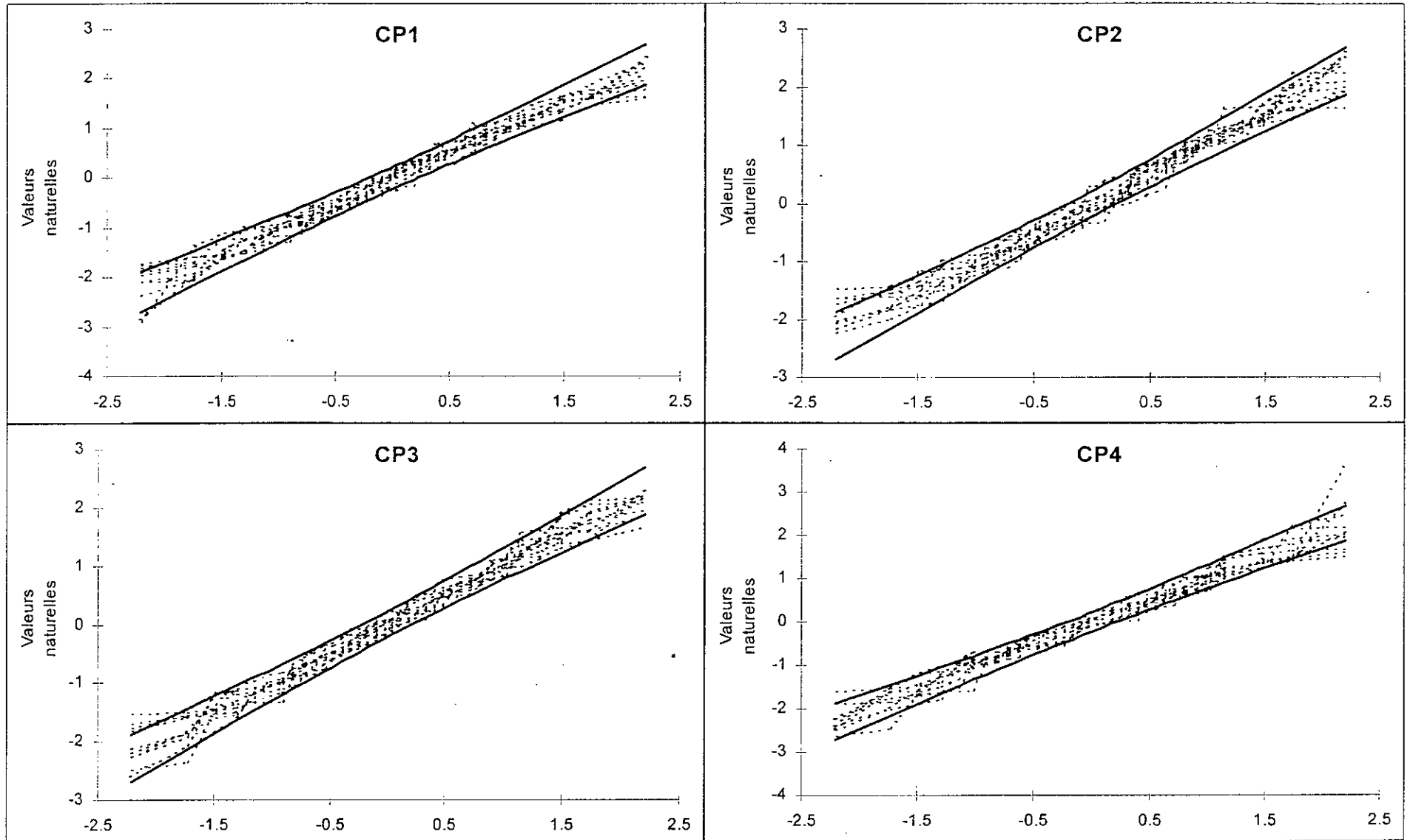


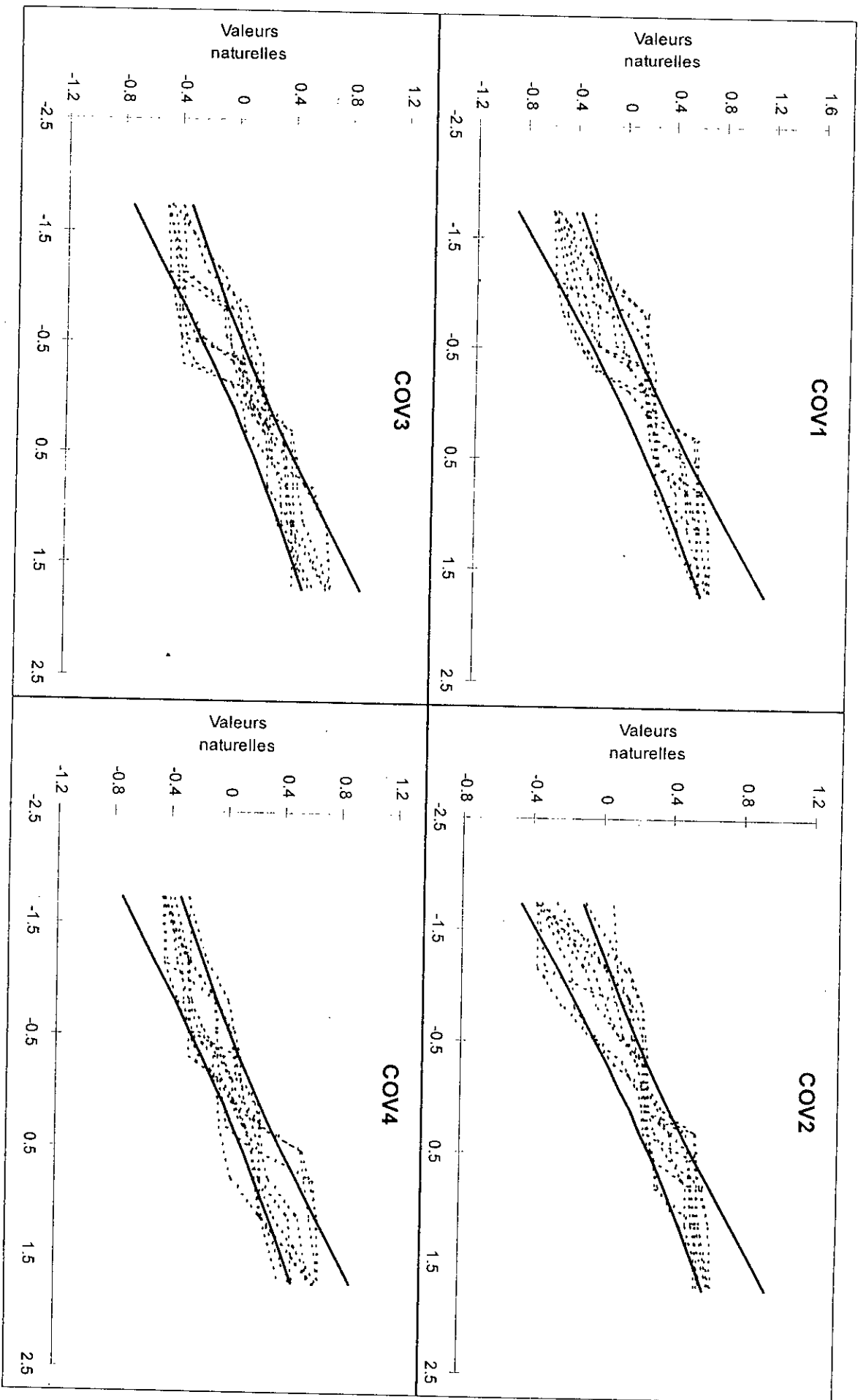


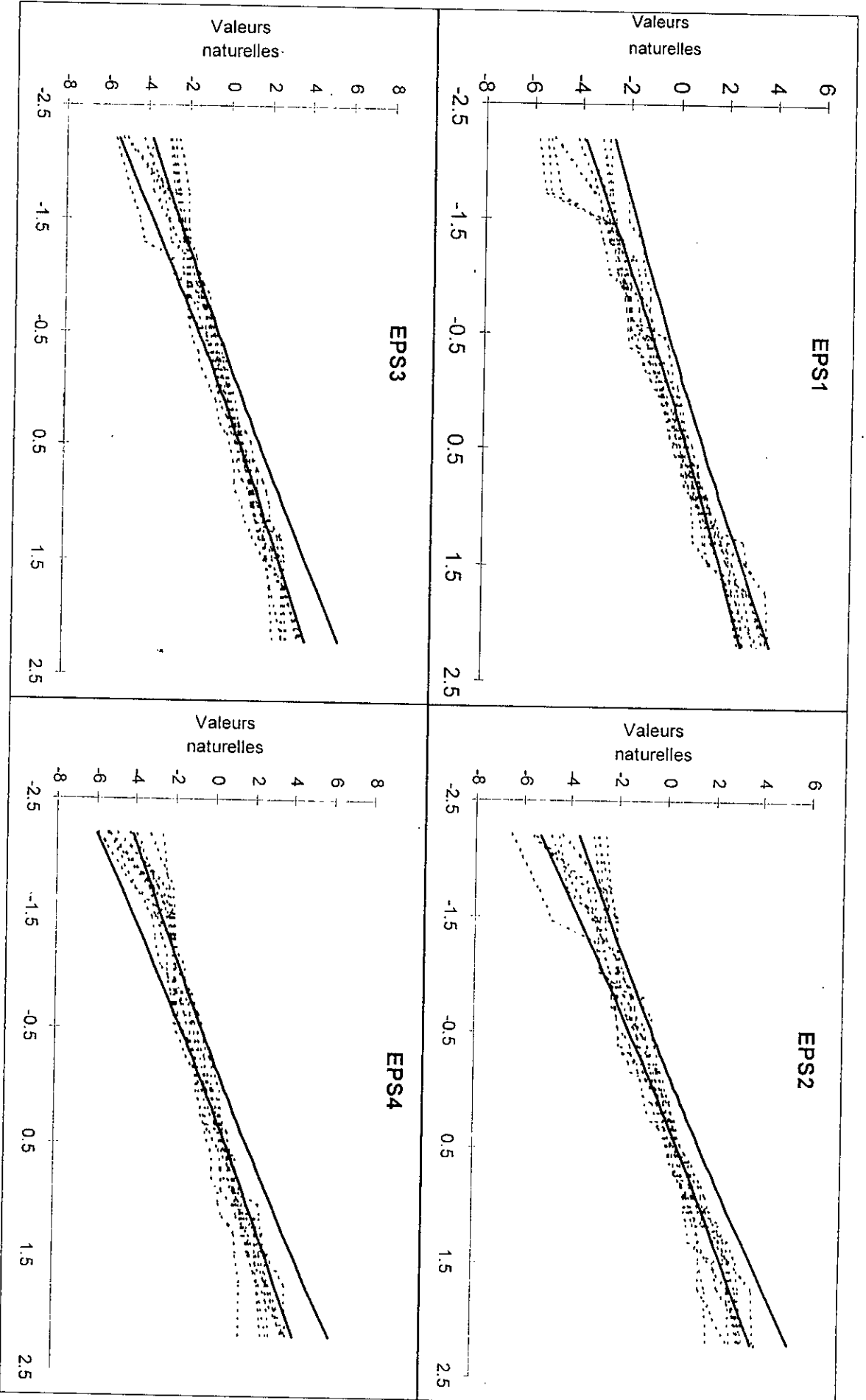




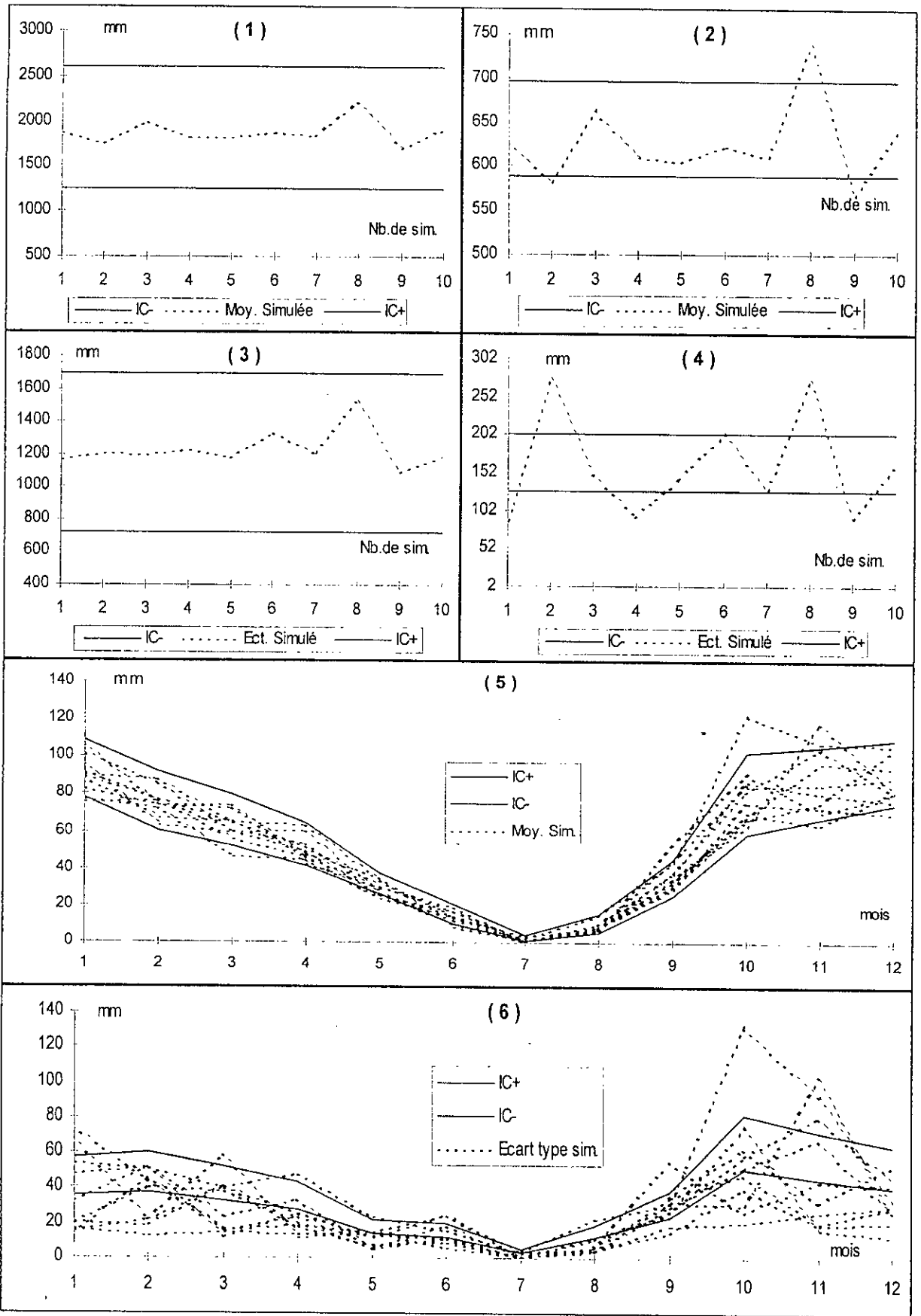
Annexe A.6 (2)







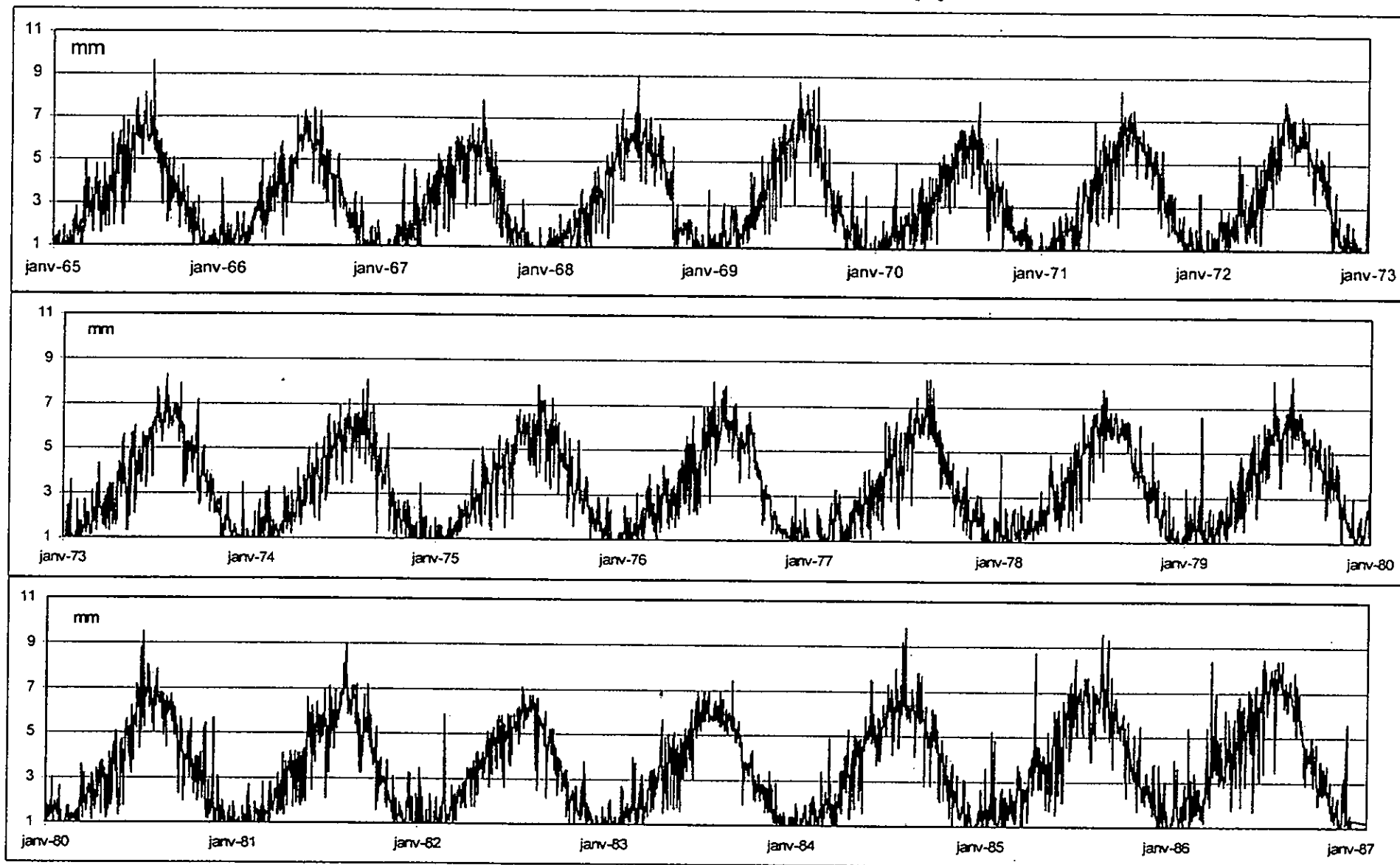
ANNEXE A.7: Tests des séries de Pluies simulées (Données transformées)



