

وزارة الجامعات والبحث العلمي
Ministère aux Universités et de la Recherche Scientifique

ECOLE NATIONALE POLYTECHNIQUE

المدرسة الوطنية المتعددة التقنيات
المكتبة - BIBLIOTHEQUE
Ecole Nationale Polytechnique

DEPARTEMENT HYDRAULIQUE

PROJET DE FIN D'ETUDES

SUJET

OPTIMISATION
D'UN RESEAU DE MESURES
PLUVIOMETRIQUES

Proposé par :

Etudié par :

Dirigé par :

Mr. A. GUEDJAL

Melle Nadia KAHLAL

Mr. M. BERMAD

PROMOTION

1991 - 1992

المكتبة - المدرسة الوطنية
BIBLIOTHEQUE - Ecole Nationale Polytechnique

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

هُوَ الَّذِي أَنْزَلَ مِنَ السَّمَاءِ مَاءً
لَكُمْ مِمَّنَّهٗ تُرَابٌ وَمِنْهُ شَجَرٌ

صِدْقًا لِلْعَالَمِينَ

ملخص :

الغرض من هذا المشروع يكمن أساسا في دراسة الجانب المنهجي الذي يفضل استعمال التحليل ذات المكونات الأساسية في جميع المراحل :

- دراسة نقد و مقارنة الأقطار السنوية الملاحظة .
- دراسة مدى الإقليمية مساقط المتغيرات على المكونات الأساسية .

ولقد افترضنا حوض "العمومام" المجهز بسبع وتسعين (97) محطة لقياس الأقطار ، و اقتضت دراستنا على واحد وأربعين (41) منها وهذا لأختوائها على فترة الملاحظات التي تفوق العشرون سنة من أجل تحليل فعائلي لشبكة المحطات المطرية في المنطقة .

Résumé :

le bassin versant de la «Soummam» est équipé de 97 stations pluviométriques. On s'intéresse aux 41 stations, disposant d'une période d'observation de 24 ans ; en vue d'une analyse spatiale du réseau pluviométrique.

l'intérêt de ce mémoire réside essentiellement dans son aspect méthodologique qui privilégie l'utilisation de l'analyse en composantes principales (A.C.P.) aux différents stades :

- Critique et homogénéisation des précipitations annuelles observées.
- Régionalisation des projections des variables sur composantes principales.

Abstract :

the catchment of «Sommam» is equipped with 97 pluviometric station ; to spatial analysis the raingauge network ; our interest in inclined to the 41 stations which dispose of observations on a period the 24 years.

the interest of the memoir lies in methodological aspect who privileged use of main components analysis at various times in study.

- the critical and homogeneous of annuals rain data.
- regionalisation of variables projections on main component analysis.

DEDICACES

A ma mère, qui a tant peiné pour faire de moi ce que je suis.

- Mon fiancé KAMEL qui m'a énormément aidé, soutenu et encouragé à aller au bout de ma tâche.

- Mes soeurs et frères que j'adore (FAIZA, KARIMA, SOFIANE et REDOUANE).

- Mes neveux que j'aime tendrement (EDDINE, YANISSE et SAHRA)

- Toute la famille BOUHAIL.

- Toute ma promotion de l'hydraulique et particulièrement: LEILA.

je dedie ce modeste travail

K. NADIA

AVANT PROPOS

La rédaction d'un mémoire de fin d'étude représente pour ces auteurs le bilan de leur formation. L'étude d'optimisation d'un réseau de mesures pluviométriques est totalement une nouvelle expérience pour moi dans le domaine de l'hydrologie ; elle m'a aidé à approfondir mes connaissances et contribuer à un travail de recherche.

je remercie avant tout, tous les enseignants de l'école nationale polytechnique (E-N-P) qui ont contribué à notre formation pendant les trois dernières années et particulièrement, Mr DECHEM pour son aide et son écoute.

Que Mr BERMAD, mon promoteur trouve ici mes vifs remerciements pour les diverses et très précieuses orientations, en espérant qu'il soit fier de mon travail.

Je tiens à présenter également mes sincères remerciements à Mr GUEDJTAL, mon copromoteur, et Mr OULD-AMARA qui m'ont accueillis dans leur bureau, guidés et conseillés tout le long de cette étude. je leur exprime ma profonde reconnaissance pour leur expérience et leur compétence.

Je ne saurais oublier Mr MAHDI le responsable du centre de calcul de l'E.N.P. ainsi que Mr ALLILI, HAMID, et LAARBI du département d'hydrologie à l'A.N.R.H. ; je signale leur énormes services et orientations.

Que l'ensemble des membres du jury trouve à travers ces quelques lignes l'expression de ma humble reconnaissance pour avoir bien voulu examiner mon mémoire.

Que toute personne ayant contribué dans l'élaboration de cette étude, de près ou de loin, trouve ici le témoignage de mes sentiments les plus forts.

SOMMAIRE

INTRODUCTION GENERALE

CHAPITRE I : CADRE DE L'ETUDE

المدرسة الوطنية المتعددة التقنيات
المكتبة — BIBLIOTHEQUE
Ecole Nationale Polytechnique

1.1 Situation générale du bassin versant de la "Soummam"

I.1.1 Situation géographique

I.1.2 Le relief du bassin

I.1.3 La structure du bassin

I.1.4 Le climat

I.2 Présentation du réseau de mesures pluviométriques

I.2.1 Historique et développement du réseau

I.2.2 Situation actuelle du réseau

1.3 La pluviométrie dans le réseau

I.3.1 La précision des pluviomètres

I.3.2 Rôle de l'observateur

I.3.3 Densité du réseau pluviométrique

1.4 Source des données

CHAPITRE II : ANALYSE DES DONNEES

II.1 Etude de la normalité

II.1.2 Vérification de l'adéquation d'une loi supposée représentative de la loi de probabilité de la population mère

II.1.2.1 Méthode graphique : ajustement par la droite de Henri

II.1.2.2 Test du Khi-deux

II.1.2.3 Test de Kolmogorov Smirnov (K-S)

II.2 Ajustement à une autre loi théorique

II.3 Conclusion

CHAPITRE III : L'ANALYSE EN COMPOSANTES PRINCIPALES (A.C.P)

III.1 Introduction

III.2 L'objectif de l'A.C.P

III.3 Rappel théorique de la méthode

III.3.1 But et cadre de l'A.C.P

III.3.2 Formulation du problème posé

III.3.3.1 La démarche en A.C.P

**CHAPITRE IV : ETUDE DE L'HOMOGENEITE DES DONNEES
PLUVIOMETRIQUES**

IV.1 Introduction

IV.2 Le test de Kendall

IV.2.1 Principe du test

IV.2.2 Analyse des résultats du test

IV.2.3 Conclusion

IV.3 La méthode du cumul des résidus de régression

IV.3.1 Introduction

IV.3.2 Principe du cumul des résidus de régression

IV.3.3 Procédé de restitution

IV.3.4 Période d'étude

IV.3.5 Recherche de la station de référence

IV.3.6 Analyse des résidus de régression

IV.4 Conclusion

**CHAPITRE V : ANALYSE DE LA STRUCTURE SPATIALE DES
POSTES PLUVIOMETRIQUES PAR A.C.P**

V.1 Introduction

*V.2. Analyse et interprétations des configurations spatiales révélées par
les axes principaux (ou composantes principales)*

V.2.1 Le comportement spatial de la première composante

V.2.2 Le comportement spatial de la 2ème composante

V.2.3 Le comportement spatial de la troisième composante

V.3. Conclusion

CONCLUSION GENERALE

ANNEXES

BIBLIOGRAPHIE

INTRODUCTION GENERALE

C'est pourquoi nous ne négligeons que d'un aspect limité et partiel de ce problème relatif à l'hydrologie relativement aux pluviométries.

Actuellement, le réseau de l'Agence Nationale des Ressources Hydrauliques (A.N.R.H) a atteint un stade de développement important qu'il devient impératif de l'analyser dans sa totalité en vue d'une gestion rationnelle et optimum afin de fournir des données de base fiables et représentatives de la vie hydrologique d'un bassin versant pour l'élaboration des projets.

Tel qu'il est conçu, le réseau pluviométrique du bassin versant de la Soummam n'est pas suffisamment adapté aux besoins des applications opérationnelles ou de développements. La qualité même des observations n'est pas uniforme et dépend du soin et de la motivation des observateurs. Pour être exploitables, les données recueillies doivent faire l'objet d'un contrôle rigoureux et permanent dont la mise en œuvre exige des moyens coûteux en personnel et en temps de calcul.

Pour pallier le mieux possible aux défauts d'homogénéité des relevés pluviométriques, il est souhaitable que la prise de mesure soit de bonne qualité, d'où l'idée de faire ressortir les faiblesses et les qualités du réseau de façon à être en mesure de proposer des changements qui assureront une information optimale, chose qui fera l'objet de notre projet de fin d'étude.

Le but de notre travail est donc d'essayer d'"optimiser" le réseau de stations de mesures pluviométriques, supposé trop dense et redondant, dont l'entretien est onéreux, par la méthode de l'analyse en composantes principales.

Il faut se garder d'éliminer toute redondance, car elle peut être utile pour atténuer l'effet d'erreurs de mesures aléatoires; lorsqu'on exploite les résultats de mesures d'un ensemble de stations, la redondance permet un lissage des erreurs accidentelles.

Il faudrait préciser ce que l'on entend par «optimiser»: par rapport à quel type d'information, sur quel pas de temps et selon quel critère. Il est toujours difficile de préjuger des besoins à venir, et même dans une situation actuelle, il est presque impossible de s'appuyer sur des considérations exhaustives pour définir rigoureusement l'optimisation.

I.1 Situation générale du bassin versant de la "Soummam"

I.1. 1 Situation géographique

Le bassin de la Soummam, avec une superficie de 9200 km² et une population d'environ 850.000 habitants, se situe dans la partie nord de l'Algérie. Il est situé entre 3° 40' - 5°45' de longitude Greenwich et 36°-36° 45' de latitude nord (voir carte n° 1).

A une partie seulement, par le Golfe de Bejaïa, il accède à la Méditerranée. Tandis que le gros de son territoire se trouve retiré dans le continent.

Il se compose de deux régions de plateaux (Bouira à l'ouest, Sétif au sud-est) et d'une vallée longue de 150 km : la vallée du Sahel-Soummam.

Au nord et à l'ouest le bassin de la Soummam est limité par les montagnes (massifs de Djurdjura), au nord-est par la Méditerranée et les chaînes côtières de la petite Kabylie (Djebel Babor).

Au sud et au sud-est, le bassin est limité par les contre-forts des Monts du Hodna. (voir carte n° 2).

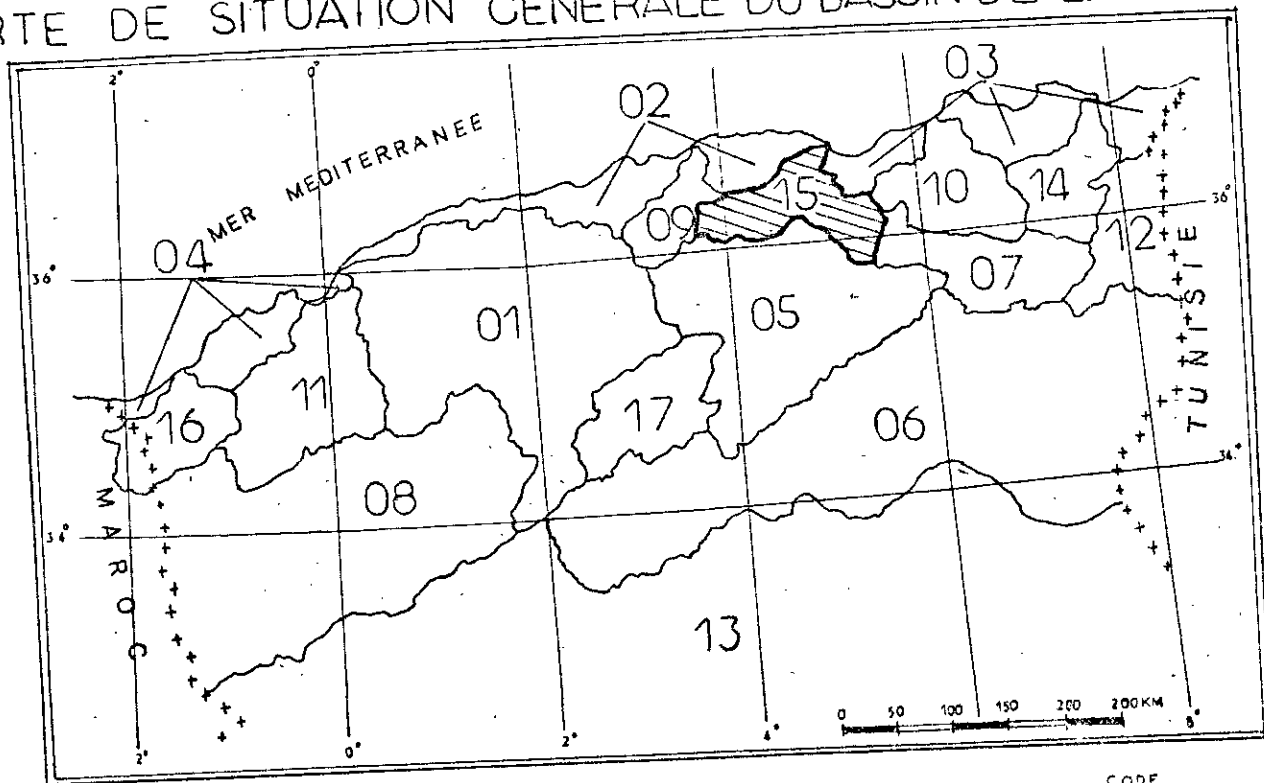
I.1.2. Le relief du bassin

Le relief de la Soummam est caractérisé essentiellement par sa plasticité accentuée. Ruisseaux, ravins et cours d'eau ont contribué à sa dislocation d'où se reflète un fort morcellement orographique.

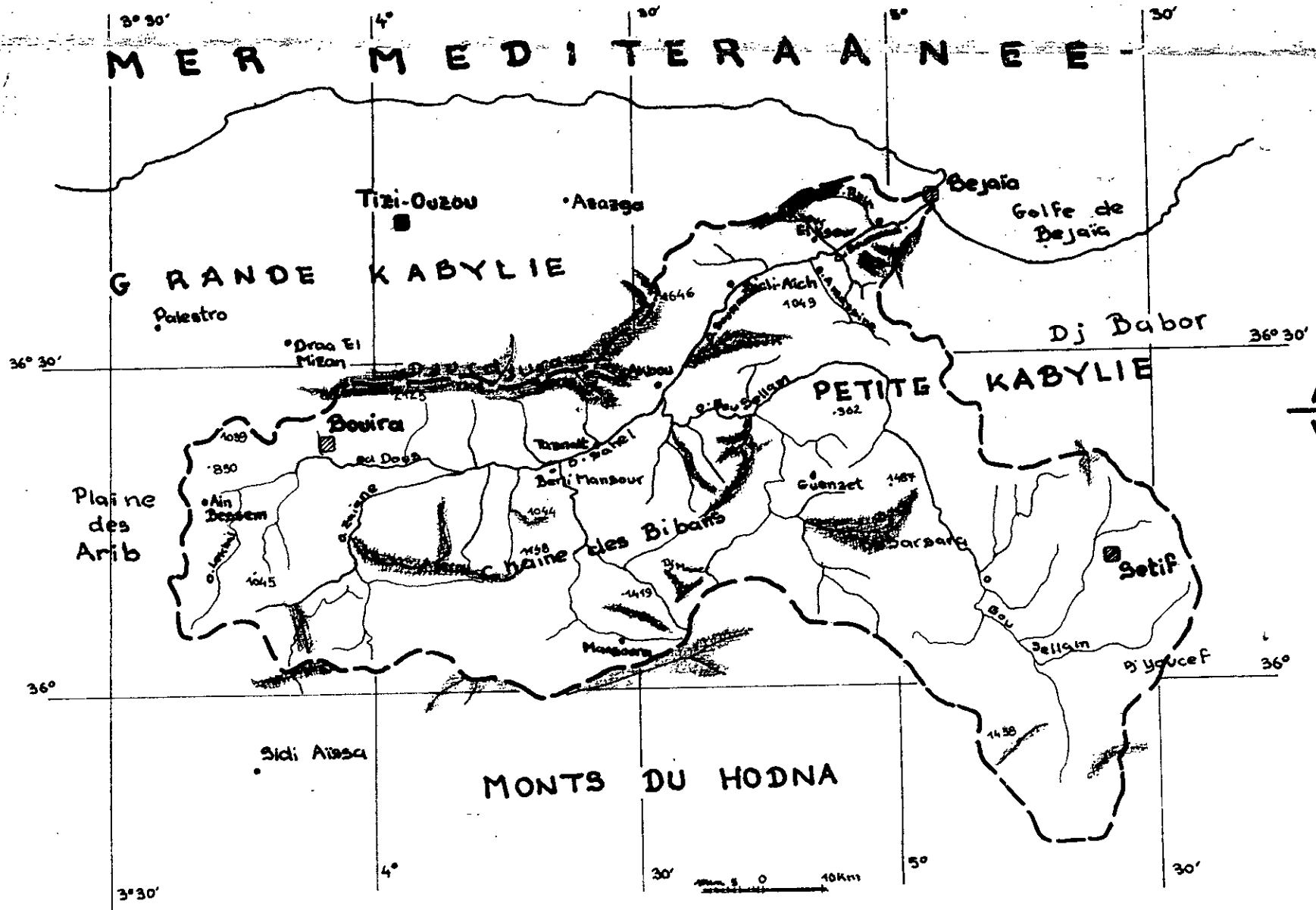
Malgré son intégrité, le relief peut être divisé en trois grandes zones distinctes :

- les versants des monts de Djurdjura (limite supérieur du bassin)
 - le plateau de Bouira (limite Ouest du bassin), faisant suite à la plaine des Arib.
 - le plateau de Setif (limite Est du bassin), cette zone constitue la partie la plus étendue du Bassin, fait partie des hauts plateaux avec une altitude moyenne de 1000m. Il est limité au Nord par les montagnes de la petite Kabylie.
-

(N°1)
 CARTE DE SITUATION GENERALE DU BASSIN DE LA SOUMMAM



CODE	BASSIN	CODE	BASSIN	CODE	BASSIN
01	CHELIFF	07	H PLATEAUX CONSTANTINOIS	13	SAHARA
02	COTIERS ALGEROIS	08	HAUTS PLATEAUX ORANAIS	14	SEYBOUSE
03	COTIERS CONSTANTINOIS	09	ISSER	15	SOUMMAM
04	COTIERS ORANAIS	10	KEBIR RHUMEL	16	TAFNA
05	CHOTT HODNA	11	MACTA	17	ZAHREZ
06	CHOTT MELRHIR	12	MEDJERDAH		



SITUATION GEOGRAPHIQUE DE LA SOUMMAM
Carte N°2

Sur le plan orographique, le bassin ne fait pas un ensemble. Il est visiblement partagé en deux régions différentes à tous les points de vue.

1/ La région du Sahel

Elle est située entre deux chaînes de montagnes : Djurdjura au Nord et les Bibans au Sud.

2/ La région de Bou-Sellam

La plus petite, est partie constituante du plateau Sétifien.

I.1.3 La structure du bassin

Du point de vue structure on distingue :

a. La chaîne du Djurdjura

Elle a une extension Est-Ouest, de Bouira jusqu'à la chaîne de Gouraya (aux environs de Bejaïa). Les hauteurs du Djurdjura ne sont en aucun endroit inférieure à 1500 m d'altitude, à partir de Akbou, elles diminuent graduellement.

b. La chaîne des Bibans

C'est une chaîne de montagnes de formation récente. Elle est étendue sur le territoire du bassin dans la direction Est-Ouest en passant par Azerou N'Saïd (1283m) et Djebel Anechar (1415m). Une ramification importante de cette chaîne s'en sépare près de Guenzet pour se diriger vers Djebel Es-Sarsara (1487m).

c. Les montagnés de la petite Kabylie

Situées du côté Est du bassin, elle sont constituées par une série de crêtes courtes différemment orientée. La hauteur moyenne des pics varie entre 1200 et 2000m.

d. Les Monts du Hodna

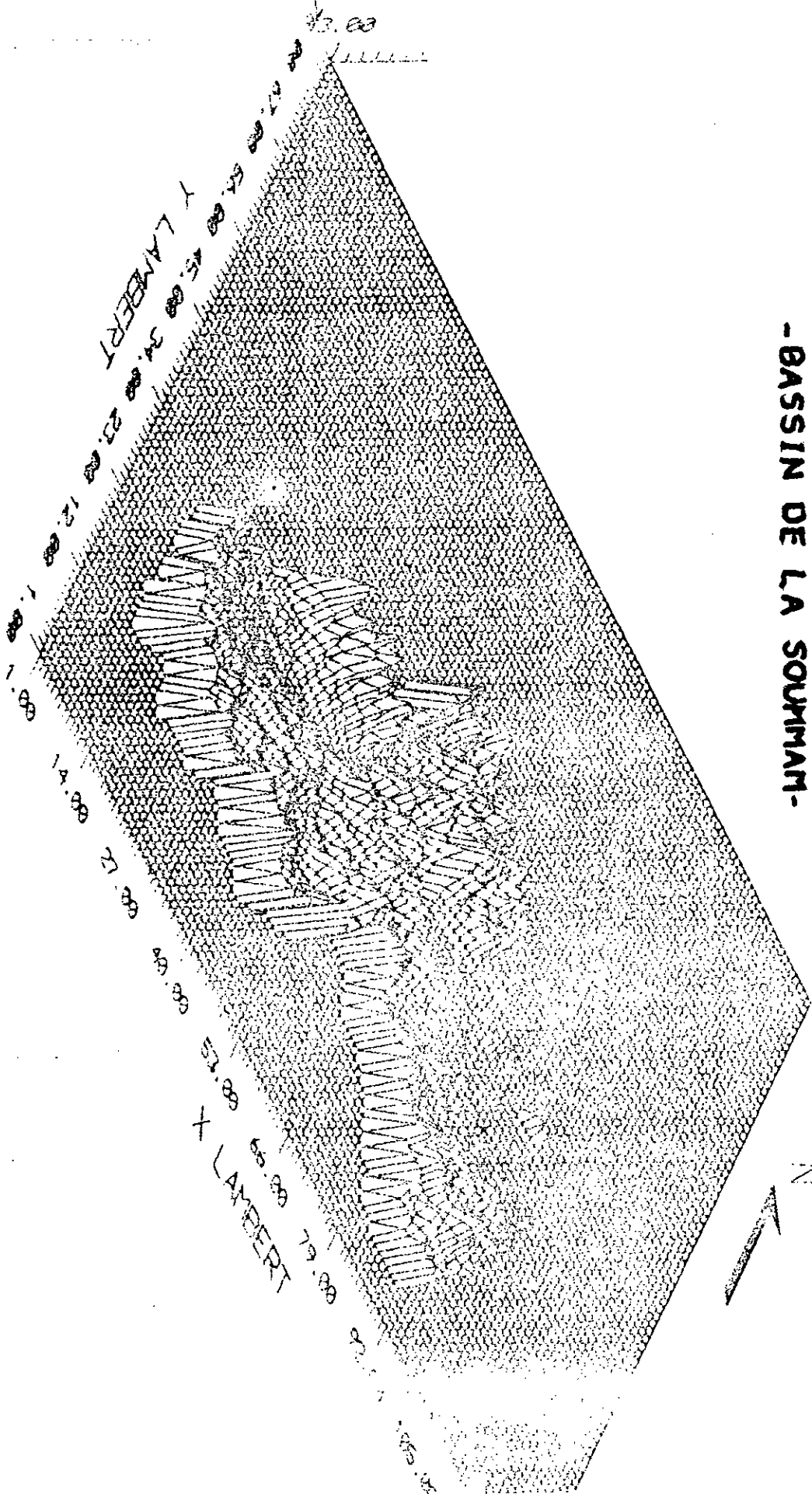
C'est la limite de la partie Sud du bassin, elle est constituée d'une série de crêtes plus petites, mutuellement séparées par :

- Djebel Dira (1810m)
- Djebel Choukchot (1832m)
- Djebel Mansoura (1862m)

Ces crêtes font un ensemble géographique délimitant nettement les terrains montagneux de l'Atlas Tellien du plateau situé au Nord. Ce dernier est connu sous le nom de la plaine du Hodna.

Nota : Sur la carte n° 3, nous remarquons bien la vallée du sahel, les reliefs du Djurdjura et le plateau Sétifien etc..., cette représentation permet de visualiser la topographie du bassin versant.

Altitudes en metres



**TOPOGRAPHIE DE TERRAIN
- BASSIN DE LA SOUMMAN -**

CARTE N° 3



I.1.4. Le climat

Les conditions climatiques dans le bassin de la Soummam ne sont pas uniformes. Les pluies sont donc réparties de façon irrégulière avec une double influence d'un vent chaud et sec (sirocco), fréquent en été et un climat doux et humide en hiver.

Il est donc régi dans ses caractéristiques générales par un climat méditerranéen du nord et par un climat brûlant et sec du Sahara.

D'après l'ouvrage "Climat d'Algérie" repris partiellement par ENERGO PROJEKT en 1965, nous déduisons que :

- le climat est humide au nord du côté du Djurdjura et Nord-Est du côté de la mer.
- Sub-humide, le long et de part et d'autre de la vallée du Sahel.
- Semi-aride à aride au niveau du plateau Sétifien et en allant vers le Sud.

Les mois les plus chauds sont en juillet et août, quand la température maximum moyenne varie entre 30°C et 37°C suivant la situation de la station.

Les mois les plus froids sont en décembre, janvier et février quand la température minimum moyenne varie entre -1,9°C à 9,2°C.

La plupart des précipitations viennent durant la période hivernale. Les mois les plus pluvieux sont : décembre et janvier, les plus secs en juillet et août.

Pendant la période (novembre à mars), il tombe plus de 61% des précipitations annuelles.

Le total des précipitations annuelles est caractérisé par de grands écarts dans ce bassin.

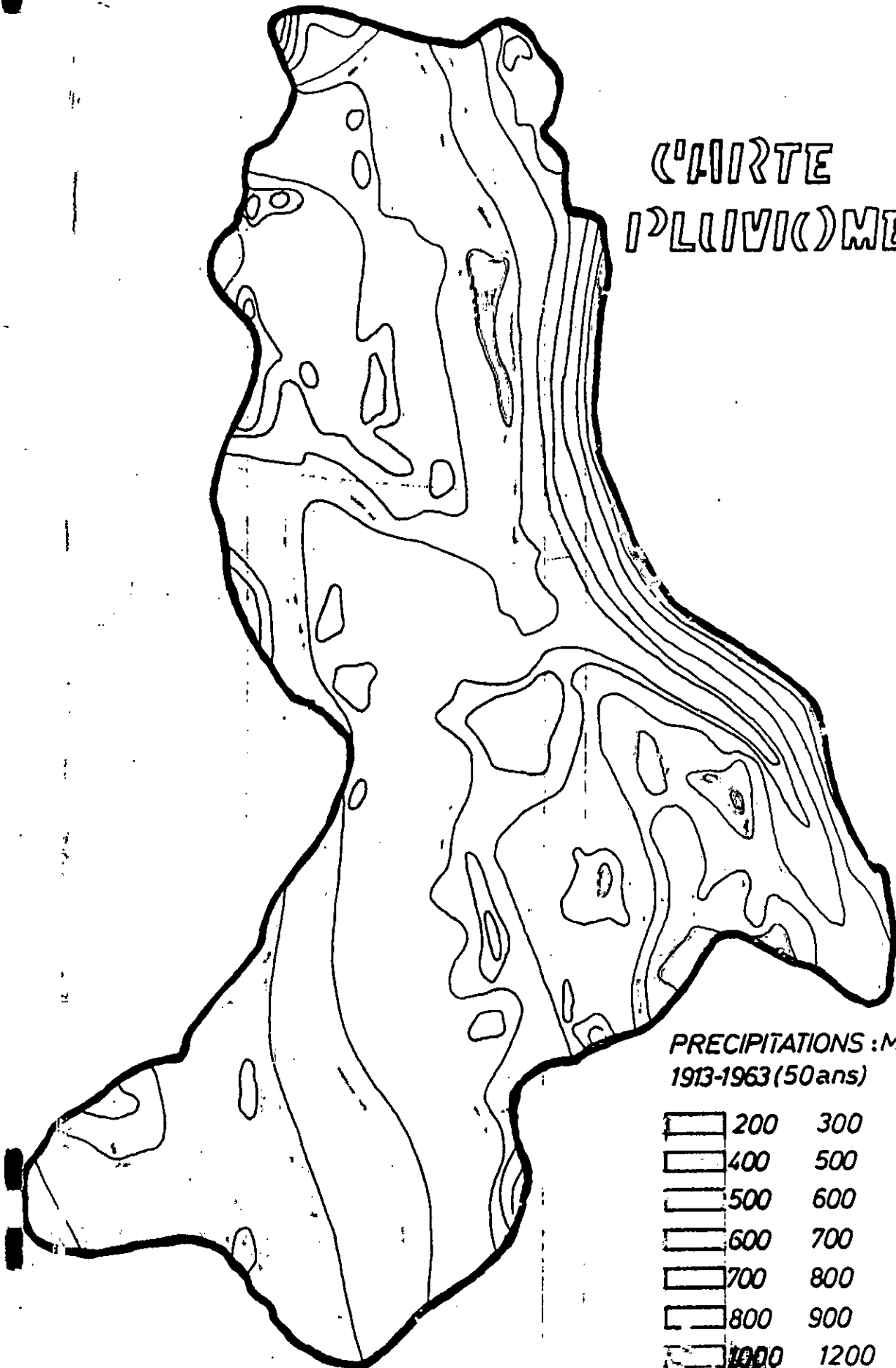
Il ressort que les années les plus humides peuvent avoir une précipitation de 3 à 4 fois supérieur à celle des années les plus sèches.

La précipitation est au maximum sur les pentes des montagnes du Djurdjura et on l'estime à environ 1500 mm/an. La côte méditerranéenne, où la pluviosité annuelle atteint 1000mm/an, est également une région de grande précipitation.

La précipitation minimum de 200 à 300 mm/an en moyenne correspond à la partie sud-est du bassin sur le plateau qui s'étend de Sétif aux Monts du Hodna (voir carte n°4).

