

Record Number: 880
Author, Monographic: Charbonneau, R.//Fortin, J. P.//Lardeau, J. P.//Morin, G.//Sochanska, W.
Author Role:
Title, Monographic: Analyse des précipitations du bassin versant de la rivière Eaton
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1978
Original Publication Date:
Volume Identification:
Extent of Work: viii, 103
Packaging Method: pages
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, Rapport de recherche
Series Volume ID: 88
Location/URL:
ISBN: 2-89146-086-3
Notes: Rapport annuel 1977-1978
Abstract: Rapport rédigé pour Environnement Canada
15.00\$
Call Number: R000088
Keywords: rapport/ ok/ dl

Analyse des précipitations du bassin
versant de la rivière Eaton

INRS-Eau
Université du Québec
C.P. 7500, Sainte-Foy
Québec G1V 4C7

RAPPORT SCIENTIFIQUE No 88
1978

Rapport rédigé pour
Environnement Canada

par

R. Charbonneau, J.P. Fortin, J.P. Lardeau, G. Morin, W. Sochanska

ISBN 2-89146-086-3

DEPOT LEGAL 1978

Tous droits de traduction, de reproduction et d'adaptation réservés

© 1978 - Institut national de la recherche scientifique

Analyse des précipitations du bassin versant de la rivière Eaton

Résumé

Le bassin versant de la Eaton, qui est équipé de 30 stations météorologiques, se prête à l'utilisation d'un modèle hydrologique. Toutefois, la précision de la simulation est très largement conditionnée par la qualité des mesures de précipitation, qu'il importe donc de critiquer rigoureusement. On s'intéresse plus particulièrement aux 14 stations qui disposent de 10 ans de relevés ininterrompus, soit du 1^{er} janvier 1966 au 31 décembre 1975. L'étude porte dans une première partie sur la qualité de la mesure elle-même: les observations sont-elles faites régulièrement, (y compris les jours fériés), avec ou sans arrondissement?

L'homogénéité spatiale des stations est analysée par la méthode des composantes principales, sur les précipitations de 10 jours, des quatre saisons: hiver, printemps, été et automne. On détermine ainsi plusieurs groupes homogènes de stations semblables sur les quatre saisons.

Les composantes principales nous fournissent l'information nécessaire pour limiter les redondances d'un réseau modifié.

L'homogénéité temporelle des précipitations de 10 jours est vérifiée pour l'ensemble des stations à l'aide de la méthode du cumul des résidus, avec pour station témoin, la projection de ces observations sur le premier axe fourni par les composantes principales. Trois stations

semblent douteuses.

La méthode d'interpolation optimale est utilisée pour déterminer l'erreur ponctuelle d'estimation des précipitations de 10 jours, et fait ressortir l'importance des erreurs de mesure et du micro-climat, particulièrement pour la saison "hiver".

L'étude de l'erreur d'estimation des précipitations de 10 jours sur la surface du bassin versant donne le nombre de stations nécessaires pour obtenir une précision donnée, pour chacune des saisons.

Mots-clé:

analyse, précipitation, réseau, composantes principales, interpolation optimum, erreur, cumul des résidus, Eaton, vérification de données, rivière, homogénéité spatiale, homogénéité temporelle, erreur d'estimation.

Référence:

CHARBONNEAU, R., FORTIN, J.-P., LARDEAU, J.-P., MORIN, G. et SOCHANSKA, W. (1978).
"Analyse des précipitations du bassin versant de la rivière Eaton".
INRS-Eau, rapport scientifique no. 88, 103 p., 2 annexes. (Rapport soumis à Environnement Canada).

TABLE DES MATIERES

	PAGE
RESUME.....	i
TABLE DES MATIERES.....	iii
LISTE DES TABLEAUX.....	v
LISTE DES FIGURES.....	vi
 CHAPITRE 1 - INTRODUCTION.....	 1
 CHAPITRE 2 - LE FICHIER DES DONNEES.....	 5
 CHAPITRE 3 - CONTROLE DE LA QUALITE DE LA MESURE.....	 11
3.1 Fréquence des événements selon les jours de la semaine.....	 13
3.1.1 Description de la méthode.....	13
3.1.2 Analyse des résultats.....	14
3.2 Occurrence des chiffres.....	16
3.2.1 Description de la méthode.....	16
3.2.2 Analyse des résultats.....	21
 CHAPITRE 4 - ETUDE DE LA VARIATION SPATIALE DES PRECIPITATIONS PAR LA METHODE DES COMPOSANTES PRINCIPALES.....	 23
4.1 Données utilisées.....	25
4.2 Application de la méthode des composantes prin- cipales.....	 26
4.3 Analyse des résultats.....	28
4.3.1 Période "HIVER".....	28
4.3.2 Période "PRINTEMPS".....	42
4.3.3 Période "ETE".....	46
4.3.4 Période "AUTOMNE".....	50
4.4 Conclusion.....	54