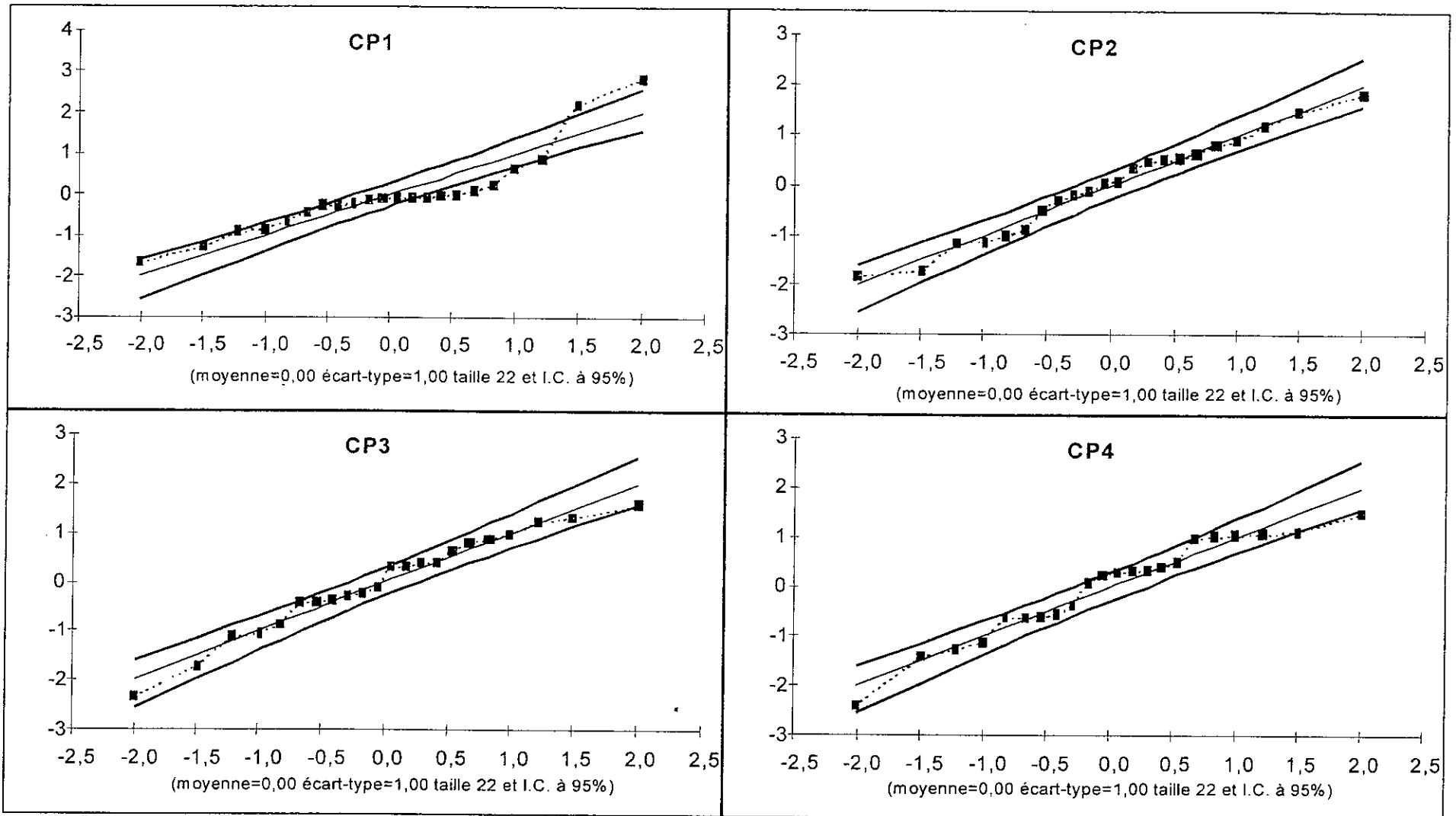
## **ANNEXE B**

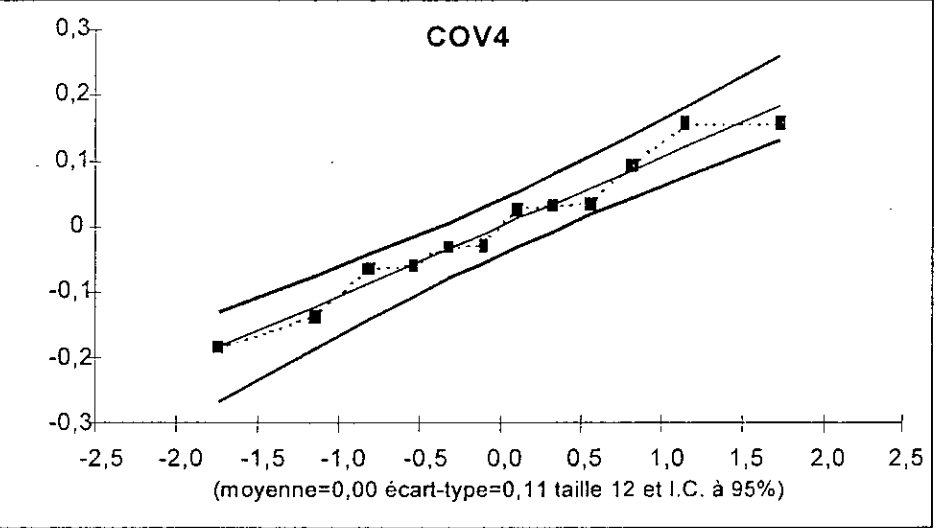
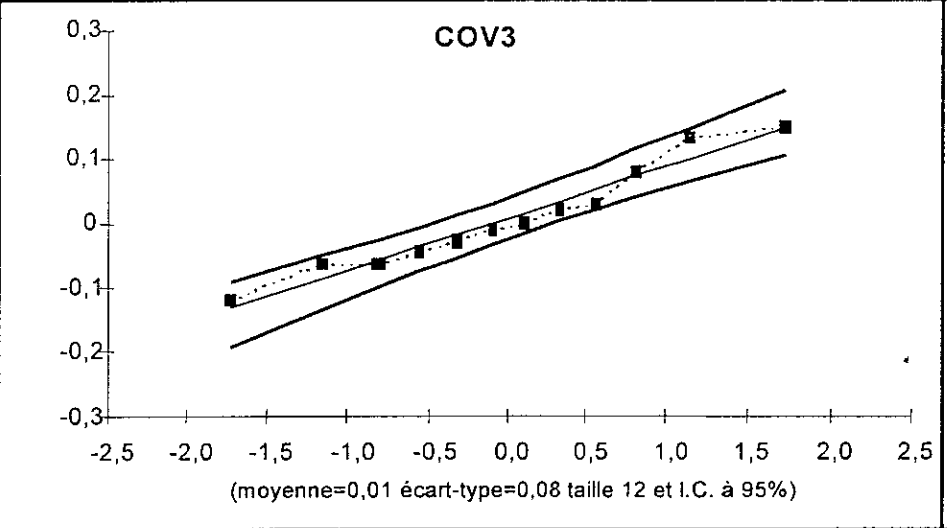
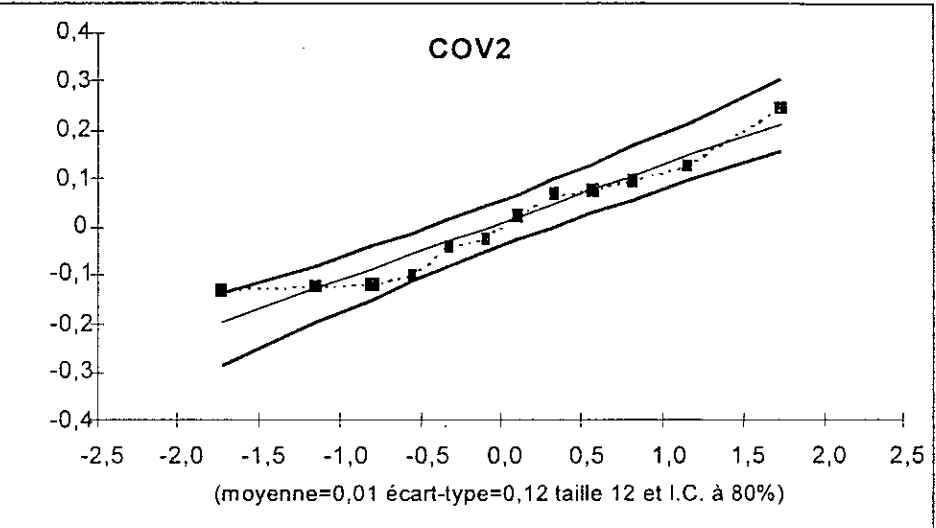
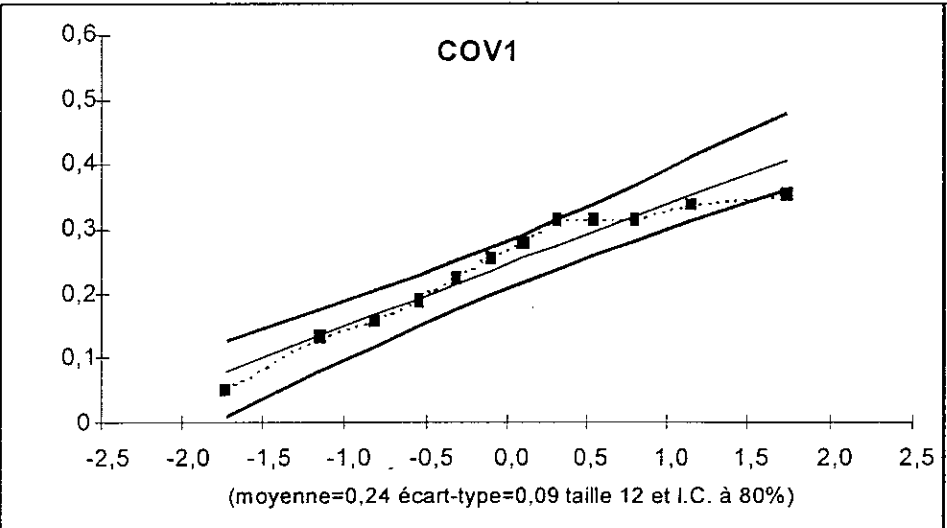
# **SIMULATION AU PAS DE TEMPS JOURNALIER**

Annexe B.1: Evolution chronologique des ETP au pas de temps journalier

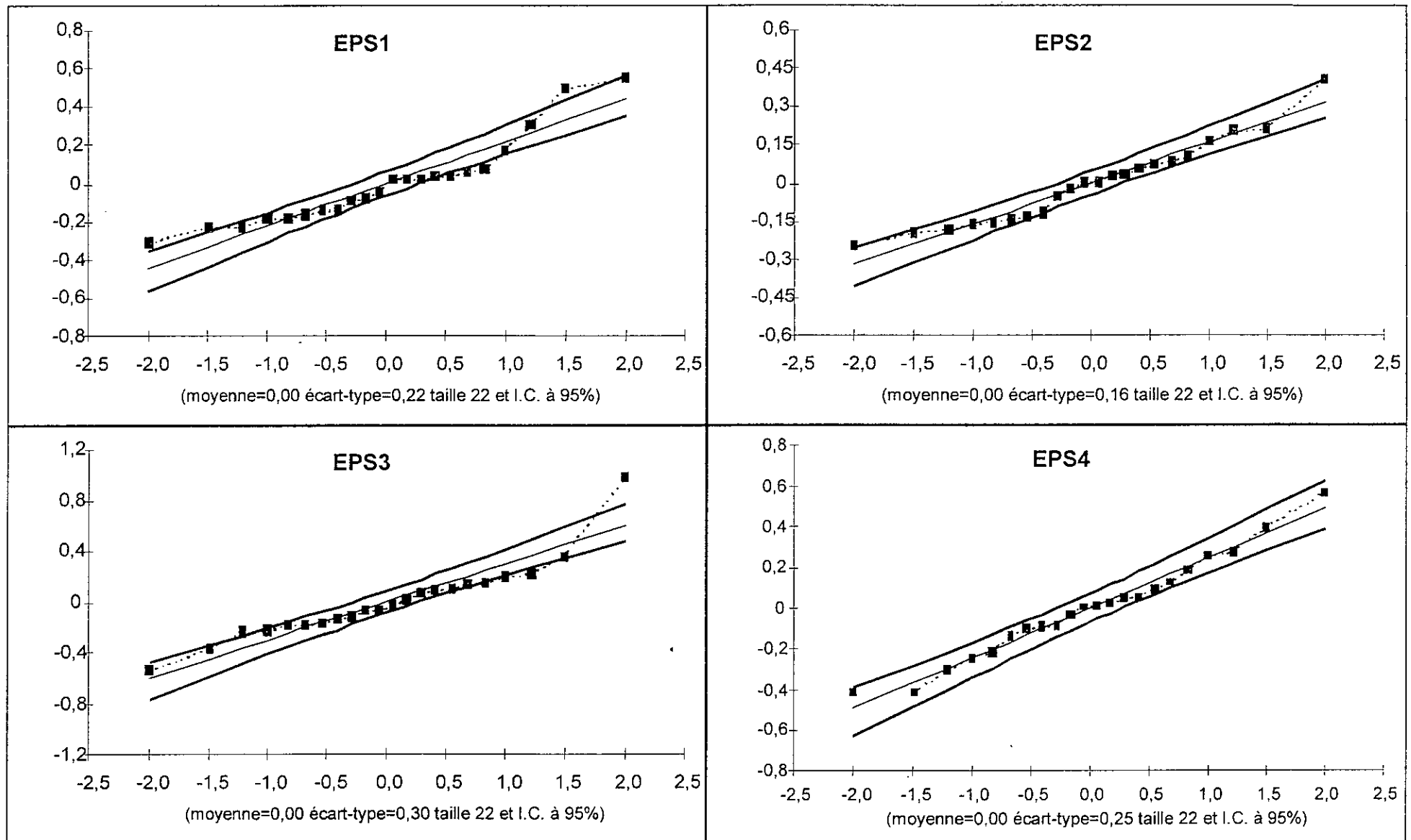


## Annexe B.2 (1) Ajustement des paramètres de simulation des ETP -

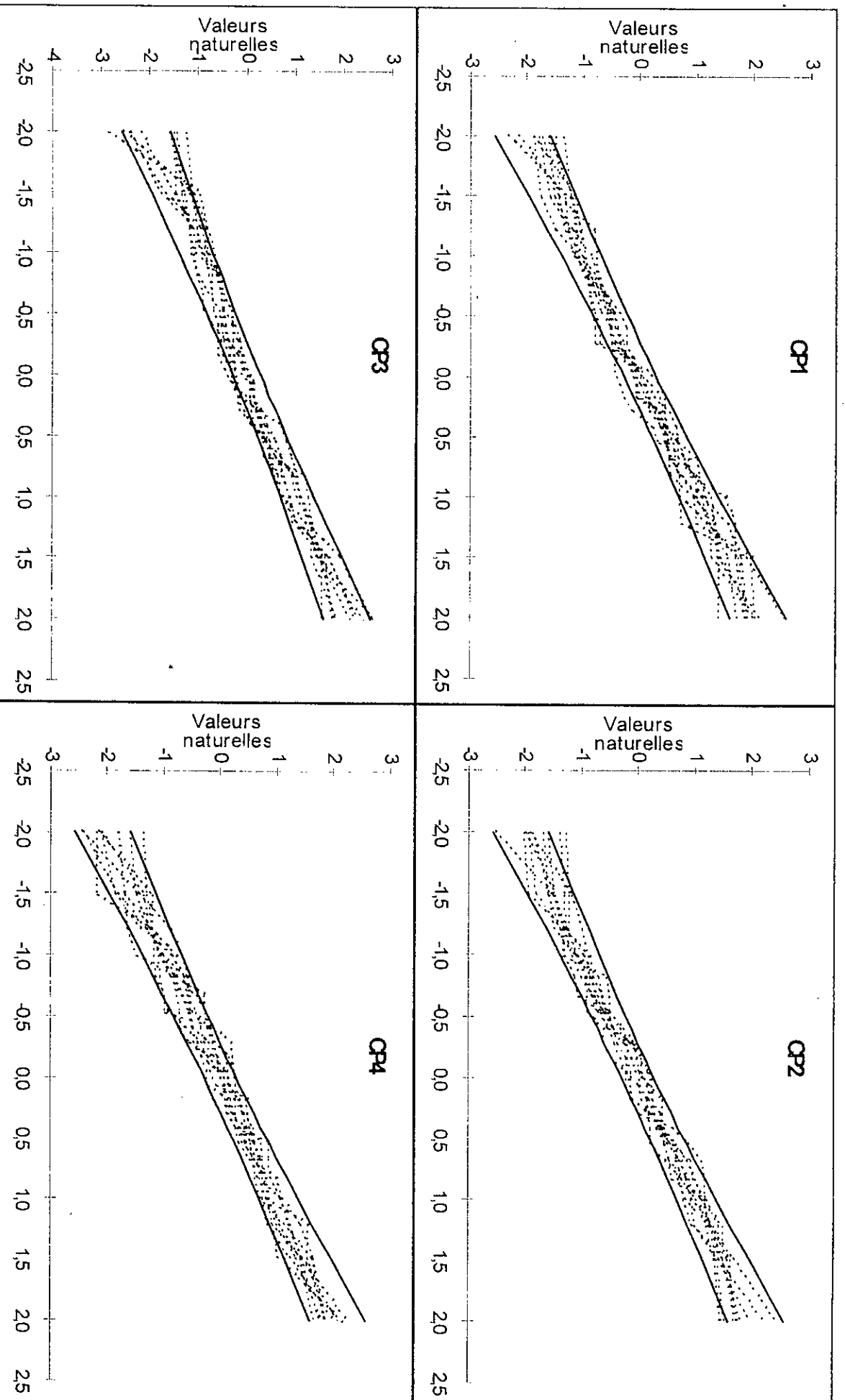


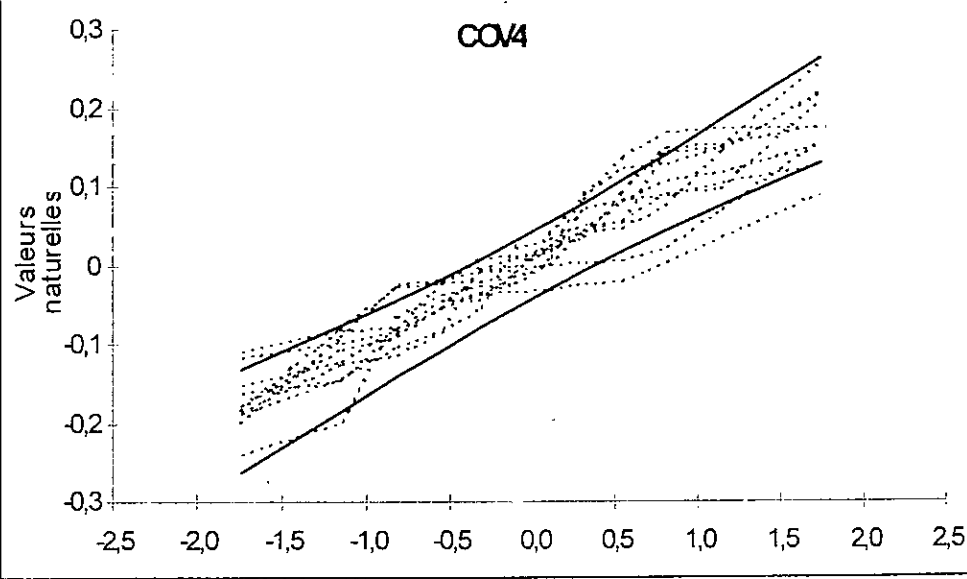
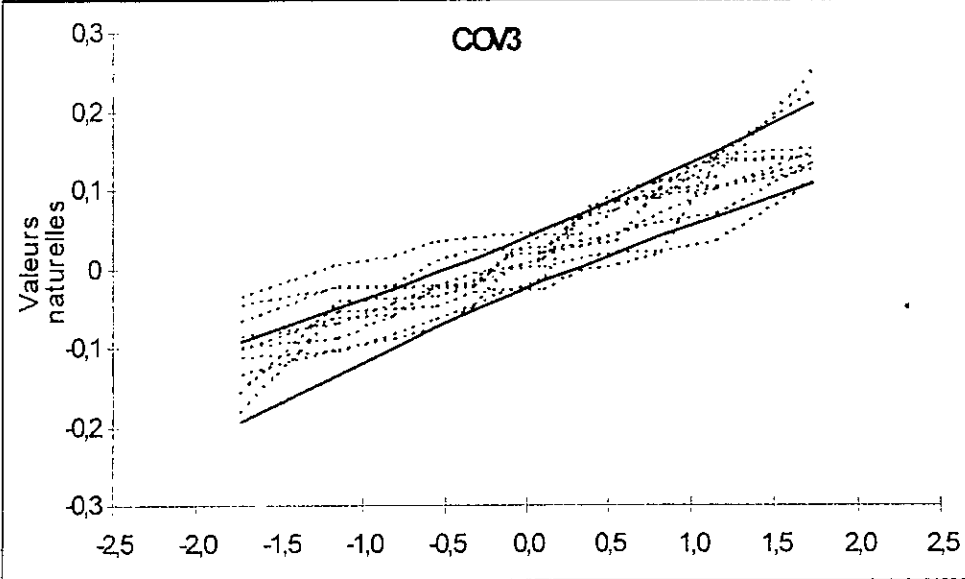
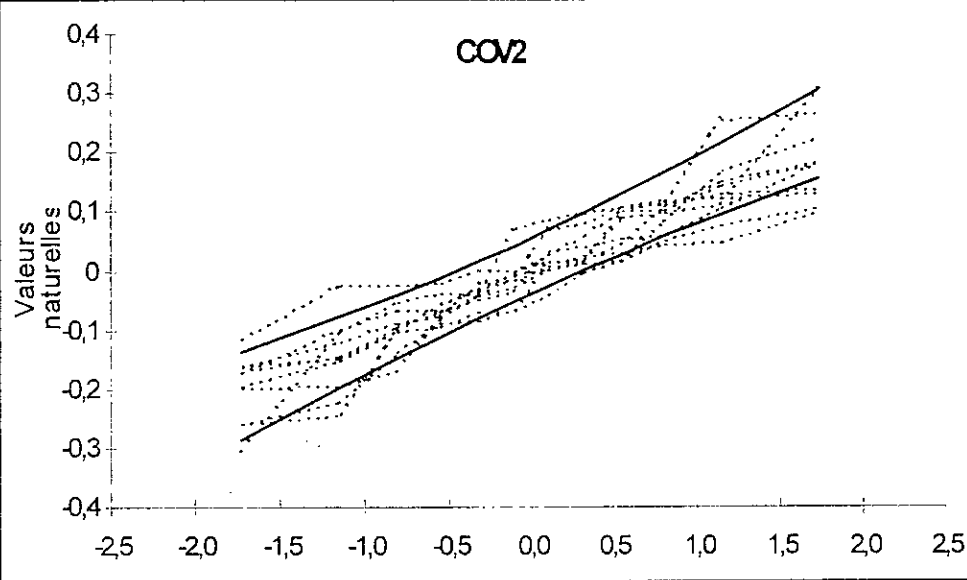
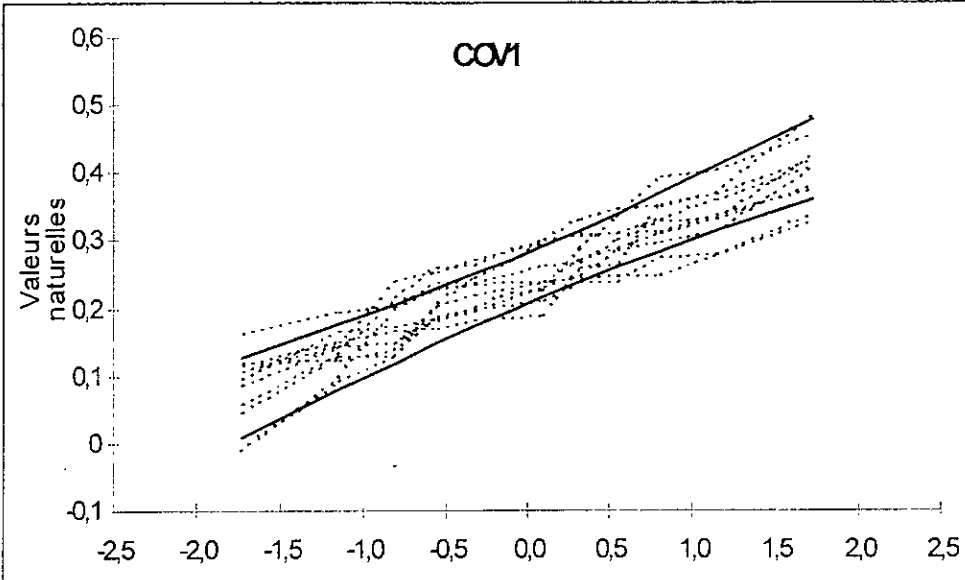


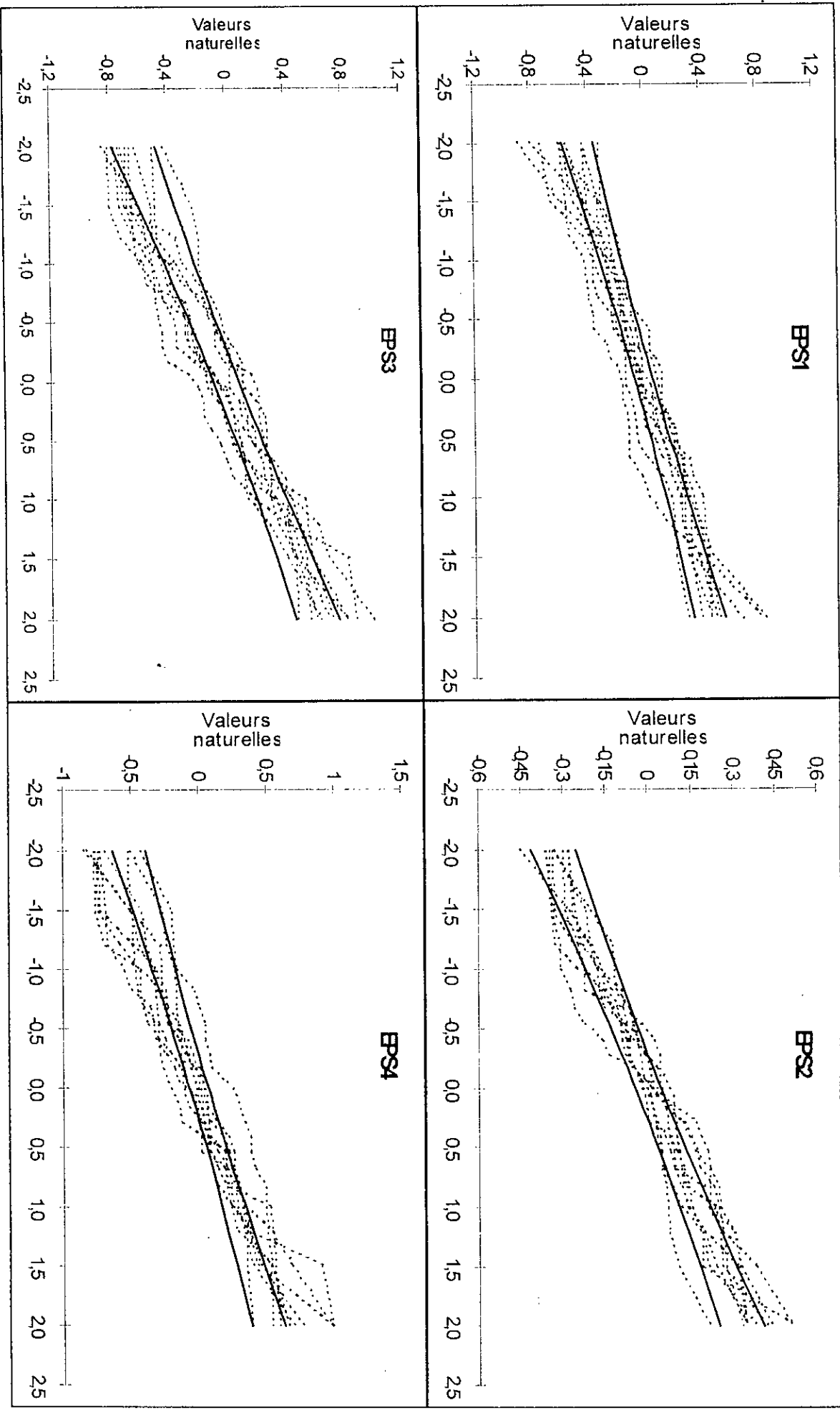




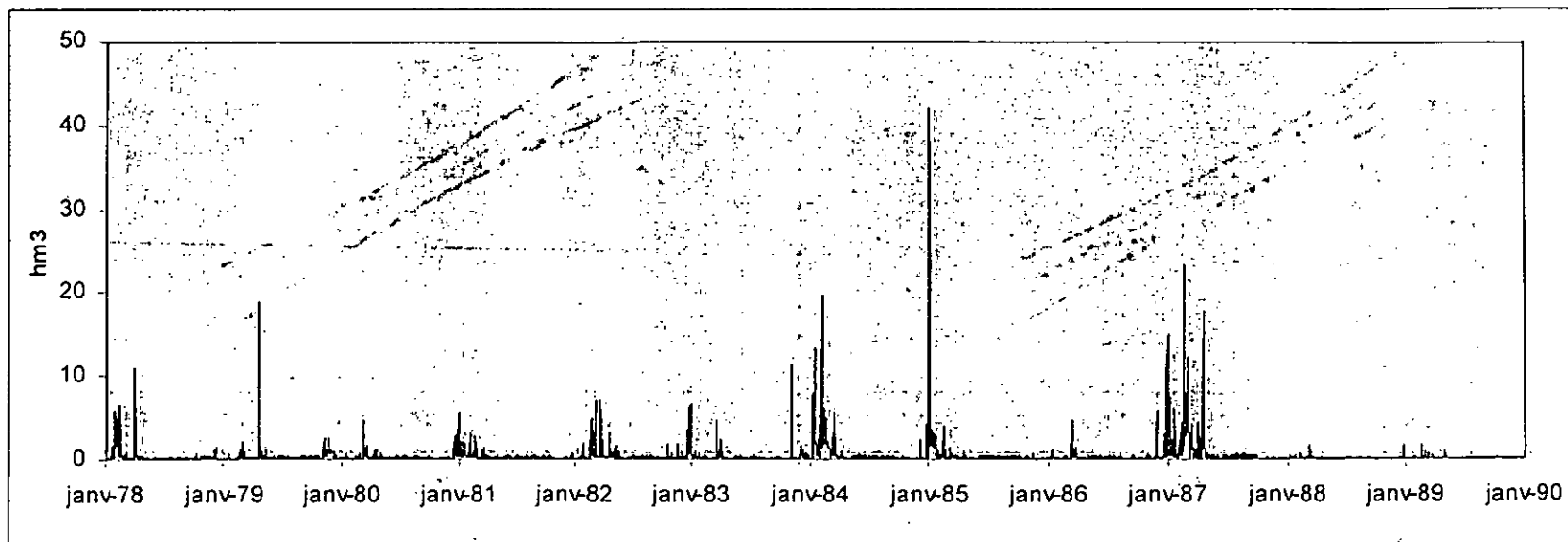
Annexe B.2 (2)