CARTE PLUVIOMETRIQUE

Carte N°4
Ech 1:500 000



PRECIPITATIONS : Moyennes annuelles
1913-1963 (50 ans) en mm

200	300
400	500
500	600
600	700
700	800
800	900
1000	1200
1200	1500

1.2. Présentation du réseau de mesures pluviométriques

1.2.1. Historique et développement du réseau

A l'origine, les stations pluviométriques sont installées dans le but de résoudre un problème bien précis concernant la construction d'ouvrages hydrotechniques. Leurs implantations répondaient surtout à un besoin de l'heure sans tenir compte de la nécessité des grandes séries d'observations dans le traitement des données statistiques, qui peuvent être de bons indicateurs pluviométriques dans la mise au point des différents modèles mathématiques, utilisés dans les études de crues, les prévisions hydrologiques, de même pour pallier à l'insuffisance ou à l'absence des données et notamment dans la fiabilité des études d'optimisation de réseaux.

A l'instar des autres réseaux de l'A-N-R-H, celui de la Soummam s'est réalisé librement. Le fonctionnement de ses stations pluviométriques a commencé en 1900 par une station; vers 1920, une dizaine fonctionnaient déjà dont certaines le sont encore. Depuis, beaucoup de stations ont été créées et d'autres arrêtées, pour diverses raisons (disponibilité de l'observateur, accessibilité du terrain, etc...).

Il faut signaler qu'avant 1960, les régions agricoles étaient parsemées de pluviomètres qui témoignaient de l'activité des régions rurales. Depuis le développement urbain, la récupération et la distribution des terres ont fait disparaître une grande partie des fermes et par conséquent des postes bénévoles.

Il faut bien admettre que le développement de ce réseau s'est fait au départ d'une façon un peu anarchique. Mais très rapidement le service gestionnaire a aperçu la nécessité de planifier un réseau pluviométrique minimal (planification basée essentiellement sur la notion de zones «hydrologiquement homogènes»).

1.2.2 Situation actuelle du réseau

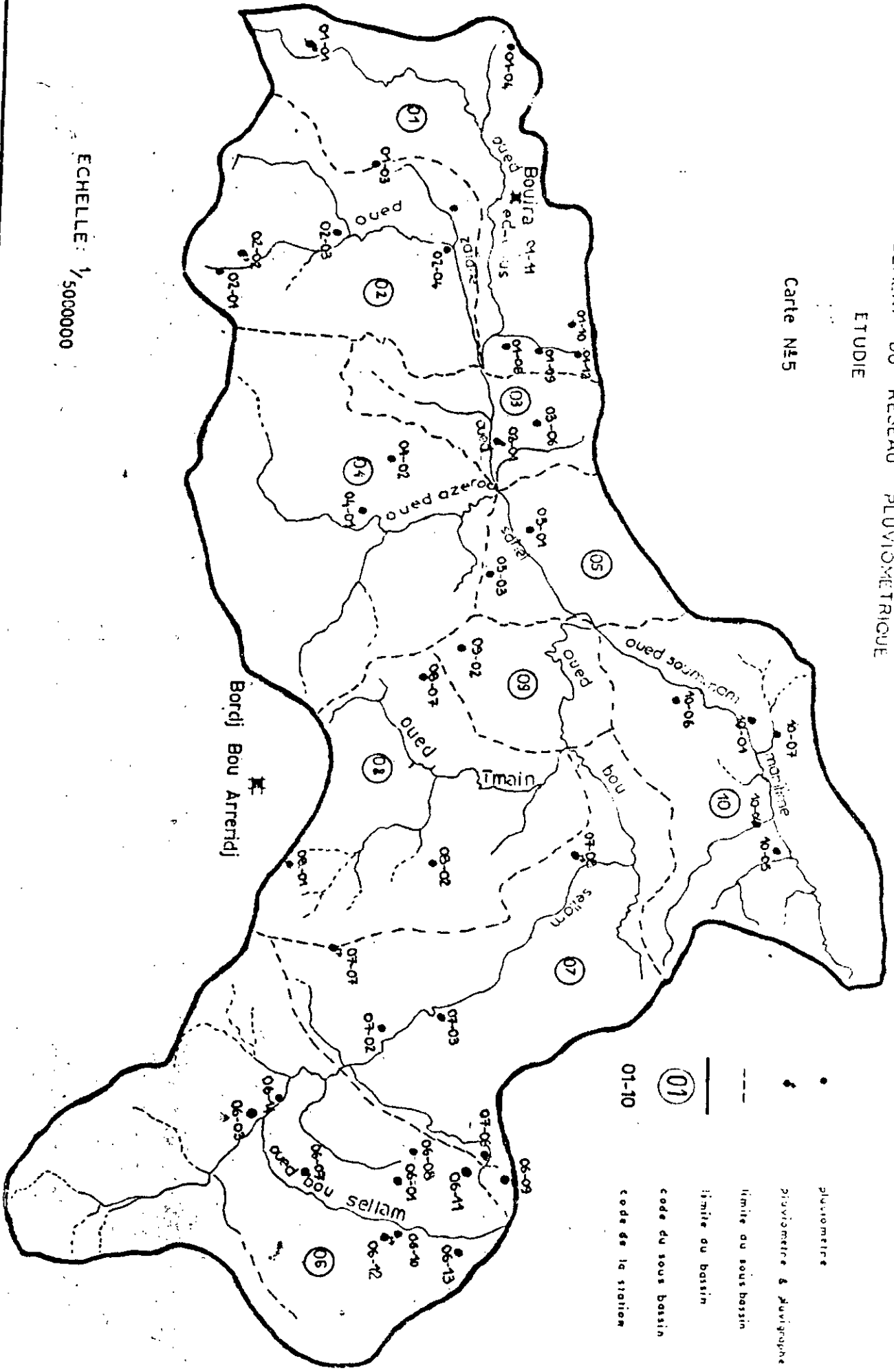
Le réseau du bassin de la Soummam est constitué de 97 stations pluviométriques dont 56 sont arrêtées ou tout au moins les données d'observations ne parviennent plus à la direction centrale.

Pour le moment 41 stations uniquement fonctionnent. Celles-ci sont réparties à travers tout le bassin, qui est codé par le numéro 15 et composé de 10 sous bassins (voir carte n°5).

Une première analyse de la carte n°5 montre un plus grands nombre de pluviomètres dans les sous-bassins (entre 8 et 9) tels ceux de Oued-Ed-Dous et Oued Bou-Sellam-Amont alors que d'autres ne comporte que deux voir même un pluviomètre; Le Sous-bassin Oued Azerou avec une superficie de 1084 km² ne possède que deux pluviomètres par comparaison à celui de Oued Ed-Dous avec une superficie de 903 km² contient 8 pluviomètres. Le tableau (I.1), met en évidence cette inégalité spatiale.

CARTE D'EMPLACEMENT DU RESEAU PLUVIOMETRIQUE
 ETUDIE
 Carte N°5

ECHELLE: 1/5000000



- pluviometre
- * pluviometre & pluviographe
- - - limite du sous bassin
- limite du bassin
- (01) code du sous bassin
- 01-10 code de la station

 * TABLEAU (N1.1) *

CODES DES STATIONS	NOM DES STATIONS	CARACTERISTIQUES GEOGRAPHIQUES DES STATIONS PLUVIOMETRIQUES			
		COORDONNEES LAMBERT		ALTITUDE	ANNEES OBSERVEES
		X(km)	Y(km)	Z(m)	
SOUS BASSIN(01) : OUED ED-DOUS / SURFACE:903km				NP = 8	
150101	SOUR EL GHOZLANE	589.30	317.15	705	52 anneéobs
150103	EL HACHINIA	600.65	325.10	719	21
150104	SIOI ATTAF	589.75	342.55	835	17
150108	ATLATA	629.05	346.15	850	14
150109	ASSANEUR	628.70	348.00	900	14
150110	AIN ALLOUANE	624.10	348.35	1025	17
150111	IRORHAT HAIZER	617.30	345.10	580	17
150112	TIKJDA	627.85	350.65	1470	13
SOUS BASSIN(02): OUED ZATANE / SURFACE: 1012km				NP = 4	
150201	TAGHDIT	616.50	303.18	1000	19 anneéobs
150202	BORDJ-OKRISS	614.70	309.90	750	14
150203	HAMMAN KSENA	613.45	321.65	600	14
150204	EL ASMAN	620.40	335.50	430	51
SOUS BASSIN(03): OUED SAHEL ANONT / SURFACE:430km				NP = 2	
150301	M'CHEDELLAH	640.90	341.70	465	23 anneéobs
150306	BENI HANMAD	635.75	346.70	546	13
SOUS BASSIN(04): OUED AZEROU / SURFACE : 1084km				NP = 2	
150401	PORTES DE FER	650.70	324.20	445	19 anneéobs
150402	BENDAOUJ	635.60	308.75	1028	16
SOUS BASSIN(05): OUED SAHEL AVAL / SURFACE :326km				NP = 2	
150501	TAZMALT	652.75	344.25	350	23 anneéobs
150503	IGHIL ALI	659.00	339.00	650	43
SOUS BASSIN(06): OUED BOU-SELLAN AMONT / SURFACE:1785km				NP = 9	
150603	GUELLAL	737.10	308.15	908	46 anneéobs
150607	AIN ARNAT	734.90	323.55	1100	31
150608	BOUIRA COLIGNY	732.00	326.65	1017	63
150609	EL-OURICIA	743.55	334.90	1089	16
150610	FERMATOU	742.00	328.60	1040	18
150611	MAHOUANE	737.75	332.95	1178	19
150612	SETIF (MRA	740.10	321.85	970	21
150613	ZEIRI	745.75	337.50	1130	19
150614	TIKTER	715.00	308.25	956	60

SUITE TABLEAU (I.1)

SOUS BASSIN(07):OUED BOU-SELLAM MOYENT / SURFACE:1234km2					NP = 5
150702	MAGRAOUA	713.75	320.80	825	15 annéesobs
150703	AIN ZADA	689.60	354.50	400	25
150706	AIN ABESSA	733.10	336.40	1112	37
150707	BIR-KASDALI	709.50	318.40	1040	59
150708	BENI-OURTILANE	693.05	351.45	1020	17
SOUS BASSIN(08):OUED EL MAIN / SURFACE:930km2					NP = 3
150801	SIDI ENBAREK	699.15	313.90	1011	20 annéesobs
150802	ZEMMORAH	692.85	331.80	950	21
150807	TENIET ENNASR	671.20	327.40	1052	13
SOUS BASSIN(09):OUED BOU-SELLAM AVAL / SURFACE:360km2					NP = 1
150902	GUELAA	669.05	334.70	1060	38 annéesobs
SOUS BASSIN(10):OUED SOUMMAN MARITIME / SURFACE:1061km2					NP = 5
151001	SIDI AICH	678.20	369.65	110	14 annéesobs
151004	DNE MAUCHE AHMED	692.75	373.40	105	15
151005	DNE ROUHA NESSAOUD	700.70	377.75	320	18
151006	SEDDOUK	678.05	362.15	484	21
151007	TIFRA	679.00	376.00	450	30

NOTA/ NP: NOMBRE DE POSTES PLUVIOMETRIQUES

I.3. La pluviométrie dans le réseau

I.3.1. La précision des pluviomètres

La mesure des précipitations, malgré son apparente simplicité, est en fait délicate point de vue qualité; On exprime la quantité d'eau tombée en millimètre (1 mm équivaut à un litre d'eau reçue sur une surface horizontale de 1 m²), la surface de réception des pluviomètres est excessivement petite par rapport à l'aire que la mesure est présumée représenter; en outre, le pluviomètre bien que dégagé de tout obstacle voisin, provoque par sa présence même, des courants artificiels locaux qui, généralement, entraînent une partie de la pluie hors du récepteur; L'appareil perturbe ainsi le phénomène qu'il a précisément pour objet de mesurer; Cet effet est encore plus marqué pour les précipitations solides : neige, grêle ou grésil.

Des améliorations ont été apportées aux pluviomètres pour réduire, sinon supprimer ces inconvénients.

Quelque soit l'appareil utilisé, la condition primordiale, pour assurer un minimum de compatibilité entre les mesures consiste à adopter un type standard unique (voir fig. I.1) et des conditions d'installations identiques (notamment en ce qui concerne la hauteur au-dessus du sol de la surface de réception).

On remarque que ces conditions n'ont pas été respectées au niveau de notre bassin versant en début du fonctionnement du réseau, ce n'est que vers les années soixante-dix qu'on adopta le type standard (pluviomètre de type association).

Lorsque ces conditions ne sont pas réunies, il est difficile de donner d'une manière générale, l'approximation sur laquelle on peut compter en matière de pluviométrie.

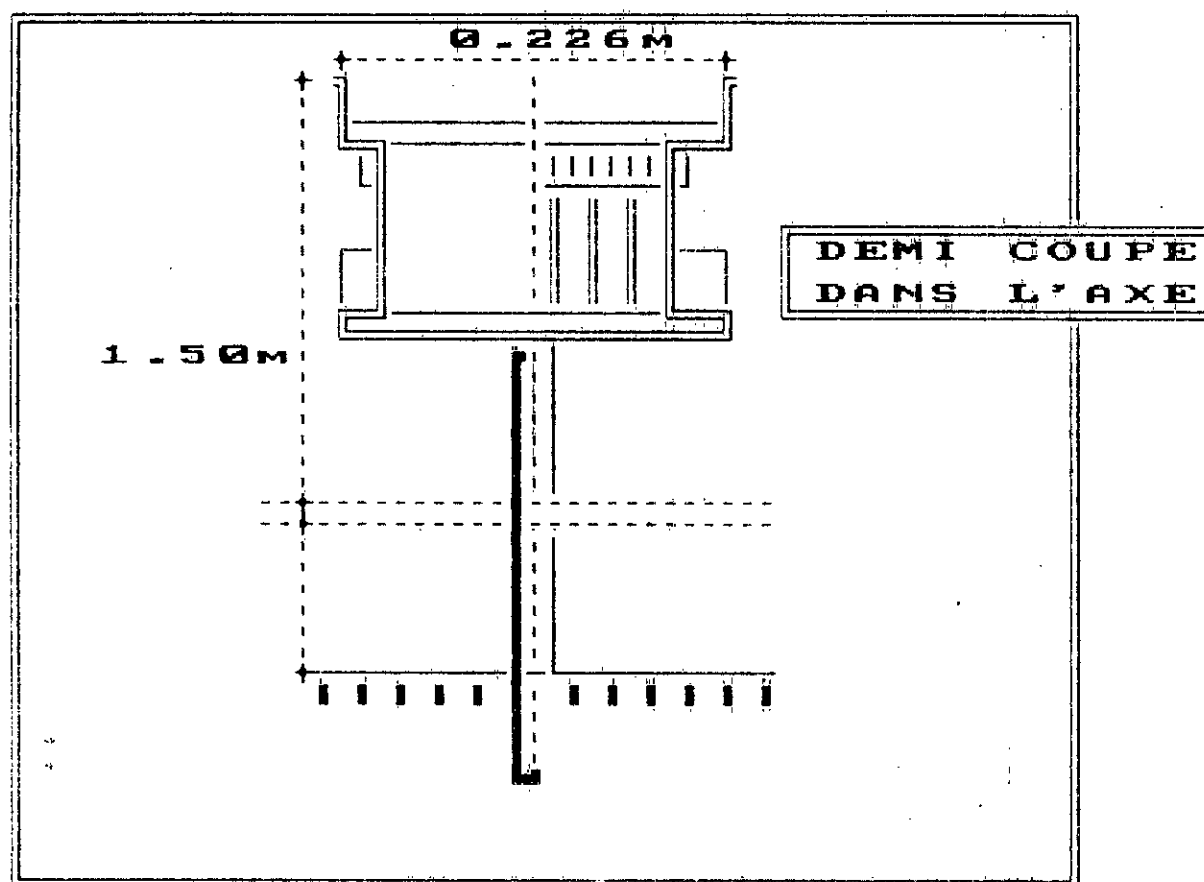
En conclusion pratique, la forme ou la taille du pluviomètre est sans importance; les mesures sont représentatives et utilisables dans les calculs hydrologiques, à la condition essentielle, qu'elles soient faites en site abrité du vent qui influence énormément sur la pluviométrie.

I.3.2. Rôle de l'observateur

On oublie trop souvent, si l'on utilise des données pluviométriques sans être directement associé à la vie quotidienne du réseau, que l'existence et l'efficacité de ce réseau sont étroitement subordonnées à l'activité et à la conscience des personnes qui font les observations.

Prendre soin d'un poste pluviométrique est, en apparence, chose simple et facile. C'est en réalité, pour qui en accepte la charge sans avoir un goût profond de ce travail, une tâche vite fastidieuse et, à la longue, rebutante et astreignante, parce qu'elle oblige à répéter deux fois par jour, aux mêmes heures (8h, 18h) les mêmes opérations : se rendre au poste, lire l'éprouvette, noter le résultat des mesures sur un imprimé, puis tout les mois, transmettre les relevés au service centrale.

Cet appareil, couramment utilisé est représenté par la Fig-11-1; a été répondu autre fois par les soins de "l'association scientifique de france"



fig(1-1) pluviomètre "Association"

40/1
C'est le nom de pluviomètre "Association" sous le quel on le désigne souvent

Il est composé de trois parties:

- 1) _ Un seau en zinc où s'accumule la pluie recueillie
- 2) _ Un entonnoir de même métal formant une surface réceptrice comportant un trou assez petit (pour diminuer les pertes par évaporation) protégé par une toile métallique fine
- 3) _ une bague circulaire à bords presque tranchants de 226 mm de diamètre environ, limitant la surface réceptrice ci-dessus, la quelle mesure exactement 400 cm^2

Cette observateur se verrait, pour ainsi dire, rivé à celui-ci finit par faillir à sa tâche. Alors compter exclusivement sur la coopération des observateurs dévoués et consciencieux, serait souvent se condamner à ne disposer plus tard que des séries pluviométriques relativement courtes. On se rappelle néanmoins que l'on peut utiliser avec fruit une telle documentation.

I.3.3. Densité du réseau pluviométrique

L'insuffisance du réseau pluviométrique est une autre lacune en matière de mesure des précipitations, puisque l'observation des précipitations à l'aide de pluviomètres correspond à un échantillonnage ponctuel, un réseau dense de pluviomètres est nécessaire pour obtenir des valeurs sûres pour un bassin versant, mais il est aussi indéniable que sans exigence de qualité des mesures ponctuelles, la densité ne suffit pas, et son importance peut même accélérer le processus de dégradation de l'information pluviométrique.

La densité optimum du réseau pluviométrique dépend évidemment du but poursuivi et de l'hétérogénéité spatiale des pluies dans la région étudiée. Elle varie selon que l'on a besoin de valeurs moyenne journalières, hebdomadaire, mensuelles ou annuelles.

Pour les analyses annuelles par exemple, il faut des réseaux moins denses; la densité de notre bassin est de l'ordre de 0.004 (nombre de pluviomètre au km²). Par rapport à d'autres bassins c'est l'une des plus forte concentrations du point de vue densité.(voir carte n° 6)

I.4. Source des données

Les données pluviométriques utilisées dans cette étude nous ont été aimablement fournies par l'A.N.R.H, elles consistent en :

- Un ensemble de registres appelés, fichiers originaux, chacun représentant une station pluviométrique et contenant les renseignements suivant :
 - code et nom de la station
 - la situation géographique en coordonnées Lambert (X,Y) et l'altitude (Z) de la station
 - les pluies journalières et mensuelles.
- des listings obtenus par ordinateurs, contenant les renseignements suivants :
 - code de la station
 - les pluies mensuelles et leurs totaux annuels

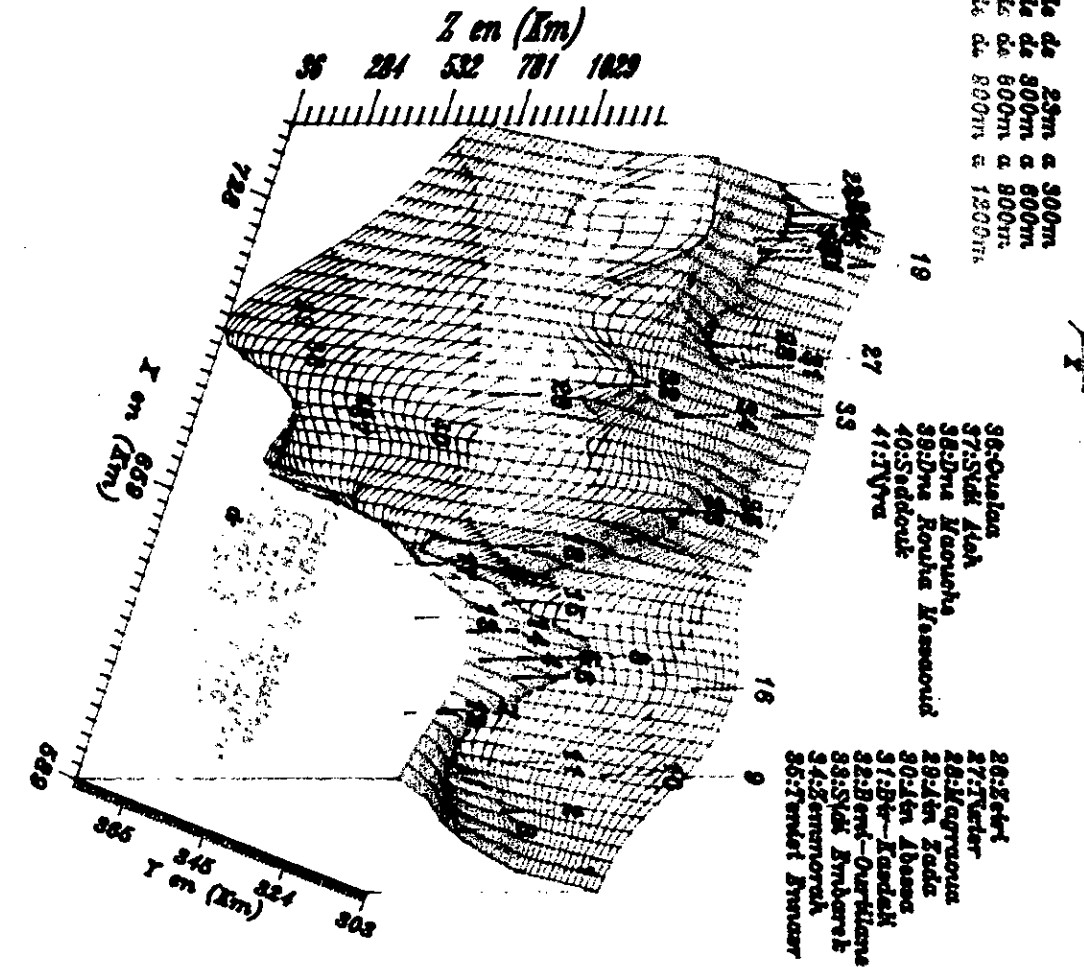
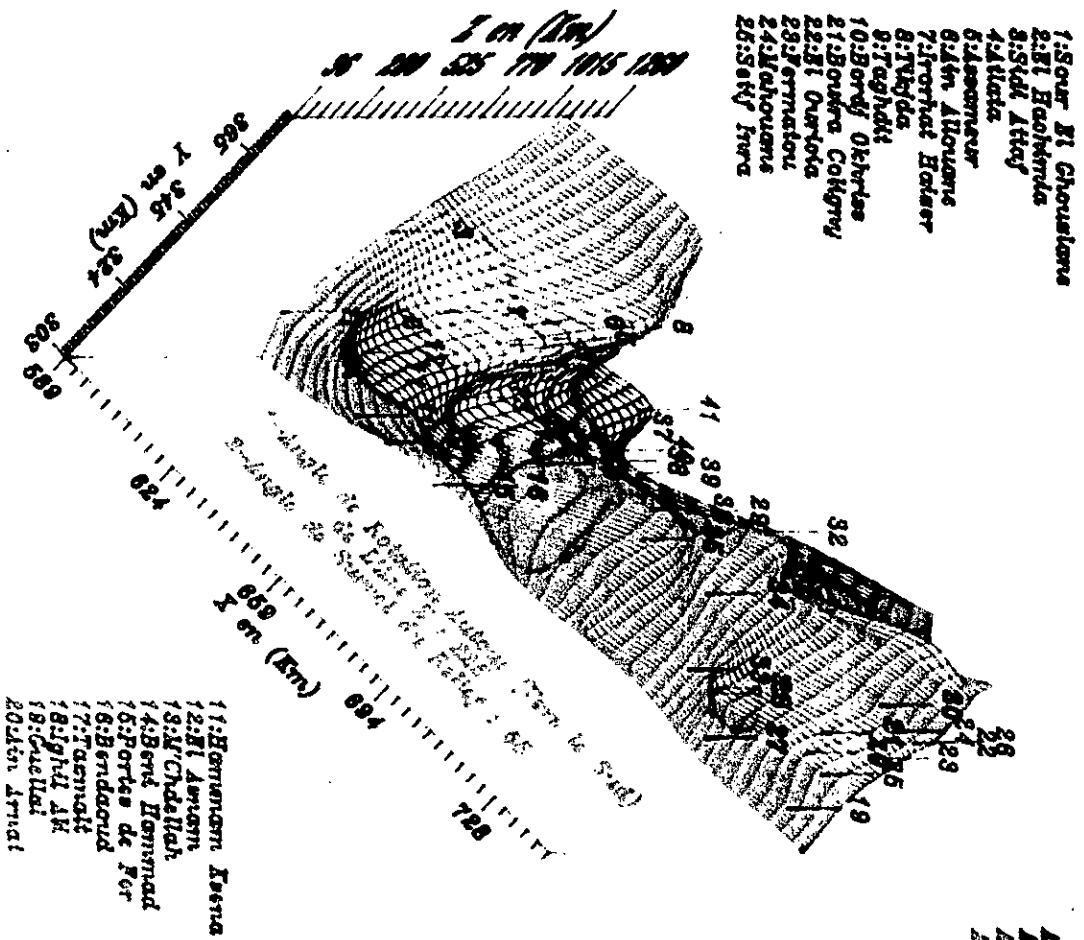
Notre première tâche consistait à confronter les données des fichiers originaux à celles des listings de l'ensemble des stations de la Soummam, c'est à dire sur les 41 stations, dans le but de détecter des erreurs de report ou de recopie, et d'en tirer le maximum d'informations.

Les principaux types de problèmes qu'on a recensé lors de la confrontation sont :

- changement fréquent de l'observateur

(N° 6)

CARTE RELIEF DU BASSIN DE LA SOUMMAH (9200 Km²)



- changement d'altitude du pluviomètre
- changement de type de pluviomètre (Déculpateur, Ville de Paris et Association)
- le non sérieux de l'observateur :
 - . par des absences fréquents
 - . des relevés non précis (il arrondi les chiffres et n'apprécie pas le 1/10 du mm)
- des données douteuses signalées par le service centrale (une donnée est douteuse lorsqu'elle est trop faible ou trop élevée comparativement aux pluviomètres avoisinants).
- deux types de lecture en mm et en cm³ étaient utilisées dans le réseau entraînant des confusions. La correction est faite en tenant compte d'un coefficient de correction en fonction de la surface réceptrice du pluviomètre.

Ces informations nous serviront par la suite dans la phase critique des données.

Les données recueillies présentent des erreurs accidentelles et parmi ces dernières nous avons :

- des erreurs de transcription sur listing :
 - valeur fausse ou virgule mal placée surestime ou sous-estime la véritable valeur

exemple : valeur fausse

Nom : Tikjda code : 15.01.12
 valeur inscrite : 61.0 mm } mois : octobre
 valeur corrigée : 16.0 mm } année : 1989

- dans le cas où il n'y a pas de relevé, par erreur au lieu de mettre le symbole (-10), ils mettent une valeur quelconque ou zéro.

exemple : valeur non mesurée

Nom : Sidi Attaf code 15.01.04
 valeur inscrite : 20.8 mm } mois : septembre
 valeur corrigée : -10 } année : 1967

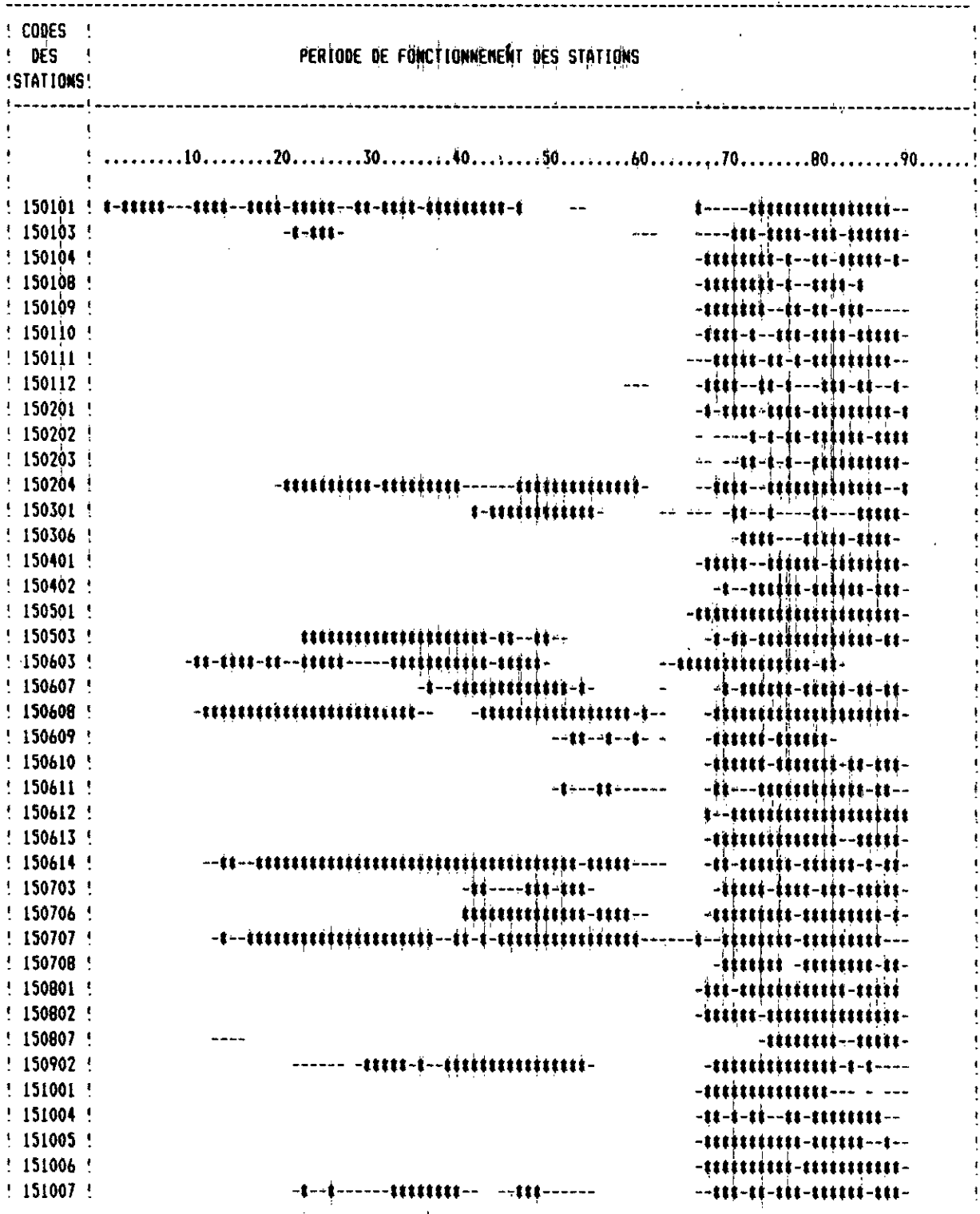
Une fois les données corrigées, on fait notre propre saisie sur support informatique en recalculant les totaux annuels. Ainsi, les données contrôlées sont prêt pour l'analyse statistique.