CHAPITRE 5 - CONTROLE DE SERIES CHRONOLOGIQUES DANS UNE REGION.	57
5.1 Description de la méthode "BOIS".....	57
5.1.1 Recherche d'une station témoin.....	60
5.1.2 Cumul des résidus.....	62
5.1.3 Comparaison avec la méthode classique des "Doubles masses".....	65
5.2 Analyse des résultats.....	66
CHAPITRE 6 - ERREURS D'ESTIMATION ASSOCIEES A L'INTERPOLATION DES PRECIPITATIONS.....	73
6.1 Données utilisées.....	76
6.2 Calcul de la fonction de structure.....	76
6.3 Interpolation optimale.....	77
6.4 Analyse des résultats.....	80
6.4.1 Période "HIVER".....	80
6.4.2 Période "PRINTEMPS".....	84
6.4.3 Période "ETE".....	86
6.4.4 Période "AUTOMNE".....	92
6.5 Erreur sur des surfaces.....	92
6.5.1 Analyse des résultats.....	98
6.6 CONCLUSION.....	101
BIBLIOGRAPHIE.....	103
ANNEXE A - METHODE DES COMPOSANTES PRINCIPALES.....	A-1
ANNEXE B - METHODE DE L'INTERPOLATION OPTIMUM.....	B-1

LISTE DES TABLEAUX

		<u>PAGE</u>
2.1	Stations météorologiques du bassin versant de la rivière Eaton.....	9
3.1	Variabilité des pluies selon les jours de la semaine...	15
3.2	Occurrence des chiffres.....	17
3.3	Occurrence des chiffres (station Maple Leaf).....	18
4.1	Corrélation entre composantes et station; saison "HIVER".....	29
4.2	Corrélation entre composantes et station après rotation; saison "HIVER".....	37
4.3	Corrélation entre composantes et station; saison "PRINTEMPS".....	43
4.4	Corrélation entre composantes et station; saison "ETE".	47
4.5	Corrélation entre composantes et station; saison "AUTOMNE".....	51
5.1	Résultats de l'analyse des 14 stations pour le cumul des résidus.....	71
6.1	Caractéristiques des précipitations de 10 jours. Bassin versant de la rivière Eaton.....	100
6.2	Erreur relative en % des précipitations de 10 jours en fonction du nombre de postes météorologiques.....	100

LISTE DES FIGURES

		<u>PAGE</u>
2.1	Emplacement sur le bassin versant de la rivière Eaton des 30 stations météorologiques.....	10
3.1	Pluie de 0.00 à 0.09 pouce.....	19
3.2	Pluie de 0.10 à 0.99 pouce.....	19
3.3	Distribution du chiffre des centièmes pour une pluie de 0.010 à 0.99 pouce	19
3.4	Fréquence d'observation des lames de pluie classées pour la station "Maple Leaf".....	20
4.1	Représentation des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours mesurées à chaque station pour la période "HIVER".....	32
4.3	Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "HIVER".....	34
4.4	Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2 ^e et 3 ^e composantes principales pour les groupes établis par la méthode varimax; saison "HIVER".....	38
4.5	Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2 ^e et 3 ^e composantes principales en changeant les groupes établis par la méthode varimax; saison "HIVER".....	41
4.6	Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "PRINTEMPS".....	44
4.7	Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2 ^e et 3 ^e composantes principales pour les groupes établis par la méthode varimax; saison "PRINTEMPS".....	45

4.8	Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "ETE".....	48
4.9	Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2 ^e et 3 ^e composantes principales pour les groupes établis par la méthode varimax; saison "ETE".....	49
4.10	Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "AUTOMNE".....	52
4.11	Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2 ^e et 3 ^e composantes principales pour les groupes établis par la méthode varimax; saison "PRINTEMPS".....	53
4.12	Répartition géographique des groupes de stations pour chaque des saisons.....	55
5.1	Sensibilité de la méthode (Multiplication par 1,1 de 1900 à 1930, $\bar{X} = 1000$, $\sigma = 160$).....	67
5.2	Double masse correspondant à la figure 5.1 (Multiplication par 1,1 de 1900 à 1930, $\bar{X} = 1000$, $\sigma = 160$).....	68
5.3	Présentation en fonction du temps du cumul des résidus.	69
5.4	Présentation en fonction du temps du cumul des résidus.	72
6.1	Fonction de structure dans l'espace des précipitations de 10 jours; saison "HIVER".....	81
6.2	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 30 stations; saison "HIVER".....	82
6.3	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 5 stations; saison "HIVER".....	83

6.4	Fonction de structure dans l'espace des précipitations de 10 jours; saison "PRINTEMPS".....	85
6.5	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 30 stations; saison "PRINTEMPS"....	87
6.6	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 5 stations; saison "PRINTEMPS".....	88
6.7	Fonction de structure dans l'espace des précipitations de 10 jours; saison "ETE".....	89
6.8	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 30 stations; saison "ETE".....	90
6.9	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 5 stations; saison "ETE".....	91
6.10	Fonction de structure dans l'espace des précipitations de 10 jours; saison "AUTOMNE".....	93
6.11	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 30 stations; saison "AUTOMNE".....	94
6.12	Ecart type (po.) d'estimation des précipitations de 10 jours. Réseau de 5 stations; saison "AUTOMNE".....	95
6.13	Exponentielle ajustée aux coefficients de corrélation des précipitations de 10 jours en fonction de la distance.....	99

CHAPITRE 1

INTRODUCTION

Dans le cadre du Programme de subvention à la recherche sur les ressources en eau, Environnement Canada a subventionné le projet ayant pour titre *"Utilisation des données du bassin représentatif de la rivière Eaton dans le but d'adapter le modèle déterministe CEQUEAU à l'analyse quantitative des répercussions sur l'écoulement, occasionnées par les modifications physiques des bassins versants"*.