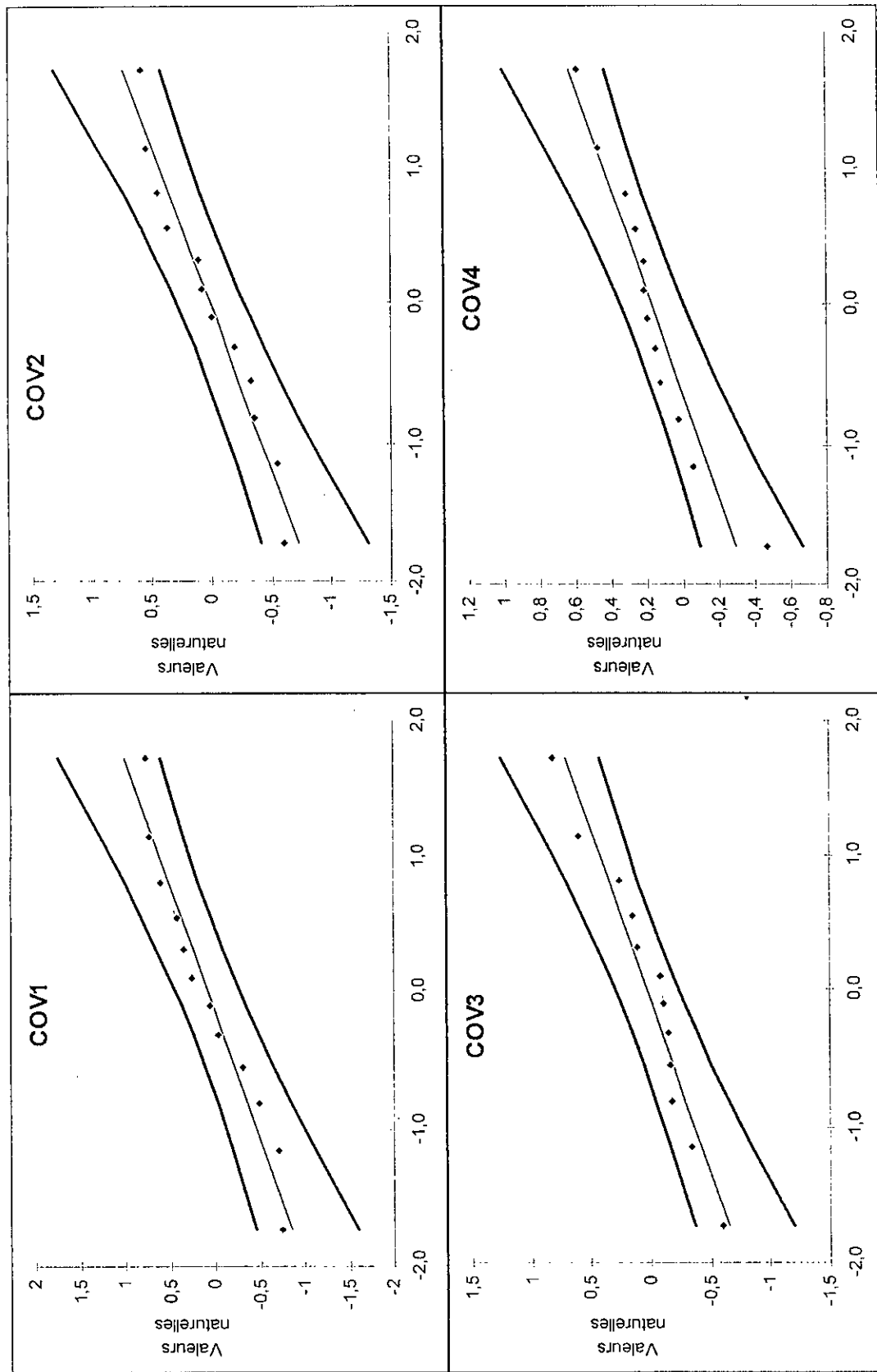


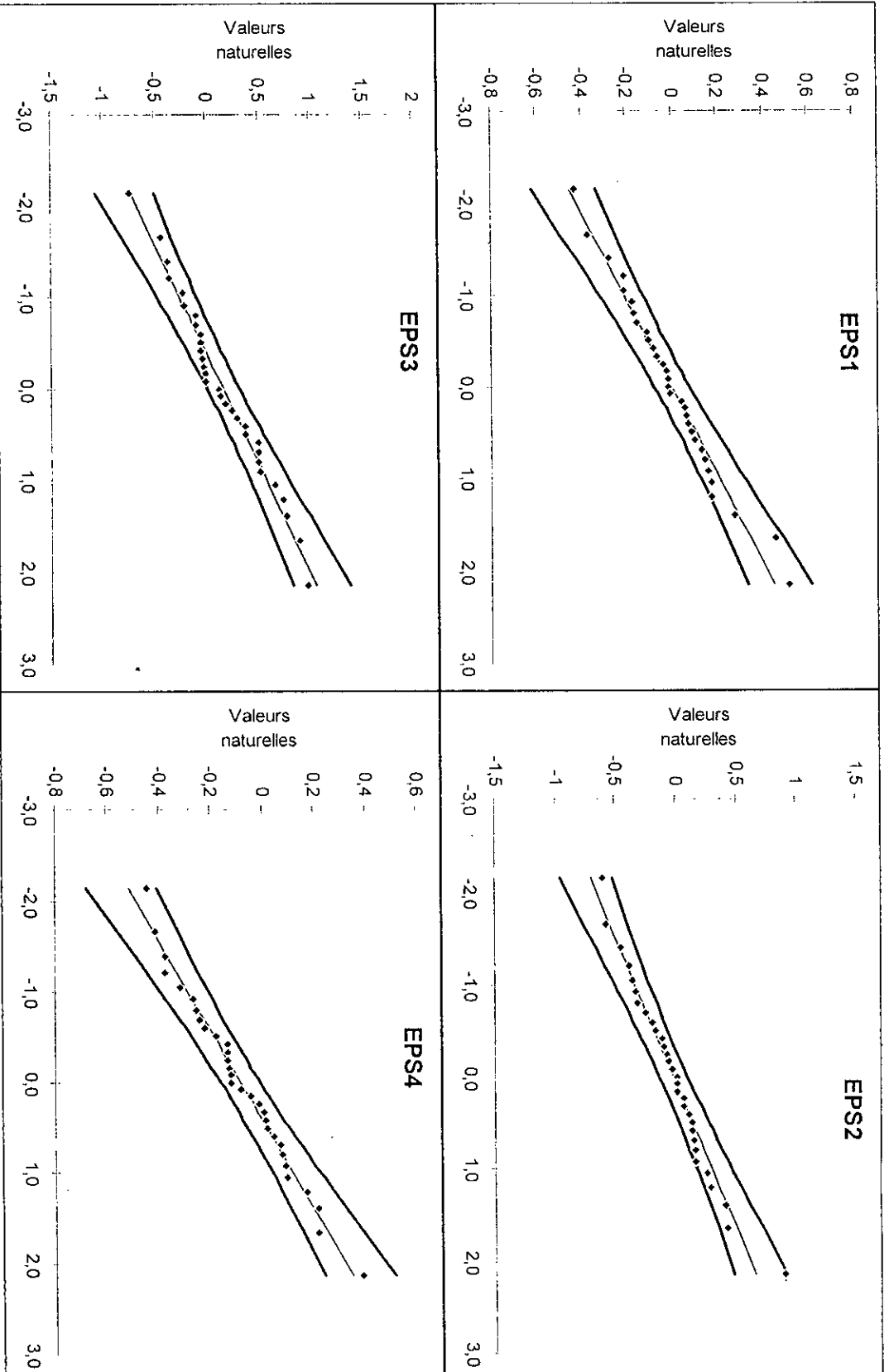




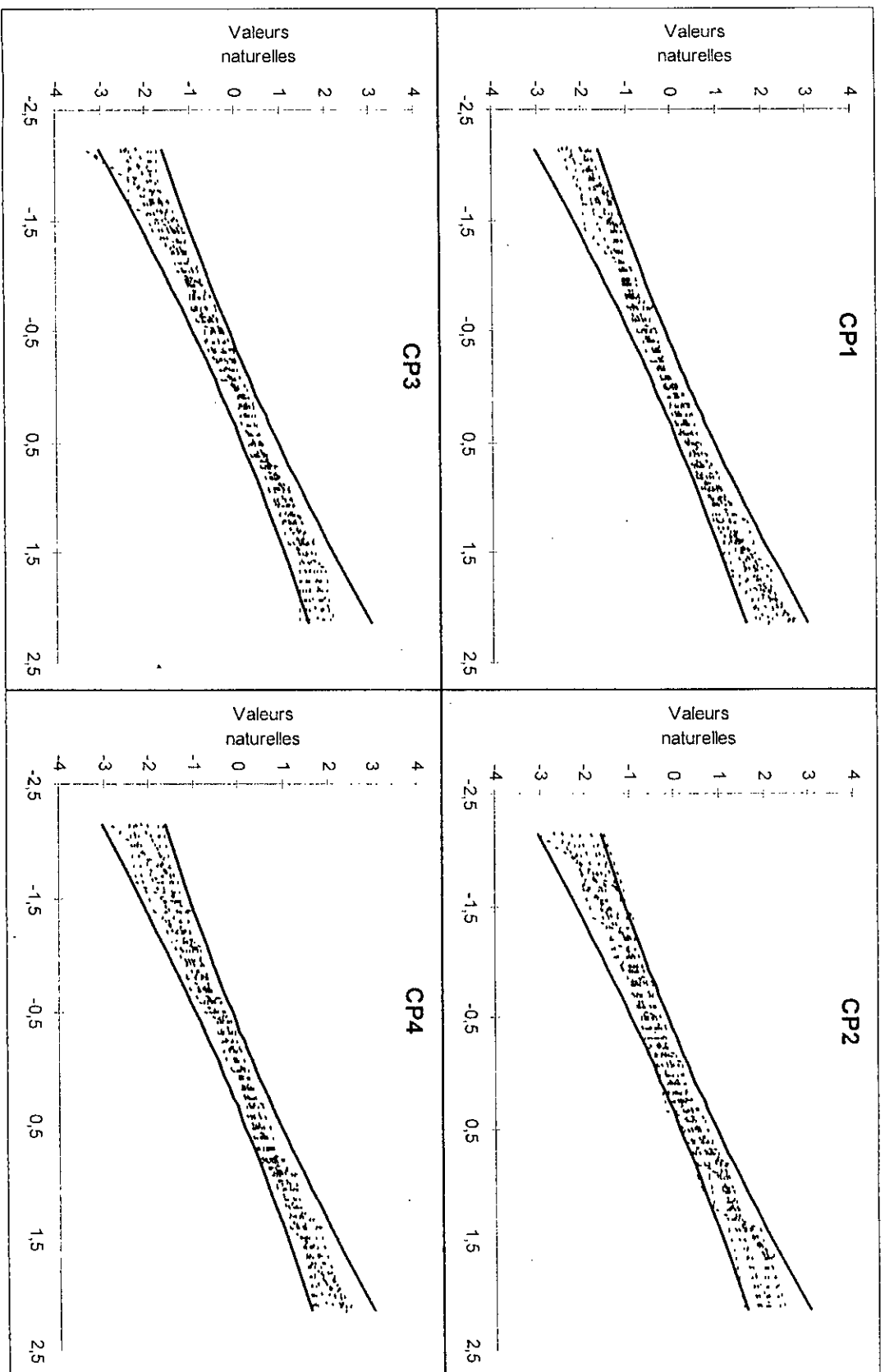
Annexe B.3: Evolution chronologique des débits au pas de temps journalier



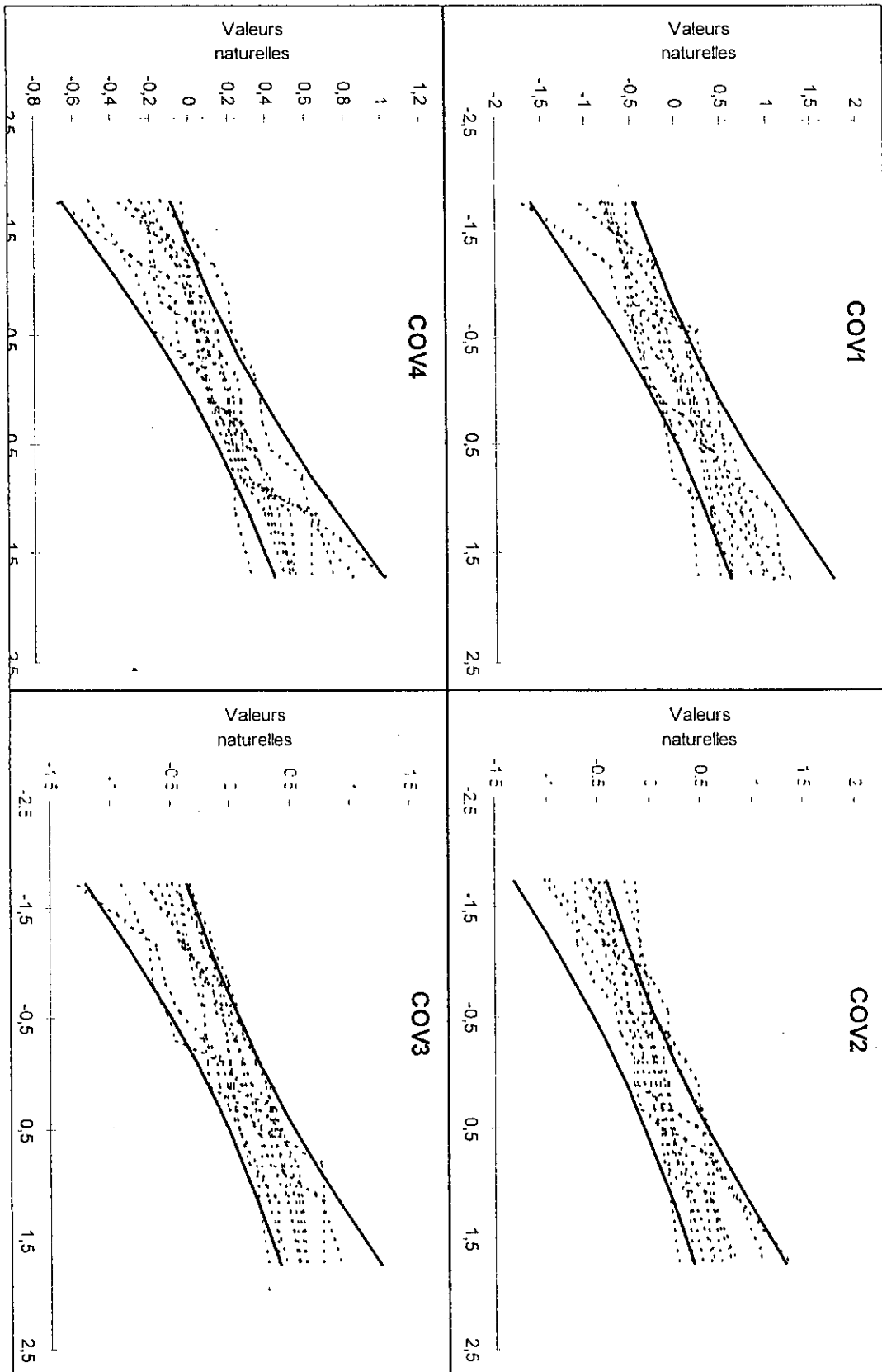


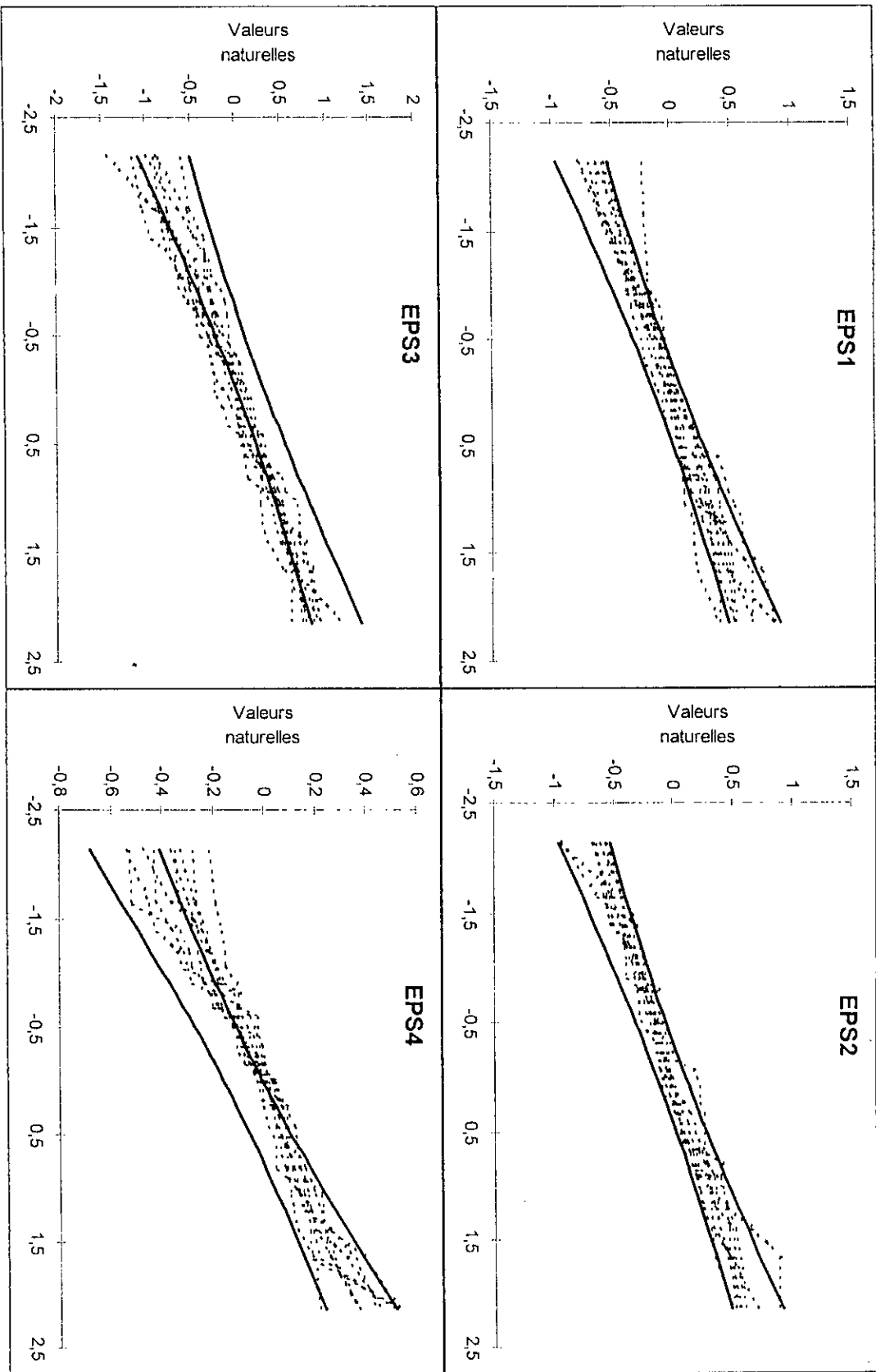


Annexe B.4 (2)