La bieveté de ce résumé, pourrait cacher l'effort et la patience déployé aux yeux des lecteurs qui n'ont point fait de confrontation de données, chose qui nous a pris beaucoup de temps pour mener à bien notre critique préliminaire.

Il se trouve que les durées d'observations de chaque station, sont très variables : de 13 ans à plus de 30 ans pour certaines et que la plupart d'entres elles présentent des lacunes à l'échelle mensuelle variant de 1 mois à plus de 6 mois, et à l'échelle annuelle s'étalant sur plusieurs années (voir tableau synoptique n° I.2).

Vu l'irrégularité des observations, on a préféré ne pas faire de choix «à priori», l'élimination de certaines stations non représentatives du milieu naturel se fera plus tard et sur des critères plus objectifs.

TABLEAU SYNOPTIQUE (N1.2)



LEGENDE
 * : ANNEE COMPLETE
 - : ANNEE INCOMPLETE
 : INTERRUPTION DE LA STATION

CHAPITRE II

ANALYSE DES DONNÉES

II.1. Etude de la normalité

Avant d'entamer la critique proprement dite des séries chronologiques, voir même leur extension, on a été conduit à étudier la loi suivie par les valeurs annuelles

II.1.1. Aspect théorique de la loi normale

La loi de Gauss, ou loi normale (loi symétrique par rapport à la moyenne) est généralement retenue pour décrire les échantillons dont les variables sont des sommes de variables aléatoires, ce qui est le cas des pluies annuelles.

Dans les études hydrologiques, on utilise la loi normale centrée réduite, loi standard de paramètre (0 et 1), c'est à dire de moyenne nulle et d'écart-type un.

La variable centrée réduite de Gauss, notée U, a pour fonction de distribution :

$$F(U) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-x^2/2} dx$$

II.1.2. Vérification de l'adéquation d'une loi supposée représentative de la loi de probabilité de la population mère

Après avoir choisi une loi de probabilité et estimé ses paramètres, on doit tester l'adéquation de cette loi, c'est à dire, voir si l'échantillon observé peut être considéré comme extrait d'une population normale.

Plusieurs tests sont disponibles et peuvent être utilisés. On s'intéressera dans cette partie aux tests suivants :

- le test graphique basé sur la droite de Henri.
- le test du Khi-deux (ou test de PEARSON), test tout à fait classique
- le test de KOLMOGOROV et SMIRNOV (plus récent)

II.1.2.1. Méthode graphique : ajustement par la droite de Henri

A - Principe de l'ajustement

Ce procédé, rapide mais sommaire, s'applique uniquement aux situations dans lesquelles on envisage d'ajuster à une distribution empirique une loi normale ou une loi assimilable comme la loi log-normale. Son champ d'application est donc restreint.

La méthode s'appuie sur le fait qu'une variable normale de paramètre m et σ se transforme en variable centrée réduite et normale de paramètre 0 et 1. En effet, grâce à un simple changement de variable $U=(X-m)/\sigma$, la courbe de la fonction de répartition de X se transforme en une droite dont l'équation est : $X=m + \sigma U$. C'est la droite de Henri représentant la courbe de Gauss.

Dès lors, cette méthode consiste à représenter la distribution expérimentale dans un système d'axes de coordonnées à échelle linéaire en X et en U et voir si les points sont suffisamment réparties de part et d'autre de cette droite pour que l'on puisse accepter l'ajustement par la loi normale.

Naturellement la réponse est incertaine puisqu'elle se fonde sur une appréciation de linéarité purement graphique et donc très subjective (chacun ayant son idée de l'ajustement à peu près), seul des tests puissants permettent par leur rigueur de faire l'unanimité.

Il se trouve que la représentation conjointe de la loi ajustée et la série d'observation possède une double utilité. Elle permet en effet :

a) lorsque la loi ajustée est rejetée par un test d'adéquation, de déterminer si la raison du rejet provient de la forme de la loi ajustée ou de la présence d'observation aberrante ou des deux.

b) lorsque la loi ajustée est acceptée, de lire aisément la probabilité F associée à une valeur particulière quelconque de la variable observée X.

B - Méthodologie adoptée :

La démarche est la suivante :

- A chaque élément de la série classée par ordre croissant, on associe une fréquence expérimentale au non dépassement égale à :

$$F = i/(N+1) \quad \begin{array}{l} i : \text{numéro d'ordre} \\ N : \text{taille de l'échantillon} \end{array}$$

- A chaque valeur de F correspond une valeur de U, pour ce faire, on peut utiliser les tables de la fonction de Gauss ou approximer la fonction de Gauss par la formule polynomiale suivante :

$$\text{pour } F \leq 0.5 \quad - U = X - \left(\sum_{i=0}^2 a_i x^i \right) / \left(\sum_{i=0}^3 b_i x^i \right)$$

$$\text{Avec : } X = \sqrt{\ln(1/F^2)}$$

pour $F > 0.5$
on prendra $F^* = 1-F$ et $U = -U$

Cette approximation permet une programmation facile avec une erreur de moins de $4.5E-5$

- la droite de Henri est calculée par des points de coordonnées (X_i^*, U_i) , $i = 1$ à N

Elle permet, en outre de vérifier l'estimation de la moyenne \bar{X} et l'écart type S, calculée sur

l'échantillon. En prenant par exemple pour :

U=0, on a $X^* = X$
 et U=2, on aura $X^* = X+2S$

- Intervalle de confiance à α % sur le quantile X^* est calculé de la façon suivante :

Les bornes de l'intervalle de confiance sont :

$$X^* \pm \frac{U_\alpha/N \sqrt{1 + U_r^2/2} \pm (U_r U_\alpha)/2N}{1 - U^2/2N} .S$$

{ 2 fois + pour borne supérieure
 { 2 fois - pour borne inférieure

Avec :

U_α : variable réduite de Gauss ayant la fréquence au non dépassement $(1 - (1-\alpha)/2)$

U_r : variable réduite de Gauss ayant la fréquence au non dépassement F.

α : risque d'erreur choisie à 5%

Un exemple d'une feuille de calcul : ajustement à une loi normale donnée par le logiciel utilisé dans cette étude est montré dans la partie annexe.

C : Analyse des résultats

Le tableau (II.1) donne une première appréciation de l'ajustement à partir des graphiques.

Les observations faites à partir de ces graphiques révèlent que parmi les 15 stations jugées d'un mauvais ajustement à la loi normale (sur leurs séries d'observations), il y en a qui présentent soit des déviations systématiques de la représentation graphique de la série d'observation par rapport à celle de la loi ajustée (exemple de la station de TIFRA, code : 151007), soit des distributions par paquets (exemple de la station de BENDAOU, code : 150402), soit par deux ou trois observations extrêmes divergentes (exemple de la station de MAHOUANE, code 150611) voir annexe.

Après vérification dans les originaux, on a constaté que les déviations des points correspondaient surtout à des années très pluvieuses par rapport à la moyenne annuelle de chaque station (par exemple les années 1968 à 1973) et la distribution par paquets de points étaient dû à des observations erronées ou entachées d'erreurs par le changement fréquent des observateurs pour le même poste pluviométrique.

Tableau (II-1)

CODE STATION	NOMS DES STATIONS	OBSERVATIONS
15-01-01	SOUR EL GHOZLANE	1
15-01-03	EL-HACHIMIA	1
15-01-04	SIDI ATTAF	0
15-01-08	ATLATA	1
15-01-09	ASSAMEUR	1
15-01-10	AIN ALLOUANE	0
15-01-11	IRORHAT-HAIZER	0
15-01-12	TIKJDA	1
15-02-01	TAGHDIT	1
15-02-02	BORDJ OKRISS	0
15-02-03	HAMMAM KSENA	1
15-02-04	EL-ASNAM	1
15-03-01	M'CHEDELLAH	1
15-03-06	BENI HAMMED	1
15-04-01	PORTES DE FER	0
15-04-02	BEN DAOUD	0
15-05-01	TAZMALT	1
15-05-03	IGHIL ALI	0
15-06-03	GUELLAL	0
15-06-07	AIN ARNAT	1
15-06-08	BOUIRA (coligny)	1
15-06-09	EL-OURICIA	0
15-06-10	FERMATOU	0
15-06-11	MAHOUANE	0
15-06-12	SETIF INRA	1
15-06-13	ZEIRI	0
15-06-14	TIXTER	1
15-07-02	MAGRAOUA	1
15-07-03	AIN ZADA	1
15-07-06	AIN ABESSA	1

Suite du tableau (II-1)

15-07-07	BIR KASDALI	0
15-07-08	BENI-OURTILANE	1
15-08-01	SIDI EMBAREK	1
15-08-02	ZEMMORAH	1
15-08-07	TENIET-ENNASR	1
15-09-02	GUELAA	0
15-10-01	SIDI AICH	1
15-10-04	Dne MAUCHE AHMED	1
15-10-05	Dne ROUHA MESSAOUD	1
15-10-06	SEDDOUK	1
15-10-07	TIFRA	0

legende:

chiffre 1 —————>————— bon ajustement
 chiffre 0 —————>————— mauvais ajustement

En principe, les observations devraient néanmoins fluctuer autour de la droite théorique dans la zone de confiance choisie.

Enfin, on peut dire que la qualité des données a une très grande importance dans la décision d'un ajustement à une loi théorique et peut donc fausser l'ajustement et mettre en doute le choix de cette loi.

Toutefois, afin de vérifier notre point de vue, nous utilisons les deux tests d'adéquation.

II.1.2.2. Test du KHI-DEUX χ^2

A - Principe du test

Le champ d'application de ce test est très vaste puisqu'il permet d'ajuster à une distribution expérimentale quelconque une distribution théorique adéquate.

Ce test d'hypothèse est une procédure rigoureuse, débouchant sur une décision simple (oui ou non) et particulièrement efficace puisqu'il se réalise uniquement à partir des effectifs.

La procédure consiste à tester l'hypothèse de similitude entre deux distributions théorique et expérimentale. Deux situations existent en théorie :

- soit l'hypothèse (notée H_0) est vraie, c'est à dire, que l'échantillon est effectivement extrait de la population mère envisagée, et l'écart entre les effectifs observés et théoriques a des fortes chances d'être plutôt faible, même s'il peut dans certains cas, très rares, être très élevé.

- soit l'hypothèse (notée H_1) est fautive, c'est à dire que l'échantillon n'est pas tiré d'une population mère envisagée, et l'écart à plutôt de fortes chances d'être important, même si le hasard peut exceptionnellement conduire, dans ce cas, à une forte proximité.

Pour prendre la décision, il faut fixer un seuil limite tel que le risque d'erreur, inévitable pour une telle décision, soit le plus faible possible (généralement de l'ordre 1%, 5% ou 10%).

Dès lors, il permet de rejeter l'hypothèse (la loi ajustée) mais pas nécessairement à la retenir, car en cas de rejet, on est en mesure de préciser le risque d'erreur (dit de première espèce ou de rejet d'une hypothèse vraie), par contre, lorsqu'on accepte l'hypothèse, on ne connaît pas le risque précis d'erreur c'est à dire celui d'acceptation d'une hypothèse fautive.

La décision dépend naturellement des caractéristiques de l'observation et plus particulièrement du nombre de classes de la variable et de la taille de l'échantillon.

Cependant, ce test accorde une grande importance aux valeurs centrales, et contrôle mal la dispersion des valeurs extrêmes, il n'est donc valable que pour de grands échantillons.

B - Le déroulement du test du χ^2

De façon pratique, l'application de ce test se divise en plusieurs étapes :

* découpage de l'échantillon des «N» totaux annuels en «K» classes d'effectifs n_1, n_2, \dots, n_k tels que :

$$\sum_{i=1}^k n_i = N$$

* Sous l'hypothèse de la loi testée (la loi normale), on calcul les effectifs théoriques (ou fréquences théoriques) notés Np_i avec $Np_i = N \times$ (probabilité que X soit compris entre les deux limites de la classe i).

Les paramètres (\bar{X} et S), de la loi testée, nécessaire au calcul de Np_i sont estimés sur la seule série des «N» valeurs.

A ce stade, il est indispensable que chacun des effectifs théoriques Np_i , soit au moins égal à 5, et si ce n'est pas le cas, on modifie la distribution initiale en regroupant les classes par voisinage. On comprend en effet, que lorsque les effectifs sont trop petits, les écarts sont trop sensibles et donc le test peu probant.

On détermine alors la distance entre les deux distributions, plus précisément entre les effectifs observés et théoriques. Cette distance, dite du KHI-DEUX s'écrit ainsi :

$$d = \sum_{i=1}^k [(n_i - Np_i)^2 / Np_i]$$

On observe que cet indice est une somme d'écarts élevés au carré et relativisés, correspond à la façon dont on a défini une variable du KHI-DEUX, on connaît donc sa loi de probabilité, cette loi du χ^2 dépend du nombre de termes de cette somme, appelé nombre de degré de liberté (ν) correspondant au nombre minimum de termes nécessaires à la définition de la distribution théorique.

On a finalement :

$$\nu = K - 1 - r$$

K = nombre de classe (une fois le regroupement d'effectifs réalisés)

r = le nombre de paramètres estimés d'après l'échantillon

pour $r = 2$ on a $\nu = K - 3$

* On conclut ce test en prenant la décision d'acceptation ou de refus de l'hypothèse envisagée. Pour cela, on détermine la valeur critique (ou valeur théorique du χ^2) notée $\chi^2_{(\alpha, \nu)}$, lue sur la table numérique de la loi du χ^2 à ν degré de liberté avec un risque d'erreur $\alpha = 5\%$.

Enfin, on compare cette valeur $\chi^2_{(\alpha, \nu)}$ avec d calculée

. si $d < \chi^2_{(\alpha, \nu)}$, on accepte l'hypothèse, considérant que l'écart est faible et imputable aux inévitables fluctuations d'échantillonnages, mais elle n'est pas forcément la meilleure, d'où l'application du deuxième test qui est plus puissant.

. si $d > \chi^2_{(\alpha, \nu)}$, on rejete l'hypothèse, car dans le cas où elle serait vraie, un tel écart est très rare, avec au plus 5 chances sur 100 de se tromper que les observations n'appartiennent pas à la loi testée.

C - Analyse des résultats du χ^2 (voir annexe, les résultats du test)

Les résultats de ce test sont résumés dans le tableau (II.2).

Le test a été effectué sur 12 stations comportant des séries d'observations relativement grandes. L'analyse des résultats est conforme à celle obtenue lors de l'interprétation des graphiques, donc pour les séries rejetés les deux tests (graphique et χ^2) sont unanimes.

II.1.2.3. Test de KOLMOGOROV SMIRNOV (K-S)

A - Principe du test

Le test de Kolmogorov Smirnov est utile surtout pour des petits échantillons pour lesquels le test du χ^2 devient peu significatif.

En revanche, ce test est sensible à toute différence éventuelle entre les distributions : valeurs centrales, dispersion, aplatissement, etc... Il ne tient pas compte de la taille de l'échantillon, il est plus générale d'où son caractère non paramétrique.

Il s'agit de comparer l'écart maximum D_n entre la fonction de répartition théorique notée $F_0(x)$ et la fonction de répartition expérimentale notée $F_n(x)$.

La statistique D_n définie comme la plus grande distance verticale entre $F_0(x)$ et $F_n(x)$ est calculée comme suit, suivant des considérations pratiques :

$$D_n = \text{Max} \{ D_n^+ ; D_n^- \}$$

$$\text{Avec : } D_n^+ = \text{Max} \{ F_n(x_i) - F_0(x_i) \}, (i = 1, 2, \dots, n)$$

$$D_n^- = \text{Max} \{ F_0(x_i) - F_n(x_i) + 1/n \}, (i = 1, 2, \dots, n)$$

noms des stations	codes des stations	d	ν	$\chi^2_{(\alpha, \nu)}$	résultats
AIN ZADA	15-07-03	0.423682	1	3.84146	01
TIFRA	15-10-07	8.36139	1	3.84146	00
AIN ARNAT	15-06-07	1.19827	1	3.84146	01
AIN ABESSA	15-07-06	3.42831	2	5.99147	01
GUELAA	15-09-02	3.16951	2	5.99147	01
IGHIL ALI	15-05-03	8.06317	3	7.81473	00
GUELLAL	15-06-03	14.1771	4	9.48773	00
EL ASNAM	15-02-04	4.80693	4	9.48773	01
SOUR EL GHOSLANE	15-01-01	4.31809	5	11.0705	01
BIR- KASDALI	15-07-07	13.2662	6	12.5916	00
TIXTER	15-06-14	7.63564	5	11.0705	01
BOUIRA	15-06-08	2.33219	5	11.0705	01

TABLEAU < II-2 >

LEGENDE

RESULTATS :

01 -----> AJUSTEMENT ACCEPTE
00 -----> AJUSTEMENT REJETE

Ces écarts maxima entre la courbe continue $F_0(x)$ et la courbe en escalier $F_n(x)$ serviront à tester l'hypothèse nulle H_0 .

B - Déroulement du test

En pratique, ce test se divise en trois étapes :

1 - D'abord les fréquences cumulées théoriques sont calculées pour chaque valeur observée et ces fréquences sont tabulées dans une colonne à côté des fréquences expérimentales déduites à partir des intégrales de la fonction de distribution de Gauss.

2 - Ensuite on calcule les différences entre ces deux colonnes pour chaque observation et le plus grand écart D_n est retenu :

$$D_n = \text{Max} \{D_n^+; D_n^-\}$$

3 - Enfin, cette valeur D_n est comparée à la valeur critique $d_{(n,\alpha)}$ pour la taille n de l'échantillon et un seuil de confiance α . Les valeurs de la probabilité $d_{(n,\alpha)}$ ont été tabulées, dans le cas d'une distribution normale définie par la moyenne et l'écart type estimés \bar{X} et S , on peut utiliser les valeurs critiques approchées suivantes, au niveaux 0.05 et 0.01.

EFFECTIF DE L'ECHANTILLON (n)	NIVEAU DE SIGNIFICATION	
	0.05	0.01
$n > 5$	$0.886 / \sqrt{(n+1.5)}$	$1.031 / \sqrt{(n+1.5)}$

On rejettera l'hypothèse nulle H_0 , respectivement aux niveaux de signification 0.05 et 0.01 lorsque l'écart maximum D_n est supérieur ou égal à l'une ou l'autre de ces valeurs.

C - Analyse des résultats du test

Les résultats de ce test effectué sur les séries d'observations de 37 stations sont resumés par les tableaux (II.3) et (II.4).

Tableau (II-3)

NOMS DES STATIONS	CODES DES STATIONS	n (Ans)	Dn	NIVEAU α		RESULTATS
				0.05	0.01	
AIN - ZADA	15.07.03	25	0.081166	0.172112	0.200279	01
AIN - ARNAT	15.06.07	31	0.085543	0.155415	0.18085	01
AIN - ABESSA	15.07.06	37	0.139878	0.14279	0.166160	01
GUELAA	15.09.02	38	0.118105	0.14097	0.16405	01
EL- ASNAM	15.02.04	51	0.0991308	0.122279	0.14229	01
SOUR-EL-GHOZLANE	15.01.01	52	0.105275	0.12113	0.14095	01
TIXTER	15.06.14	60	0.107125	0.11298	0.131468	01
BOUIRA	15.06.08	63	0.110113	0.110319	0.128374	01

LEGENDE

Chiffre 01 —————> Ajustement accepte

TABLEAU (II-4)

NOMS DES STATIONS	CODES DES STATIONS	N (ans)	DN	NIVEAU α		RESULTATS
				0.05	0.01	
TIKJDA	15.01.12	Cult.	0.131689	(II-1)		01
TENIET-ENNASSR	15.08.07	13	0.155002	0.232675	0.270754	01
BENI HAMMAD	15.03.06	13	0.128252			01
ATLATA	15.01.08	14	0.136178			01
ASSAMEUR	15.01.09		0.144562			01
BORDJ-OKHRISS	15.02.02		0.245251	0.255044	0.261874	00
HAMMAM KSENA	15.02.03		0.116911			01
SIDI AICH	15.10.01		0.167378			01
DNE MAUCHE AHMED	15.10.04	15	0.112842	0.218118	0.253815	01
MAGRAOUA	15.07.02		0.140494			01
BENDAOU	15.04.02	16	0.1957470	0.211795	0.246456	01
EL OURICIA	15.06.09		0.214406			00
AIN ALLOUANE	15.01.10	17	0.21783	0.205991	0.239703	00
IRCRHAT HAIZER	15.01.11		0.189471			01
SIDI ATTAF	15.01.04		0.2579			00
BENI OURTILANE	15.07.08		0.159586			01
FERMATOU	15.06.10	18	0.147791	0.20064	0.233475	01
DNE ROUHA MESSAOU	15.10.05		0.159569			01
TAGHDIT	15.02.01	19	0.108087	0.195685	0.22771	01
PORTES DE FER	15.04.01		0.2162332			00
MAFOUANE	15.06.11		0.198275			00
ZEIRI	15.06.13		0.23595			00
EL HACHIMIA	15.01.03	20	0.153858	0.19108	0.22235	01
SIDI-EMBAREK	15.08.01		0.150788			01

Les séries d'observations des stations, dont l'ajustement à été accepté par le test χ^2 est confirmé par ce dernier voir tab. (II.3).

Pour les autres séries, le test de K-S a rejeté 7 ajustements à la loi normale voir tab. (II.4).

En définitive, on a constaté que ce test n'est pas sensible aux valeurs aberrantes, car il a accepté trois ajustement à la loi normale, or suivant les graphiques, on voit bien que ce n'est pas conforme

II.2 Ajustement à une autre loi théorique

L'ajustement à la loi normale a fourni des résultats pas très satisfaisants pour 15 séries pluviométriques; pour cela on a tenté l'ajustement à la loi Log-normale.

A - Rappel théorique de la loi Galton : (Loi Log-normale)

La loi Log-normale donne d'assez bon résultats en climat méditerranéen. (Chaumont/MEYLAN).

La loi de Galton est la loi de distribution de la variable aléatoire X, lorsque cette dernière est transformée en :

$Y = \log (X - X_0)$ suit une loi normale "loi à trois paramètres".

Sa fonction de répartition est la suivante

$$F(X) = 1/ \sqrt{2\pi} \int_{-\infty}^u e^{-\frac{t^2}{2}} dt$$

Avec : $U = a \log (X - X_0) + b$

U : variable réduite de Gauss

X_0 : paramètre de position, borne inférieur de la variable aléatoire X
 $X \in]X_0, + \infty [$

a et b : respectivement paramètre d'échelle et paramètre de forme.

Les paramètres de la droite d'ajustement sont calculés par les formules suivantes :

Avec : $a = 1.517 / (\text{Log} (1 + S^2 / (\bar{X} - X_0)))^{0.5}$

$b = 1.1513/a - a \text{Log} (\bar{X} - X_0)$

où $\frac{S}{\bar{X}}$: écart - type
 \bar{X} : moyenne

B - Analyse des résultats :

On a utilisé d'abord une loi à deux paramètres, celle ci donne des droites théoriques traversant un champ de points expérimentaux dont la forme est une courbe à concavité vers le bas (exemple des stations Bir Kasdali, code : 15.07.07 et Guellal, code : 15.06.03). Des écarts importants peuvent être visualisés.

Nous avons, de ce fait, essayer la loi à trois paramètres qui permet, grâce au déplacement de la borne inférieur X_0 , de rabâtrer les faibles valeurs avec les plus fortes. L'application aux échantillons cités ci-dessus, montre ce cas de figure.

Les figures en annexe : illustrent les ajustements graphiques de la droite $U = a \log (X - X_0) + b$. L'ajustement est assez bon.

II.3. Conclusion

D'après les résultats obtenus graphiquement, l'ajustement à une loi de Galton pour l'ensemble des 15 stations est de qualité nettement supérieur à celui de la loi normale, ceci reste toujours vrai en effectuant les tests d'adéquation.

Il reste cependant à signaler, que les tests d'adéquation qu'on a utilisé pour valider notre travail, sont insensibles aux anomalies de mesures. Par exemple, en comparant l'ajustement à la loi normale de la station code : 15.06.13 avec l'ajustement à la loi log normale, on remarque que la déviation par petit paquet de points correspondant aux données douteuses signalées par le service centrale (en revenant aux fichiers originaux) est visible sur les deux graphiques.

CHAPITRE III

L'ANALYSE EN COMPOSANTES PRINCIPALES A.C.P

III - 1 Introduction

L'étude des phénomènes climatiques et plus généralement des phénomènes naturels traite des situations très complexes, ou interviennent des données nombreuses.

En possession d'une masse importante de données, l'utilisateur est amené à créer des moyens de traitements efficaces et rapides, sinon comment peut-on analyser efficacement des séries chronologiques de plus d'une décennie, des précipitations, des températures ou autres ...

Parmi les techniques d'analyses factorielles qui sont des techniques d'analyses multidimensionnelles, l'analyse en composantes principales (souvent désigné par l'abréviation A.C.P.) est un outil puissant, efficace et assez objectif pour traiter des tableaux de données à caractère quantitatives et particulièrement des phénomènes physiques homogènes.

L'A.C.P. a été mise au point au début du siècle par des statisticiens (HOTELLING ET PEARSON), dépendant de la disponibilité de moyens de calculs lourds, n'est devenue d'une application courante, que grâce au développement de l'informatique.

Dans la suite de l'étude, on utilisera l'ACP (du point de vue plus récent) en tant que technique de représentation de données, ayant un caractère optimale, selon certains critères algébriques et géométrique spécifiés, et que l'on utilise en générale sans référence à des hypothèses de nature statistique.

Son domaine d'application est très vaste ; elle est très utilisée en hydrologie notamment en prévision et simulation.

III. 2 L'objectif de l'A.C.P.

Dans notre étude, l'A.C.P. permet de répondre à un double objectif :

D'abord elle nous permet de faire ressortir une station de référence, qui servira de support dans la phrase critique des données. Ensuite à faire une analyse spatiale sur la base des stations plus ou moins homogènes, c'est à dire, de voir s'il y a existence ou non de groupes homogènes de variables vis à vis de l'ensemble.

III. 3 Rappel théorique de la méthode.

L'exposé que nous allons donner de l'A.C.P., ne correspond pas à une présentation complète des principes de l'analyse, mais simplement pour aider à comprendre la signification des opérations qui permettent le passage d'un tableau de données à un ensemble d'axes et de plans factoriels, puis, pour essayer de savoir ce que représentent les points figuratifs des variables par rapport à leur proximité et à leur éloignement, et comment interpréter les axes.

III. 3.1 But et cadre de l'ACP.

Elle s'applique à des tableaux rectangulaires croisant des individus et des variables. selon un usage bien établi, les lignes du tableau représentent les individus et les colonnes représentent les variables.

la figure 1 : illustre ces notions et complète les notations.

		Variables			
		1	j	p	
Individus	1				X_{ij} : valeur de la variable j pour l'individu i
	i		x_{ij}		n : nombre d'individus et nuage des individus.
	n				p : nombre des variables et nuage des variables.

fig-1- tableau des données

les termes d'individus et de variables recouvrent des notions différentes. par exemple, dans notre cas d'étude, les individus sont les années d'observations et les variables sont les stations pluviométriques dont les éléments (X_{ij}) représentent les pluies annuelles.

Son but est donc d'obtenir une représentation du nuage de points dans un espace de dimension réduite avec un minimum de perte d'information.

Elle permet cependant une synthèse et une meilleur appréhension de l'information (volume d'observation).

Elle peut également orienter la recherche d'une typologie, aussi bien dans le domaine des variables que dans celui des individus, ceci en faisant un bilan de liaison entre variables et un bilan des ressemblances entre individus.

III. 3.2. Formulation du problème.

Considérons un "profil" qui est l'ensemble des valeurs numériques prises par les p. variables sur les n. individus.

le profil est une description qui n'est quelquefois pas très satisfaisante, en raison par exemple:

- du fait que le nombre des descripteurs (c'est à dire les variables) est très élevé, ce qui entraîne des difficultés pour en avoir une bonne vue d'ensemble (le nombre total des «données» est $n \times p$);

- ou encore du fait que la description est trop redondante, en ce sens que certaines des valeurs numériques constituant le profil individuel, pourraient « presque » se déduire des autres valeurs numériques.

D'où le souhait de substituer à ce profil initial un autre profil à la fois plus simple et moins redondant, tout en conservant le plus possible d'information.

III.3.3. Résolution du problème posé.

La solution au problème, consiste à créer à partir des variables initiales : X_1, X_2, \dots, X_p , de nouvelles variables : Z_1, Z_2, \dots, Z_r , que l'on appelle composantes principales :

- en nombre r aussi petit que possible (afin de permettre une meilleure visualisation)
- chaque composante principale est une combinaison linéaire des p variables de départ, c'est à dire, doivent tenir compte de toutes les variables X_i , donc de la forme :

$$Z = a_1 X_1 + a_2 X_2 + \dots + a_p X_p$$
- deux à deux orthogonaux, c'est à dire, non corrélées (ce qui, faute d'annuler toute interdépendance, annulera au moins toute dépendance linéaire)

NOTA :

- obtenir un nouveau profil moins redondant que le profil initial, c'est faire en sorte que ces variables artificielles Z_s ($s = 1$ à r) soient aussi peu liées que possible entre elles.
- Perdre le moins possible d'information dans cette opération, c'est obtenir le moyen de reconstituer (si possible exactement, sinon avec une bonne approximation) le profil initial à partir du nouveau.

III 3.3.1 La démarche en ACP.

A - recherche des composantes principales ou des axes principaux

La recherche du sous espace $[Z]$ de dimension plus réduite r , revient à ce que les projections X_i sur ce sous espace, soient les plus proches possibles des X_i ainsi:

- chaque Z_i sera une combinaison linéaire des X_i
- chaque X_i sera une combinaison linéaire des Z_i .
- en tant que vecteur de $[Z]$ le plus proche de X_i ; la projection \hat{X}_i sera une «reconstitution» de X_i , sous la forme recherchée d'une combinaison linéaire de r vecteurs deux à deux orthogonaux, constituants des plans principaux ou plans factoriels.

Ainsi la détermination des vecteurs Z_1, Z_2, \dots, Z_r revient à déterminer les éléments propres de la matrice M de dimension $(p \times p)$, des produits scalaire de ces variables.

B - Le choix de la matrice à diagonaliser

Dans la plupart des applications, l'analyse est pratiquée non pas sur des variables «brutes», mais sur des variables obtenues après transformation de ces variables « brutes».

on distingue deux cas :

1 - ACP de variables centrées

La transformation consiste à soustraire à chacune des valeurs observées de la variable, la valeur moyenne de celle-ci.

Elle est utilisée lorsque les variables sont mesurées dans des unités identiques et ont des variances de même ordre.

2 - ACP de variables centrées et réduites

C'est la transformation la plus courante qu'on appelle standardisation des variables, c'est-à-dire, leur centrage suivi de leur réduction.

Elle est utilisée lorsque les variables sont mesurées dans des unités différentes et plus généralement lorsqu'elles ont des variances notablement différentes.

Il s'agit donc de rechercher les composantes principales de l'ensemble $\{ 1/s_1(X_1 - m_1), 1/s_2(X_2 - m_2), \dots, 1/s_p(X_p - m_p) \}$

où m_i : la moyenne de X_i
et s_i : son écart-type

il s'en suit de ces deux cas :

- la matrice des produits scalaire des variables cas n° 1 n'est autre que la matrice des covariances des variables initiales

- la matrice des produits scalaire des variables cas n°2 n'est autre que la matrice de corrélation de variables initiales.

Ainsi la recherche des valeurs et vecteurs propres ne demande pas explicitement la standardisation ou le centrage des colonnes du tableau des données, mais le calcul direct des matrices, soit de covariances, soit de corrélation (matrice symétrique), puis de la diagonaliser, ce qui permet d'utiliser des algorithmes de diagonalisation relativement simples.

ainsi :

* la valeur propre représente la part de la variation du nuage des points expliquée par l'axe correspondant.

* Les vecteurs propres définissent les axes principaux par rapport aux variables initiales (ils fournissent des coefficients qui pondèrent les variables initiales pour le calcul des composantes principales).

Remarque :

On utilisera dans notre analyse la matrice de corrélation, ceci revient à analyser le nuage centré réduit.

Ce choix rend les résultats de l'analyse indépendant des unités de mesures choisies par les variables. Ce faisant, on utilise comme unité de mesure pour la variable son écart-type. Toutes les variables présentent alors la même variabilité, et de ce fait, la même influence dans les calculs des liaisons entre variables.

C - procédé de calcul de l'A.C.P.

Le calcul de la première composante revient à rechercher l'axe tel que la somme des carrés des différences entre valeur vraie (X_j) et valeur reconstituée (\hat{X}_j), porté par cet axe soit minimale, ou encore que la somme des carrés des projections des points sur cet axe soit maximale.

Ce procédé de calcul se répète p.fois pour le calcul et la conservation des r composantes principales, ceci revient à effectuer une rotation du système d'axe d'origine.

l'axe le plus près de tout les points au sens des moindres carrés a pour vecteur directeur un des vecteurs propres de la matrice des coefficients de corrélation et L est la valeur propre correspondante.

L'abscisse d'un point variable j sur un axe factoriel est le coefficient de corrélation entre cette variable et la composante à cet axe; c'est également le cosinus de l'angle des deux vecteurs par rapport au centre du nuage.

$$R_{jk} = a_{jk} \sqrt{L_k}$$

avec :

a_{jk} = coefficient directeur du vecteur propre
 L_k = k éme valeur propre du vecteur propre.

Ainsi, les coordonnées des points variables sur le 1er axe ne sont autres que les coefficients de corrélations entre les diverses variables et la 1er composante principale.

D - aides à l'interprétation.

L'analyse des résultats et graphiques de l'ACP est une phase délicate et demande la subtilité du chercheur et la connaissance du problème.

IL est donc nécessaire de construire des indicateurs permettant de mesurer la qualité de la représentation du nuage de points par les projections orthogonales des points- variables, dans le sous -espace de dimension réduite.

* Qualité de la représentation

$$\text{soit : } d^2(X_i, \hat{X}_i) = \sum_{i=1}^p \|X_i - \hat{X}_i\|^2 = \text{tr}M - \sum_{i=1}^p \|X_i\|^2$$

$$\text{Et } \sum_{i=1}^p \|X_i\|^2 = \text{tr}XX = \text{tr}M = \sum_{i=1}^p L_i \quad (1)$$

$$\sum_{i=1}^p \|\hat{X}_i\|^2 = L_1 + L_2 + \dots + L_r \quad (2)$$

X : matrice des données brutes de dimension (n x p)

'X : matrice transposée de X

M : matrice de corrélation

L : valeur propre

la quantité (1) mesure la variabilité globale de l'ensemble des variables initiales.

La quantité (2) mesure la variabilité globale de l'ensemble des projections des variables.

Le rapport des deux

$$\text{soit : } \sum_{i=1}^p L_i / \text{tr}M$$

est le taux de reconstitution des valeurs des variables initiales à partir des variable artificielles.

dans la pratique, ce rapport est multiplié par 100 et exprimé en %. Il peut être cumulé sur plusieurs axes, on parle alors du pourcentage d'inertie ou de variance extrait par un plan ou par les S premiers facteurs.

exemple :

$L_1 / \text{tr}M$ est la part de variance porte par le premier axe factoriel.
 $(L_1 + L_2) / \text{tr}M$ est la part de variance porte par le plan factoriel

NOTA :

-- L est classé par ordre croissant

soit : $L_1 > L_2 > L_3 \dots > L_r$

- Par construction, la trace de la matrice de corrélation est égale au nombre de variable, c'est-à-dire, p .

Ce rapport est d'autant plus proche de 100% que $\sum \|X_i - \hat{X}_i\|^2$ est faible, c'est-à-dire que les valeurs reconstituées des variables X_i sont globalement proches des vraies valeurs de ces variables.

La corrélation au carrée (cosinus au carrée de l'angle) est une autre mesure de qualité de la représentation.

$$\text{soit } \cos^2 \alpha_i = \|X_i\|^2 / \|\hat{X}_i\|^2$$

Plus cette valeur est proche de 1, ce point est bien représenté sur l'axe.

Plus cette valeur est proche de 0, la projection du point est déformante et le point est mal représenté.

De manière générale :

- La proximité entre points d'observations signifie un même comportement de ces observations aux p variables analysées. L'éloignement d'un point permet éventuellement de découvrir une anomalie.

- Plus deux points variables sont proches, plus ces variables sont corrélées et ceci est d'autant plus vrai que ces points sont éloignés de l'origine (les coordonnées étant des coefficients de corrélations).

CHAPITRE IV

**ETUDE DE
L'HOMOGENEITE
DES DONNEES
PLUVIOMETRIQUES**

IV.1. Introduction

C'est une condition fondamentale pour aboutir à des résultats fiables et représentatifs car les paramètres climatiques, comme les autres données expérimentales, comportent un risque d'erreur.

Ce risque est plus ou moins grand suivant la diversité, la multiplicité des appareils, le changement de site et du personnel.

Les tests de contrôle ont pour finalité la recherche et la détection des observations aberrantes et erronées. Elles transforment l'ensemble des données utilisées en un ensemble de données contrôlées et prêtes au traitement afin de confirmer, de corriger ou de rejeter les séries brutes. Parmi les méthodes de contrôle on a utilisé :

- le test de KENDALL
- la méthode du cumul des résidus de régression

IV.2. Le Test de KENDALL

IV.2.1. Principe du test

C'est un test non paramétrique, utilisé pour mettre en évidence l'existence éventuelle d'une tendance significative de la série d'observations et de situer l'époque ou l'année à partir de laquelle cette tendance s'est manifestée. Il est basé sur la statistique de rang «t» de KENDALL.

Le test se déroule comme suit :

- 1 - Classement des observations par ordre chronologique
- 2 - On affecte à chaque valeur observée X_i le rang Y_i qu'elle occupe lorsque la série est rangée par ordre de grandeur croissant.
- 3 - On calcule pour chaque Y_i le nombre n_i d'éléments Y_j qui le précèdent tels que :

pour $i > j$ on a $Y_i > Y_j$
 en commençant de $i = 1$ à N

avec N : nombre d'observations

- 4 - On calcule le coefficient de KENDALL «t» tel que :

$$t_i = \sum_{i=1}^N n_i \quad (\text{cumul de } i=1 \text{ à } N)$$

La loi de répartition de cette statistique «t» suit asymptotiquement une loi normale de moyenne et de variance :

$$E(t) = N(N-1)/4 \quad \text{et} \quad \text{Var}(t) = N(N-1)(2N+5)/72$$

5 - On calcul les variables réduites $U(t_i)$: série progressive de la statistique du test de tendance

$$\text{pour } i = 1 \text{ à } N \quad U(t_i) = (t_i - E(t)) / \sqrt{\text{var}(t)}$$

6 - On en déduit la série rétrograde par la relation :

$$U'(t_i) = -U(t_i)$$

* **Interprétation du test :**

En absence de toute tendance dans la série d'observations la représentation graphique de U_i et de U'_i en fonction de i , donne généralement des courbes qui s'enchevêtrent par contre dans le cas d'une tendance significative, l'intersection de ces courbes permet de situer approximativement le début du phénomène.

IV.2.2. Analyse des résultats du test

Le test a été effectué sur la totalité des stations à étudier, dans ce qui suit, on présente l'interprétation des résultats obtenus de quelques stations.

Un exemple d'une feuille de calcul est donné en annexe.

Exemples :

* Station : Bendaoud
Code : 15.04.02

Le test de KENDALL a mis en évidence une tendance significative . Après retour aux originaux, la consultation du fichier de cette station, nous a permis de savoir que durant la période de 1970 à 1978, il y a eu divers observateurs et à partir de 1980, les observations sont douteuses en les comparant avec les stations avoisinantes, ce qui relève le non sérieux de l'observateur.

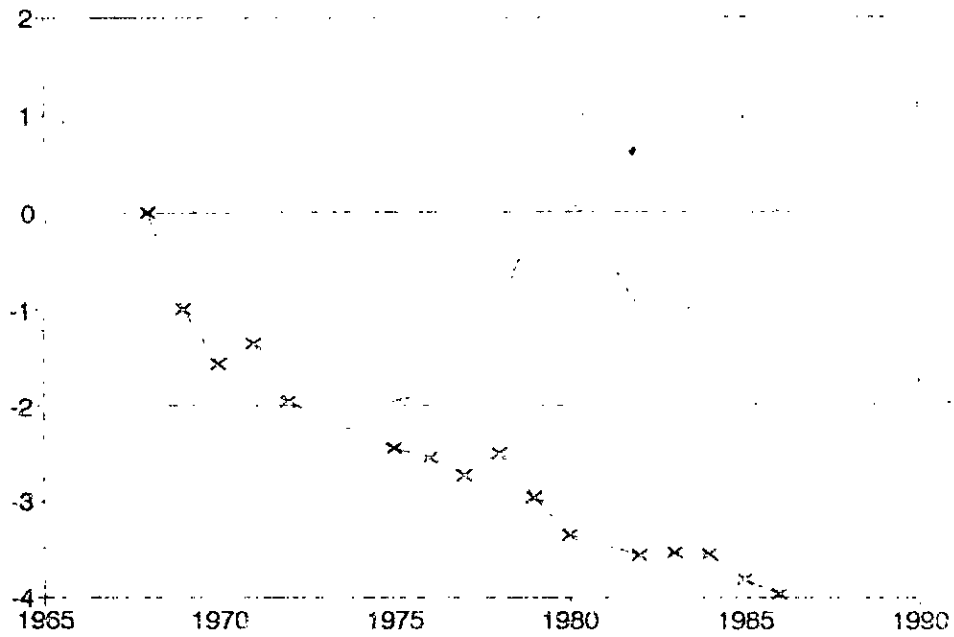
Tout ceci peut être la cause de cette hétérogénéité dans la série.

* Station : Tazmalt
Code : 15.05.01

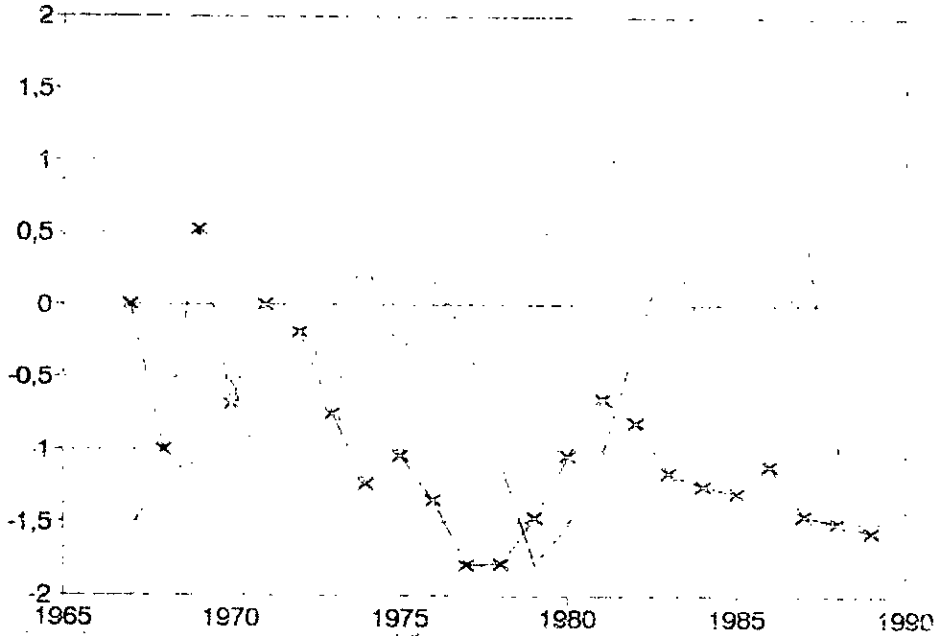
D'après le test, cette station présente une tendance à partir de 83; l'original de cette station a montré un changement dans les modalités de relevé (un relevé par jour au lieu de deux) donc ça peut sous-estimé le total de la pluie dans la mesure où il y a risque d'évaporation qui influe sur la valeur de la mesure. Ceci, peut-être la cause de cette tendance.

TEST DE KENDAL

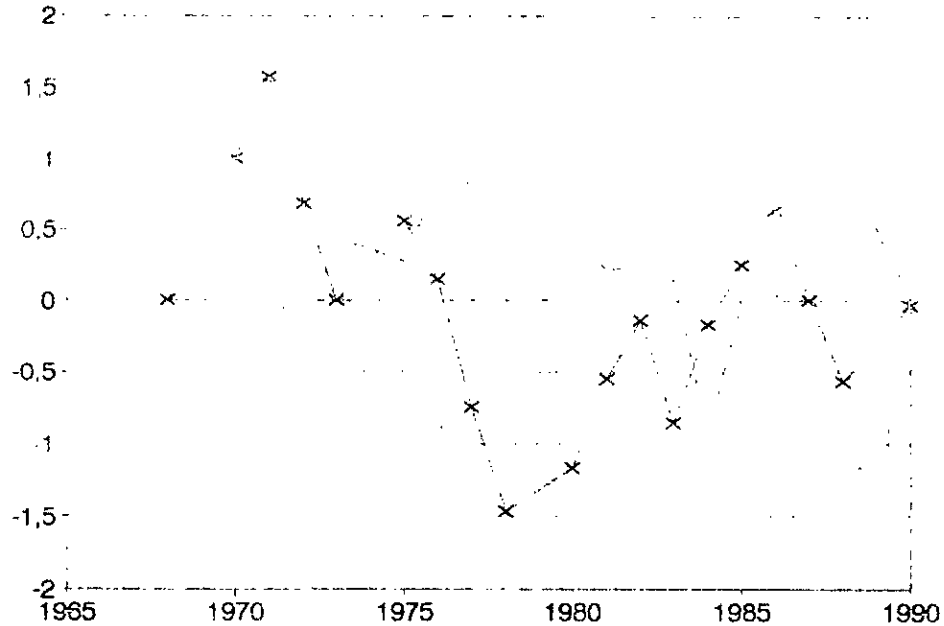
STATION : BEN DAOUD CODE:150402



Station: Tazmalt Code:150501



Station: Taghdit Code:150201



*Station : Taghdit
Code : 15.02.01

D'après le test, cette station ne présente pas de tendance, on dira qu'elle est homogène avec un risque d'erreur de 5%.

IV.2.3. Conclusion

Le test statistique de KENDALL a mis en évidence des hétérogénéités pour la plupart des stations, pour certaines l'hétérogénéité est due à la mauvaise dispersion des valeurs observées par rapport à la pluie moyenne annuelle de la station. Ce problème de dispersion se pose pour des séries courtes où l'on passe d'un cycle sec à un cycle humide.

En général, ce test de tendance permet de vérifier l'homogénéité des séries pluviométriques, seulement, il présente des difficultés qui résident dans le fait que les séries sont discontinues et possède beaucoup de lacunes. C'est pourquoi il semble nécessaire d'utiliser la technique de l'analyse en composante principale (ACP) pour vérifier l'homogénéité des séries d'observations par rapport à une station représentative de la région d'étude.

IV.3. La méthode du cumul des résidus de régression.

IV.3.1. Introduction

Cette méthode beaucoup plus puissante que la méthode des doubles masses (ou double cumul), nécessite cependant des calculs un peu plus longs; Elle a moins pour objectif de détecter une anomalie importante mais isolée, que d'analyser des tendances ou dérivé à long terme.

Elle repose pour l'essentiel sur une idée de cohérence spatiale entre stations voisines : une station ne saurait avoir un comportement très différent de ses voisines, à partir d'une période donnée sauf pour des raisons explicites.

Elle est donc utilisée pour détecter des anomalies systématiques ou accidentelles conséquence de l'hétérogénéité.

La méthode a été longuement testée en simulation pour ph. Bois (1976) sur les différents types d'erreurs courantes dans les données hydrométéorologiques puis appliquée à des cas concrets et implantée opérationnellement dans des services.

Elle semble cependant sous-utilisée, compte tenu des possibilités graphiques désormais disponibles.

Un problème se pose parfois de définir une station témoin; la moyenne arithmétique des stations peut être prise, cependant une station très erronée peut y être incluse et biaise la référence.

Pour cela, on a l'avantage d'utiliser la technique de l'analyse en composante principale pour former la station témoin; une station présentant éventuellement des valeurs erronées à toute la chance d'être filtrée des anomalies de mesures dans l'analyse des cumuls des résidus par rapport à la référence.

Elle sera confrontée avec chacune des stations de la zone pour vérifier leur homogénéité temporelle.

Cependant on se retrouve confronté à un problème étroitement lié à celui de la critique des données et l'estimation des valeurs manquantes car l'A.C.P qui est un moyen de critique ne peut être utilisé que sur une matrice de données complètes, sans aucun manque.

Vu l'irrégularité des séries d'observations sur le réseau de la Soummam, nous amène au départ à combler les lacunes.

IV.3.2. Principe du cumul des résidus de régression

Il s'agit d'une extension et d'une formulation de l'idée des doubles cumuls. Elle correspond à l'étude du cumul dans le temps des résidus (écart entre valeurs calculées et valeurs observées) d'une corrélation entre station à contrôler et une station de référence.

Les résultats des développements théoriques qui ont été menés par Ph.Bois sont :
soit deux séries d'observations :

$$\left. \begin{array}{l} - Y_i \text{ (série à contrôler)} \\ - X_i \text{ (série témoin)} \end{array} \right\} \quad i = 1 \text{ à } n \quad , \quad n: \text{ nombre d'année commune}$$

La série témoin et celle de la station à critiquer étant supposée suivre la loi normale, de moyennes respectivement X et Y et d'écart-type σ_x et σ_y , sont liées par une relation linéaire de coefficient de corrélation "r".

Tels que :

$$\begin{cases} \hat{Y}_i = a X_i + b \\ Y_i - \hat{Y}_i = \xi_i \end{cases}$$

Avec :

- \hat{Y}_i : série calculée à partir de la station témoin
- a, b : constantes de régression
- ξ_i : erreur résiduelle ou écart à la droite de régression

où ξ_i est une variable gaussienne de moyenne nulle et d'écart type σ_ξ tel que:

$$\sigma_\xi = \sigma_y \sqrt{1 - r^2}$$

Le cumul des i premiers résidus sur un total possible de n est noté Z_i :

$$Z_i = \sum_{j=1}^i \xi_j = \xi_1 + \xi_2 + \dots + \xi_i$$

où Z_i est une variable aléatoire de moyenne nulle et d'écart type :

$$\sigma_{z_i} = \sigma_{\xi} \sqrt{i(n-i)/(n-1)}$$

On montre que la ligne brisée (chemin suivi par les cumuls des résidus pour chaque i , c'est à dire, année) doit être incluse (avec une probabilité de non dépassement) dans une ellipse dite «ellipse de bois» de grand axe (n) et de petit axe ($\pm U_{\alpha}, \sigma_{z_i}$).

Ainsi pour $\alpha=5\%$, et $U_{\alpha} = 1,96$

- Il y a 95% de chance que l'écart à la droite de régression entre une variable et la station témoin soit compris entre :

$$-1,96 \sigma_{\xi} < \xi_i < 1,96 \sigma_{\xi}$$

on détecte ainsi des anomalies accidentelles qui ont moins de 5% de chance d'être dues au hasard.

- pour ce qui concerne les cumuls des résidus, il y a 95% de chance pour avoir :

$$-1,96 \sigma_{z_i} < Z_i < 1,96 \sigma_{z_i}$$

on détecte ainsi des anomalies systématiques qui ont moins de 5% de chance d'être dues au hasard.

IV.3.3. Procédé de restitution

A l'échelle annuelle, l'absence d'observations pendant quelques mois fait que le total annuel réel est inconnu. Deux démarches sont alors possibles pour reconstituer celui-ci.

- Reconstituer pour chaque mois manquant le total mensuel probable par sa moyenne mensuelle interannuelle, puis faire la somme des douze pluies mensuelles réellement observées ou reconstituées.

- Reconstituer pour l'ensemble de l'année le total probable par régression à l'échelle annuelle.

Le choix entre ces deux méthodes dépend essentiellement du nombre de mois manquants et on a pensé raisonnable de n'appliquer la première méthode que lorsqu'il n'y a pas plus de trois mois manquants.

A - Comblement à l'échelle mensuelle

Dans ce cas la démarche «normale» aurait été de combler chaque lacune mensuelle par régression linéaire avec la station ayant fonctionné ce mois et présentant le meilleur coefficient de corrélation pour le mois considéré.

Cette démarche est quasiment inutilisable devant le nombre de corrélations à envisager (23616) et la multiplicité des tests pour déterminer la meilleure solution (valeur intrinsèque du coefficient de corrélation et effectif commun).

On a donc choisi une démarche beaucoup plus simple qu'on pense être suffisante pour combler sans erreur majeure un total annuel connu par au moins 9 mois sur 12, elle consiste à remplacer les lacunes mensuelles par la moyenne mensuelle interannuelle du mois considéré.

Le seuil du nombre de mois à remplacer a été fixé, au maximum et à priori, à trois par année incomplète, représente 25% du total annuel (3 mois sur 12), peut être considéré comme acceptable car il n'influe que peu sur le total.

La méthode envisagée a été au préalable testée sur des stations prises au hasard auxquelles on a ôté trois mois sur des années réellement observées et qu'on a substitué par la moyenne mensuelle interannuelle de chaque mois considéré, les résultats étaient conforme avec une erreur de 5% sur le total réel. Et qu'au delà de 3 mois, on surestime ce total réel.

L'exemple de la station de Ighil Ali, montre que les années 1944, 1947, 1948, 1951, 1968, 1970, 1987 et 1990 sont susceptible d'être reconstituées ainsi.

B - Comblement des lacunes à l'échelle annuelle

En l'absence d'observations complètes pour une année donnée, ou pour des lacunes de plus de 3 mois, il n'était pas raisonnable de procéder ainsi. C'est donc classiquement par régression linéaire qu'on a reconstitué des années manquantes.

Pour chaque station on a donc étudié les corrélations avec les 41 stations et noté les coefficients de corrélation ainsi que l'effectif commun.

 * EXEMPLE DE DONNÉES DE BASE POUR UNE STATION PUVIOMETRIQUE *

** NOM: IGHIL ALI CODE: 150503 **

	sept.	oct.	nov.	dec.	jan.	fev.	mars	avril	mai	juin	juil	août		
MOYENNE:	2784	3729	4374	6856	5806	5467	5211	4454	3832	1312	455	1018		
Année	sept.	oct.	nov.	dec.	jan.	fev.	mars	avril	mai	juin	juill	août	Total	Nb
1923	149	389	505	1143	769	539	555	384	104	191	44	4	4860	
1924	19	688	428	671	15	224	715	306	295	274	35	84	3754	
1925	985	558	776	390	192	441	360	413	183	42	98	40	4478	
1926	59	78	209	663	1346	402	396	93	758	21	10	0	4035	
1927	139	218	1024	601	911	1204	979	205	415	2	507	100	6305	
1928	220	748	771	661	461	1340	659	135	911	8	5	27	5946	
1929	980	285	711	343	479	454	233	652	287	191	11	0	4826	
1930	122	318	10	1857	554	1044	410	419	278	64	0	6	5082	
1931	200	150	341	1218	555	687	494	168	199	68	92	0	4122	
1932	150	485	218	646	574	303	621	178	488	253	148	78	4137	
1933	625	50	862	844	1394	387	741	615	434	43	0	1376	7371	
1934	212	152	645	532	919	241	1128	664	747	138	78	215	5666	
1935	0	949	435	476	327	576	481	351	804	168	0	20	4587	
1936	62	420	496	562	25	381	614	405	302	45	0	405	3717	
1940	88	707	342	864	545	738	254	428	367	39	71	0	4444	
1941	822	188	425	177	1089	1567	457	220	288	0	0	320	5553	
1942	697	80	487	853	110	230	593	549	355	69	32	12	4067	
1943	75	424	351	494	65	260	508	135	10	244	0	32	2598	
1944	958	435	47	720	1415	20	56	50	92	105	-10	0	-10	-1
1945	180	190	487	632	570	98	395	1595	735	0	0	0	4802	
1946	170	30	722	1578	284	791	77	165	125	0	19	200	4161	
1947	147	-10	-10	1225	1140	324	388	291	374	335	18	43	-10	-2
1948	33	706	62	204	730	618	832	358	891	-10	-10	-10	-10	-3
1949	193	11	461	556	501	284	642	251	606	19	28	85	3635	
1950	298	497	48	848	1540	439	405	363	368	148	47	87	5088	
1951	379	1204	410	530	860	1128	386	894	-10	5	-10	-10	-10	-3
1952	-10	174	281	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10
1968	-10	-10	537	731	394	340	822	813	150	332	57	29	-10	-2
1969	667	1331	746	2945	715	130	485	710	405	58	20	0	8212	
1970	25	748	135	620	810	40	1010	325	272	0	-10	123	-10	-1
1971	180	130	695	745	1580	510	615	913	745	331	0	0	6444	
1972	245	374	205	762	1096	997	619	331	0	312	0	0	4941	
1973	-10	35	58	539	107	1108	1134	816	170	324	0	0	-10	-1
1974	102	288	124	173	130	625	827	133	601	174	0	72	3249	
1975	223	0	443	452	84	1223	461	889	644	191	130	85	4825	
1976	214	961	259	605	528	102	12	675	272	38	17	0	3683	
1977	7	110	674	46	532	300	502	484	436	90	0	0	3181	
1978	0	443	718	147	713	250	312	657	361	81	0	84	3766	
1979	536	304	426	402	338	279	283	578	359	0	0	0	3505	
1980	103	180	610	1417	156	598	511	613	185	50	0	140	4563	
1981	108	275	85	638	1300	775	385	240	790	0	75	10	4681	
1982	150	1345	850	450	20	425	180	315	150	0	111	220	4216	
1983	0	65	165	140	605	890	290	500	190	275	0	65	3185	
1984	55	365	260	565	560	220	1305	75	465	40	0	0	3910	
1985	315	320	560	155	290	525	1275	45	65	350	0	0	3900	
1986	355	285	595	1555	435	1165	440	145	270	85	35	30	5395	
1987	-10	453	752	153	240	240	640	505	325	120	25	0	-10	-1
1988	240	20	530	1130	125	120	165	995	125	225	50	80	3805	
1989	505	220	360	115	410	0	285	365	965	205	300	30	3740	
1990	220	120	355	700	175	940	910	180	290	100	112	-10	-10	-1

N.B:*(Nb)donné le nombre de mois manquant pour chaque année
 *le nombre(-10) correspond a une valeur non observé

A titre d'exemple, on trouvera ci-dessous, un extrait du tableau auquel on aboutit après avoir rangé les stations explicatives par ordre décroissant des coefficients de corrélation. (voir matrice des corrélations en annexe):

SOUR EL GHOZLANE CODE 15.01.01	CODE	15.01.03	15.02.04	15.05.03	15.06.08
	r	0.92	0.89	0.82	0.72
	k	17	29	32	38
	E	0.954	0.974	0.980	0.987
	a'	0.172	0.222	0.346	0.502
SIDI ATTAF CODE 15.01.04	CODE	15.03.01	15.04.02	15.01.11	15.05.01
	r	0.90	0.73	0.81	0.68
	k	9	12	13	17
	E	0.922	0.963	0.955	0.976
	a'	0.240	0.539	0.396	0.589
IGHIL ALI CODE 15.05.03	CODE	15.01.09	15.05.01	15.06.08	15.09.02
	r	0.95	0.87	0.63	0.78
	k	11	18	36	12
	E	0.926	0.961	0.989	0.956
	a'	0.118	0.270	0.628	0.455

La valeur du coefficient de corrélation n'est pas le seul critère de choix de la variable la plus explicative, en effet l'effectif commun joue un rôle dans l'efficacité de la corrélation. On sait que l'efficacité E d'une corrélation est donnée par :

$$E = 1 + (1 - k/n) (1 - (k-2)/r^2) / (k-3)$$

(k effectif commun, n effectif de la série explicative et r coefficient de corrélation).

Et que la taille équivalente n' d'un échantillon comportant k valeurs réellement observées et (n-k) valeurs reconstituées est donnée par :

$$n' = k/E$$

Si l'on pose $n = k+1$, la valeur $a' = n - n'$ est un indicateur de la confiance que l'on peut accorder à la première année reconstituée.

D'après le tableau précédent on constate par exemple que le meilleur coefficient de corrélation entre la station de Sour-El-Ghozlane et toutes les autres stations est de 0.92 (avec El-Hachimia code 15.01.03), mais qu'il est préférable de reconstituer cette station à partir de celle de Bouira (code 15.06.08) puisque r passe de 0.172 à 0.502 du fait de l'augmentation d'effectif de 17 à 38.