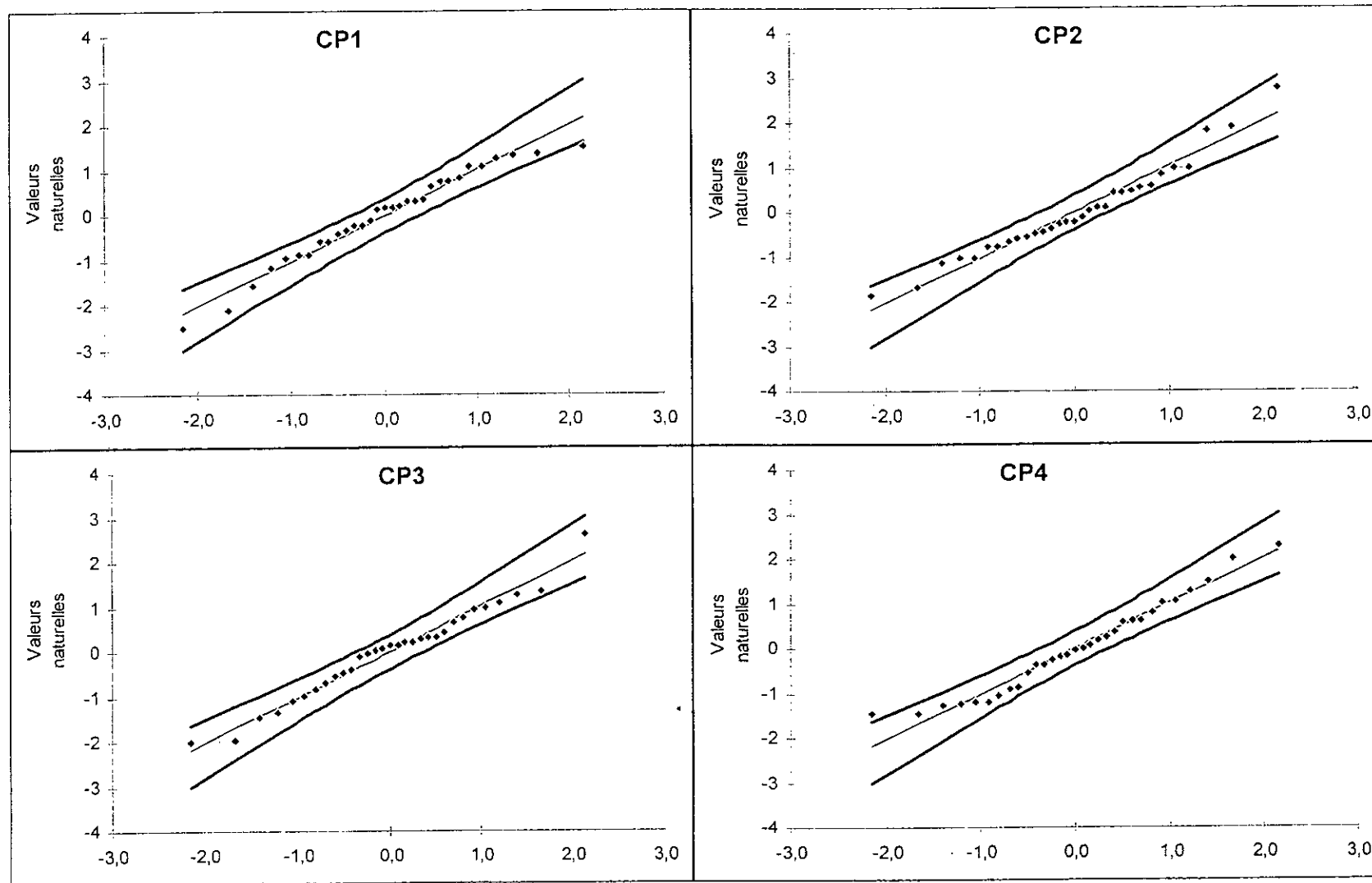


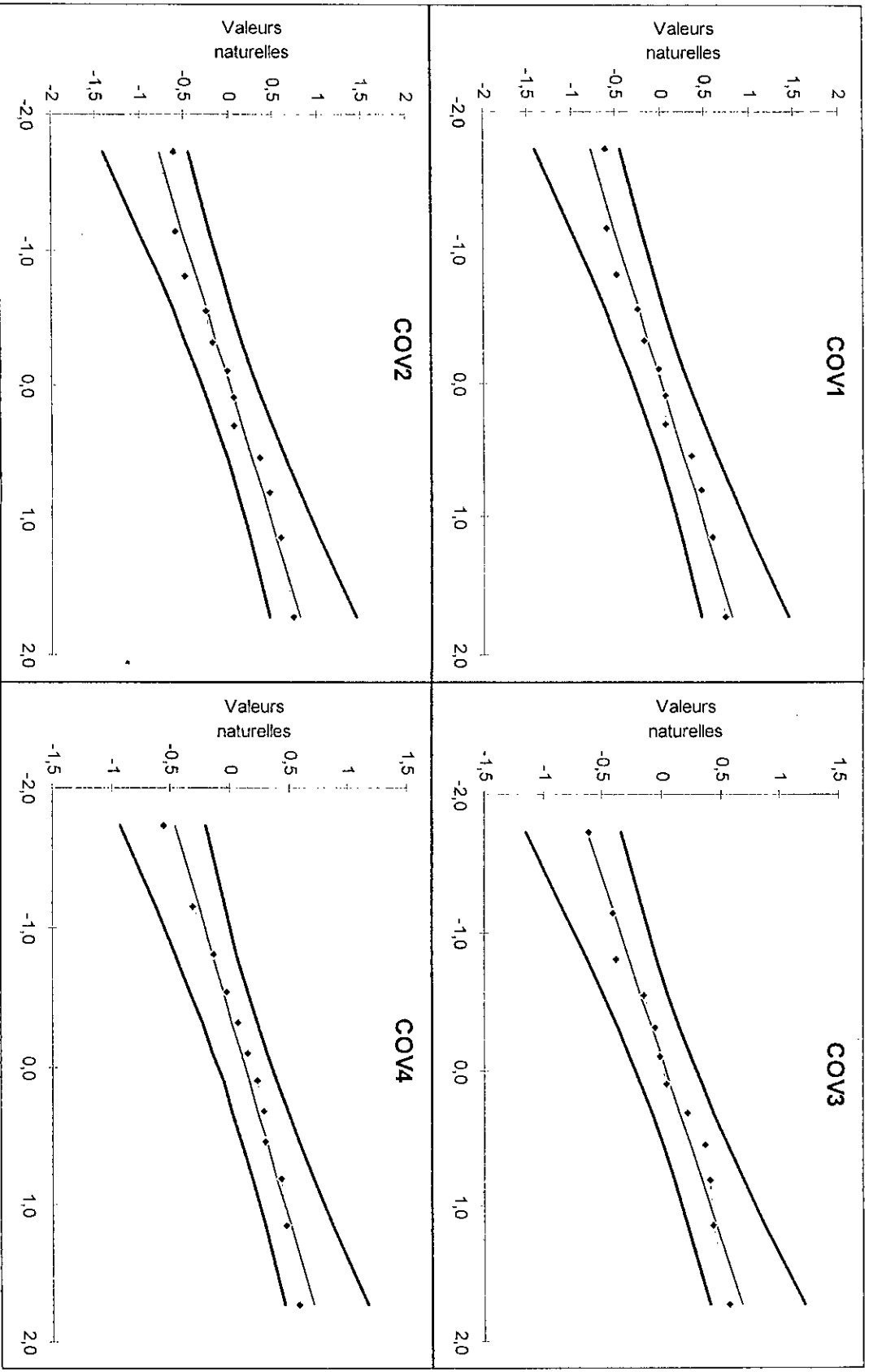


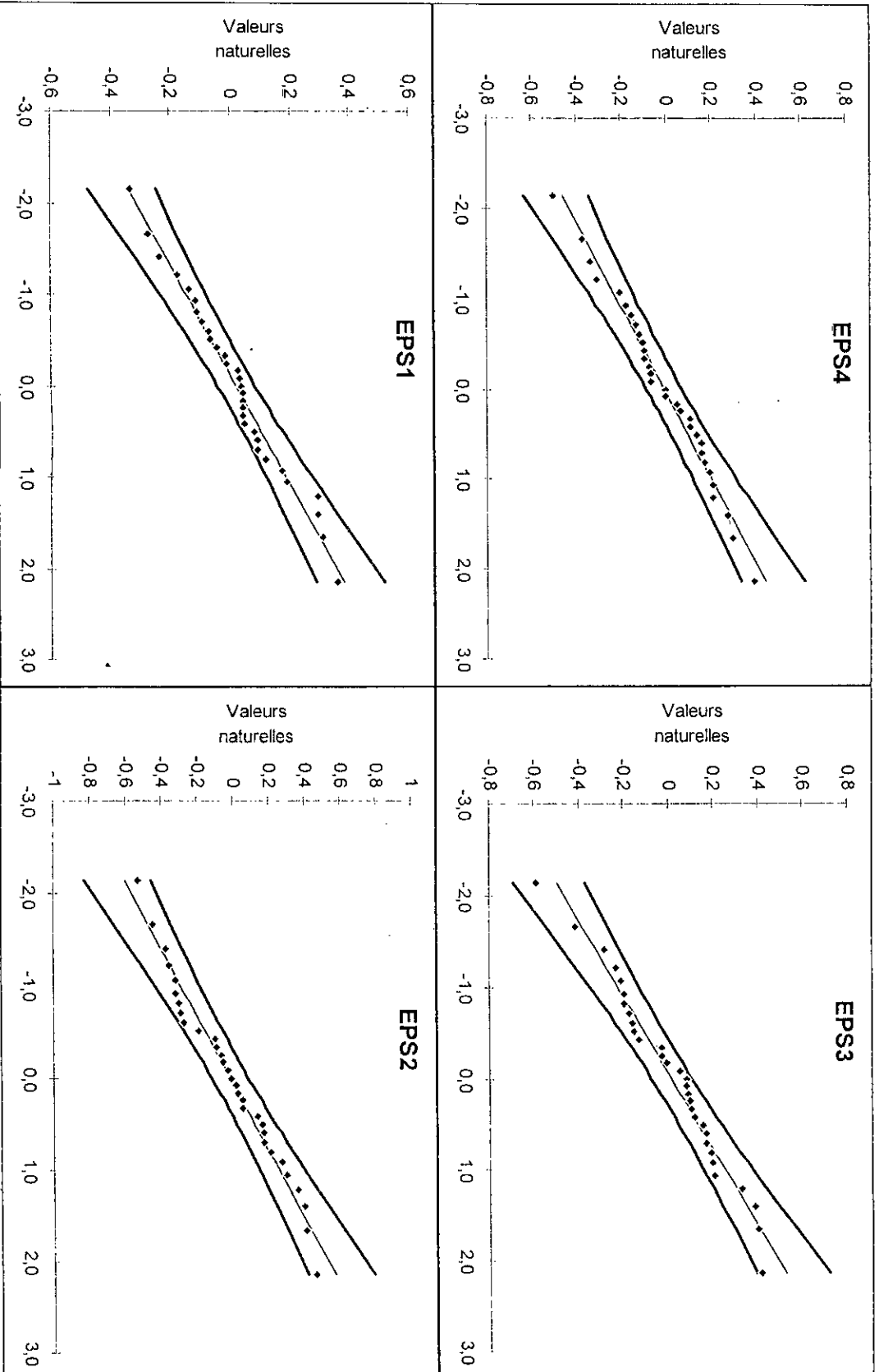




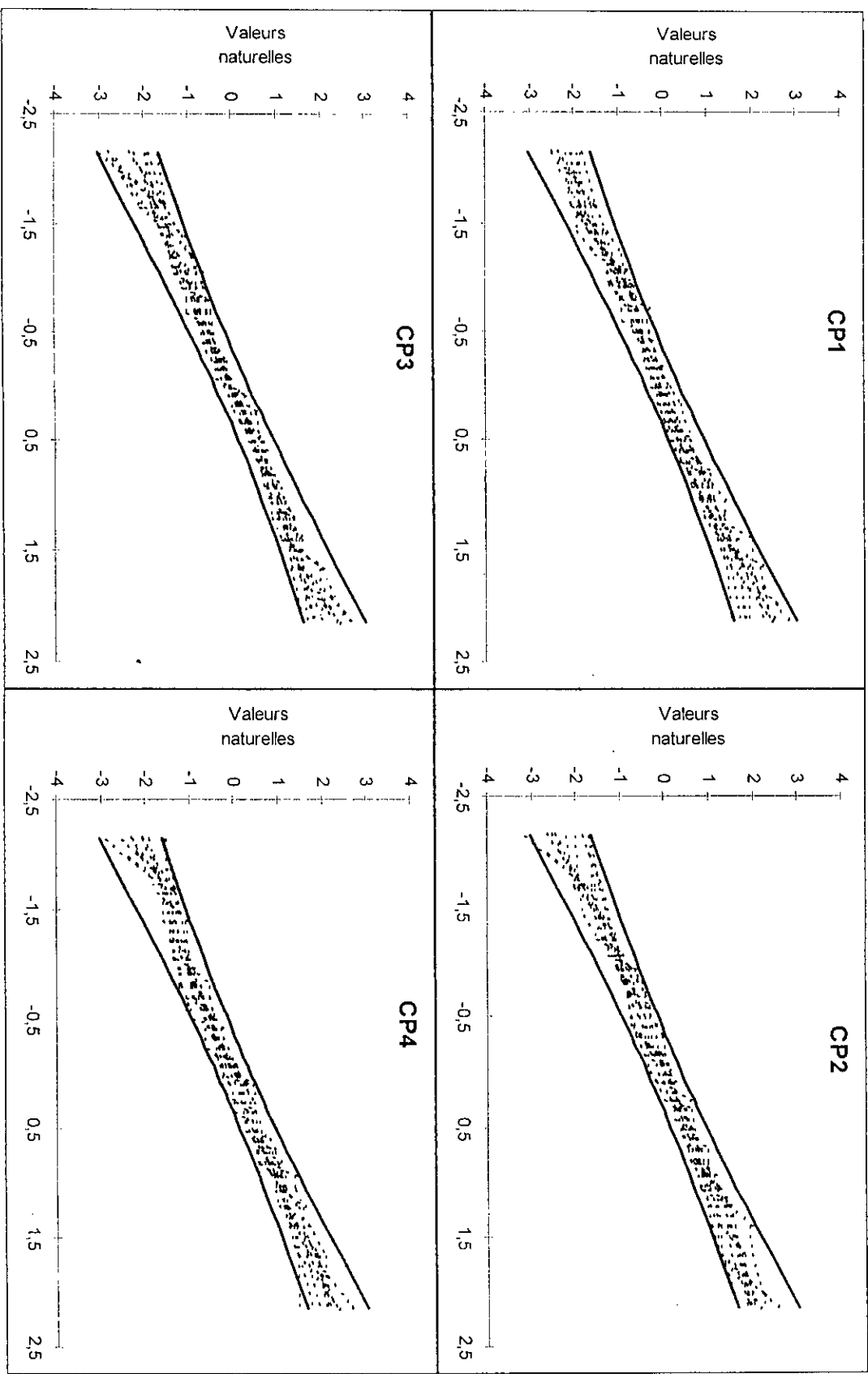
Annexe B.5 (1): Ajustement des paramètres de simulation des Débits (Juillet)

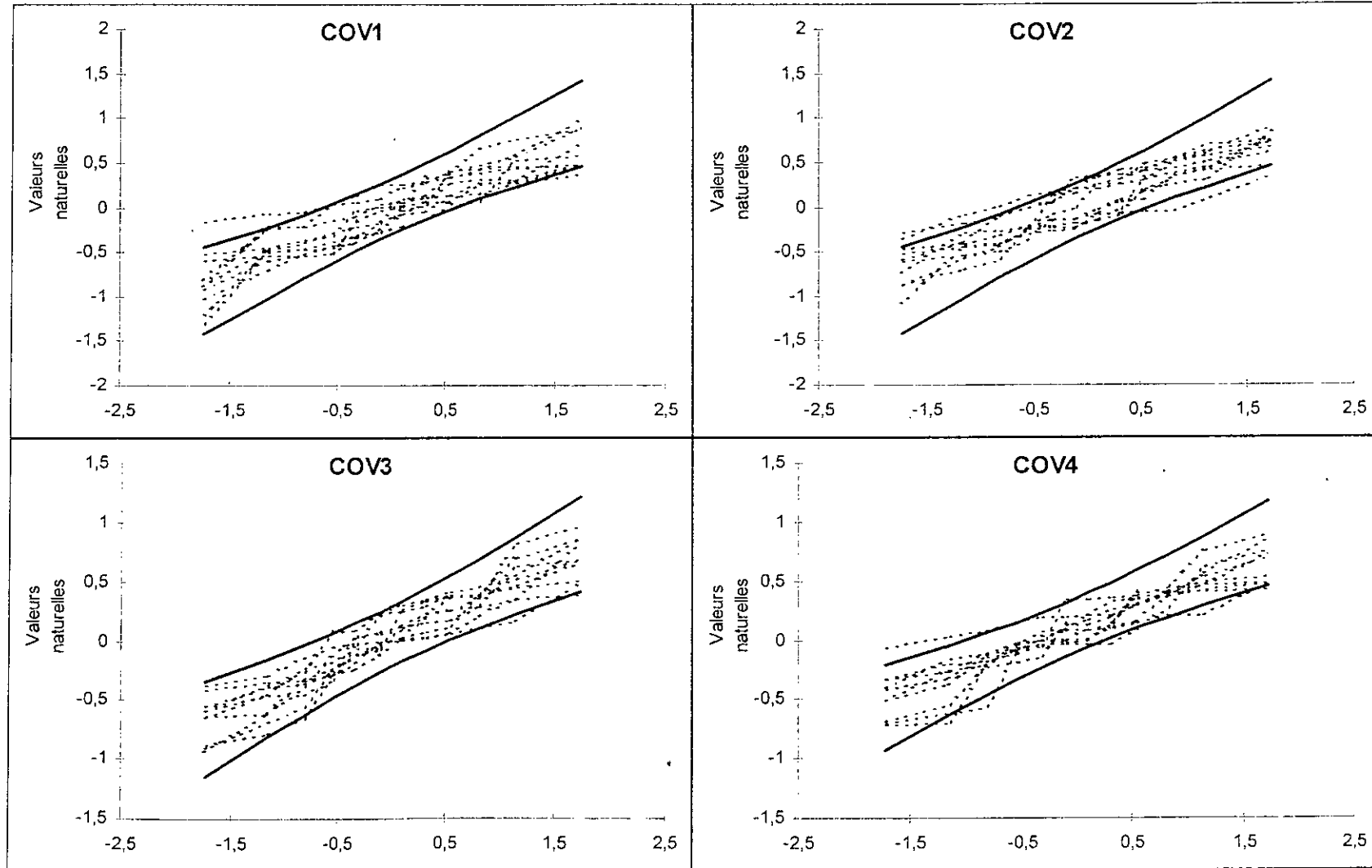


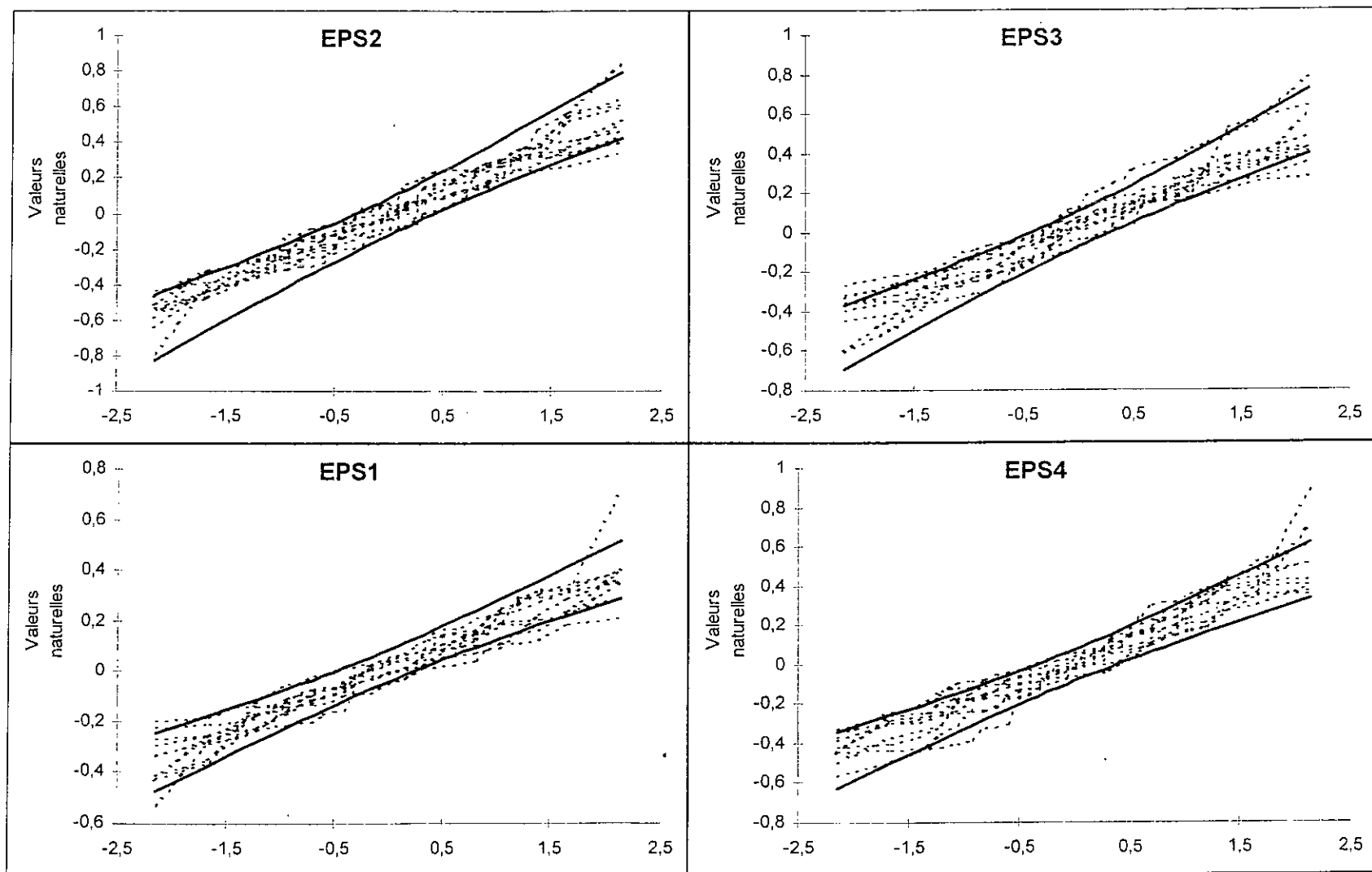




Annexe B.5 (2)

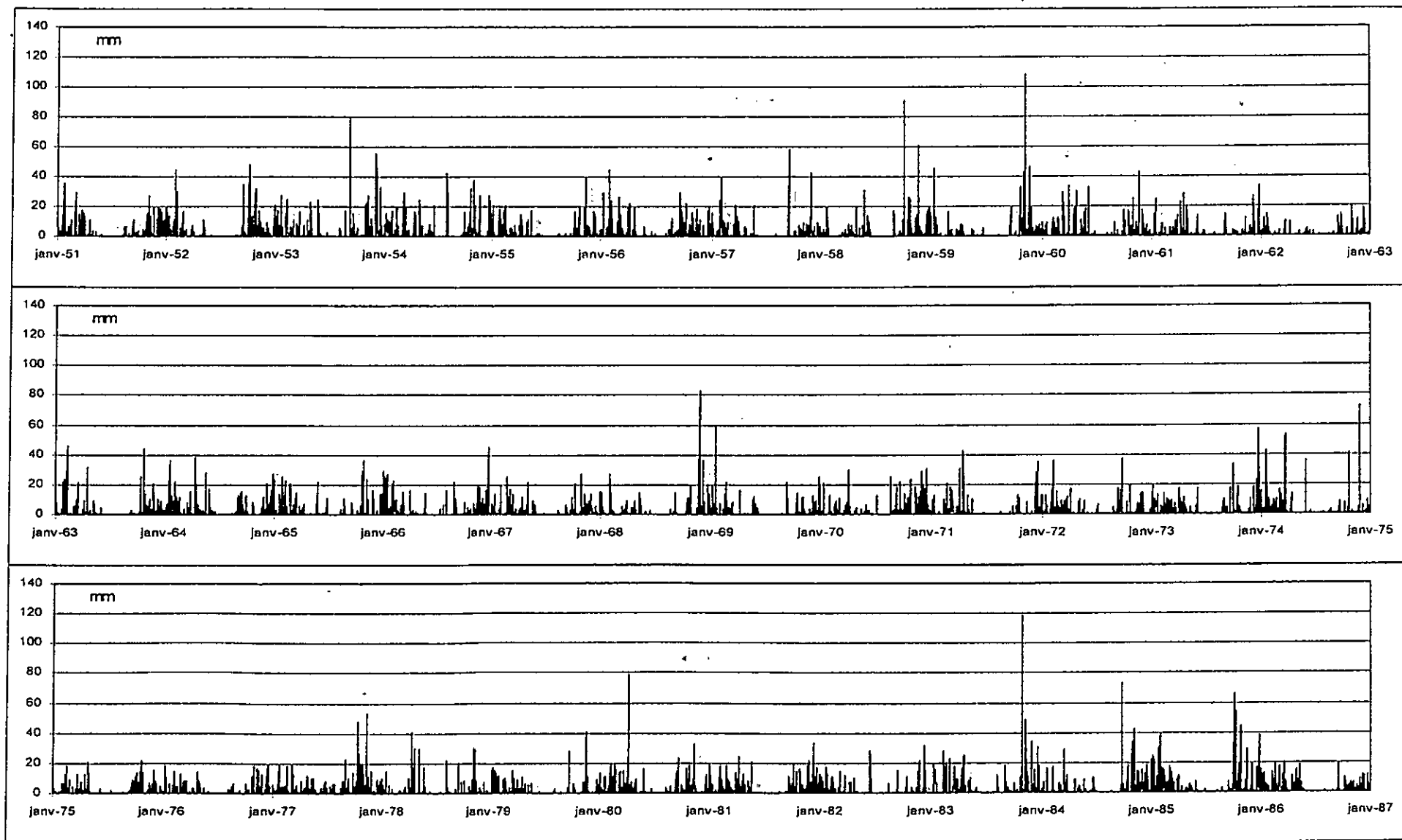




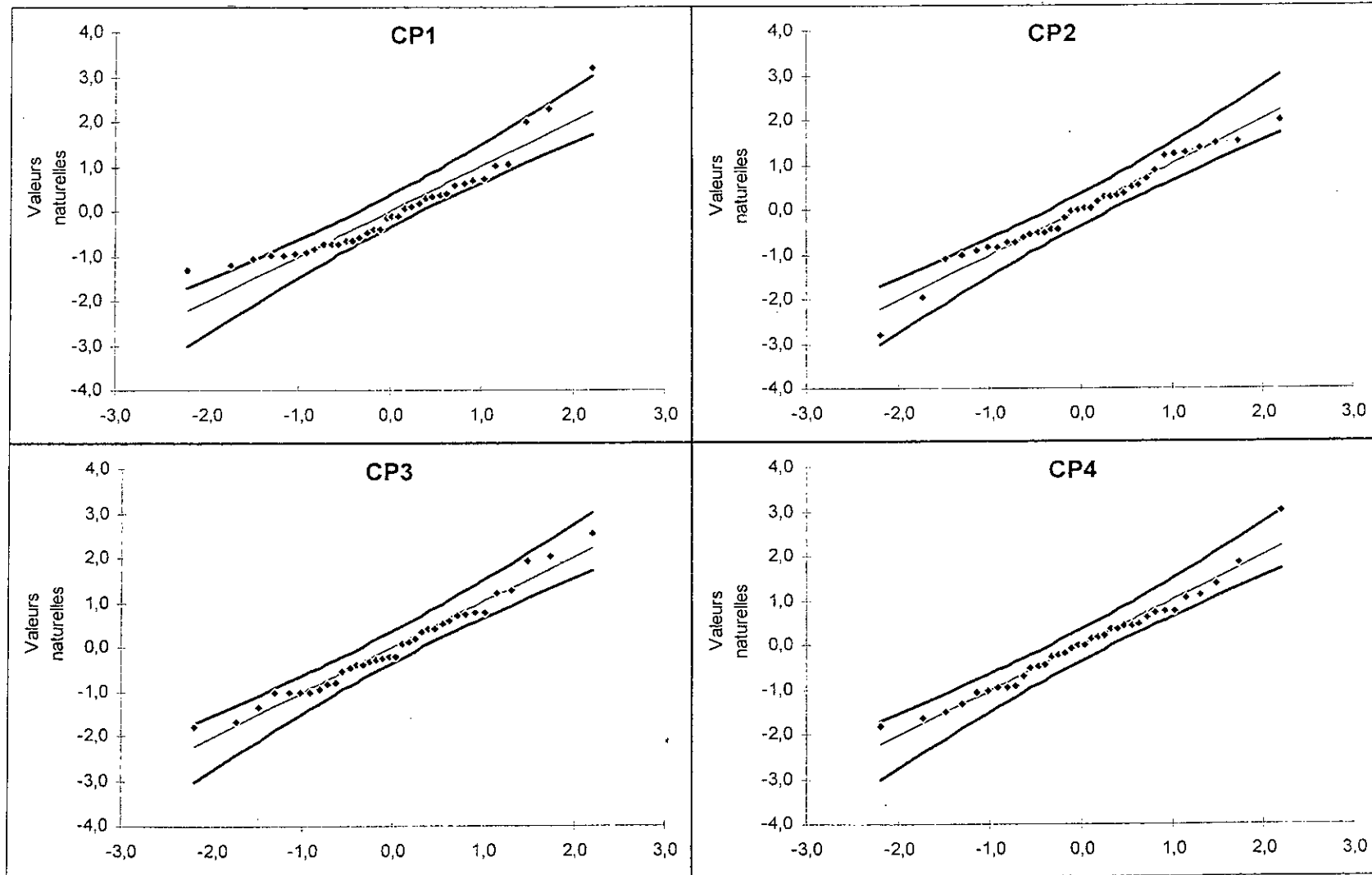


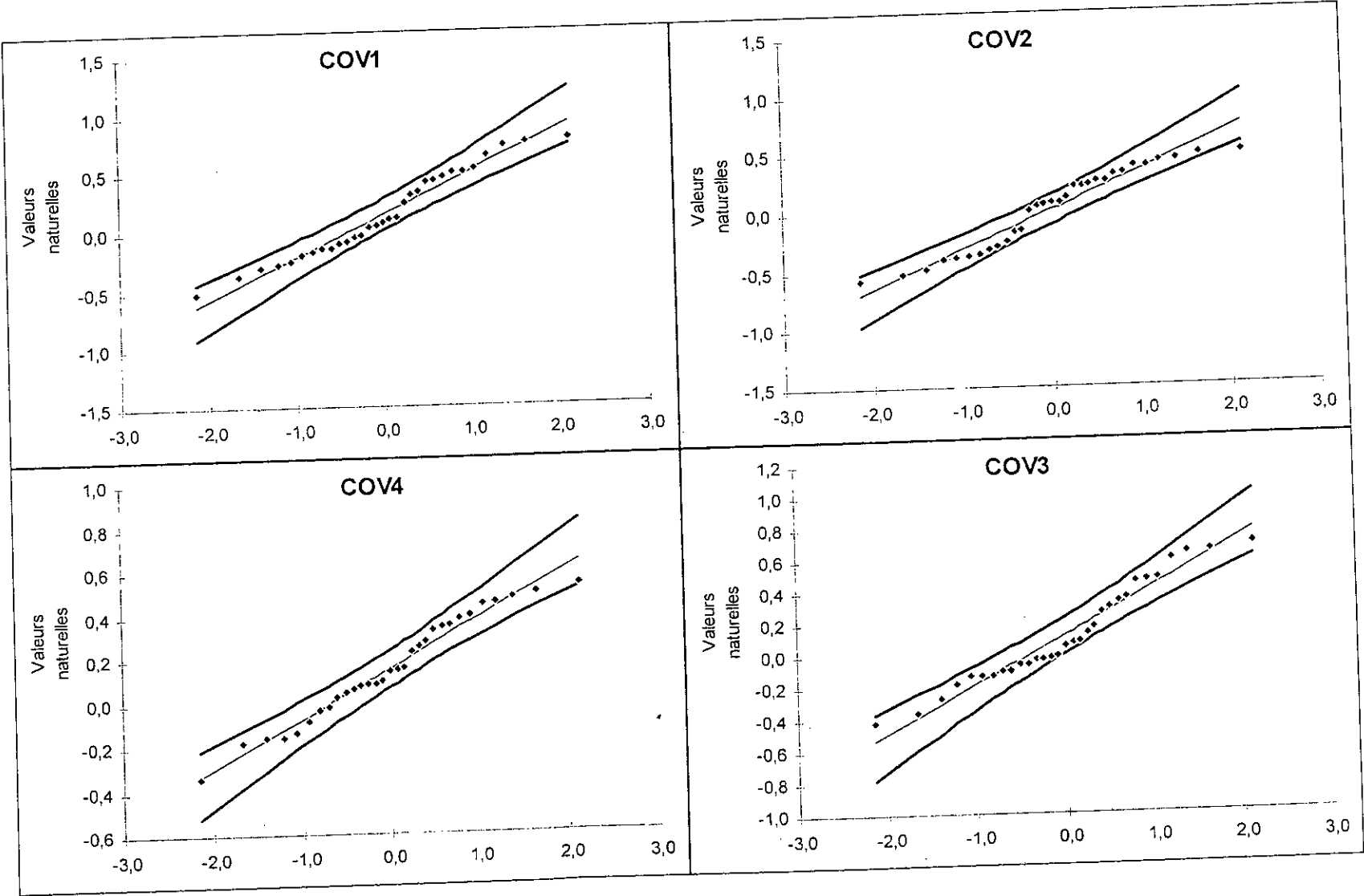


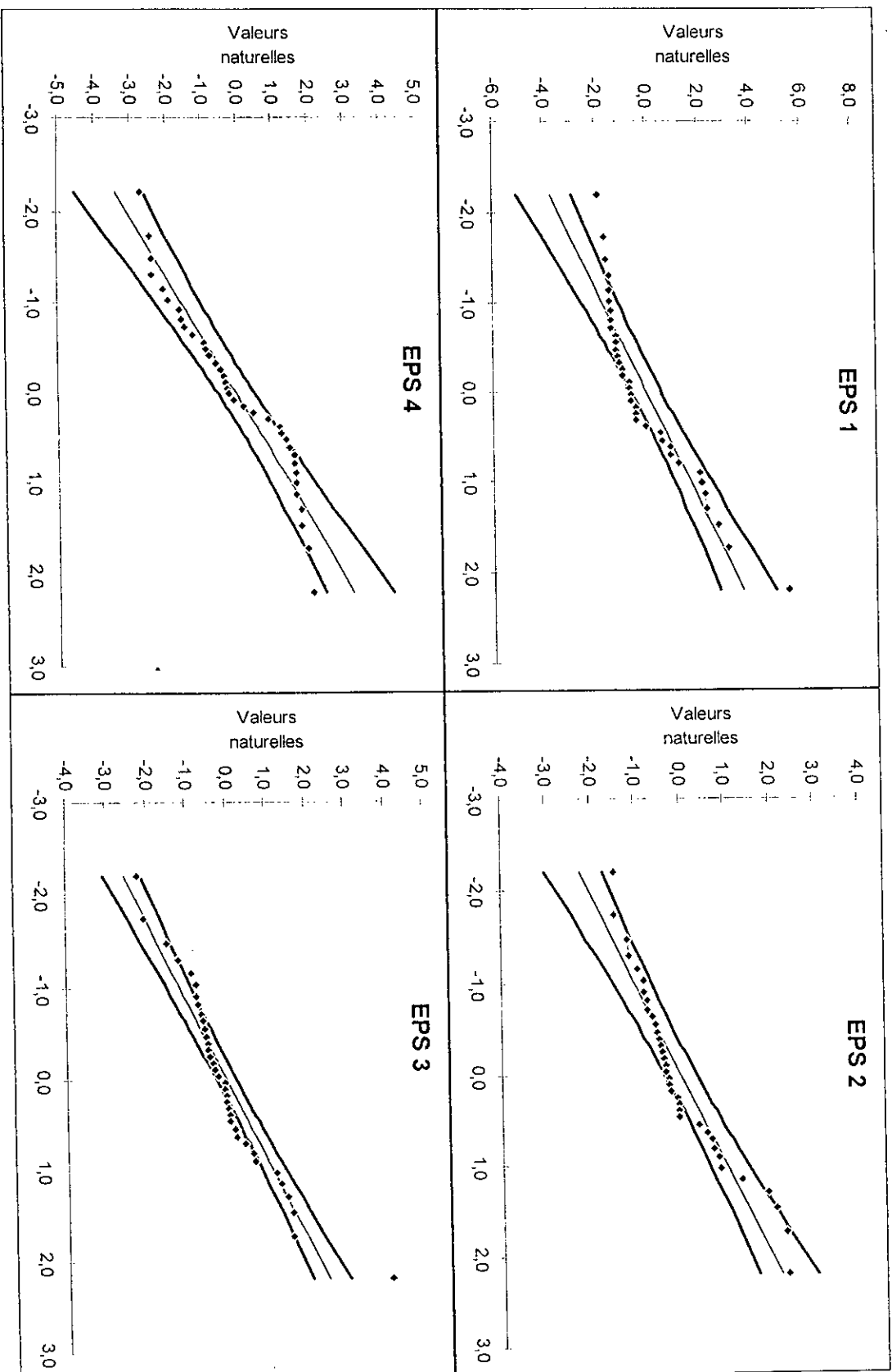
Annexe B.6: Evolution chronologique des pluies au pas de temps journalier



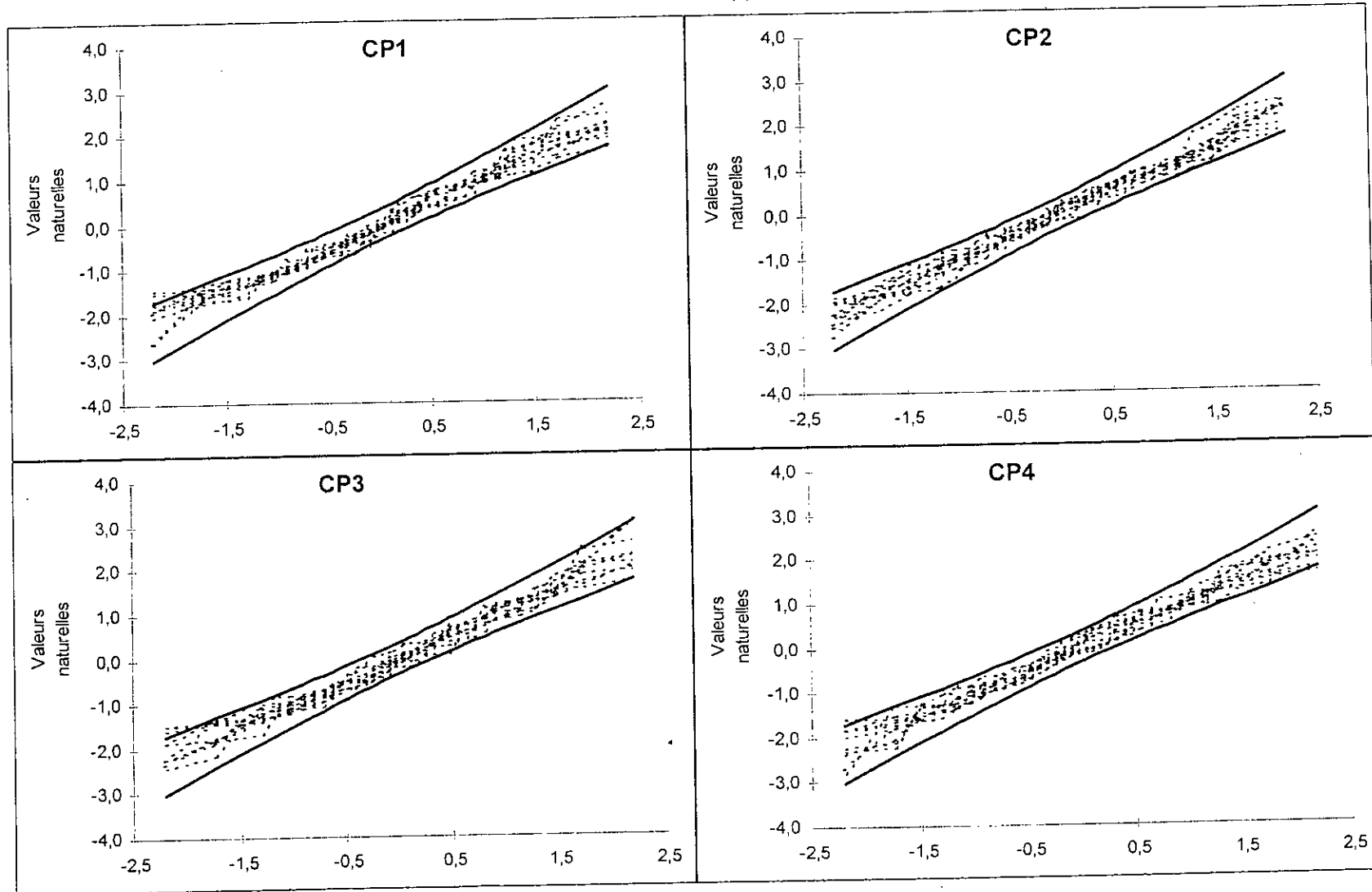
Annexe B.7 (1): Ajustement des paramètres de simulation des Pluies (Décembre)