C'est ainsi qu'on a reconstitué les données manquantes au niveau de deux stations en prenant comme limites de ne pas reconstituer des années où r était inférieur à 0.7 ou k inférieur à 20, mais seulement quand pour une même année une observation a été faite sur au moins une des deux stations. Dans le cas contraire, on cherchera une autre station et on vérifie les mêmes critères.

Sachant que les séries des stations suivent globalement une loi de Gauss, la formule de restitution est donnée par la droite de régression des moindres carrés entre deux stations.

La droite de régression $Y = AX + B$ a pour expression :

$$Y = (S_y/S_x) X + (\bar{Y} - (S_y/S_x) r \bar{X})$$

Les paramètres A et B de la droite sont regroupés dans des matrices relatives aux stations prises deux à deux. (voir annexe)

Les variances et covariances des séries de X et de Y s'expriment :

$$S_x^2 = \sum_{i=1}^n ((X - \bar{X})^2)/n \quad , \quad S_y^2 = \sum_{i=1}^n ((Y - \bar{Y})^2)/n$$

$$S_{xy} = \sum_{i=1}^n ((X - \bar{X})(Y - \bar{Y}))/n$$

Le coefficient de corrélation r est :

$$r = S_{xy} / S_x.S_y$$

IV.3.4. Période d'étude

Prévoyant de traiter les données par ACP sur la matrice des coefficients de corrélation, il nous était nécessaire de travailler sur un ensemble de stations disposant d'une période commune d'observations.

Après comblement nous nous trouvons alors devant un choix difficile entre période longue et très représentative, mais avec peu de stations utilisables ou une période plus courte, moins représentative, mais permettant d'intégrer à l'étude l'ensemble des stations (voir tableau n° I.2).

Les normes de l'OMM (Organisation Météorologique Mondiale), préconise pour les études hydrologiques, des périodes de 30 ans, compte tenu des événements liés à l'indépendance de l'Algérie, il n'y a que très peu de stations ayant fonctionnées correctement de 1961 à 1966. Nous sommes donc contraint de retenir une période d'étude allant de 1967 à 1990.

Sur cette période, les stations disposent d'une série complète ou complétée, globalement sur un total de 984 station-année, 784 ont été réellement observées (79,7%) et (20,3 %) non observées, ont été complétées en partie par reconstitution mensuelles et le reste par reconstitution à l'échelle annuelle.

IV.3.5. Recherche de la station de référence

Après restitution des mesures manquantes, on obtient un tableau complet des observations sur la période d'étude pour l'ensemble des stations. A ce niveau, on a effectué une analyse en composantes principales sur une matrice de (41x24) c'est à dire la série de 41 stations sur 24 années d'observation, les cinq composantes ont été évaluées.

Composantes N°	C1	C2	C3	C4	C5
Valeurs propres	17,37	3,72	2,77	2,15	1,88
Variances expliquées	46,9%	10,1%	7,5%	5,8%	5,1%
Variances cumulées	46,9%	57%	64,5%	70,3%	75,4%

Si nous analysons les résultats de l'A.C.P, on constate qu'une bonne part de la variance (57%) est accaparée par les deux seules premières composantes et qu'à partir de la 3ème, la part expliquée devient faible.

Nous pouvons penser que seules les deux composantes traduisent le comportement naturel des précipitations et qu'au delà nous avons affaire à un «bruit» imputable aux anomalies de mesure.

La station de référence en question est donc le résultat de la projection des observations (ou individus) sur C1 et C2, telles quelles figurent dans le tableau (IV.1), résumant la part essentielle des fluctuations naturelles des précipitations sur la zone d'étude, on considère donc que les observations en un poste quelconque ne dépendent que des valeurs prise par C1 et C2.

En supposant que les distributions de C1 et C2 sont normales, nous sommes donc dans les conditions où l'on peut préjuger que les régressions sont linéaires.

On a ensuite évalué les paramètres de la régression multiples entre la variable X et les deux premières composantes C1 et C2 d'après les n observations réellement disponibles.

Ainsi pour chaque observation X, on peut retrouver son estimation X_e à partir de ces deux composantes :

$$X_e = a C1 + b C2 + c$$

IV.3.6. Analyse des résidus de régression

Pour détecter les anomalies systématiques ou accidentelles, on a procédé à l'analyse des résidus de régression déjà décrite, mais en considérant que la variable X était la variable à tester et X_e la variable de référence.

1 - Les anomalies accidentelles détectées

Ce travail a été fait sur toutes les stations et on trouvera ci-dessous quelques exemples. Dans l'ensemble on pourra retenir les faits marquants suivants :

* Les anomalies accidentelles détectées sont parfois le fait de phénomènes locaux et exceptionnels.

A titre d'exemple le total annuel observé de Sétif INRA en 1969 est de 404.3 mm alors que la valeur la plus probable donnée par la station de référence est de 281 mm. En revenant aux originaux, on constate qu'en fait, en septembre il est tombé en un jour 50 mm suivi d'une autre pluie de 30mm, et cet orage ou ces orages n'ont pas affecté les autres stations. De même en 1980-81.

2 - Les anomalies systématiques détectées

De la même façon, nous avons détecté toutes les anomalies systématiques n'ayant pas plus de 5% de chance d'être dues au hasard.

On trouve ainsi des anomalies que l'on peut parfois expliquer.

Par exemple :

1 - Le graphique de la station Bendaoud montre une rupture vers 1968.

En se reportant à l'originale de la station, on constate qu'en effet la période de 1968 à 1970, il y a eu divers changements :

En 1968 :

- changement d'altitude de la station de 650 m à 670 m
- changement d'observateur
- changement de type de pluviomètre (déculpeur en ville de Paris)

En 1969 :

- relevés sur originaux sont raturés, arrondis au mm.

En 1970 :

- changement de type pluviomètre (Ville de Paris en Association)

A partir de 1979, les relevés sont douteux et présentent de nombreuses lacunes ou cumulés. Cette série pluviométrique est quasiment irrécupérable, nous l'avons éliminé.

2 - Le graphique de la station de Bir Kasdali signale une anomalie en 1987.

Le total annuel observé est 157,8mm alors que la valeur la plus probable est de 295,8mm.

En revenant à l'original de la station, on constate que le mois de janvier de l'année en question est douteux (1,5mm).

On a corrigé ce mois, en le remplaçant par la moyenne mensuelle interannuelle du mois de janvier.

On a trouvé (40,5mm); ensuite on a recalculé le total annuel, il devient, de 196,8mm (résultat plus proche de la valeur probable).

IV.4 Conclusion

L'analyse des résidus de régression faite sur l'ensemble des stations, nous a permis d'éliminer quatre stations dûe à de nombreuses valeurs douteuses et entachées d'erreurs, suite au changement fréquent des observateurs.

Les quatre stations éliminées en questions, dont on a pas pu corriger les valeurs sont :

- Beni-Hammed code : 15.03.06
- Bendaoud code : 15.04.02
- Mahouane code : 15.06.11
- Magraoua code : 15.07.02

Le test de Kendall effectué au préalable sur ces quatre stations a fait ressortir une hétérogénéité générale.

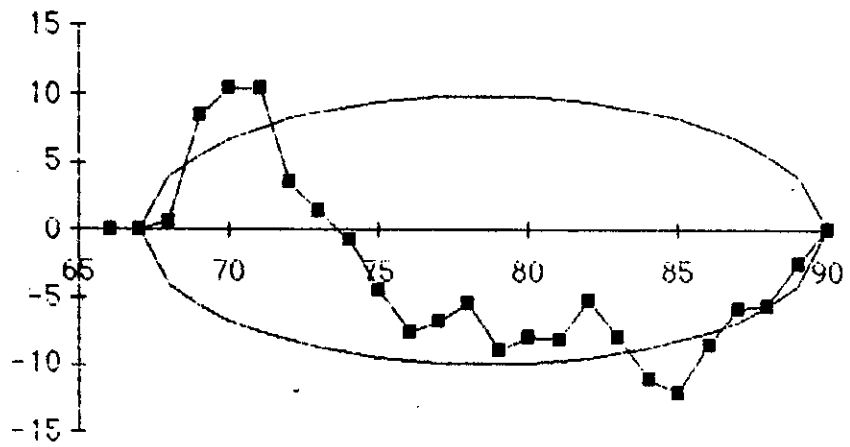
A ce niveau, on a effectué une analyse en composantes principales sur une matrice de données (37x24), c'est à dire, la série de 37 stations retenues sur 24 années d'observations, les quatre composantes principales ont été évaluées (voir chap. suivant).

*** TABLEAU (NIM.1) ***

PROJECTIONS DES OBSERVATIONS (24 années) SUR LES COMPOSANTES

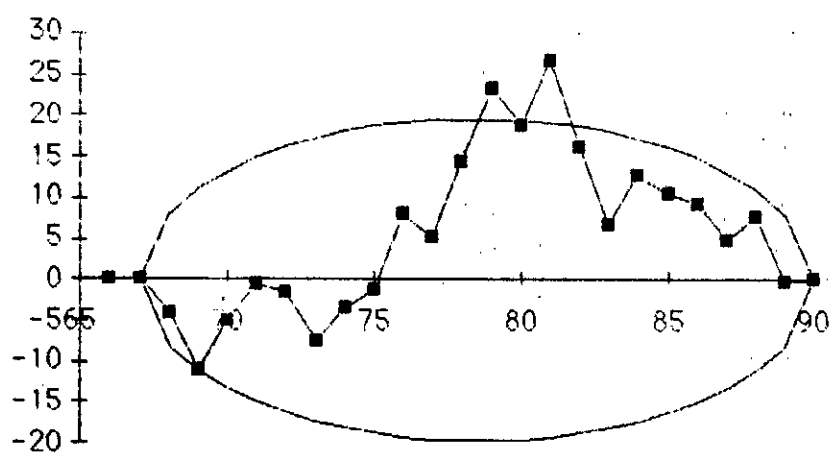
ANNÉES	COMPOSANTES PRINCIPALES			
	C1	C2	C3	C4
1967	,8942025	2,29027	,8693382	,3889004
1968	,3932171	1,668805E-03	-3,540433	2,005241
1969	2,34864	-,2158149	,3881138	-,4283772
1970	-,2715034	-2,087916	-1,088378	-,5365685
1971	2,116471	-1,355898	-,7186717	-1,467579
1972	1,579527	-,5985261	1,358065	1,578907
1973	,67158	-1,47078	1,374774	,661517
1974	-,8323025	-1,145332	,2324946	-,9069186
1975	,8611712	1,002741	4,520234E-02	-1,31293
1976	-,3101637	-6,441818E-02	,7118394	1,188559
1977	-1,360694	-,6900327	1,073826	-5,267728E-02
1978	-,7822604	-,4909538	,7222095	,7241639
1979	-,4241973	-,3601435	-,2599887	1,069522
1980	7,978828E-02	-9,230267E-02	,1652422	,1871125
1981	,195033	1,388896	-,4418778	-,6629798
1982	-,5514085	-2,247072E-02	-,1166791	-1,473248
1983	-1,00839	,4409243	,282519	-1,014287
1984	-,1789531	1,131711	,2078853	-,477117
1985	-,4056187	,256977	-,7857233	-,2230168
1986	,4211177	,9650695	-,1972682	-,1629152
1987	-1,043932	-,1659251	-,2742162	-,8531508
1988	-,6743159	,4329191	,5768023	1,718616
1989	-1,087072	-,221588	4,205584E-02	-5,906207E-02
1990	-,628938	1,07073	-,6271341	,0971879

CUMUL DES RESIDUS DE REGRESSION



STATION : BENDAOUY / CODE : 150402

CUMUL DES RESIDUS DE REGRESSION

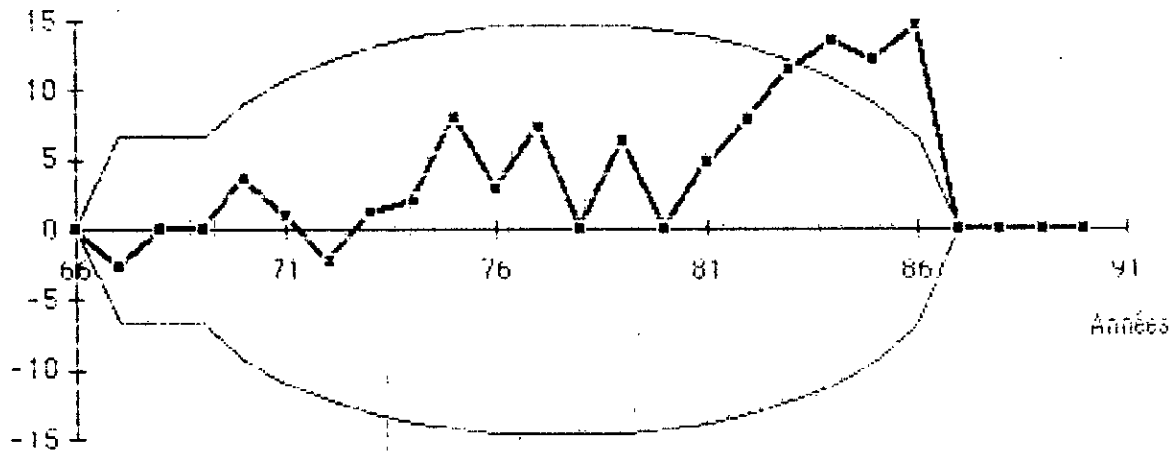


STATION : SETIF INRA / CODE 150612

CUMUL DES RESIDUS DE REGRESSION

Résidus de régression

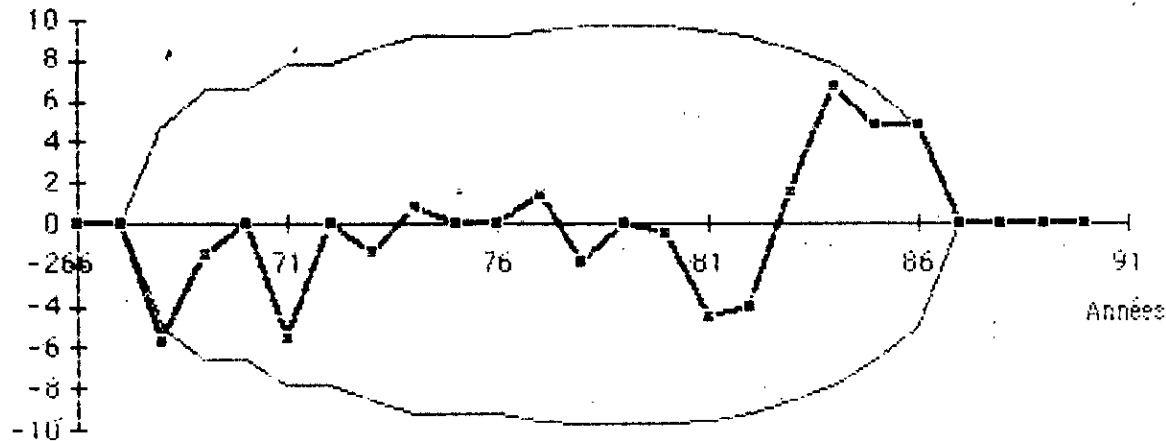
Station : Bir - Kasdali Code : 150707



CUMUL DES RESIDUS DE REGRESSION

Résidus de régression

Station : One Rouhs Messaoud Code : 151005



CHAPITRE V

**ANALYSE DE LA
STRUCTURE SPATIALE
DES POSTES
PLUVIOMETRIQUES
PAR A.C.P**

I.1- Introduction

l'intérêt de l'A.C.P. réside dans le fait qu'elle détermine des variables explicatives et qu'elle autorise de nombreuses interprétations à partir d'un nombre faible de résultats.

En effet, l'A.C.P. obtenue après critique des données, nous donne deux types de résultats :

- Pour chaque station, elle fournit la valeur de la projection des variables sur les composantes principales par rapport à chacune des 4 composantes voir tableau n° V.2

- nous obtenons également les valeurs propres, le pourcentage de la variance expliquée et le pourcentage cumulé pour les 4 composantes voir tableau N° V 1

Ainsi les résultats sur le pourcentage de la variance expliquée vont nous montrer le poids de chacune des composantes et nous aident à choisir les composantes principales ayant un sens et par conséquent limitent le nombre de variables ; tandis que les valeurs des projections nous permettent de regrouper les stations pluviométriques dans l'espace.

Par la suite, on va effectuer un classement par ordre croissant des valeurs prises par chaque composante principale. Ainsi, la valeur de la projection des variables nous permet de juger de la fiabilité des stations selon quelle est forte (station fiable) ou faible (station suspecte).

A partir de ces résultats, une série de graphiques et de cartes est construite. Il en existe deux types :

- les cercles des corrélations avec les projections des variables sur les composantes principales qui peuvent montrer des regroupements entre variables.

- les cartes qui permettent de visualiser les stations pluviométriques en fonction de leur coefficient de corrélation avec les composantes principales, exécutées sur la carte de la zone d'étude.

V.2- Analyses et interprétations des configurations spatiales révélées par les axes principaux (ou composantes principales).

V.2.1 Le comportement spatial de la première composante

Le tableau n° V.1, montre l'importance de la première composante principale qui représente 59% de la variance et au total, les quatre premiers facteurs cumulent un pourcentage de 85,79% de la variance totale du nuage. On a donc que 14% de perte d'informations.

Les pourcentages de la variance expliquée deviennent faibles à partir de la quatrième composante principale (5%), cependant, il semble que seule les trois premières composantes aient un réel intérêt (exprimant au total un peu plus de 80% de la variance totale).

TABLEAU (N V.1)

! Valeurs Propres (variances sur les axes principaux) :				
	1	2	3	4
	14,19698	2,152225	1,882675	0,806671
! Variance Expliquée (pourcentages expliqués par les axes principaux) :				
	1	2	3	4
	,5899138	,1434302	,0742585	,0502851
! Variances Cumulées :				
	1	2	3	4
	,5899138	,733344	,8076025	,8578875

TABLEAU (N V.2)

! VARIABLES !	! COMPOSANTES PRINCIPALES !									
	! AXE 1 !		! AXE 2 !		! AXE 3 !		! AXE 4 !		! AXE 5 !	
! st0101 !	,7547	,5895	-,3584	,1284	,2264	,0512	-,1201	,0144	-,0341	,0012
! st0103 !	,6349	,4031	-,3739	,1398	-,1772	,0314	,1716	,0294	-,2411	,0581
! st0104 !	,7843	,6152	-,2076	,0430	,2422	,0586	,0383	,0015	-,4523	,0246
! st0108 !	,8527	,7271	-,1541	,0237	,0322	,0010	-,3375	,1139	,0575	,0033
! st0109 !	,8506	,7234	-,0796	,0063	,0504	,0025	,0108	,0001	,0129	,0002
! st0110 !	,7025	,4935	-,3534	,1249	,1613	,0260	,0123	,0002	,5094	,2595
! st0111 !	,7439	,5534	-,3216	,1034	,1293	,0167	-,3030	,0918	-,4591	,2108
! st0112 !	,7330	,5373	-,0629	,0039	,3530	,1246	,1049	,0110	,0055	,0000
! st0201 !	,6815	,4645	,1733	,0300	-,1254	,0157	-,0317	,0010	,3997	,1597
! st0202 !	,5915	,3499	-,4703	,2111	-,1820	,0331	-,3527	,1244	,0788	,0062
! st0203 !	,5454	,2975	-,6147	,3778	-,0480	,0023	-,2868	,0823	-,3407	,1161
! st0204 !	,8813	,7767	,0818	,0066	-,1870	,0349	-,1778	,0316	,2603	,0677
! st0301 !	,7546	,5694	-,4488	,2014	,0353	,0012	,2368	,0561	,1326	,0176
! st0401 !	,6985	,4879	-,0704	,0049	-,0778	,0060	-,0257	,0007	,3119	,0973
! st0501 !	,8556	,7320	,2938	,0863	-,0240	,0005	,0011	,0000	,2866	,0821
! st0503 !	,9136	,8347	,0619	,0038	-,1144	,0130	-,2596	,0674	,0707	,0050
! st0603 !	,5600	,3136	,1901	,0361	-,5871	,3448	-,2228	,0496	-,0892	,0080
! st0607 !	,4355	,1897	,6322	,3996	-,5375	,2889	,2232	,0498	,0672	,0045
! st0608 !	,5391	,2906	,6144	,3775	-,3308	,1094	,2683	,0720	,0197	,0004
! st0609 !	,8358	,6986	-,1688	,0284	,0995	,0099	,0926	,0086	,1319	,0174
! st0610 !	,6399	,4094	,5340	,2850	-,5760	,3317	,2202	,0485	,0912	,0083
! st0612 !	,4727	,2235	,1643	,0269	-,1880	,0353	,6056	,3667	-,1266	,0160
! st0613 !	,5309	,2818	,5508	,3033	-,3972	,1577	-,2373	,0563	,1609	,0259
! st0614 !	,5087	,2588	,2625	,0689	-,3529	,1245	-,4148	,1721	,0816	,0067
! st0703 !	,7463	,5569	,2484	,0617	-,1930	,0372	,3623	,1313	-,1624	,0264
! st0706 !	,3947	,1558	,5841	,3410	-,6544	,4282	,3413	,1165	-,1396	,0195
! st0707 !	,5252	,2758	,4295	,1844	-,3593	,1291	,0125	,0002	,1588	,0252
! st0708 !	,8537	,7288	,0111	,0001	,2860	,0818	,3633	,1320	,0642	,0041
! st0801 !	,6541	,4279	,3087	,9536	-,2830	,0800	-,1539	,0237	-,0281	,0008
! st0802 !	,6907	,4771	,3761	,1415	-,0207	,0004	,2283	,0521	-,3381	,1143
! st0807 !	,4628	,2142	-,1612	,0259	-,0930	,0086	,4024	,1619	,3190	,1017
! st0902 !	,7991	,6386	-,4234	,1793	,0495	,0025	-,2689	,0723	,0442	,0020
! st1001 !	,8430	,7106	-,0624	,0038	,3787	,1434	,0381	,0015	,1116	,0125
! st1004 !	,8014	,6422	,2990	,0894	,3297	,1087	-,0450	,0020	-,3505	,1228
! st1005 !	,8472	,7177	,0423	,0017	,3153	,0994	-,0078	,0001	-,1933	,0374
! st1006 !	,9145	,8363	,2666	,0710	,1885	,0355	,0997	,0099	-,0377	,0014
! st1007 !	,8370	,7005	-,1463	,0214	,2606	,0679	-,0398	,0016	-,1919	,0368

NOTA / 1ere colonne : corrélations entre variables et les axes principaux
2eme colonne : corrélations au carré

1 - Analyse sur le 1er cercle de corrélation

Le premier cercle de corrélation (voir fig. V.1) est la projection des variables sur la première et la deuxième composante principale, dont les axes sont représentés par ces composantes.

Ce cercle des corrélations nous montre deux groupes de stations, s'organisant sur le côté positif de l'axe C1.

Lorsque l'on identifie certaines stations dans les deux groupes, on remarque que les stations du groupe (exemple : TIFRA (1007), EL ASNAM (0204) et M'CHEDELLAH (0301)) situées en bas du cercle appartiennent à la région NE-NW, alors que les stations du reste de la zone d'étude se retrouvant dans le groupe situé en haut du cercle des corrélations (exemple : Sidi Embarek (0801), Ain Arnat (0607) et Taghdit (0201)) appartiennent à la région SE-SW.

2 - Analyse sur la carte des coefficients de corrélation avec la 1ère composante

de façon à mieux visualiser la caractéristique de la première composante, on a classé les valeurs des coefficients de corrélation par ordre croissant et on les a reporté sur la carte de la zone d'étude en fonction des postes pluviométriques.

Cette carte nous montre un regroupement des stations fortement positives dans la zone nord du bassin et les stations moins corrélées avec la première composante se regroupent dans la partie sud du bassin.

On peut donc interpréter la première composante comme étant la caractéristique d'une opposition Nord Sud (découpant approximativement le bassin en son milieu, séparant le nord et le sud suivant un axe E-W)

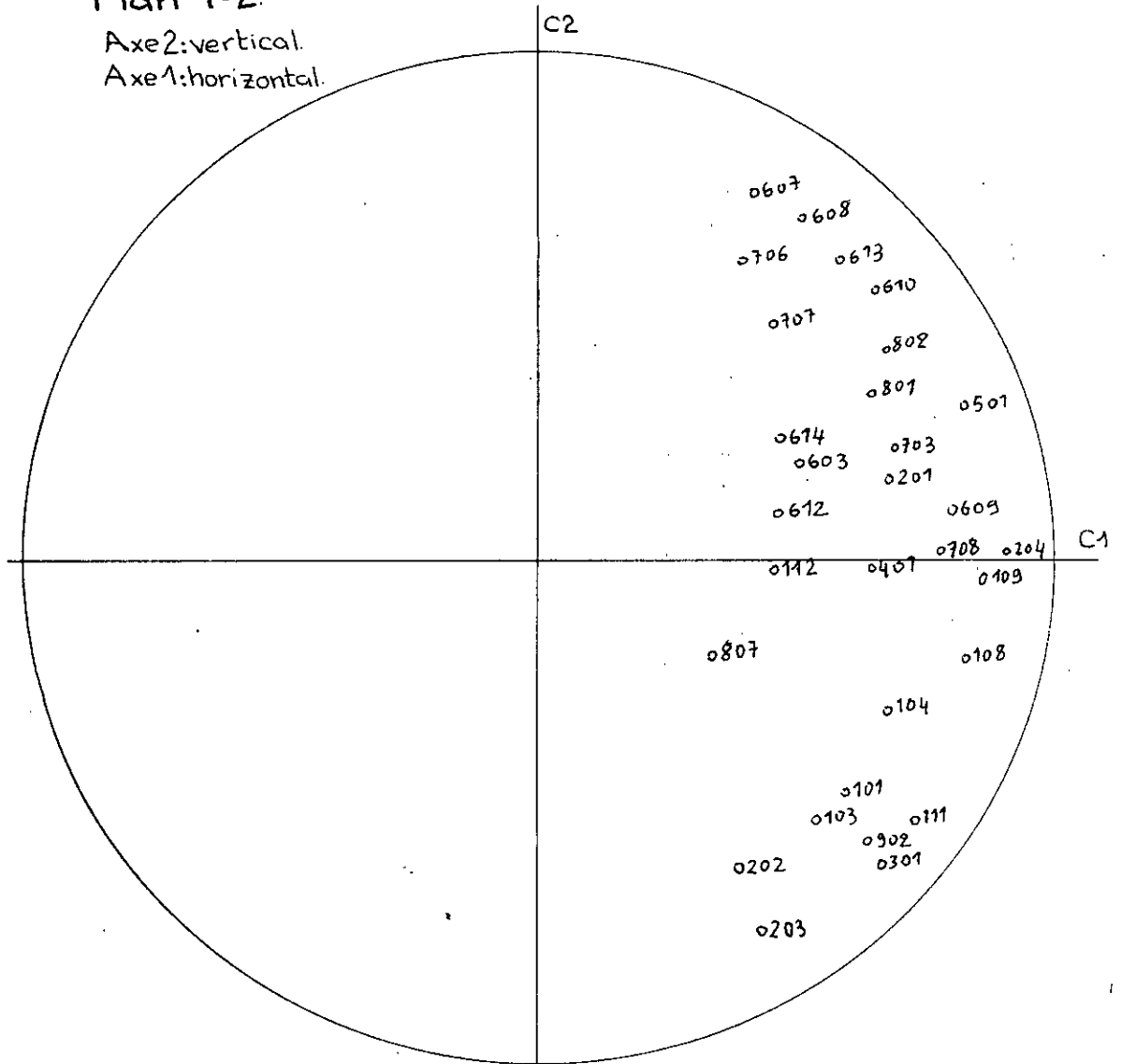
Cela paraît normal à l'échelle géographique car la quantité des précipitations qui tombe au nord est plus importante qu'au sud ; car la chaîne de montagne du Djurdjura avec des altitudes qui ne sont en aucun endroit inférieure à 1500 m d'altitude constitue une barrière naturelle aux masses d'air humide venant du Nord. Ces masses d'air deviennent sèches dès qu'elles traversent ces hauteurs, et se réchauffent dynamiquement en descendant vers le sud, conduisant ainsi à une faible précipitation.

Pour illustrer cette variance spatiale des précipitations annuelle du bassin de la Soummam, on a utilisé, la carte d'isoprécipitations moyenne, tracée sur la période d'étude (24 ans). On remarque une baisse très nette au fur et à mesure que l'on se déplace du Nord-Est au Sud-Est et du Nord-Ouest au Sud-Ouest ; ainsi on a deux ensembles plus ou moins homogènes.

Cercle des corrélations

Plan 1-2.

Axe 2: vertical.
Axe 1: horizontal.