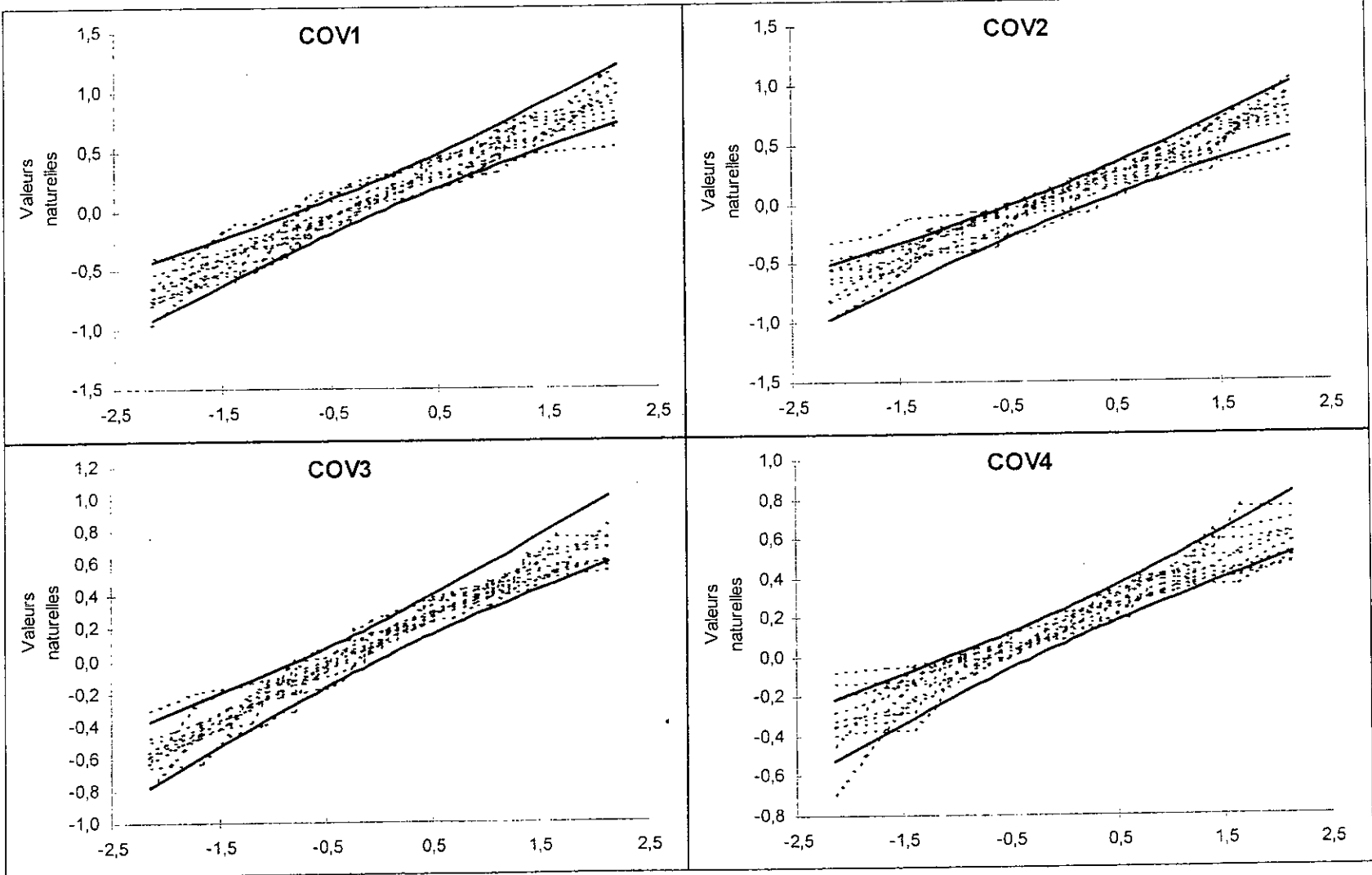


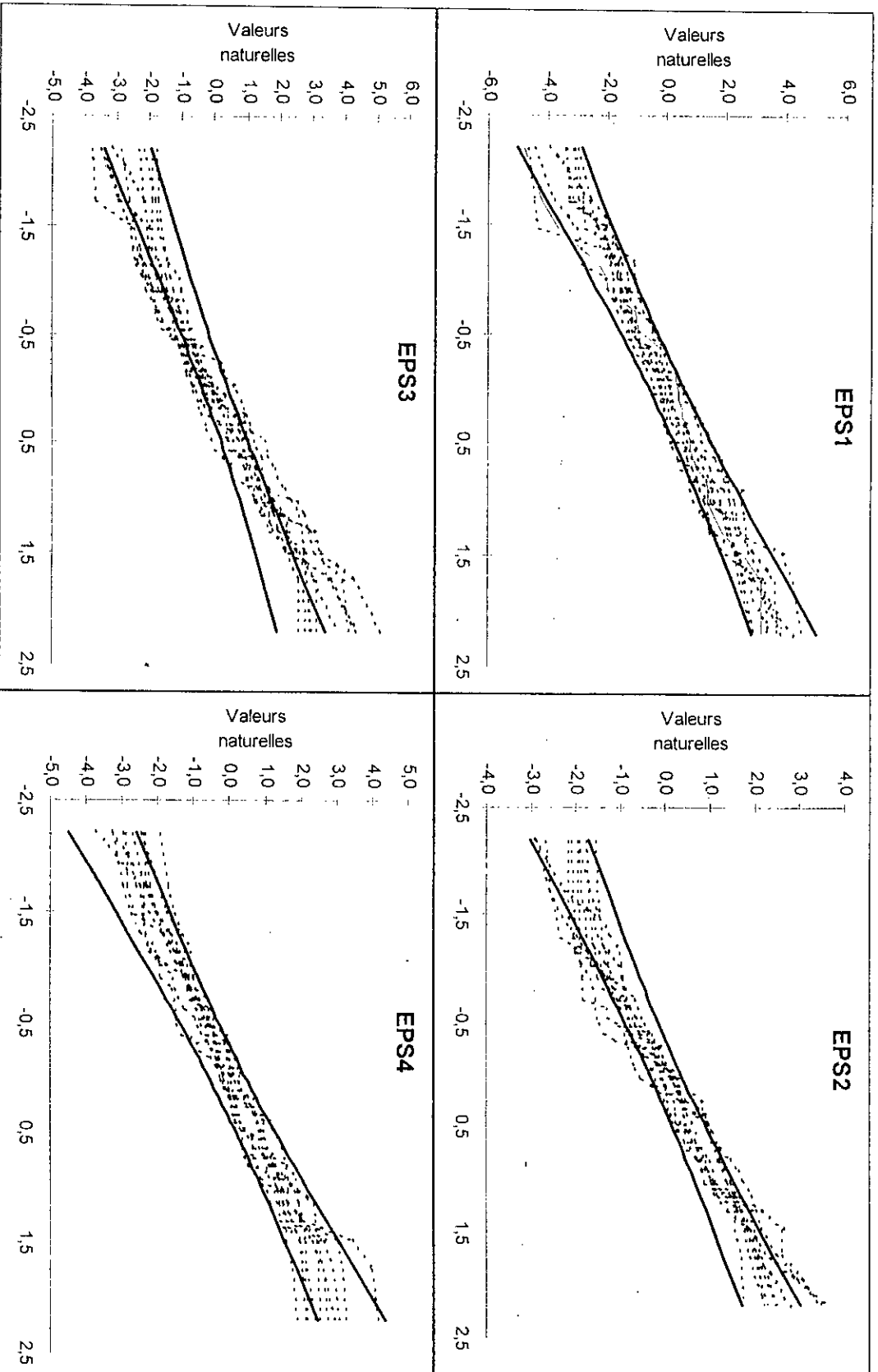




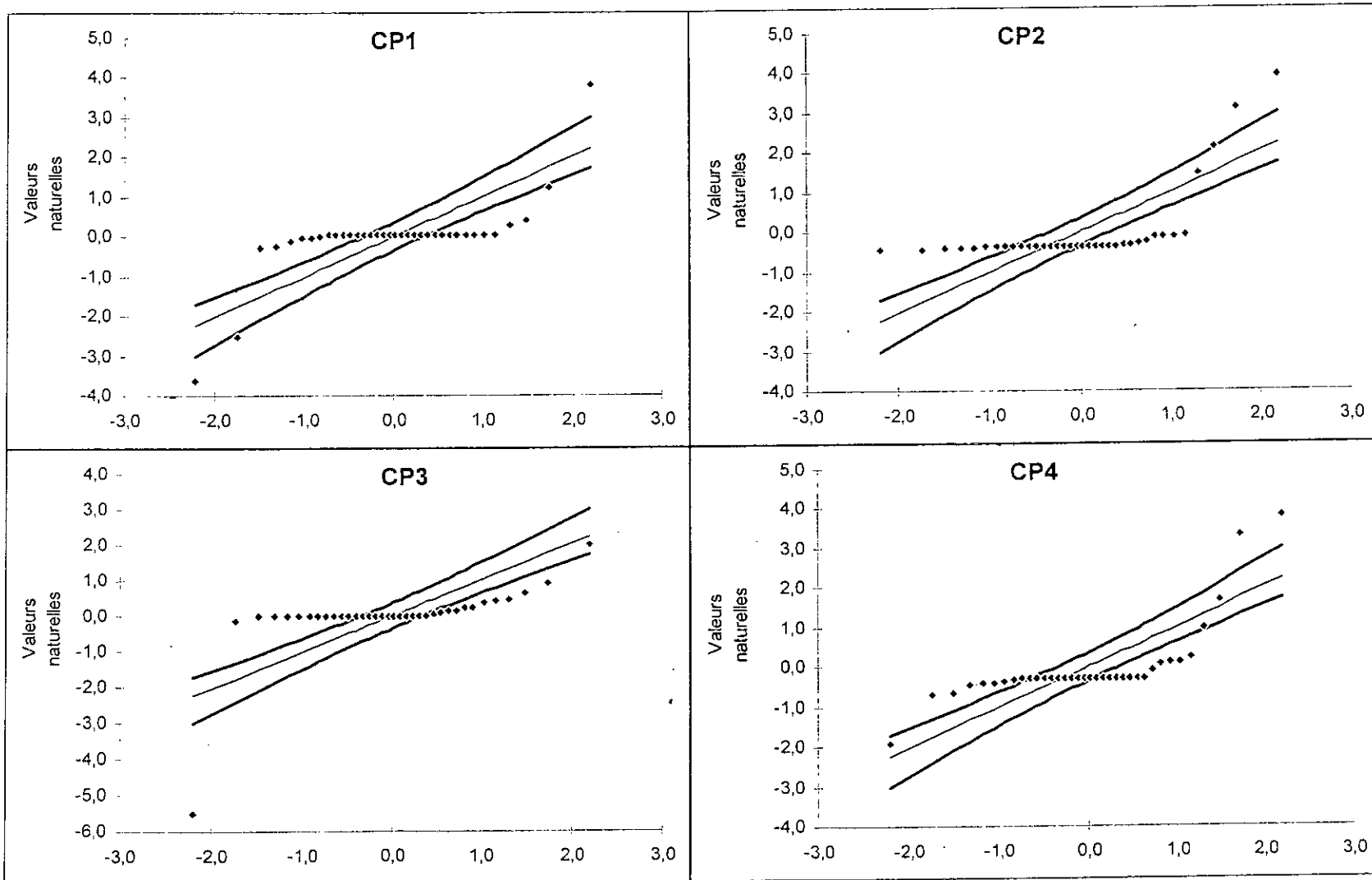
Annexe B.7 (2)



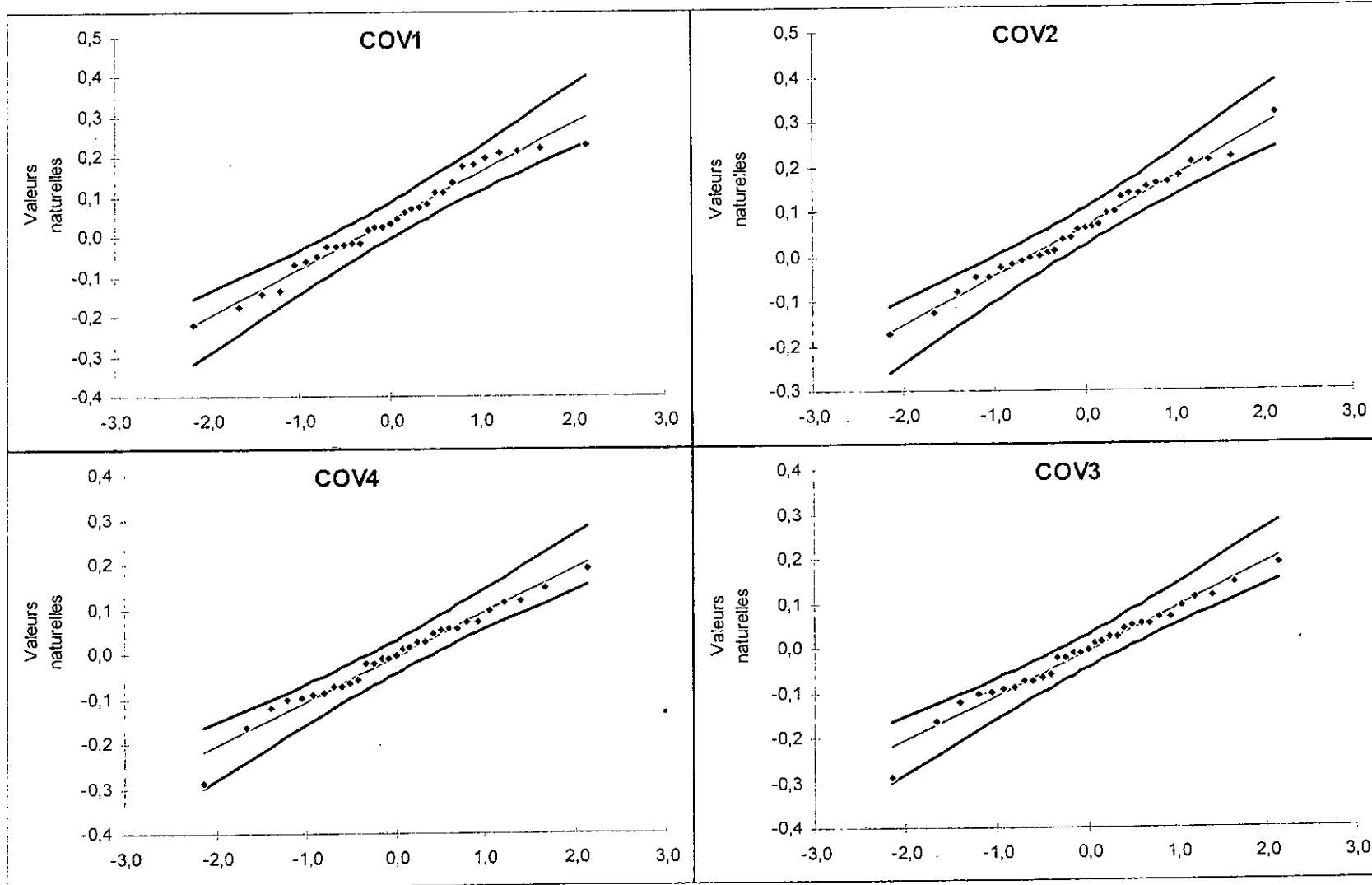




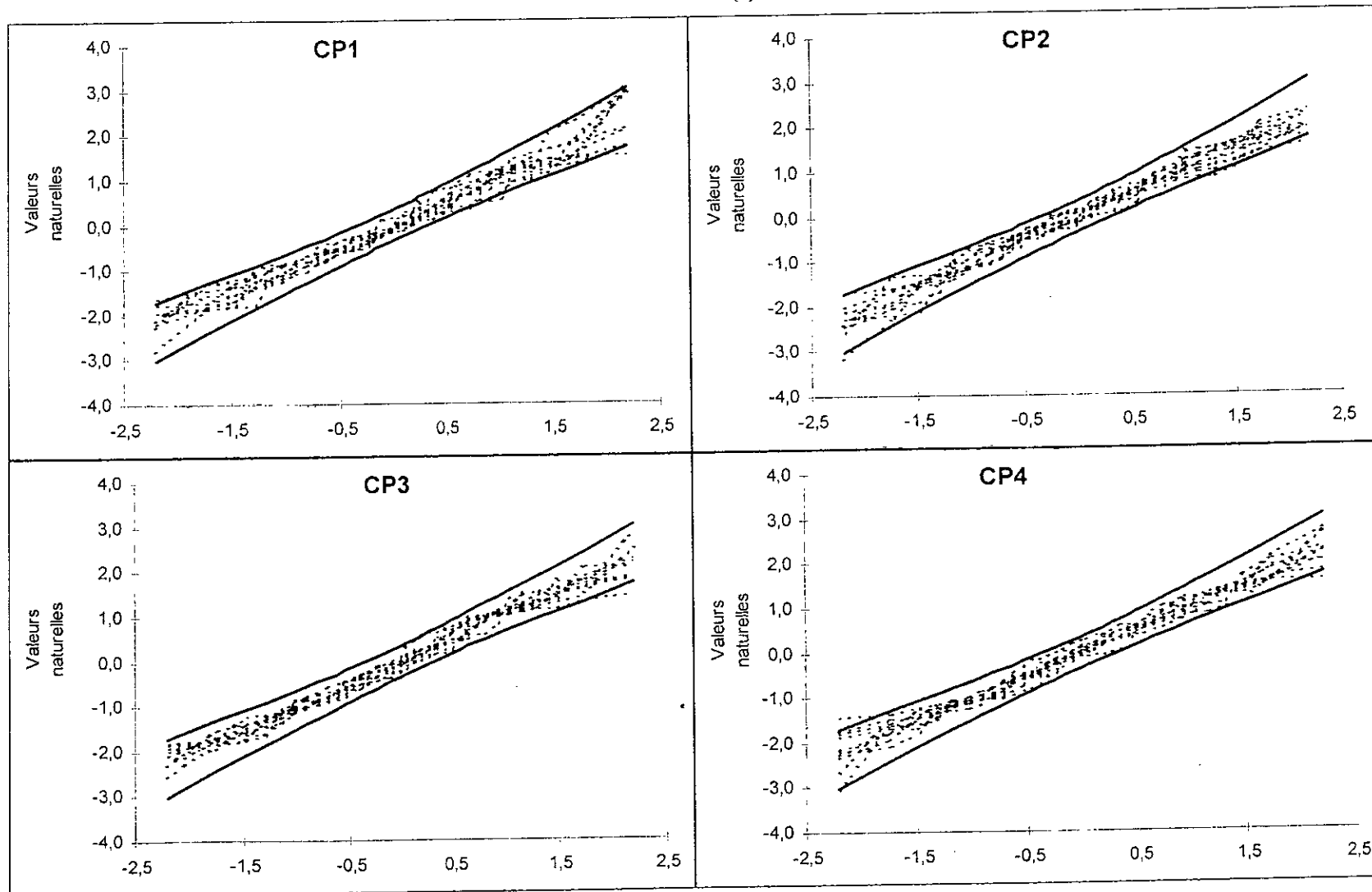
Annexe B.8 (1): Ajustement des paramètres de simulation des pluies (Juillet)

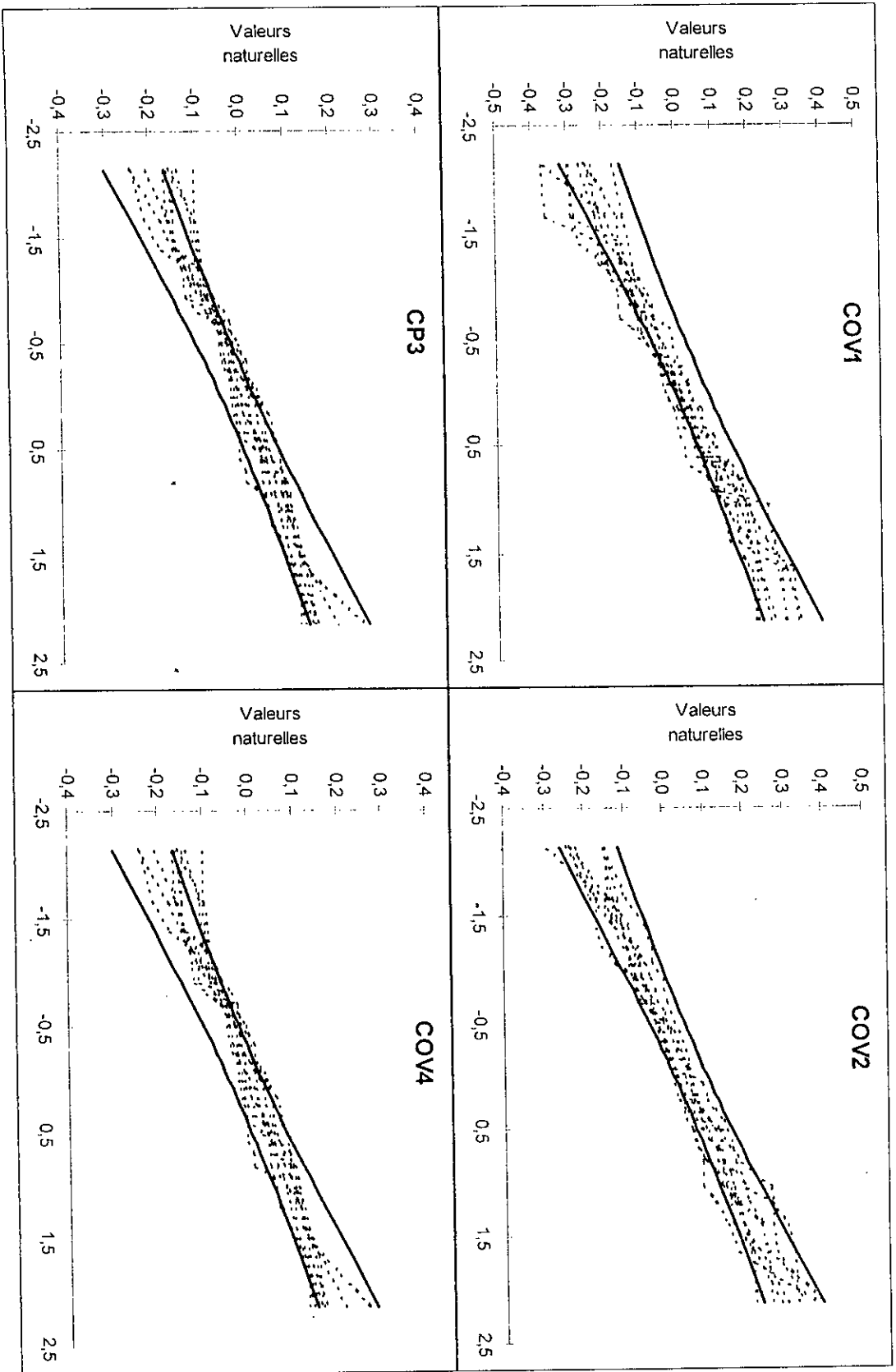


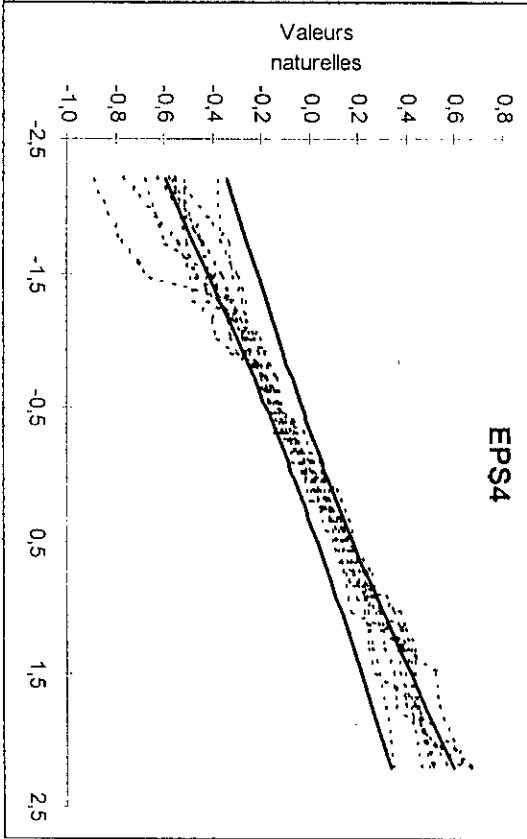
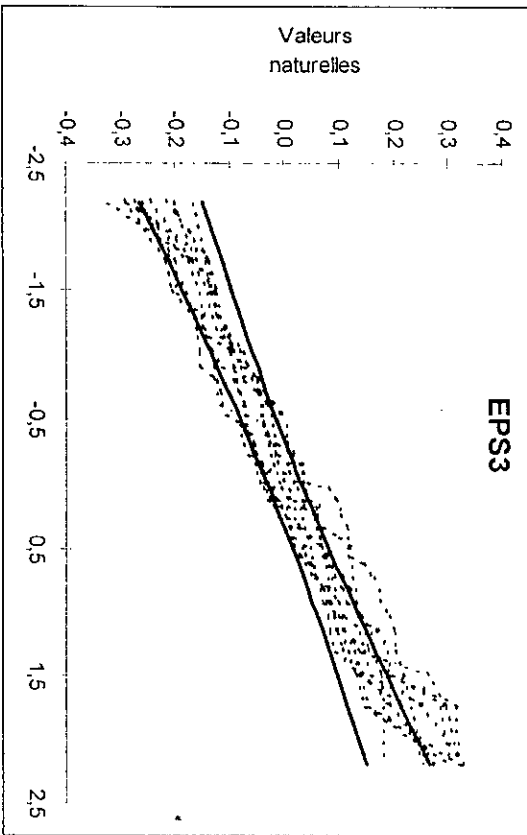
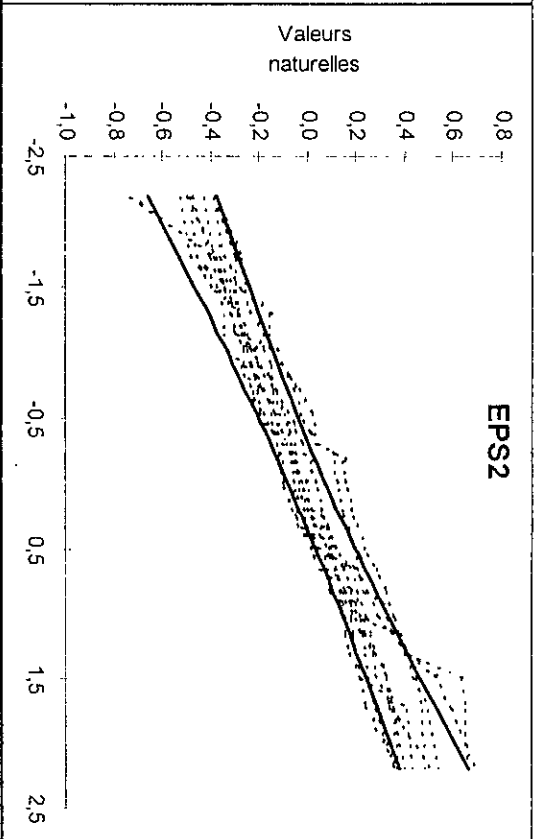
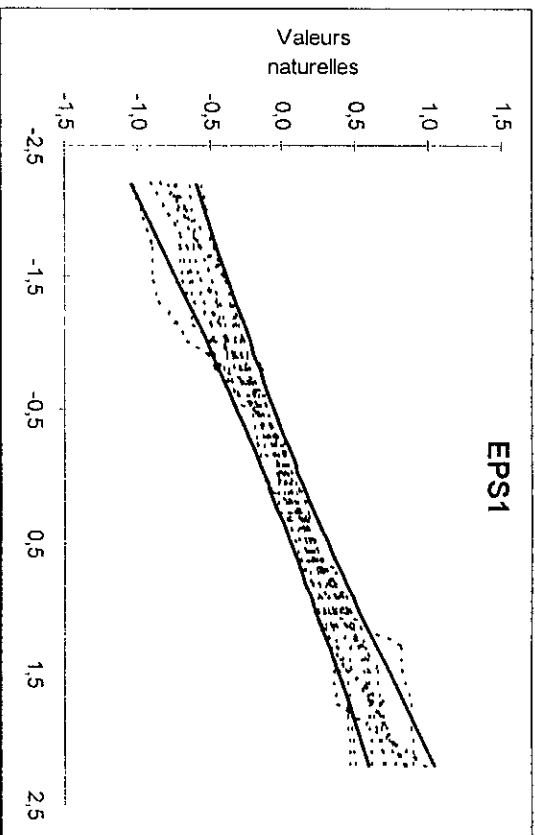


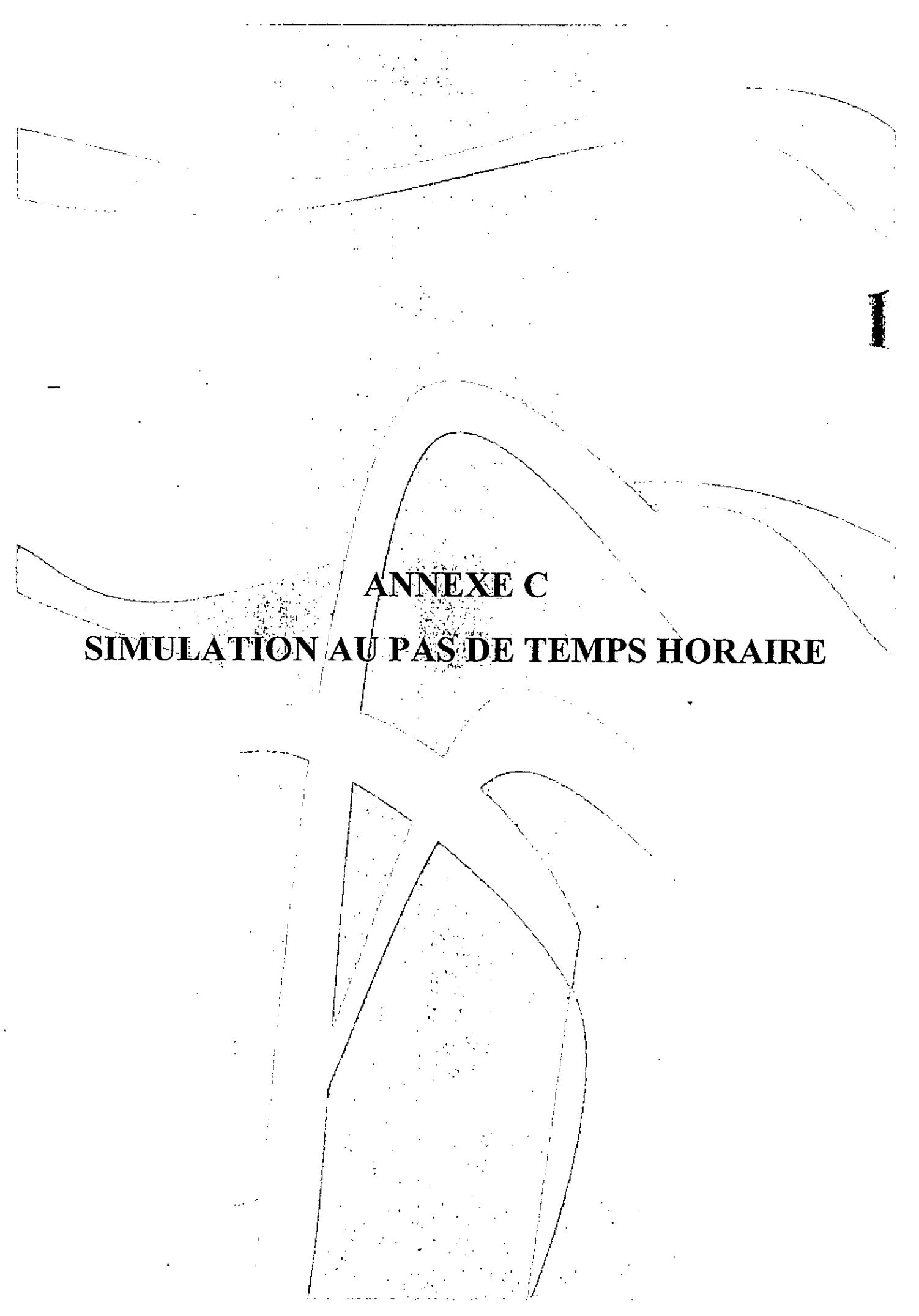


Annexe B.8 (2)



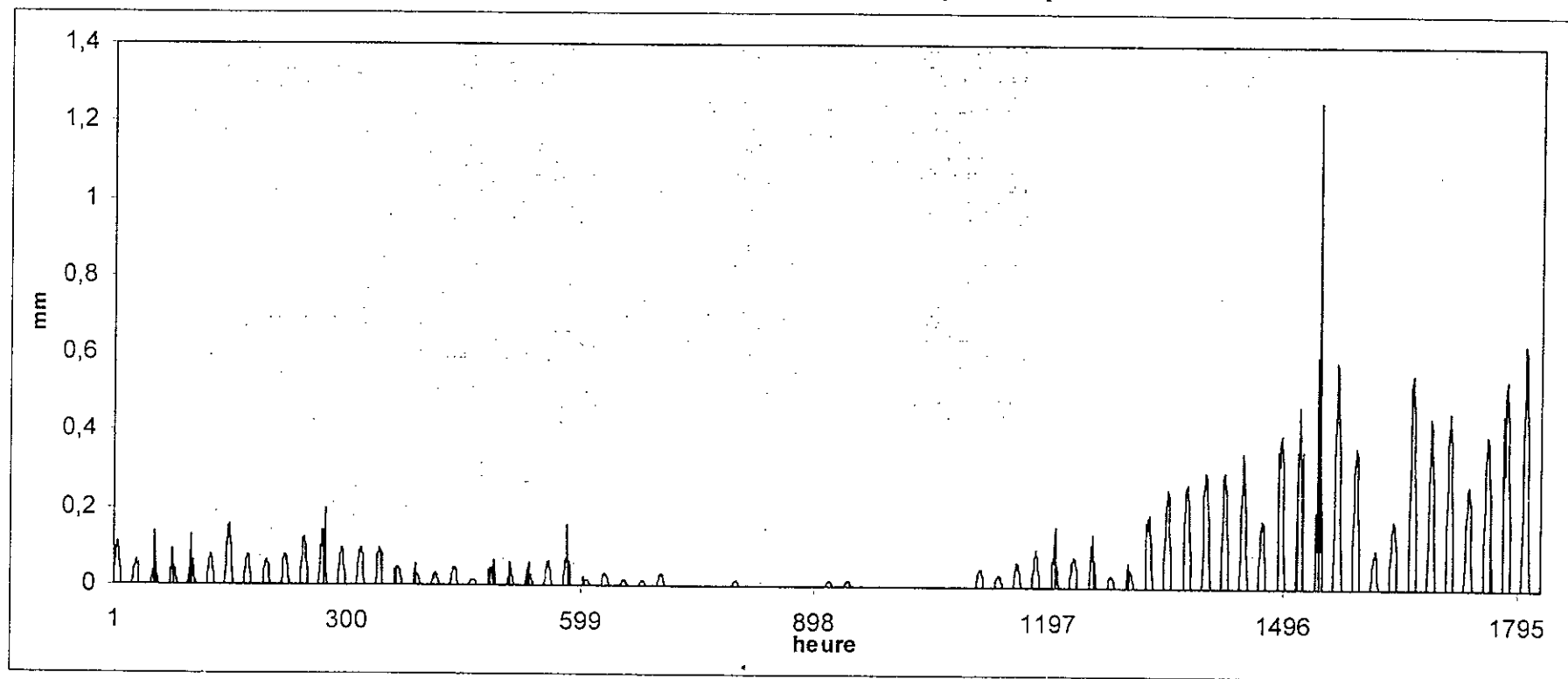




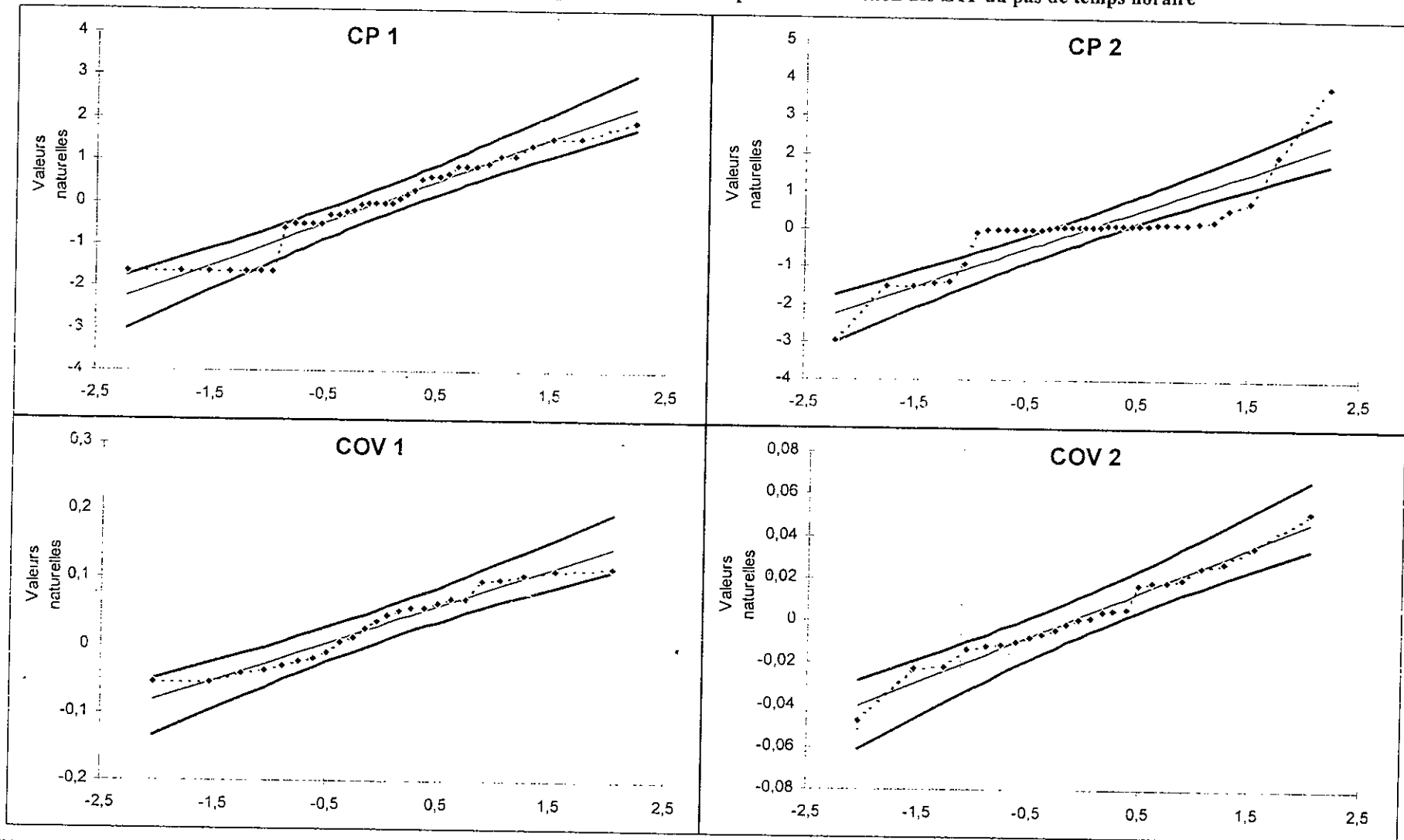


**ANNEXE C**  
**SIMULATION AU PAS DE TEMPS HORAIRE**

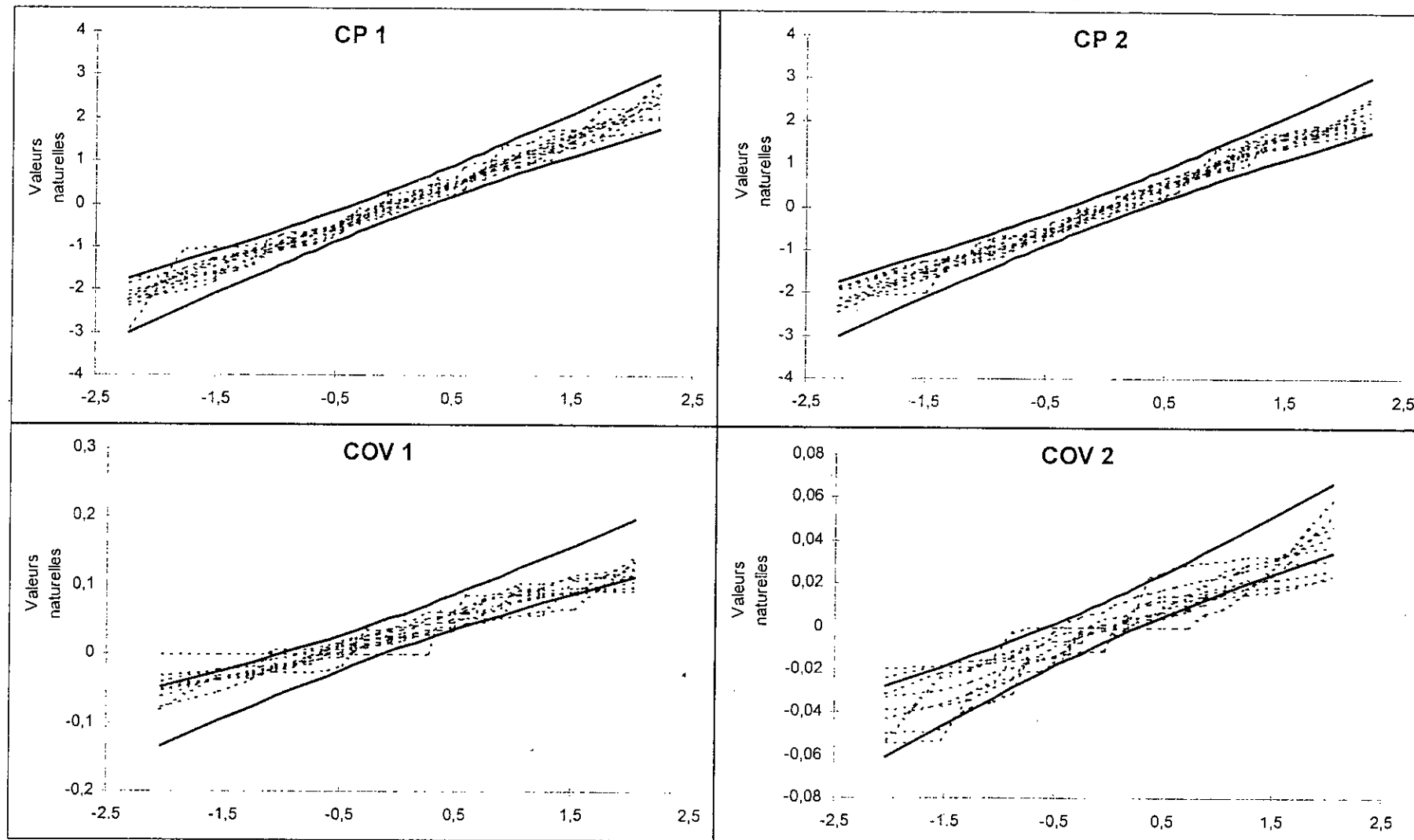
Annexe C.1: Evolution chronologique des ETP au pas de temps horaire



Annexe C.2: Ajustement des paramètres calculés pour la simulation des ETP au pas de temps horaire

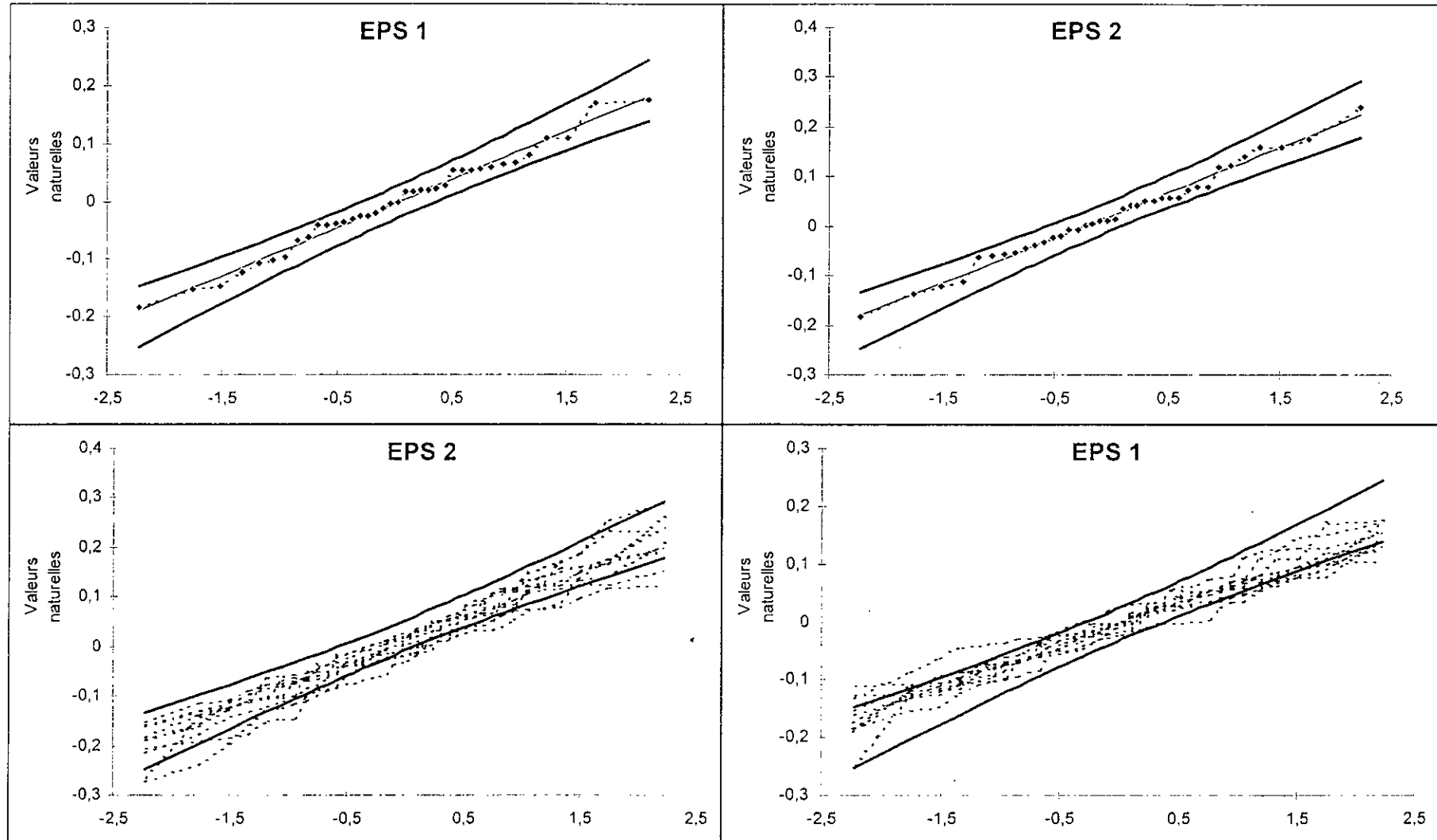


Annex C.3 Ajustement des paramètres générés pour la simulation des ETP au pas de temps horaire

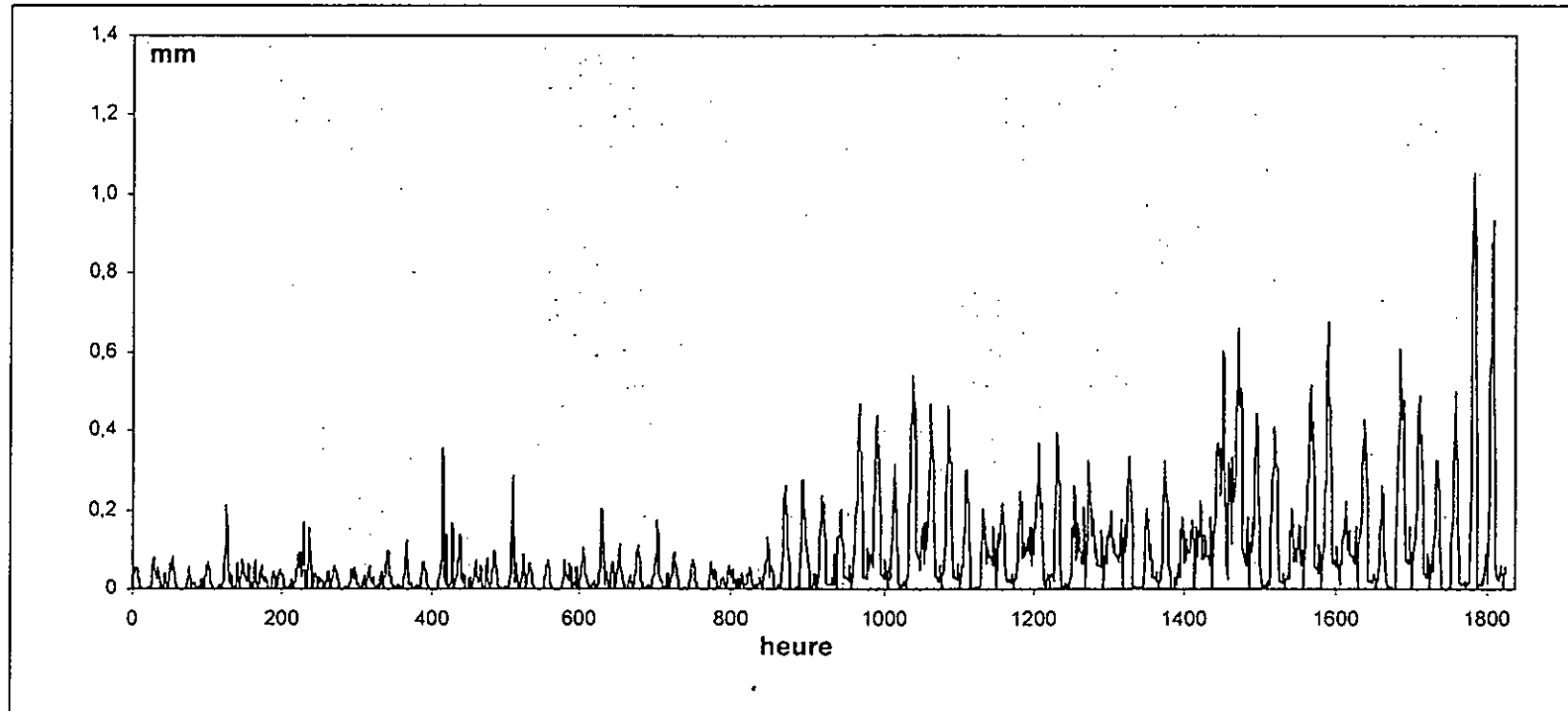




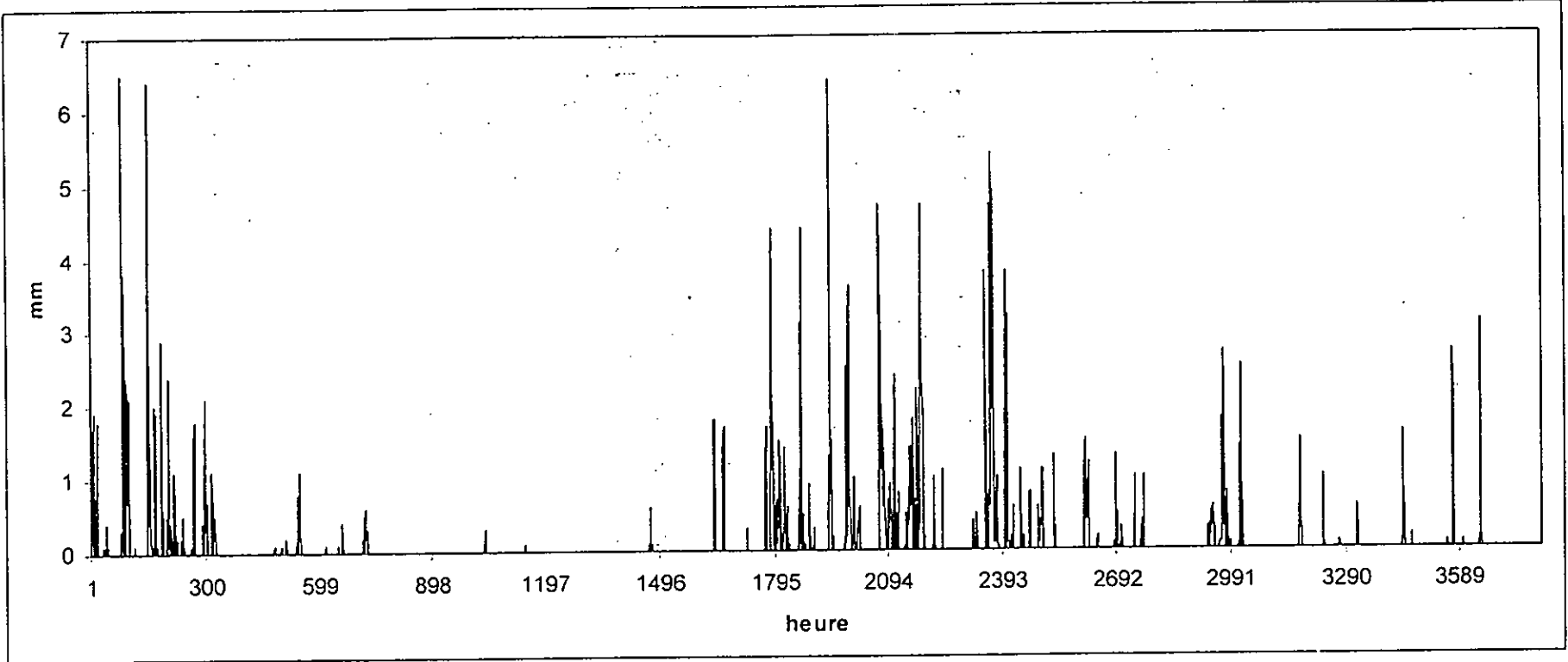
Annex C.3 (suite) Ajustement des résidus calculés et générés pour la simulation des ETP au pas de temps horaire