Point vu : 0204

Point caché : 0503

Point vu : 0109

Point caché : 1001

Point vu : 0801

Point caché : 1004

Point vu : 0708

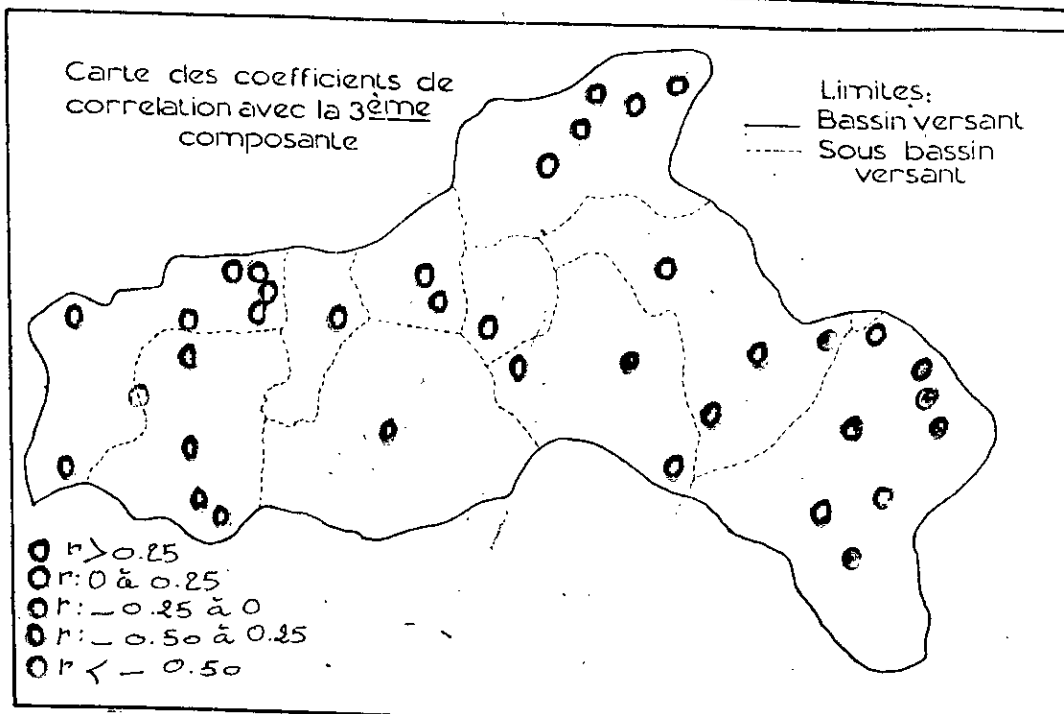
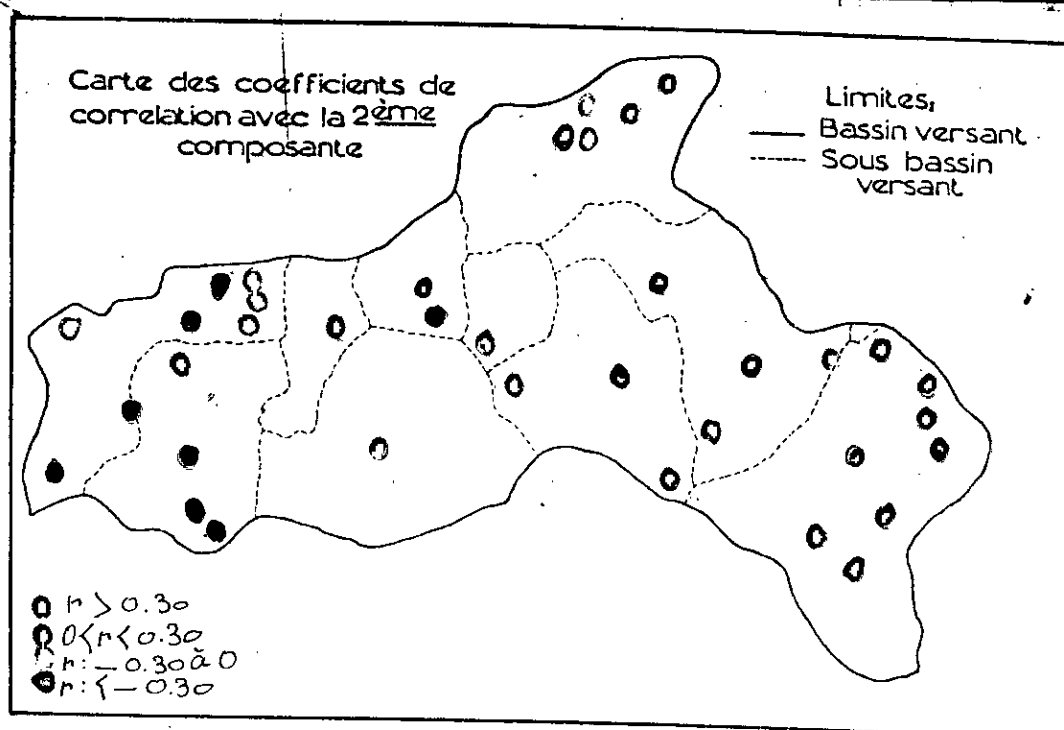
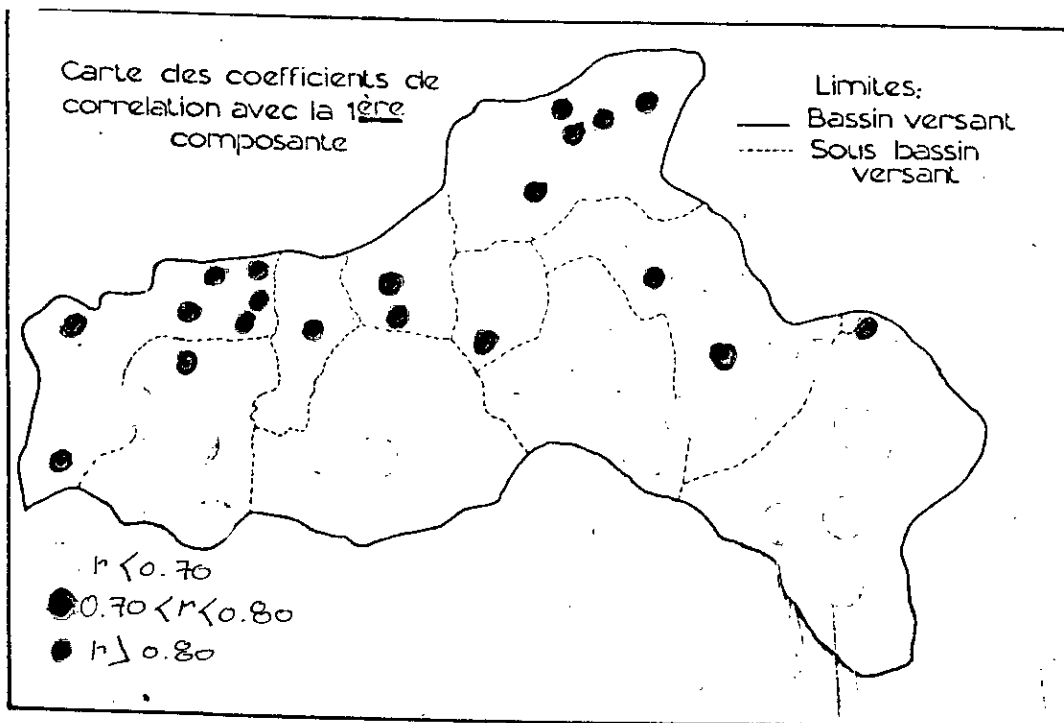
Point caché : 1005

Point vu : 0603

Point caché : 0110

Point vu : 0401

Point caché : 0712



V. 2.2 Le comportement spatial de la deuxième composante

1- Analyse sur le plan de la deuxième composante

Le second rôle des corrélations... Le second rôle des corrélations est de montrer la répartition du nuage (stations) sur l'axe c2. On remarque en effet une opposition entre les stations de l'ouest et celles de l'est... Les stations de l'ouest sont regroupées à gauche du cercle et celles de l'est sont groupées à droite du cercle.

2- Analyse des coefficients de corrélations avec la deuxième composante.

Les stations corrélées positivement se trouvent dans la région Ouest et les stations corrélées négativement se trouvent dans la région Est. Les coefficients négatifs sont le lien entre l'Est et l'Ouest. Ces stations sont situées près du centre de l'axe c2.

Comme exemple, on a :

- la station GUELIA, code : 0802
- la station TENESSASSAR, code : 0807
- la station PORT-DE-FER, code 0401

L'étirement du nuage sur l'axe c2, révèle deux remarques importantes :

- Les stations de l'Est sont plus dispersées que les stations de l'Ouest, compte tenu de la faible cohérence spatiale entre la plupart de ses stations.
- par contre les stations de l'Ouest se distinguent très légèrement en formant deux sous groupes bien organisés :

* le premier sous groupe est constitué de stations qui sont :

- Sour El Cozane (0101)
- El Achimia (0103)
- Bordj - Okriss (0104)
- Hammam Ksen (0110)
- Ain Allouane (0111)
- Irohat Haizer (0111)
- M'chedallah (0301)

1	Sour El Cozane
2	El Achimia
3	Bordj - Okriss
4	Hammam Ksen
5	Ain Allouane
6	Irohat Haizer
7	M'chedallah
8	Sour El Cozane
9	El Achimia
10	Bordj - Okriss
11	Hammam Ksen
12	Ain Allouane
13	Irohat Haizer
14	M'chedallah
15	Sour El Cozane
16	El Achimia
17	Bordj - Okriss
18	Hammam Ksen
19	Ain Allouane
20	Irohat Haizer
21	M'chedallah
22	Sour El Cozane
23	El Achimia
24	Bordj - Okriss
25	Hammam Ksen
26	Ain Allouane
27	Irohat Haizer
28	M'chedallah
29	Sour El Cozane
30	El Achimia
31	Bordj - Okriss
32	Hammam Ksen
33	Ain Allouane
34	Irohat Haizer
35	M'chedallah
36	Sour El Cozane
37	El Achimia
38	Bordj - Okriss
39	Hammam Ksen
40	Ain Allouane
41	Irohat Haizer
42	M'chedallah
43	Sour El Cozane
44	El Achimia
45	Bordj - Okriss
46	Hammam Ksen
47	Ain Allouane
48	Irohat Haizer
49	M'chedallah
50	Sour El Cozane
51	El Achimia
52	Bordj - Okriss
53	Hammam Ksen
54	Ain Allouane
55	Irohat Haizer
56	M'chedallah
57	Sour El Cozane
58	El Achimia
59	Bordj - Okriss
60	Hammam Ksen
61	Ain Allouane
62	Irohat Haizer
63	M'chedallah
64	Sour El Cozane
65	El Achimia
66	Bordj - Okriss
67	Hammam Ksen
68	Ain Allouane
69	Irohat Haizer
70	M'chedallah
71	Sour El Cozane
72	El Achimia
73	Bordj - Okriss
74	Hammam Ksen
75	Ain Allouane
76	Irohat Haizer
77	M'chedallah
78	Sour El Cozane
79	El Achimia
80	Bordj - Okriss
81	Hammam Ksen
82	Ain Allouane
83	Irohat Haizer
84	M'chedallah
85	Sour El Cozane
86	El Achimia
87	Bordj - Okriss
88	Hammam Ksen
89	Ain Allouane
90	Irohat Haizer
91	M'chedallah
92	Sour El Cozane
93	El Achimia
94	Bordj - Okriss
95	Hammam Ksen
96	Ain Allouane
97	Irohat Haizer
98	M'chedallah
99	Sour El Cozane
100	El Achimia

La proximité de leurs projections sur l'axe c2, correspond bien à une proximité réelle sur le plan géographique, constitue ainsi un ensemble homogène.

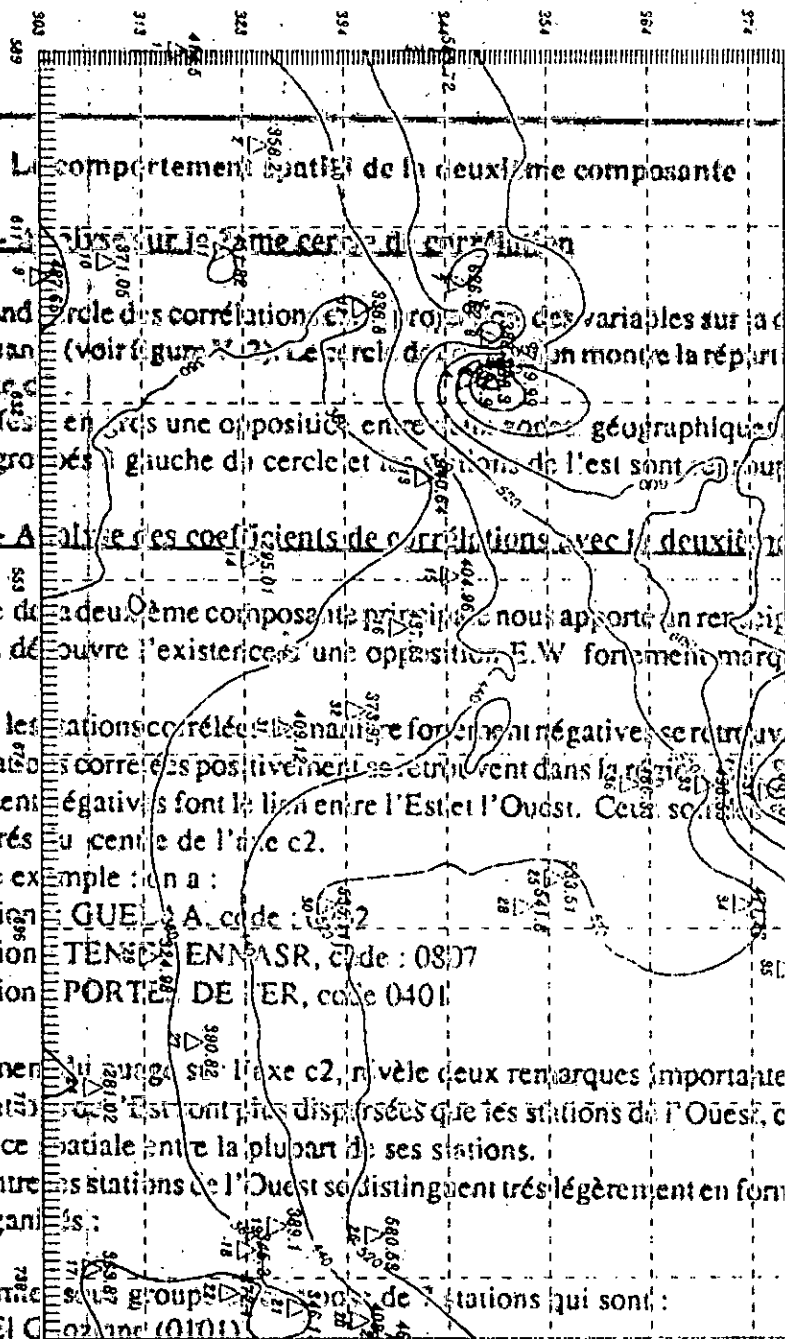
La station la plus représentative de ce groupe correspond au code (0203)

* le deuxième sous groupe constitue les autres stations près de l'axe c2.

La proximité de leurs projections ne correspond pas à une proximité réelle sur le plan géographique, constitue ainsi un ensemble moins homogène que le premier sous groupe.

CARTE D'ISOPRESCRIPTIONS MOYENNES

SCALE 1 cm = 17,38 km



Ainsi le groupe ouest semble plus homogène que le groupe Est, ce dernier subit l'influence du caractère ouest par trois stations : (0201), (0204), (0501), ceux-ci est probablement lié aux anomalies signalées dans les séries d'observations ou par un effet de relief ou autres paramètres physiques.

V.2.3 Le comportement spatial de la troisième composante

1- Analyse sur le troisième cercle de corrélation

Le troisième cercle est la projection des variables sur la première et la troisième composante (fig V.3)

Cette dernière projection nous montre deux groupes nettement divisés et homogènes. Après identification de certaines stations s'alignant de bas en haut, nous remarquons que les stations situées en bas du cercle appartiennent à la région sud Est, alors que les stations situées en haut du cercle correspondent à la région Nord Est.

2- Analyse sur la carte des coefficients de corrélations avec la 3ème composante

Sur la carte de cette composante, semble se dessiner, une opposition entre le NE et le SE. En effet, les stations corrélées de manière fortement négatives se retrouvent dans la région Sud Est et les stations des mieux corrélées positivement se situent dans la région Nord Est.

Ainsi, la station la plus représentative du groupe SE correspond au code (0706) et celle du groupe NE est (1001)

V.3 Conclusion

L'A.C.P a permis de montrer à partir des résultats obtenus avec les représentations cartographiques et les cercles de corrélations, que les groupes de stations sont effectivement distinguées essentiellement par les positions relatives EW et SE-NE ainsi que par l'apposition Nord-Sud.

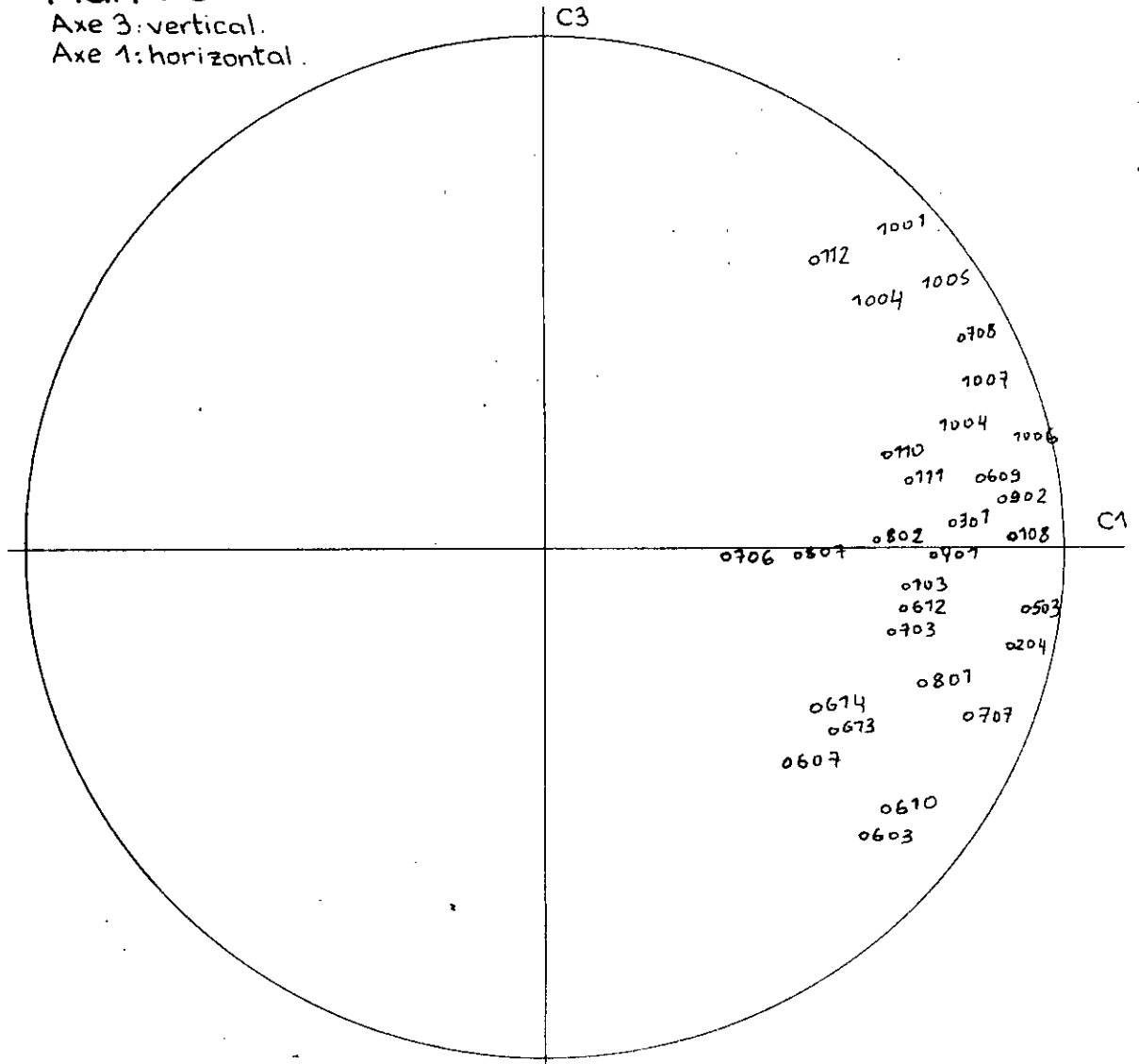
De cette façon, on peut représenter chacun des groupes par une ou deux stations les plus représentatives de ces groupes. Ainsi on aura réduit le nombre de stations par des stations types représentant le comportement moyen des différentes stations.

Cercle des corrélations

Plan 1-3.

Axe 3: vertical.

Axe 1: horizontal.



point vu : 0614

point caché : 0608

point vu : 0612

point caché : 0202

point vu : 0103

point caché : 0201

point vu : 0802

point caché : 0203

point vu : 0108

point caché : 0501

point vu : 0902

point caché : 0109

point vu : 0104

point caché : 0101

- fig. V - 3 -

Conclusion Générale

Au terme, de ce travail mené sur le bassin de la Soummam, on a essayé de montrer selon quel critère les 37 stations pluviométriques retenues se regroupent dans le domaine d'étude.

Pour arriver à ce but, la méthode de l'analyse en composantes principales des précipitations annuelles a été utilisée; cependant elle nécessitait tout d'abord de disposer d'un ensemble cohérent et homogène de données pluviométriques.

Avant, de pratiquer cette analyse, quatre traitements préalables ont donc été effectués sur les données observées afin d'en tester la fiabilité et de corriger les erreurs, ou de combler les lacunes.

En premier lieu on a effectué une critique préliminaire des données en confrontant les fichiers d'origine pris aux archives de l'A.N.R.H. avec ceux des listings en vue de corriger les erreurs de reports (transcriptions).

En second lieu la recherche des anomalies ou des hétérogénéités des séries a fait appel, outre le test, de Kendall, à une autre méthode celle du cumul des résidus de régression.

La méthode du cumul des résidus nous semble être plus fiable car elle relève plus finement les hétérogénéités au niveau des postes pluviométriques.

Pour mener à bien cette tâche, deux problèmes se sont posés :

- dans le choix d'une stations de référence ;

à ce niveau on a utilisé l'avantage de l'A.C.P dans la phase critique par le cumul des résidus.

- L'A.C.P qui est un moyen de critique, ne peut être utilisée que sur une matrice de données complète (sans aucune marque)

A ce niveau, on a comblé les observations manquantes par la méthode de régression linéaire simple en sachant que les séries suivent globalement une loi de Gauss.

A la fin de cette critique on a éliminé 4 stations dont les données sont trop suspectes et qu'il était impossible de les corriger.

L'analyse en composantes principales a été appliquée à un ensemble de données réduit de 37 stations (au lieu de 41) et globalement homogènes.

Cette analyse fait émerger plusieurs ensembles spatiaux :

- La première composante met en évidence une opposition entre le nord et le sud.

- La deuxième composante révèle une opposition entre l'est et l'ouest ; le groupe est, se divise en deux sous groupe plus homogène que le groupe lui même.

- La troisième composante nous a permis de voir une autre opposition (NE - SE)

Comme il est nécessaire de limiter l'étendue d'un domaine de recherche et qu'il faut bien s'arrêter dans un développement; on recommande afin de mieux connaître les différents groupes de stations pluviométriques et pour vérifier si ces groupes sont bien homogènes, de faire une étude de chacun de ces groupes en réalisant différentes A.C.P. sur ces groupes, ainsi on pourra affiner ces premiers résultats et voir s'il n'est pas possible de détailler encore plus les groupes de stations.

On pourra également comparer la méthode de l'A.C.P. avec la méthode de classification automatique qui permettra de mettre en évidence une partition de n variable en K classes; ainsi elle permet de :

- choisir le nombre de K classes pour partitionner N variables.
- affecter chaque variable à une des classes.
- définir le comportement moyen de chaque classe.

On a fait largement appel à l'outil informatique au cours de notre travail, on a en effet utilisé des logiciels de traitement de données récents et efficaces.

Parmi ces logiciels, on a utilisé :

- les tableurs Lotus et Excel,
 - Surfer, qui nous a permis la représentation graphique des données sous forme de cartes à deux dimensions et à trois dimensions.
 - Statgraphe (pour le test du khi-deux et K-S).
-

ANNEXES

 * FEUILLE DE CALCUL: AJUSTEMENT A UNE LOI NORMALE *

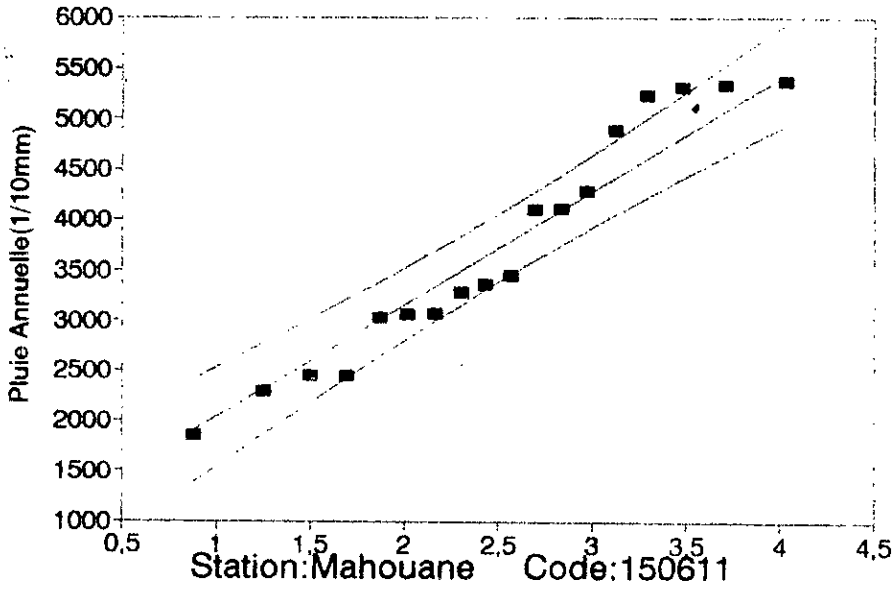
STATION:AIN ARNAT code:150607

N= 31 X̄=valeurs theoriques
 MOY =3609,387 R.T.=1056,514 F̄=frequences experimentales

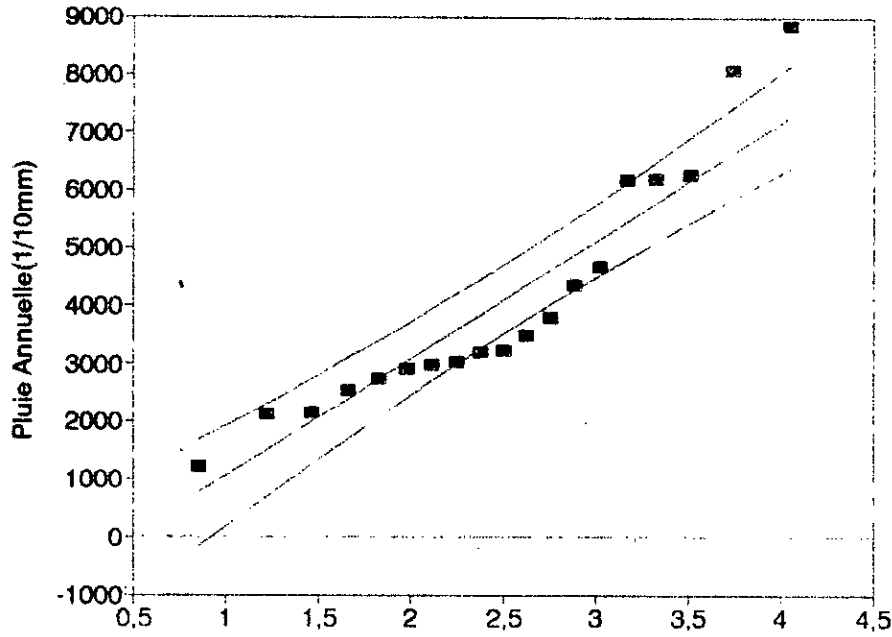
ANNEE	i	P.clas	F=i/N+1	F*	X	A	B	U	X*	Borne inf	Borne sup	U+2.5	
1977	1	1362	0,031	0,031	2,633	4,701	6,108	-1,86	1640,9	402,27	1238,675	2043,217	0,636852
1974	2	1592	0,063	0,063	2,355	4,463	5,441	-1,53	1988,2	358,87	1629,362	2347,109	0,965565
1973	3	1939	0,094	0,094	2,176	4,311	5,027	-1,32	2216,7	332,49	1884,193	2549,165	1,181789
1970	4	2459	0,125	0,125	2,039	4,196	4,72	-1,15	2393,9	313,52	2080,412	2707,46	1,349564
1950	5	2555	0,156	0,156	1,927	4,101	4,473	-1,01	2542,3	298,87	2243,476	2841,21	1,490033
1949	6	2580	0,188	0,188	1,830	4,019	4,263	-0,89	2672,2	287,1	2385,142	2959,333	1,612979
1989	7	2930	0,219	0,219	1,743	3,947	4,08	-0,78	2789,3	277,44	2511,88	3066,757	1,723797
1940	8	3081	0,250	0,250	1,665	3,881	3,917	-0,67	2897,1	269,44	2627,66	3166,534	1,825811
1979	9	3104	0,281	0,281	1,593	3,821	3,768	-0,58	2997,9	262,79	2735,121	3260,706	1,921235
1943	10	3105	0,313	0,313	1,525	3,764	3,63	-0,49	3093,4	257,31	2836,121	3350,734	2,011639
1952	11	3220	0,344	0,344	1,461	3,711	3,502	-0,40	3184,9	252,84	2932,031	3437,71	2,098191
1954	12	3245	0,375	0,375	1,401	3,66	3,382	-0,32	3273,2	249,29	3023,911	3522,498	2,1818
1983	13	3287	0,406	0,406	1,342	3,612	3,267	-0,24	3359,2	246,6	3112,607	3605,805	2,263201
1982	14	3375	0,438	0,438	1,286	3,565	3,158	-0,16	3443,5	244,71	3198,823	3688,237	2,343015
1980	15	3513	0,469	0,469	1,231	3,519	3,053	-0,08	3526,7	243,59	3283,165	3770,335	2,421783
1941	16	3543	0,500	0,500	1,177	3,475	2,951	0,00	3609,4	243,21	3366,174	3852,601	2,5
1979	17	3590	0,531	0,469	1,125	3,432	2,853	0,08	3691,9	243,58	3448,354	3935,523	2,578136
1985	18	3740	0,563	0,438	1,073	3,389	2,756	0,16	3774,9	244,7	3530,194	4019,596	2,656654
1944	19	3760	0,594	0,406	1,021	3,346	2,662	0,24	3858,8	246,58	3612,184	4105,339	2,736035
1942	20	3802	0,625	0,375	0,970	3,304	2,568	0,32	3944,1	249,24	3694,842	4193,323	2,816792
1946	21	3890	0,656	0,344	0,918	3,261	2,476	0,40	4031,5	252,73	3778,735	4284,198	2,899502
1988	22	4176	0,688	0,313	0,866	3,218	2,383	0,48	4121,6	257,11	3864,511	4378,729	2,984833
1937	23	4240	0,719	0,281	0,813	3,175	2,29	0,57	4215,4	262,46	3952,944	4477,858	3,073598
1972	24	4307	0,750	0,250	0,759	3,13	2,196	0,67	4313,9	268,89	4044,991	4582,778	3,166813
1947	25	4500	0,781	0,219	0,703	3,085	2,101	0,77	4418,5	276,58	4141,9	4695,067	3,265817
1981	26	4767	0,813	0,188	0,644	3,037	2,002	0,87	4531,1	285,77	4245,37	4816,913	3,372449
1986	27	4887	0,844	0,156	0,583	2,987	1,9	0,99	4654,7	296,82	4357,86	4951,508	3,483383
1948	28	5025	0,875	0,125	0,517	2,933	1,791	1,12	4793,5	310,33	4483,193	5103,858	3,620798
1945	29	5180	0,906	0,094	0,444	2,874	1,673	1,27	4955,3	327,34	4627,948	5282,625	3,773906
1975	30	5240	0,938	0,063	0,359	2,805	1,539	1,46	5155,3	349,97	4805,343	5505,277	3,96323
1951	31	5897	0,969	0,031	0,252	2,718	1,373	1,73	5434,9	383,99	5050,89	5818,88	4,227851