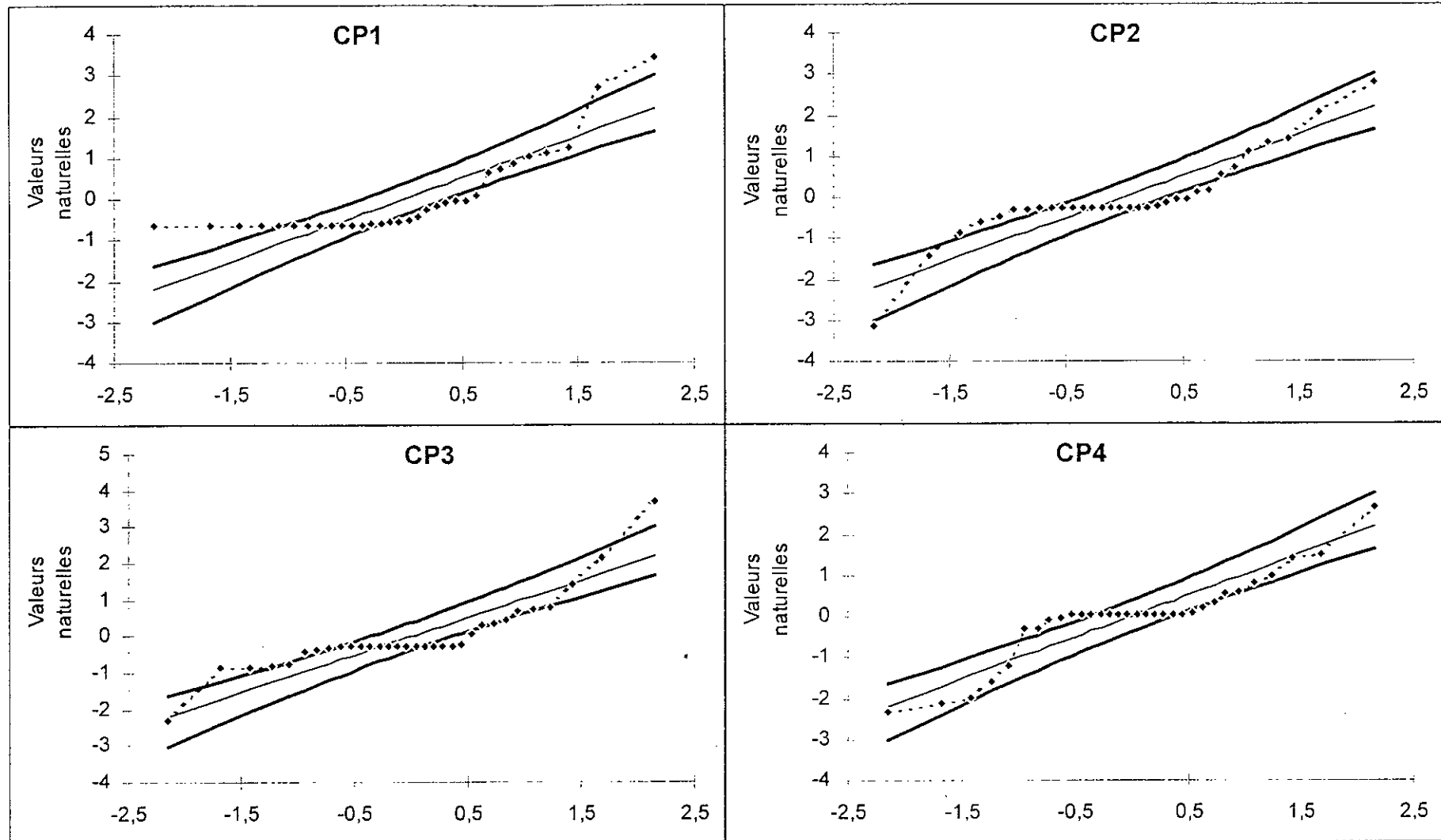
Annexe C.4: Evolution chronologique des ETP simulées au pas de temps horaire



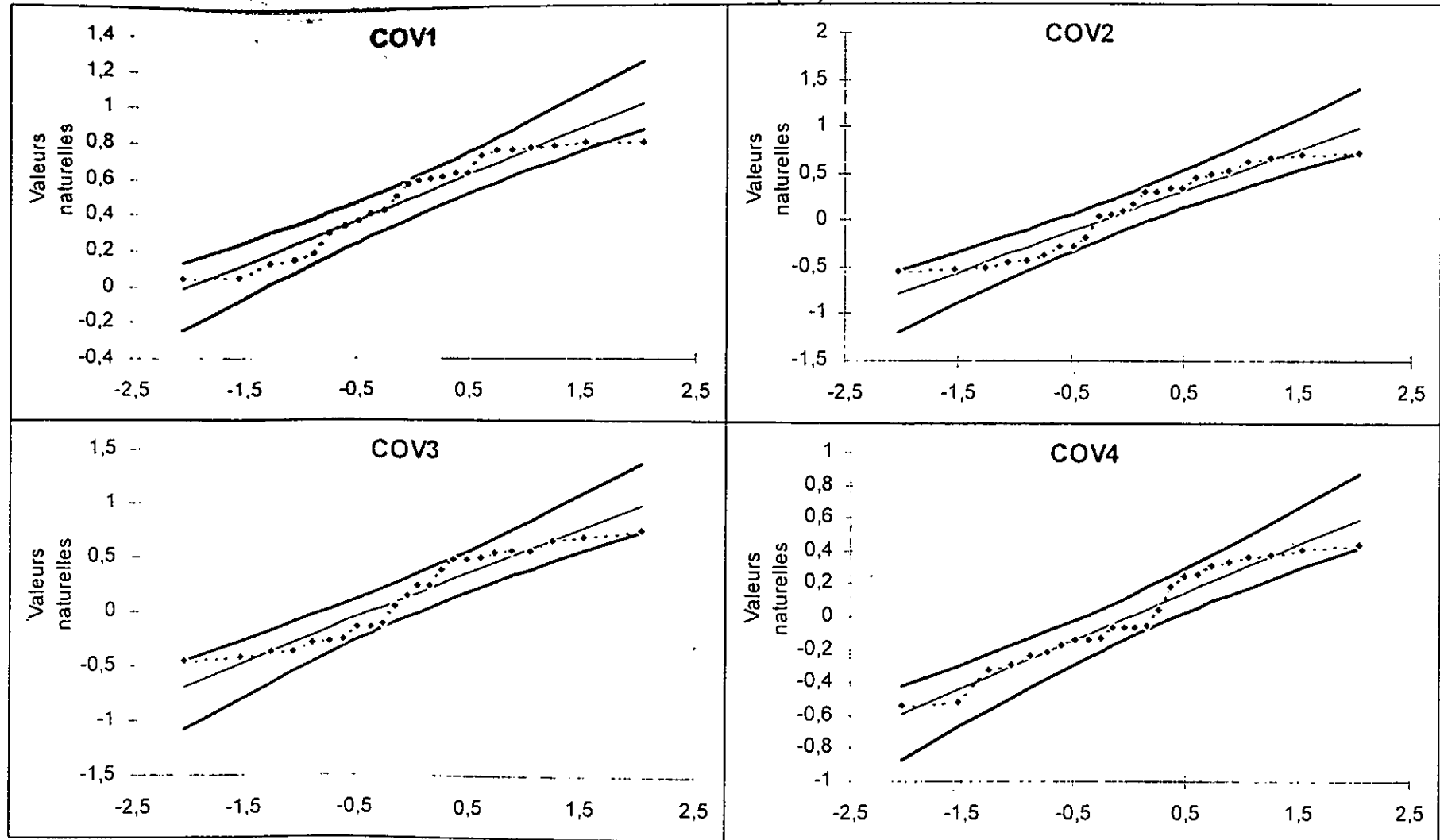
Annexe C.5: Evolution chronologique des pluies au pas de temps horaire



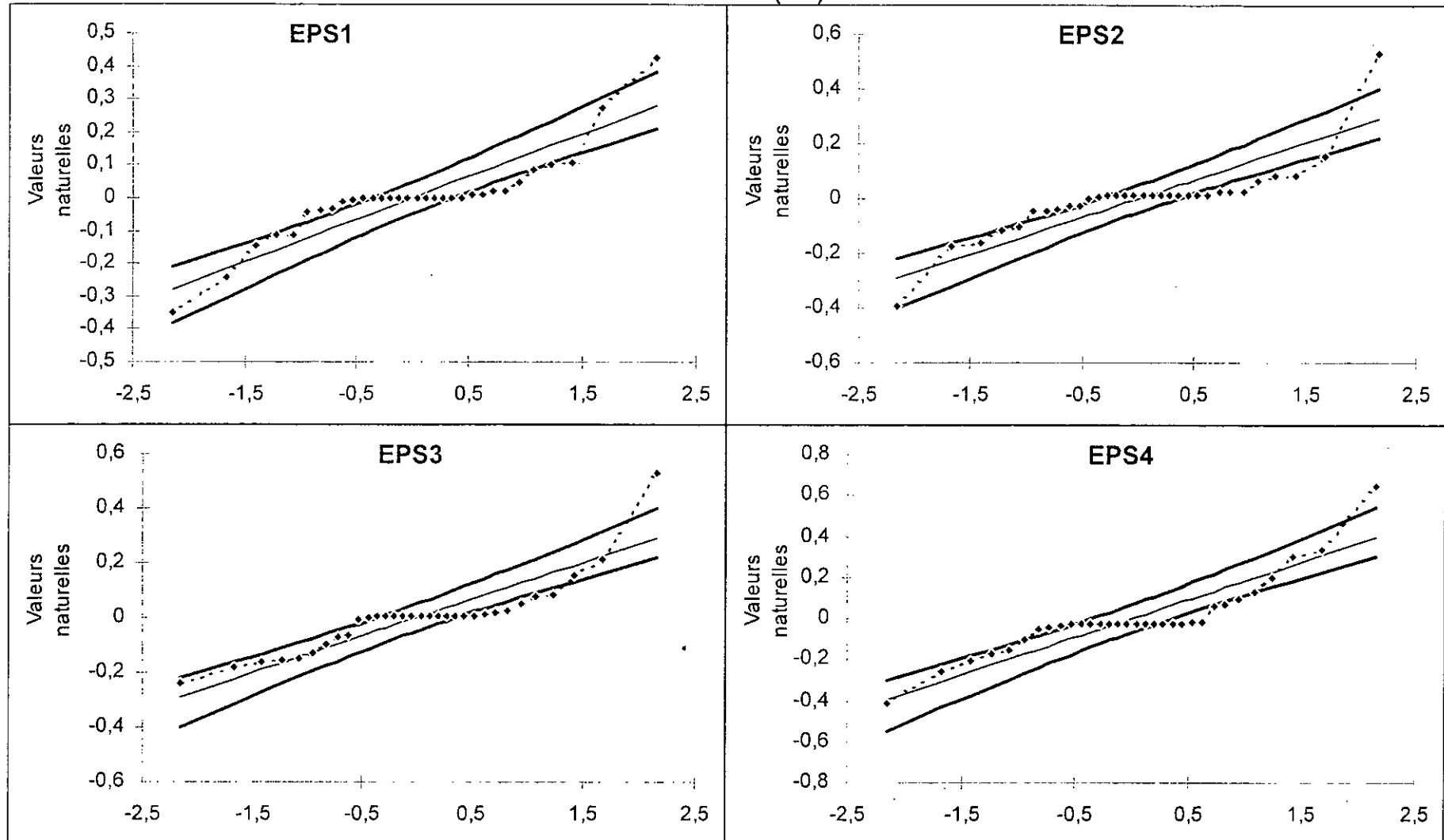
Annexe C.6 (1/3) : Ajustement des paramètres de simulation des pluies au pas de temps horaire



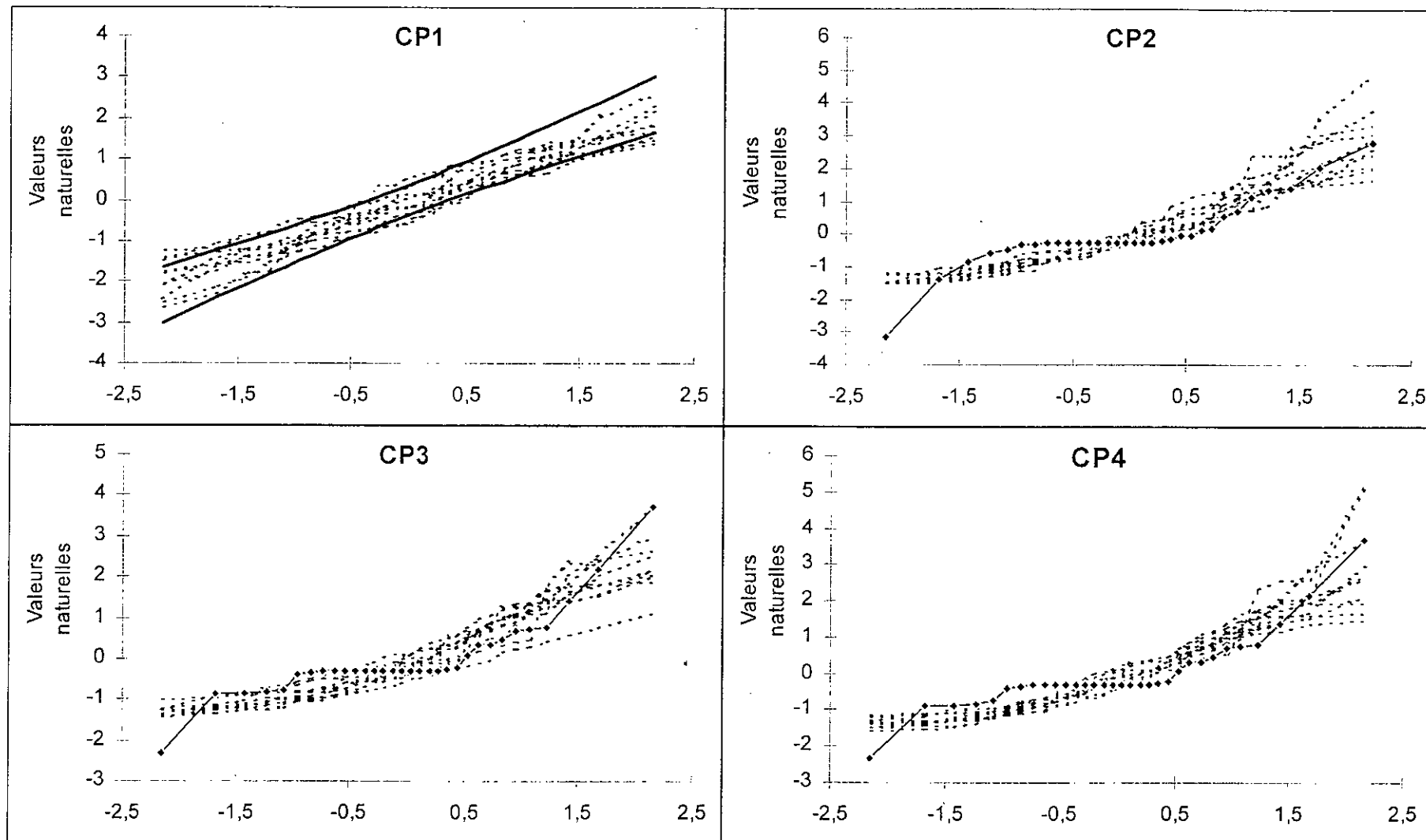
Annexe C.6 (2/3)



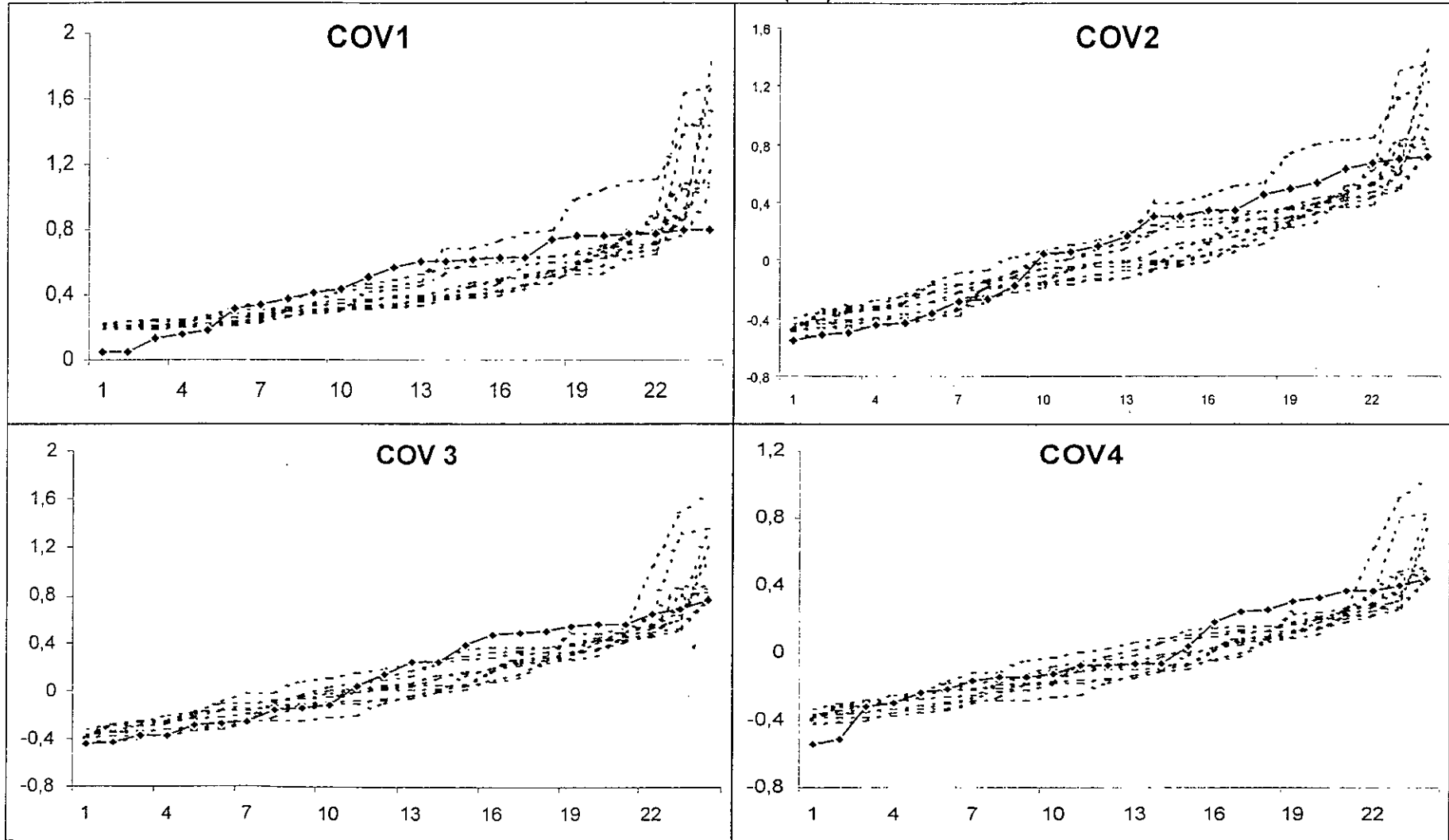
Annexe C.6 (3/3)



Annexe C.7 (1/3): Ajustement des paramètres générés pour la simulation des pluies au pas de temps horaire

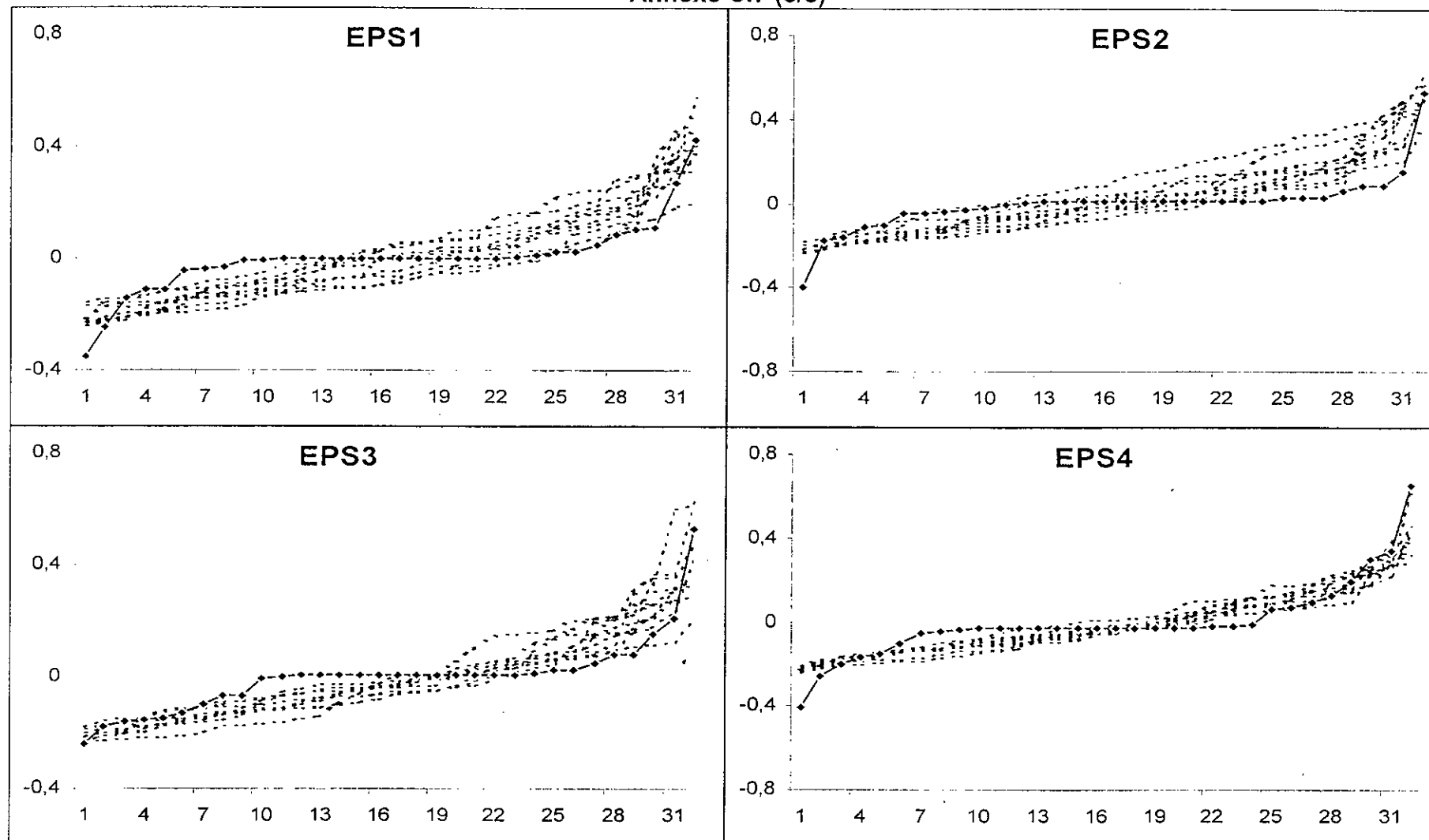


Annexe C.7 (2/3)





Annexe C.7 (3/3)



Annexe C.8: Evolution chronologique des pluies simulées au pas de temps horaire