AJUSTEMENT A UNE LOI NORMALE

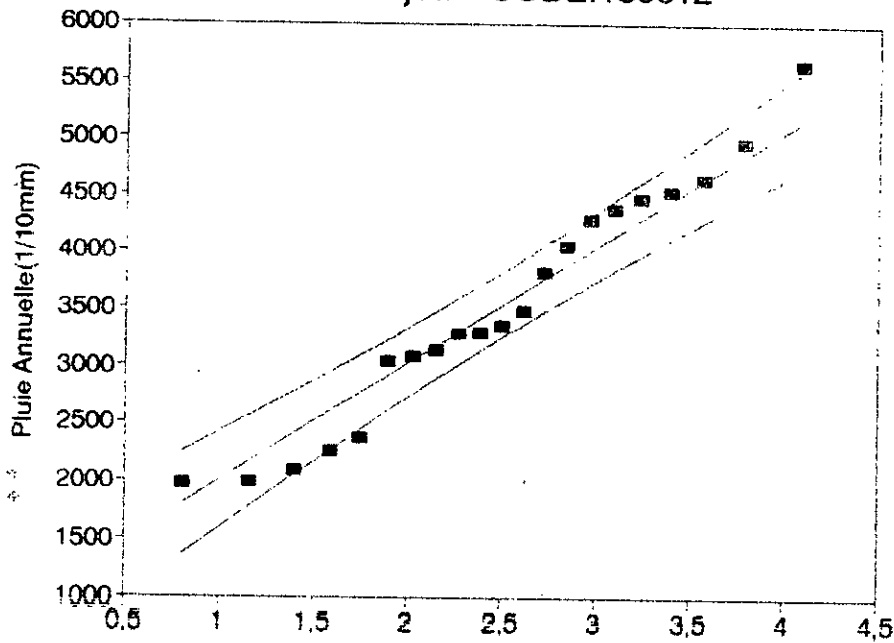
Station:Fermatou Code:150610



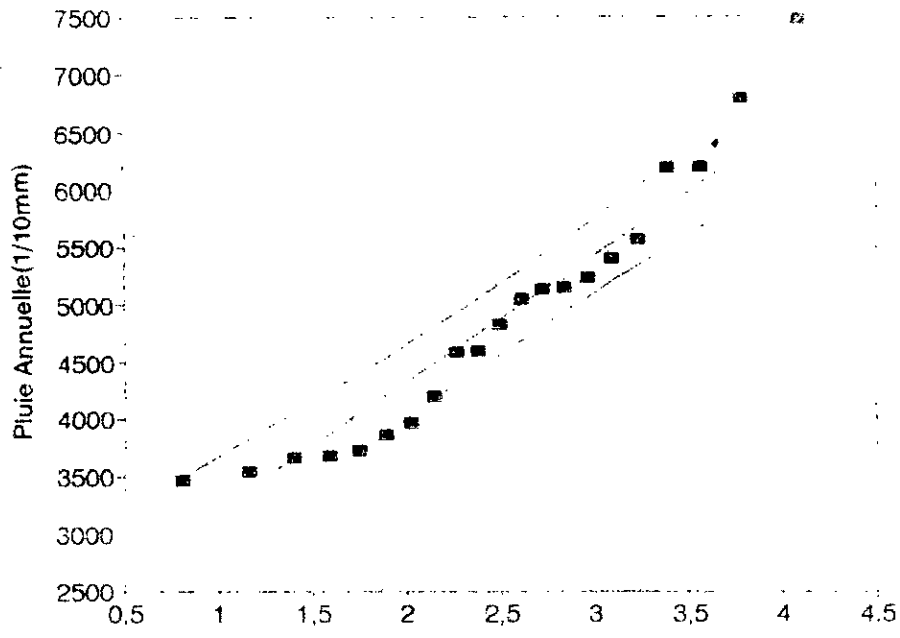
Station:Mahouane Code:150611



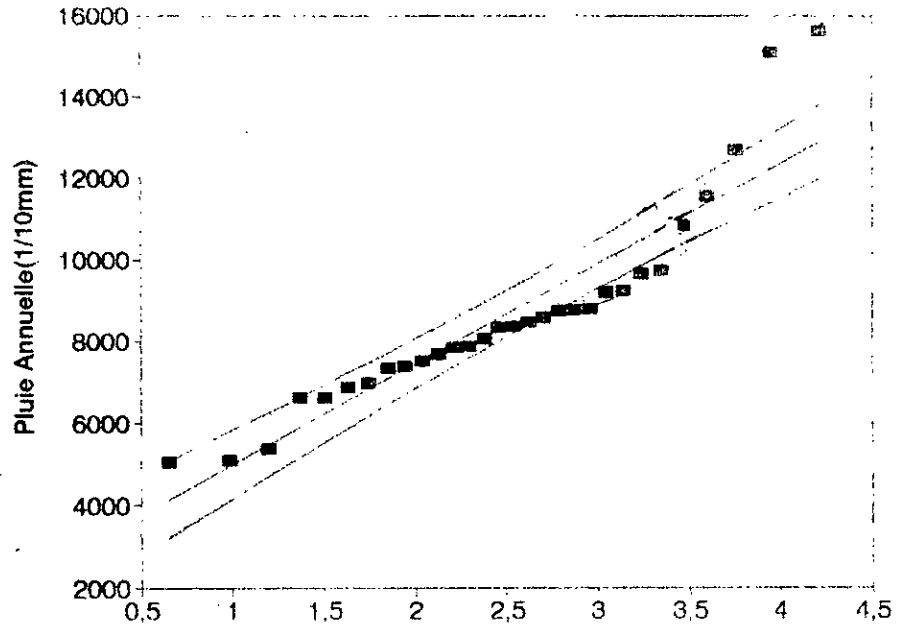
Station:Tikjda CODE:150612



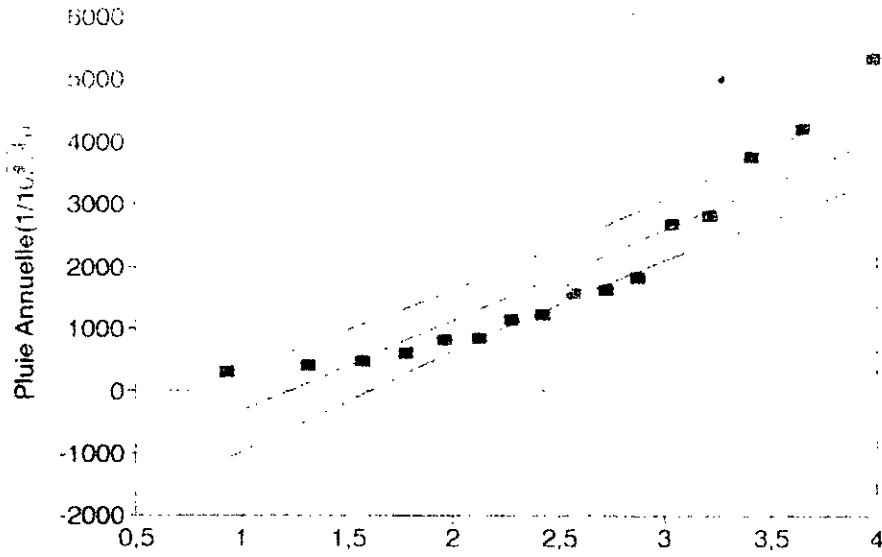
Station: Seddouk Code: 151006



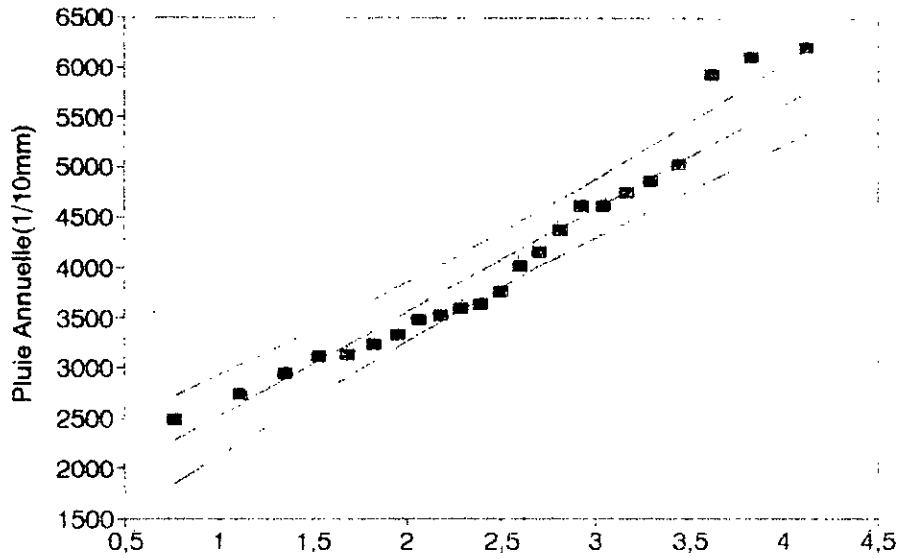
Station: Tifra Code: 151007



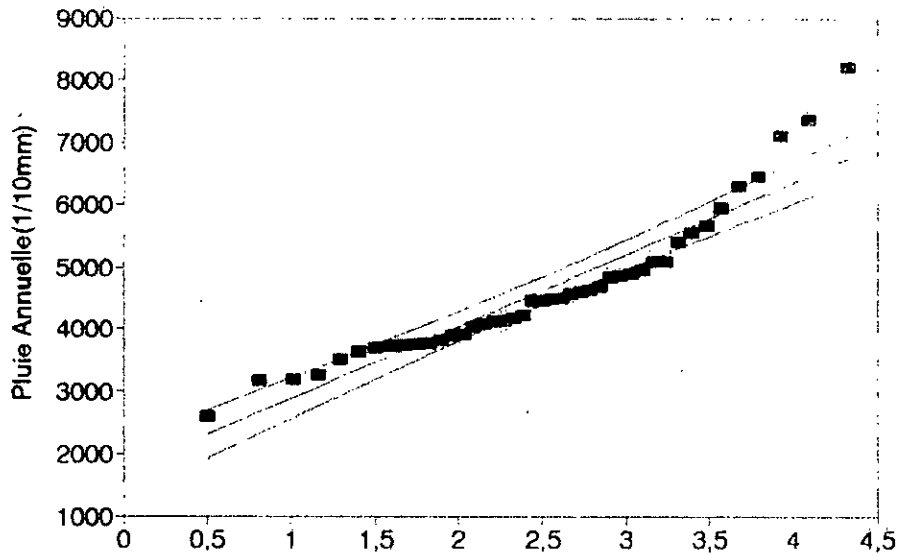
Station: Bendaoud Code: 150402



Station: Tazmalt Code: 150501

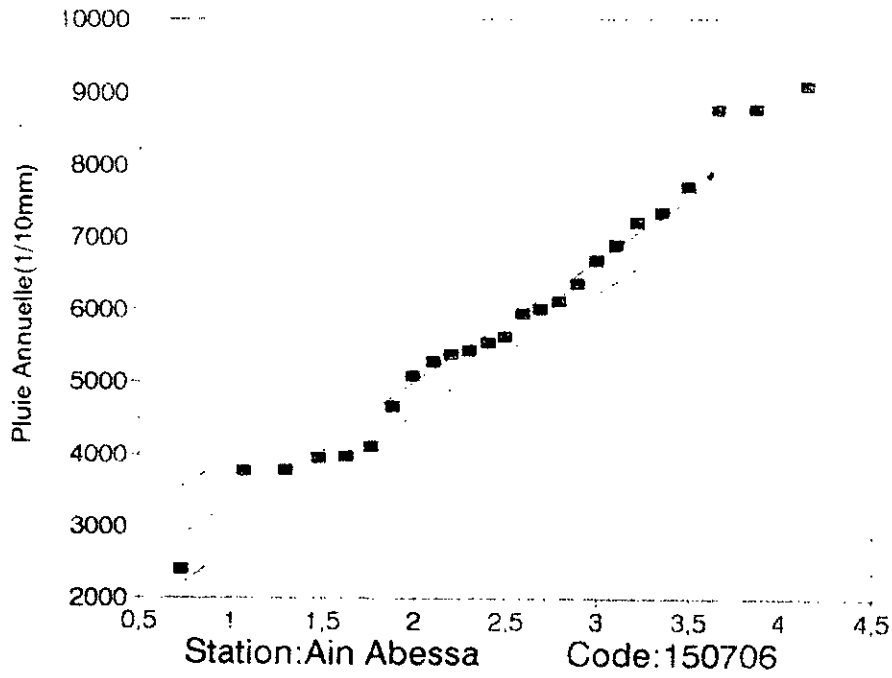


Station: Ighilali Code: 150503



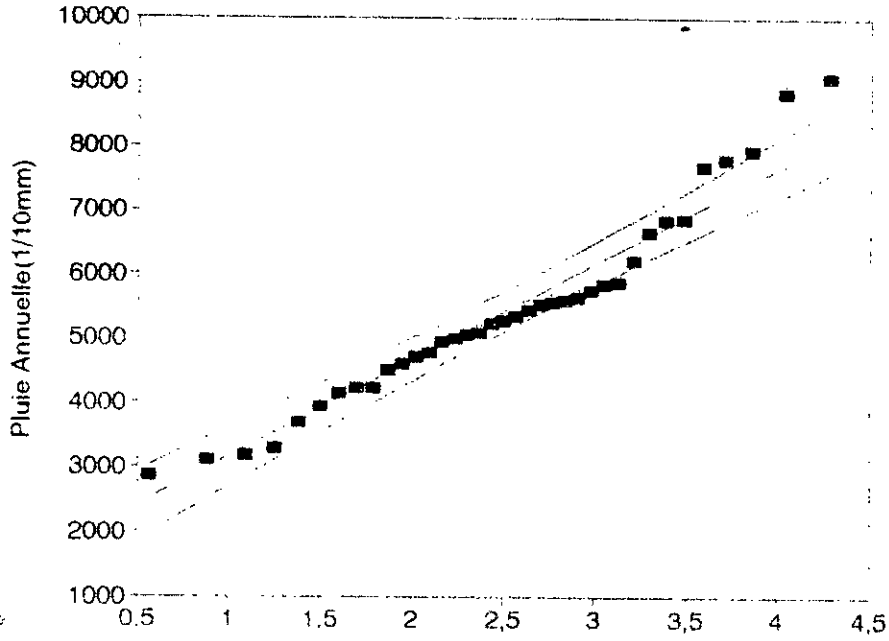
Station:Ain Zada

Code:150703



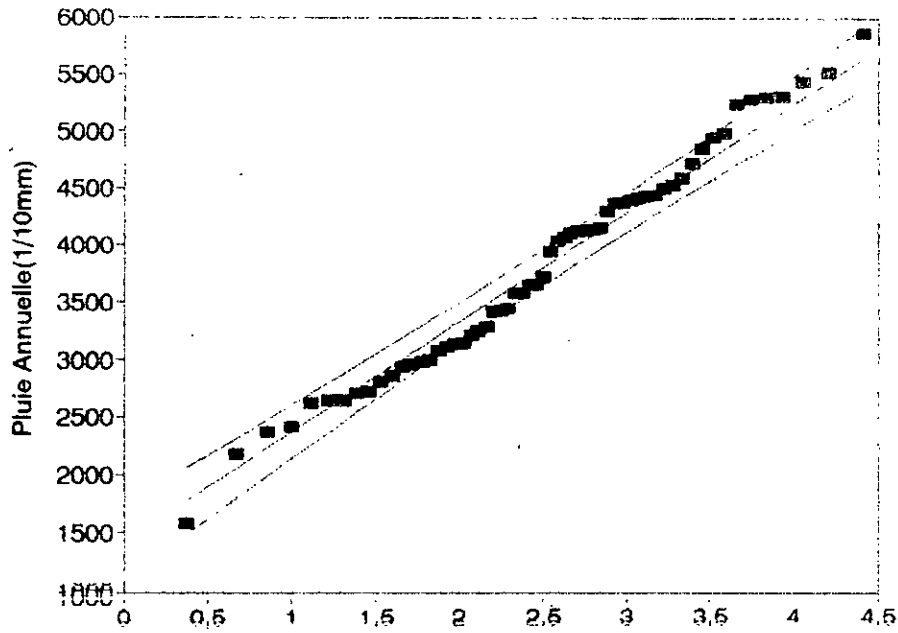
Station:Ain Abessa

Code:150706

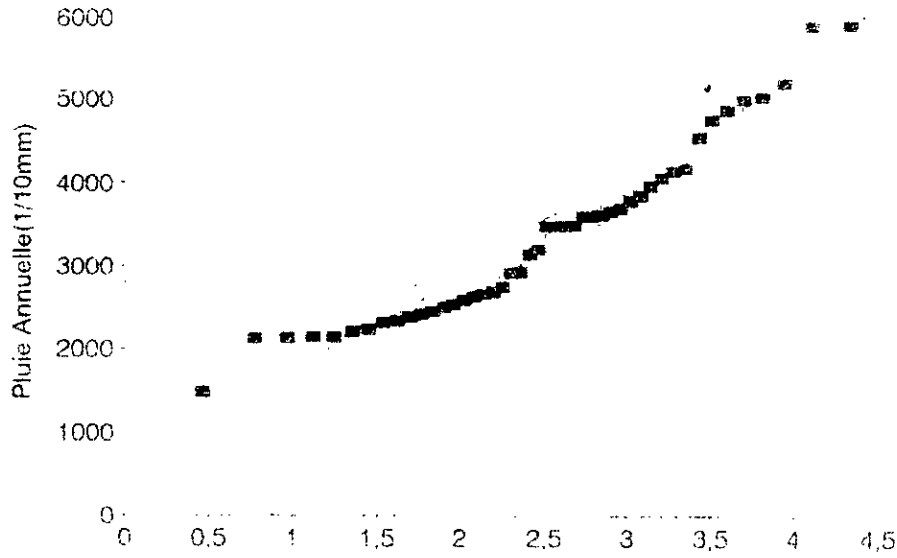


Station:Bir Kasdalli

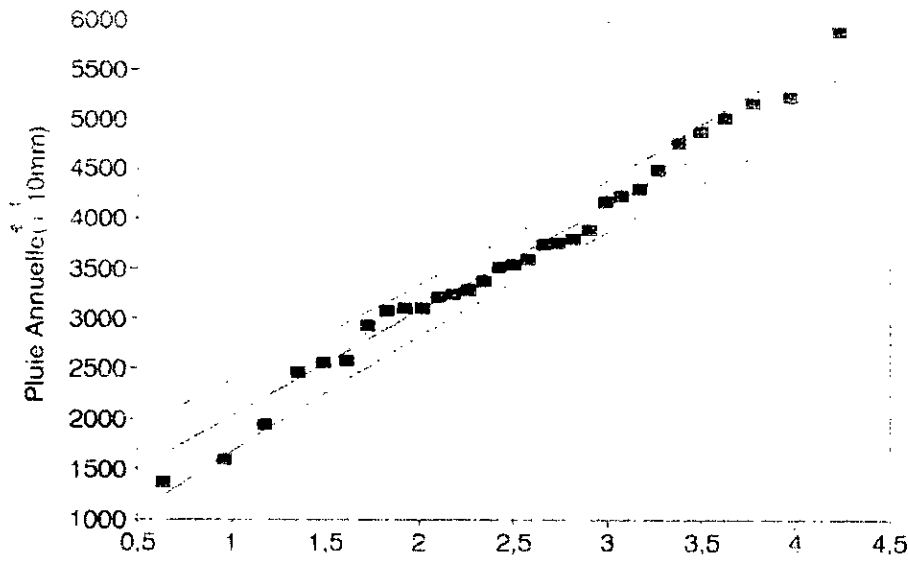
Code:150707



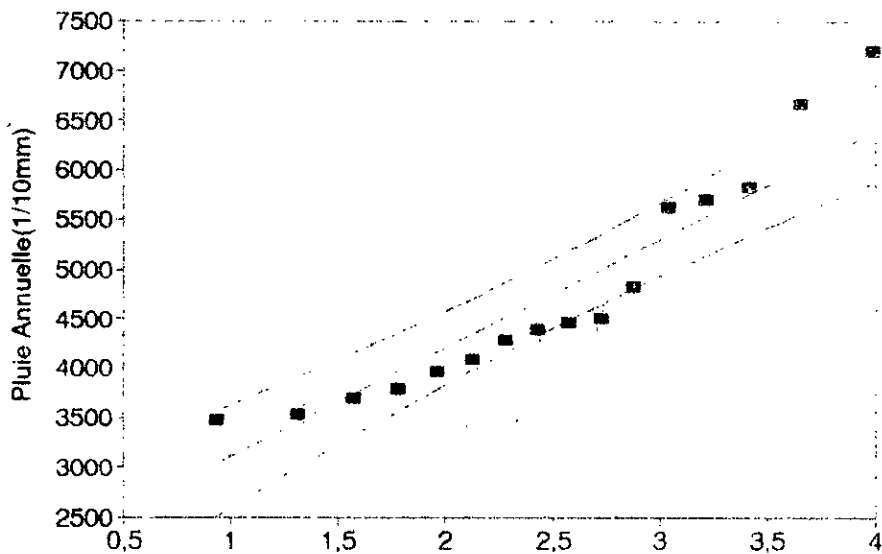
Station:Guellal Code:150603



Station:Ain Arnat Code:150607

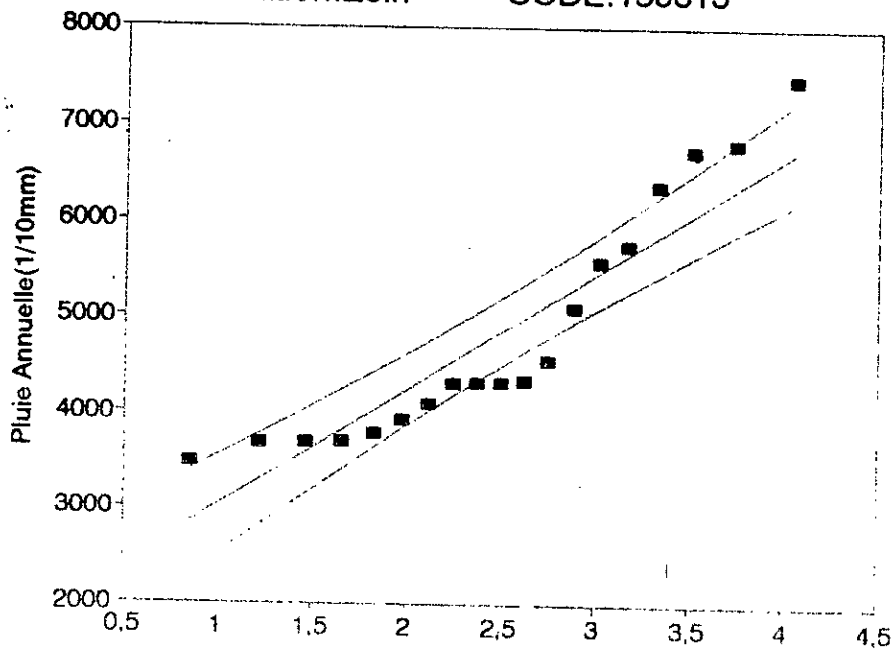


Station:El Ouricia Code:150609



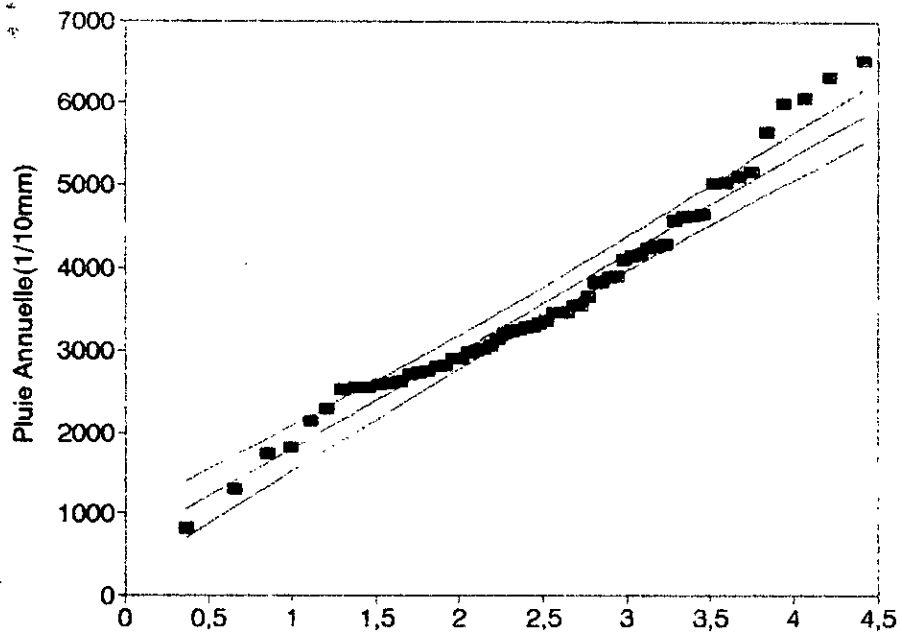
Station:Zeiri

CODE:150613



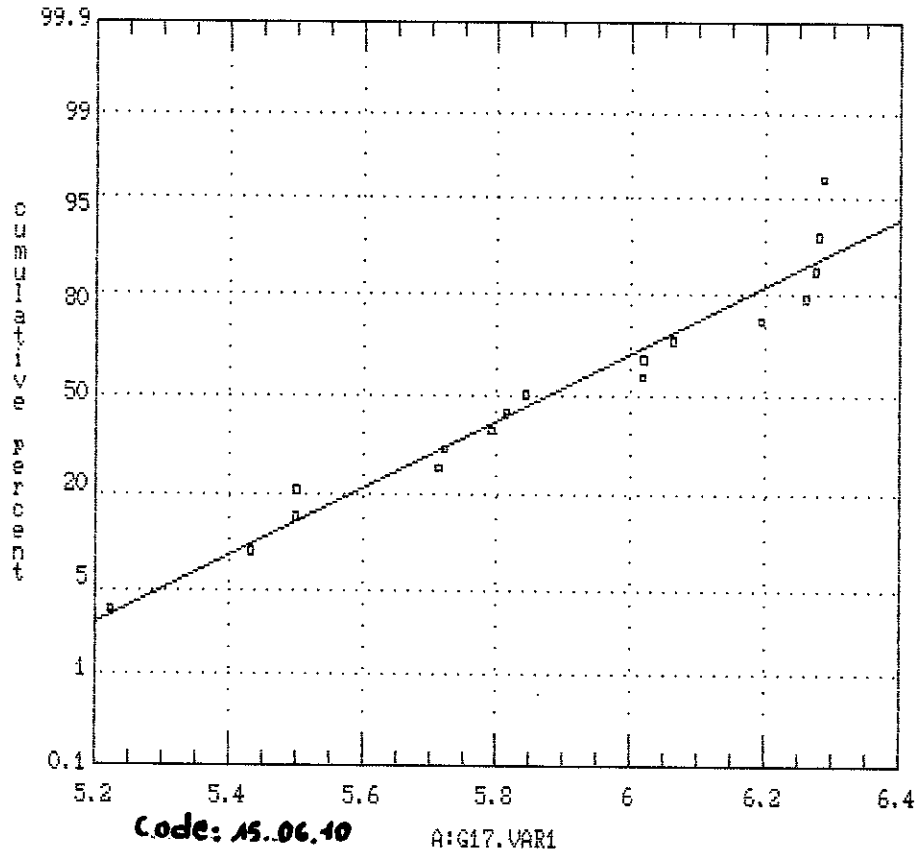
Station:Tixter

CODE:150614

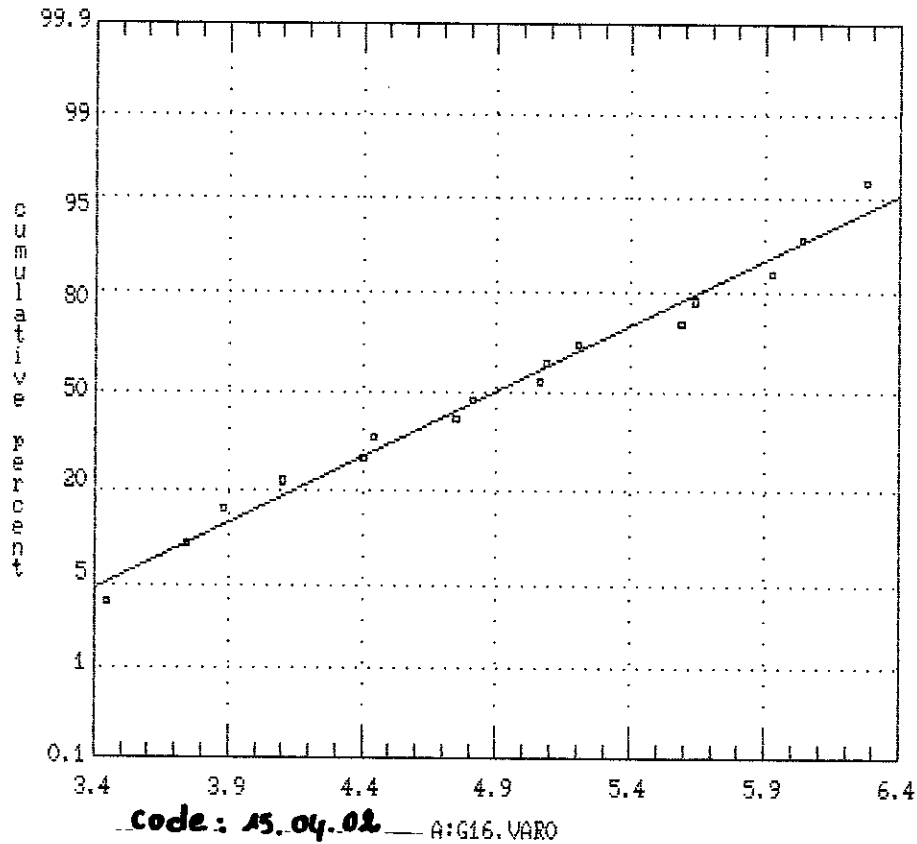


AJUSTEMENT A UNE LOI GALTON

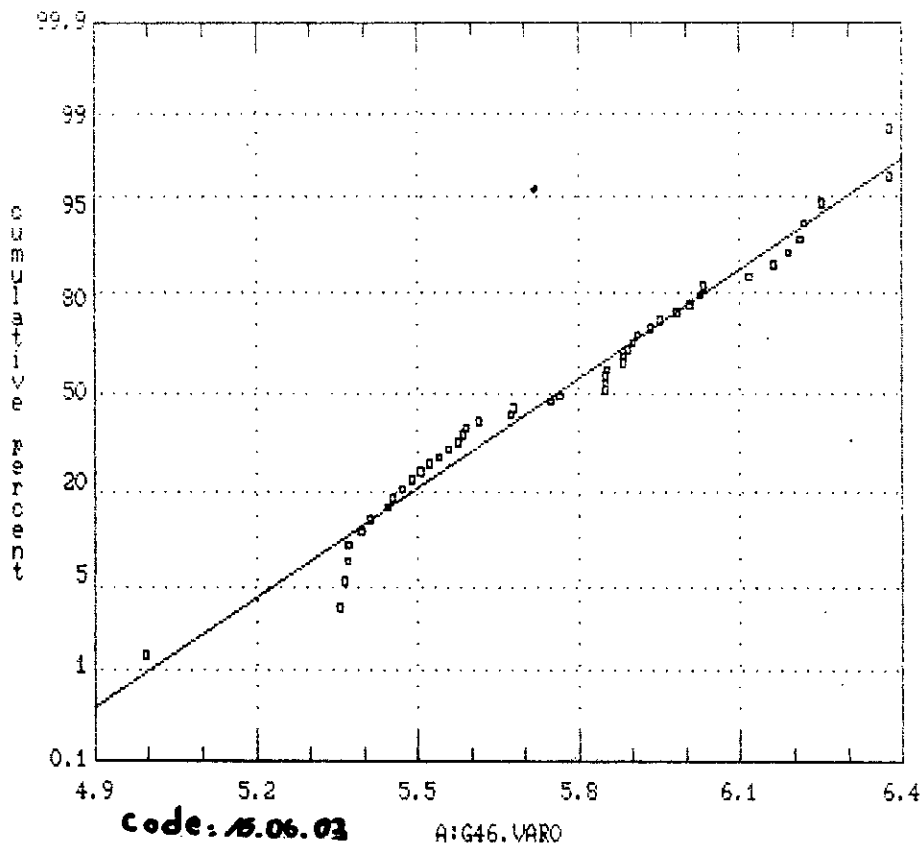
Normal Probability Plot



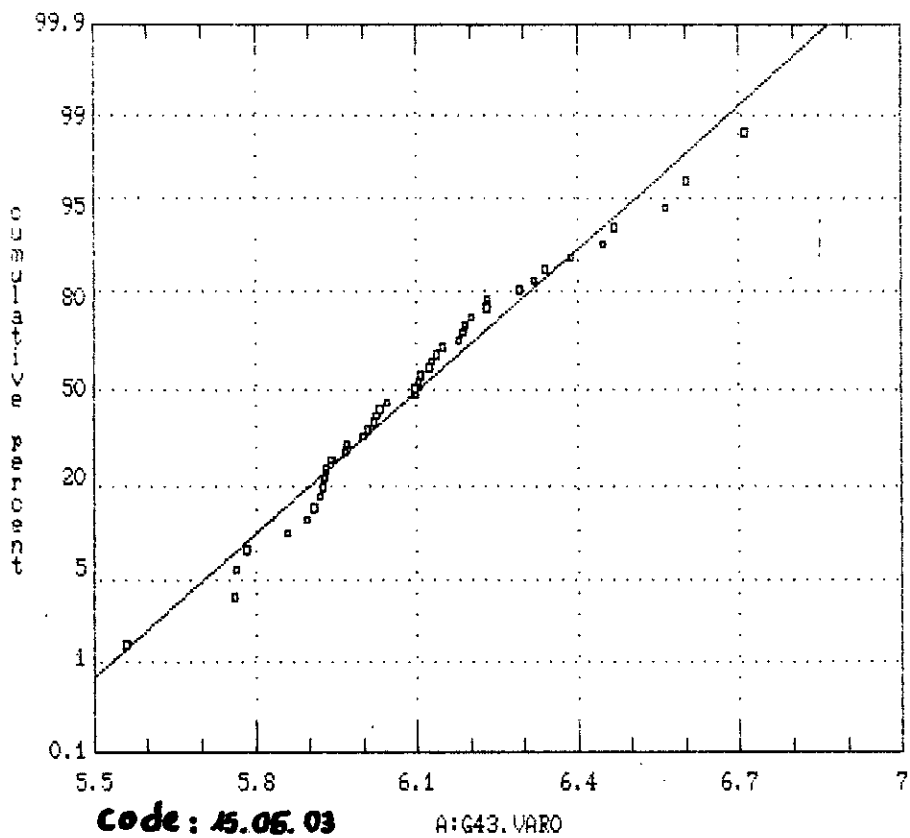
Normal Probability Plot



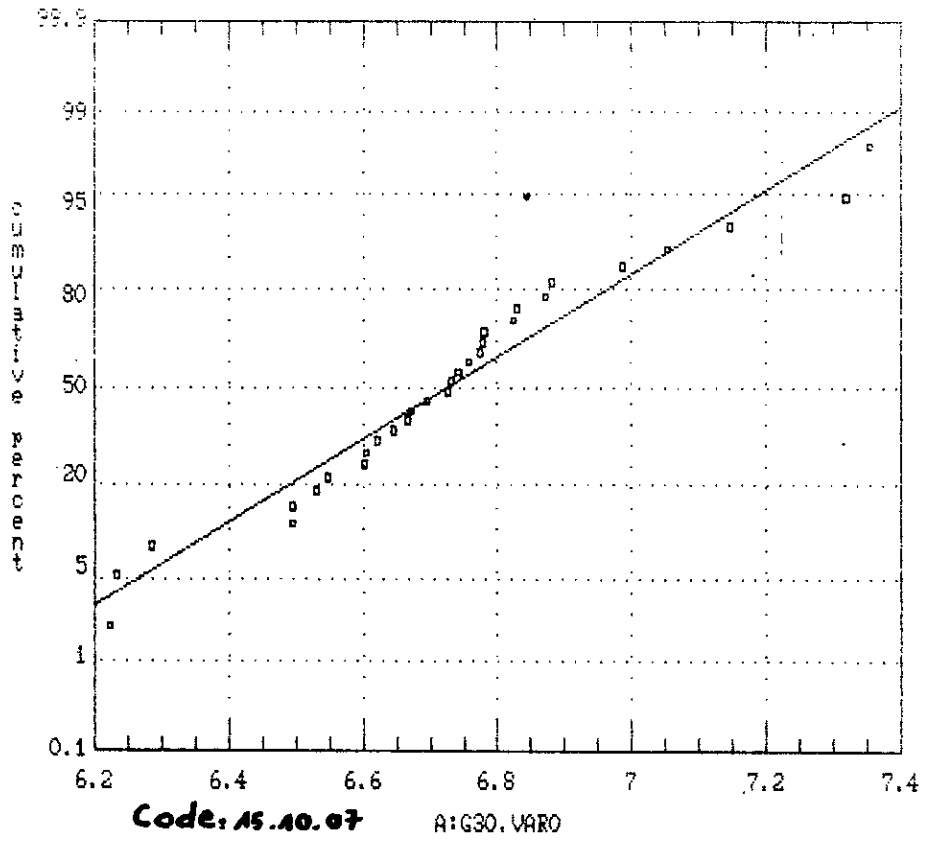
Normal Probability Plot



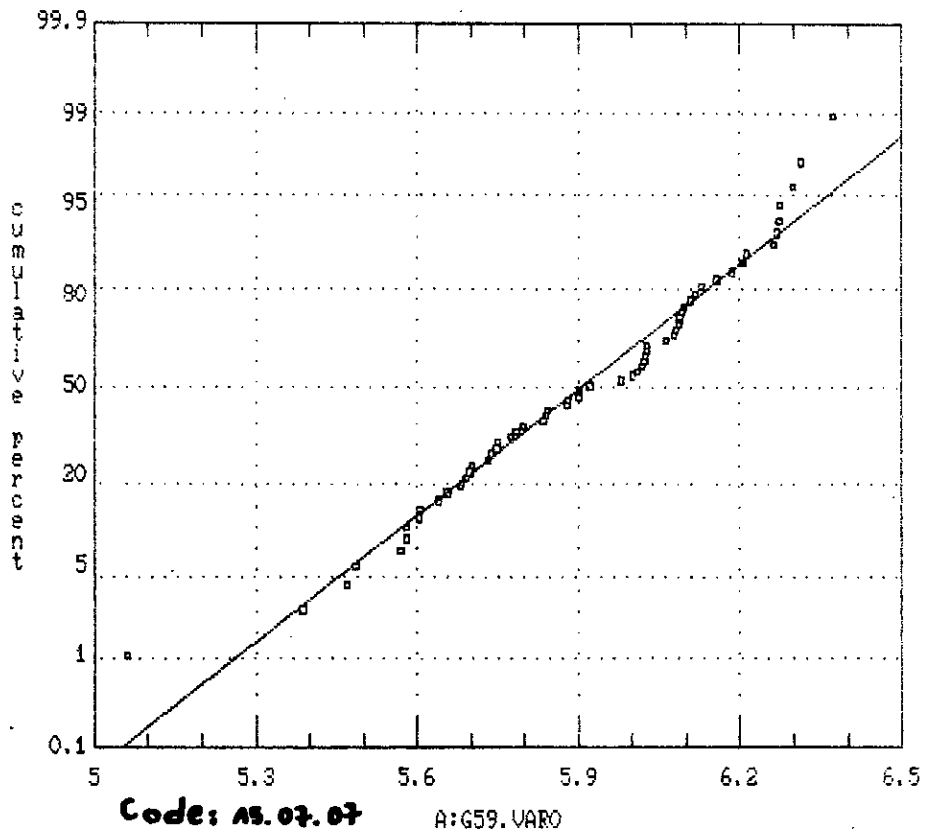
Normal Probability Plot



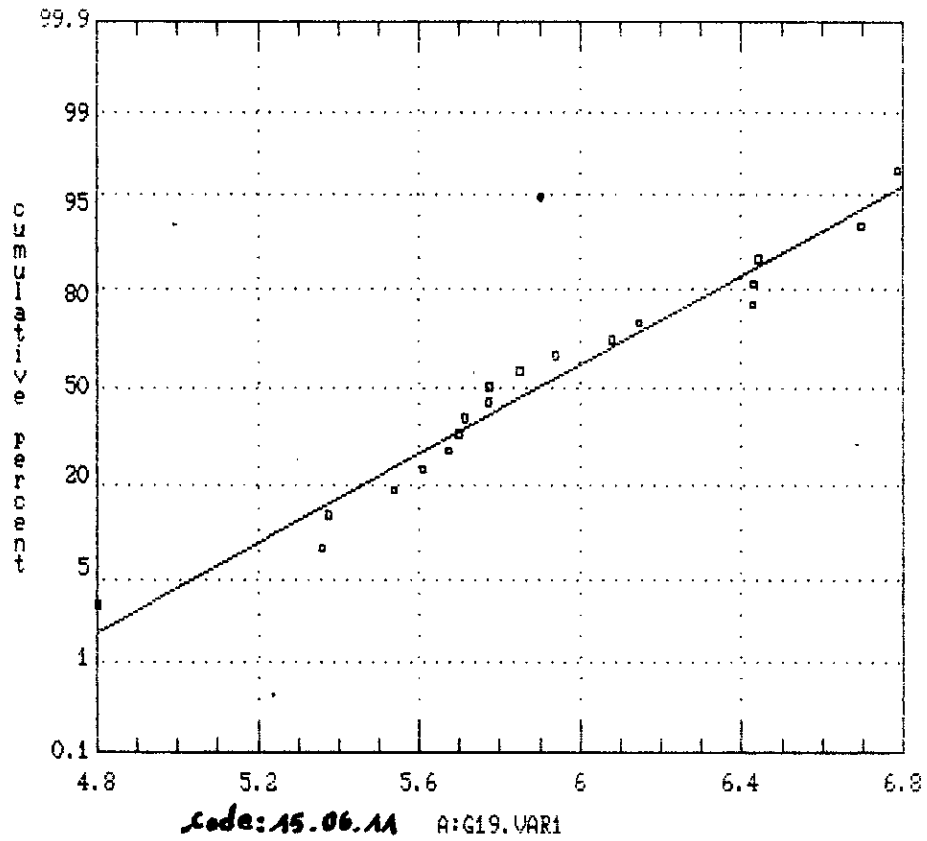
Normal Probability Plot



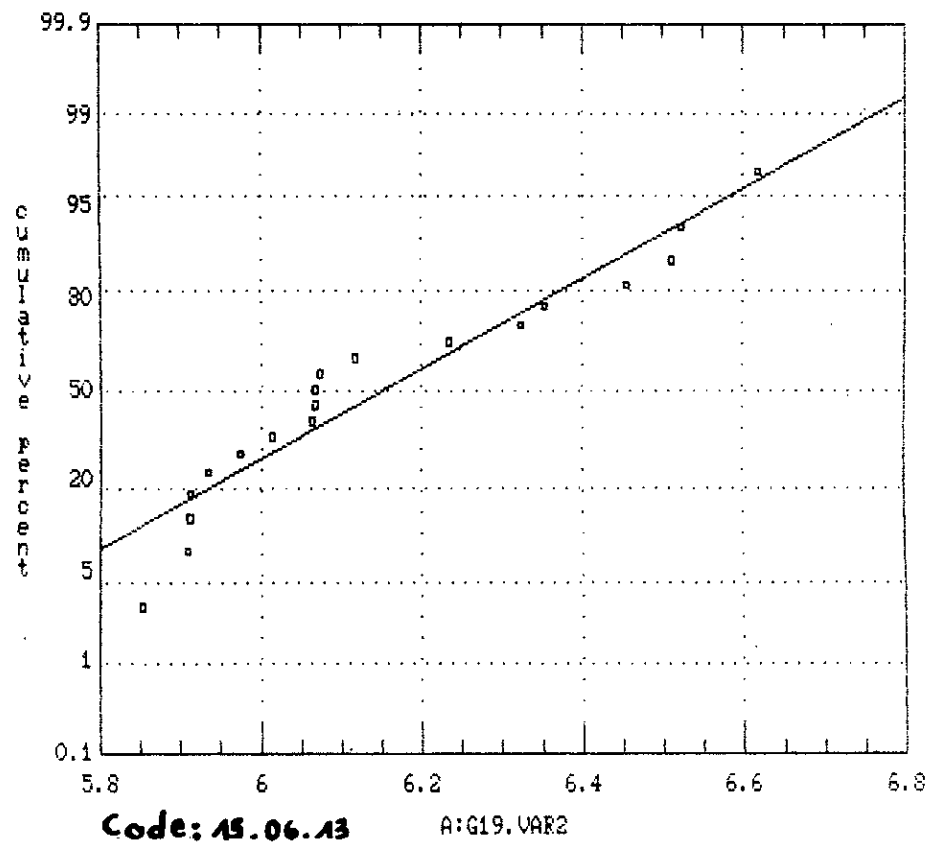
Normal Probability Plot



Normal Probability Plot



Normal Probability Plot



TEST χ^2 ET K-S

code: 15.07.03

Chisquare Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chisquare
at or below		4585.71	6	6	.00429
	4585.71	5728.57	7	6	.14462
	5728.57	6871.43	5	6	.26853
above	6871.43		7	7	.00625

Chisquare = 0.423682 with 1 d.f. Sig. level = 0.515106

code: 15.07.03

Estimated KOLMOGOROV statistic DPLUS = 0.081166
Estimated KOLMOGOROV statistic DMINUS = 0.0753297
Estimated overall statistic DN = 0.081166
Approximate significance level = 1

code: 15.05.03.

Chisquare Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chisquare
at or below		3558.82	5	8	-1.059
	3558.82	4264.71	15	9	4.816
	4264.71	4617.65	6	5	.156
	4617.65	4970.59	6	5	.158
	4970.59	5676.47	5	9	1.462
above	5676.47		6	8	.413

Chisquare = 8.06317 with 3 d.f. Sig. level = 0.044724

code: 15.05.03.

Estimated KOLMOGOROV statistic DPLUS = 0.133509
Estimated KOLMOGOROV statistic DMINUS = 0.086947
Estimated overall statistic DN = 0.133509
Approximate significance level = 0.427464

TEST K.S

code: 15. 01. 08

Estimated KOLMOGOROV statistic DPLUS = 0.136178
Estimated KOLMOGOROV statistic DMINUS = 0.101348
Estimated overall statistic DN = 0.136178
Approximate significance level = 1

Code: 15. 01. 09

Estimated KOLMOGOROV statistic DPLUS = 0.137678
Estimated KOLMOGOROV statistic DMINUS = 0.144562
Estimated overall statistic DN = 0.144562
Approximate significance level = 1

Code : 15. 02. 02

Estimated KOLMOGOROV statistic DPLUS = 0.174372
Estimated KOLMOGOROV statistic DMINUS = 0.245251
Estimated overall statistic DN = 0.245251
Approximate significance level = 0.368831

code: 15. 02. 03

Estimated KOLMOGOROV statistic DPLUS = 0.116911
Estimated KOLMOGOROV statistic DMINUS = 0.085268
Estimated overall statistic DN = 0.116911
Approximate significance level = 1

code : 15. 10. 01

Estimated KOLMOGOROV statistic DPLUS = 0.167378
Estimated KOLMOGOROV statistic DMINUS = 0.12386
Estimated overall statistic DN = 0.167378
Approximate significance level = 0.999986

Code: 15. 06. 10

Estimated KOLMOGOROV statistic DPLUS = 0.147791
Estimated KOLMOGOROV statistic DMINUS = 0.12726
Estimated overall statistic DN = 0.147791
Approximate significance level = 0.999986

 * TEST DE TENDANCE (de kendall) *

STATION:AIN-ARNAT / CODE:150607

N = 31
 MOY= 3609.387 mm
 VAR= 1153429
 E.T.= 1073.978 mm

P(i)	U(ti)	U(ti)
4240	0	-.3909176
3081	-1	-.1605693
3543	-.5222329	-.4501935
3802	0	-.4741565
3105	-.489898	-.3543975
3760	-.1878673	-.5951219
5180	.7509392	-.5138093
3890	.9897433	-.0496088
4500	1.459601	.1320523
5025	1.878297	.4793648
2580	.8563488	.0266695
2555	0	.7786628
5897	.7321067	.4548124
3220	.3832142	1.174209
3245	.1484615	1.071008
2459	-.5402702	.9904954
4307	-.1647705	.6433331
1939	-.7954317	1.149643
1592	-1.364437	.7321067
5240	-.7137743	.1371451
3104	-.9059077	1.012049
1362	-1.438094	.6260991
3590	-1.294112	-.2085144
3513	-1.24022	-.2474358
4767	-.7940688	-.4505635
3375	-.8155374	.1878673
3287	-.8547233	0
3740	-.7112347	-.6793662
4887	-.300129	-1.566699
4176	-8.920515E-02	-1
2930	-.3909176	0

 * MATRICE DES COEFFICIENTS B (de la droite de regression) *

	st 1	st 2	st 3	st 4	st 5	st 6	st 7	st 8	st 9	st10	st11	st12	st13	st14	st15	st16	st17	st18	st19	st20	st21	st22	st23	st24	st25	st26	st27	st28	st29	st30	st31	st32	st3	
0101	.0																																	
0103	348.9	.0																																
0104	550.9-4117.7	.0																																
0108	1564.7	-670.3	2308.1	.0																														
0109	2289.0-2280.3	2405.6	1887.2	.0																														
0110	1336.0-1832.3-3910.7	436.9-2693.4	.0																															
0111	1838.8-2405.7	-286.5	1876.3	-91.4	2828.0	.0																												
0112	-8186.4-6174.5	9.8	4889.0	2644.1	2197.5	3512.7	.0																											
0201	2635.0	2539.3	3128.1	1770.2	1778.5	3257.2	2793.6	3589.9	.0																									
0202	2456.9	2261.4	1894.0	2057.6	1368.7	2458.5	1853.9	1728.6	1589.5	.0																								
0204	253.3	912.4	1765.8	2269.4	772.8	2233.7	1967.2	3528.3	1628.5	-839.4	.0																							
0301	2854.1	41.5	1037.1	1283.5-2566.7	2307.2	1910.5	2855.8	1805.1	2083.5	214.0	.0																							
0306	1651.3	1384.8	35.3	960.6-2072.6	905.5	2531.2	2584.7	3074.3	2604.9	651.6-1425.6	.0																							
0401	1802.3	1480.7	1961.0	2810.2	949.0	2150.2	2047.5	4066.3	1505.6-1068.0	469.1	3287.5	2519.3	.0																					
0402	1182.5	1813.8-1757.4	80.5	-950.7	314.1	-347.8	114.2	705.8	.8	184.3	741.4	985.7-1381.6	.0																					
0501	1821.3	1089.3	2025.4	1672.0	1770.9	3152.0	2531.8	4174.5	2375.3	2662.7	925.2	2163.9	4114.5	1013.4	3360.2	.0																		
0503	1039.7	2273.7	2570.3	1530.4	1343.0	3116.5	3121.7	3301.5	2882.2	2511.0	1447.2	4209.7	3802.2	1488.9	3468.7	1104.3	.0																	
0603	1500.9	1650.5	2453.1	2007.5	2196.0	2708.2	2205.1	2570.5	1402.3	2538.6	1925.7	3186.5	3534.6	1150.8	2701.9	1055.7	1045.6	.0																
0607	1519.8	1111.7	2195.3	2265.1	2175.9	3988.3	4391.0	2614.6	2181.8	2397.9	1129.9	3660.5	4280.6	1803.8	3348.4	-732.6	250.4	1840.0	.0															
0608	1675.2	3963.0	3605.7	4078.9	4018.5	4232.5	4323.3	3772.4	3630.7	3772.7	2184.0	4673.7	4388.0	3440.5	4161.4	2829.4	2557.6	2276.2	1908.7	.0														
0609	4015.4	3090.6	2892.9	2525.2	2494.8	3668.4	3418.3	4175.4	3199.5-2592.3	2094.3	3157.2	4354.8	2815.9	4084.1	2487.1	2261.3	3277.7	3317.4	2242.7	.0														
0610	1807.9	901.4	3167.3	2914.4	2954.8	3877.6	3988.3	3436.1	2051.6	4998.7	1672.9	3957.4	4918.5	2845.7	4396.1	1379.8	1936.3	1328.1	1175.4	-36.0	1300.5	.0												
0611	2443.1	6060.8	1402.9	2771.1	3363.8	2675.3	3319.0	2052.0	3223.4	4558.4	783.6	3781.4	2086.5	4882.0	3131.0	4937.9	1779.0	5300.4	3032.0	1182.6-4904.3	4493.2	.0												
0612	509.1	1933.2	1783.7	1957.1	1988.2	2691.0	2751.0	1868.5	1983.5	1023.8	1626.0	2136.9	3419.9	2095.9	2552.4	293.9	542.9	2388.5	1518.1	708.7	-158.2	1826.6	574.8	.0										
0613	2642.9	3036.6	1938.0	3839.1	3591.0	4809.2	3600.7	3849.3	3554.7	2147.8	3733.6	2692.7	4788.6	3254.6	3424.9	2308.8	1771.6	3790.0	3197.6	2659.1	445.8	2930.4	2727.2	495.1	.0									
0614	978.6	2496.5	1850.6	2299.8	1512.1	2404.9	2203.4	1872.1	2085.8	1101.8	876.0	1804.8	1923.5	2710.0	2323.8	2115.3	817.8	2006.9	983.8	-239.7	-398.5	2972.9	1241.2	1772.0	1756.5	.0								
0702	4062.1	-286.0	1428.4	1145.0	2083.6	3177.5	3070.6	2580.8	3317.8	6199.2	1635.1	3517.9	2837.2	2926.8	3881.9	982.1	533.3	3255.0	2126.5	2341.0-2570.7	1425.4	3141.2	1245.1	1738.8	2528.8	.0								
0703	6975.8	2548.0	1982.2	786.0	1801.4	3776.4	2982.8	4049.3	3047.3	3280.1	1940.1	478.5	4179.4	1770.5	3832.7	262.4	2523.4	2662.3	4582.4	5003.3-2307.6	4964.3	3407.2	3419.3	2174.1	2955.3	3590.5	.0							
0706	4794.7	4266.4	4116.8	5083.7	5909.0	5523.6	5185.5	4868.4	4641.1	3581.9	4935.7	2978.5	6083.7	4322.8	4895.6	3559.9	3912.7	2087.1	2759.1	2346.0	3832.6	3387.8	5269.8	3022.1	2230.4	4910.9	5421.7	5204.9	.0					
0707	1832.4	1756.6	2347.1	3401.2	3049.8	3339.0	3060.0	3598.5	2200.1	343.1	1775.8	2903.1	3699.8	2504.7	3268.6	1430.0	1251.1	2180.8	1695.6	831.5	3137.9	2030.8	3129.2	2055.4	1576.7	2243.9	1513.6	2269.9	1694.6	.0				
0708	4491.8	3856.7	1845.6	1592.6	1482.5	4185.5	3481.4	3709.2	2457.8	2682.1	2476.4	3056.4	4825.3	2518.0	4480.5	-154.5	1214.5	4971.8	3399.1	5950.4	-841.6	4857.8	3550.5	3041.5	2746.8	4413.3	2673.2	1762.7	3333.1	2575.7	.0			
0801	652.1	1459.8	2667.9	2183.5	1574.7	3041.9	3180.8	2529.9	1806.2	2106.4	1609.7	2775.4	2774.2	2029.6	2948.8	1312.5	321.9	1687.2	1547.8	622.7	833.4	2317.2	2134.1	2336.6	2145.0	1287.6	3206.4	2371.6	1642.8	99.4	4497.3	.0		
0802	4406.3-1407.1	945.1	2367.8	1828.7	3903.9	2969.1	4862.9	5076.0	4960.4	2454.1	1388.1	5174.2	2373.3	4133.9	809.0	1721.2	2002.6	4079.5	2373.6	-324.8	2724.4	5405.7	3963.6	2255.1	5958.2	4730.9	2022.9	2373.0	3514.1	4561.9	2796.6			

BIBLIOGRAPHIE

I- SOURCES ET DOCUMENTS (ANRH)

- Ressources en eaux de surface et possibilités de leur aménagements. BV de la Soummam. hydrotechnique corporation. Ingénieurs-conseil NEW-YORK / Madrid (1969).
- La carte de l'Algérie du Nord
- La carte pluviométrique de l'Algérie du Nord fait par Monsieur CHAUMONT (1913_1963).
- Les séries pluviométrique mensuelles et annuelles

II OUVRAGES ET ARTICLES SUR LA CRITIQUE DES DONNEES

ALLILI D. (1989) : étude générale pluviométrique (vérification, de la normalité des totaux annuels pluviométriques dans la région du SEBAOU). pub. int n° 070. 40 p.

ALLILI D. (1989) : vérification de l'homogénéité des séries pluviométriques du BV SEBAOU. pub.int. 58p.

ARLERY R. , GRISOLLET H. et GUILMET B. (1973) : Climatologie méthodes et pratiques 2 éd ; GAUTHIER - villars. Paris. p. 59 -72

BRUNET-MORET Y. (1980) : Retour sur l'homogénéisation des pluies annuelles par vecteur régional. La météorologie VI série n° 20-21. p.61-63

BRUNET-MORET Y. (1974) : Distribution gaussio-logarithmique.cah.O.R.S.T.O.M., série hydro.,vol XI, n°4.

DAGNELIE P. (1975) : Théorie et méthodes statistiques.application agronomique. tome II ; 2ème éd.

DUBREUIL P. (1974) : Initiation à l'analyse hydrologique. MASSON et CIE ; O.R.S.T.O.M., Paris.

JOSE ROSE (1988) : Initiation aux probabilités/ estimations/testes / sondages. presses universitaires de nancy. p.103 -111.

LABORDE J.P. (1991) : analyse et traitement des données pluviométriques. rapport d'activité- A.N.R.H.

LABORDE J.P. (1982) : Eléments hydrologie de surface. tome 2 : Critique et analyse statistique des données hydrologiques.

LANG-MICHANT C. (1990) : Pratique des tests statistiques, éd.DUNOD. p.30- 107.

MURRAY R. SPEIGEL (1984) : Théorie et application de la statistique série SCHAUM ; Paris.

NORDENSON T.J. (1971) : Préparation de cartes coordonnées des précipitations de l'écoulement et de l'évaporation. O.M.M., rapport n°6. p 2-3.

REMENIERAS G. (1980) : Hydrologie de l'ingénieur. 2 éd. revue et argumentée.Eyrolles; Paris. p 94-96.

ROCHE P.A : Guide de prévision de crues. tome I.p.92-122

SNEYERS R. (1975) : L'analyse statistique des séries d'observations. O.M.M., note tech. n° 143. p 6-84.

SNEYERS R. (1969) : On the climatological analysis of local series of observations, WMO technique. note n° 100. p.93-102.

III - OUVRAGES ET ARTICLES SUR L'APPLICATION DE L'ANALYSE EN COMPOSANTES PRINCIPALES (A.C.P.)

AÏSSANI B. : Application de l'ACP à la critique et à l'homogénéisation pluviométrique : exemple des hauteurs de pluies mensuelles sur le bassin de RHUMEL. thèse de D.E.A.

AURAY J.P. (1990) : Analyse des données multidimensionnelles.

BOUKHAROUBA M. et AGGAD N. (1990) : Etude hydrologique de régularisation du barrage de AIN - YUCEF sur l'oued SIKKAK. P.F.E. p. 8 -15

BOUROCHE J.M et SAPORTA G. (1983) : L'analyse des données. presses universitaires de France, collection que sais-je ? n° 1854.

DE LAGARDE J.(1983) : Initiation à l'analyse des données, éd. DUNOD

DIDAY et COLL (1989) : Classification automatique des données ; environnement statistique. DUNOD. p.196-211.

DIDAY E., LEMAIRE J., POUGET J. et TESTU F. (1982) : Eléments d'analyse des données DUNOD. p. 196-213.

DUBAND D. (1980) : Rationalisation d'un réseau de mesure de la précipitation. météorologie VI série n° 20 -21 .p.219 - 222.

ESCOFIER B. (1988) : Analyse factorielle simple. p.11 24.

FOUCART T. (1985) : L'analyse factorielle et programmation. 2^{ème} éd : MASSON.

LAADOUA A. (1987) : Les variations spatio - temporelles des précipitations au Maroc septentrional. Thèse 3^{ème} cycle, université des sciences et techniques de Lille.

LABORDE J.P. (1984) : Analyse des données et cartographie automatique en Hydrologie : Elément d'Hydrologie Lorraine. Thèse Doctorat es - science, I.N.P. de Lorraine.

PONTIER J. , DUFOUR A.B. et NORMAND M. (1990) : Le modèle euclidien en analyse des données, éd. S.M.A. p. 153 - 180.

TROUDE C. et PASSOUANT M. (1992) : Analyse des données. Manuel rédigé ; cycle S.E.A. p. 1 - 8.