La précision d'une simulation est fonction de la qualité du modèle hydrologique utilisé et de la qualité des données d'entrée. Il est par conséquent essentiel d'effectuer l'analyse critique des données météorologiques du bassin de la rivière Eaton afin d'identifier les limites au-delà desquelles il est inutile sinon désavantageux de raffiner la précision des simulations.

Comme les anomalies sont très fréquemment liées aux données de précipitation, nous nous attacherons, dans ce rapport à l'analyse de ces données.

Nous utiliserons les trois méthodes complémentaires suivantes:

- Contrôle de la qualité de la mesure;
- Analyse en composantes principales;
- Contrôle de séries chronologiques (cumul des résidus).

L'étude se termine par une application de la méthode d'interpolation optimale de L.S. Gandin (1970) dont le but est de calculer la distribution spatiale des erreurs d'interpolation et des quantités de précipitation. Cette méthode est également utilisée pour calculer l'influence d'une réduction du nombre de postes météorologiques, sur l'estimation de la précipitation des bassins versants.

CHAPITRE 2

LE FICHIER DES DONNEES

Le fichier obtenu du Service de Météorologie du ministère des Richesses naturelles est constitué d'"images de carte" sur bande magnétique. Chaque "image de carte" contient:

- l'identification de la station;
- l'identification de la date (an, mois, jour);
- pluie du jour;
- code - pluie (estimée, cumulée ou supérieure à 10 pouces);
- neige du jour;
- code - neige (estimée ou cumulée);
- précipitation totale du jour;
- code - précipitation (estimée, cumulée ou supérieure à 10 pouces);
- température maximum;
- code - maximum (estimée);
- température minimum;
- code - minimum (estimée).

Quant au reste des informations, elles ont trait au vent ou à l'humidité relative. Dans les champs réservés aux précipitations (pluie, neige et précipitation totale) il peut exister une double perforation pour signifier:

- --- un manque
- T une trace

- C précipitation avec quantité inconnue;
- L présence de précipitation inconnue; cumulée sur le Lendemain.

A partir de ce fichier de base, nous avons formé un fichier de travail où seules les données de précipitation et de température apparaissent, sans les "codes" et les doubles perforations.

La liste des 30 stations météorologiques et les périodes d'observations sont présentées à la figure 2.1. Comme les données n'étaient pas toutes disponibles sur support informatique et compte tenu du temps nécessaire à la perforation de ces données à partir des manuscrits originaux, nous avons fait perforer en priorité les 4 stations suivantes: Clifton River South, Eaton River North, Sherman Brook North et Birchton, de façon à mieux couvrir l'ensemble du bassin. La figure 2.1 montre l'emplacement sur le bassin versant de la rivière Eaton des 30 stations météorologiques.

TABLEAU 2.1 Stations météorologiques du bassin versant de la rivière Eaton.

NUMERO FEDERAL	NOM	PRECIPITATION	TEMPERATURE	PERIODE																							
				52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
702BC0E	Eaton	x														----- NON DISPONIBLE SUR SUPPORT INFORMATIQUE											
702GGR7	Sawyer Brook	x														-----											
702HACH	Sherman Brook North	x														-----											
702HBHA	Statton Brook North	x														-----											
702HBHG	Sucker Pond North	x														-----											
702K302	Eaton River North	x														-----											
7020785	Birchton	x														-----											
7021460	Christmas Brook North	x														-----											
7021718	Clifton River	x														-----											
7021719	Clifton River South	x														-----											
7021918	Cookshire	x	x													-----											
7022288	East Clifton North	x														-----											
7022306	Eaton Second Branch	x	x													-----											
7022374	Flanders	x														-----											
7023312	Island Brook	x	x													-----											
7023313	Island Brook North	x														-----											
7024263	Lawrence	x	x													-----											
7024270	Learned Plain	x														-----											
7024623	Maple Leaf	x	x													-----											
7024624	Maple Leaf East	x	x													-----											
7024625	Maple Leaf South	x														-----											
7025425	New Mexico	x														-----											
7026005	Petit Lac	x														-----											
7026048	Pine Brook North	x														-----											
7026345	Randboro South	x														-----											
7027372	St-Isidore d'Auckland	x	x													-----											
7027550	St-Mathias	x														-----											
7027800	Sawyerville	x														-----											
7027802	Sawyerville North	x	x													-----											
7028906	West Ditton	x	x													-----											

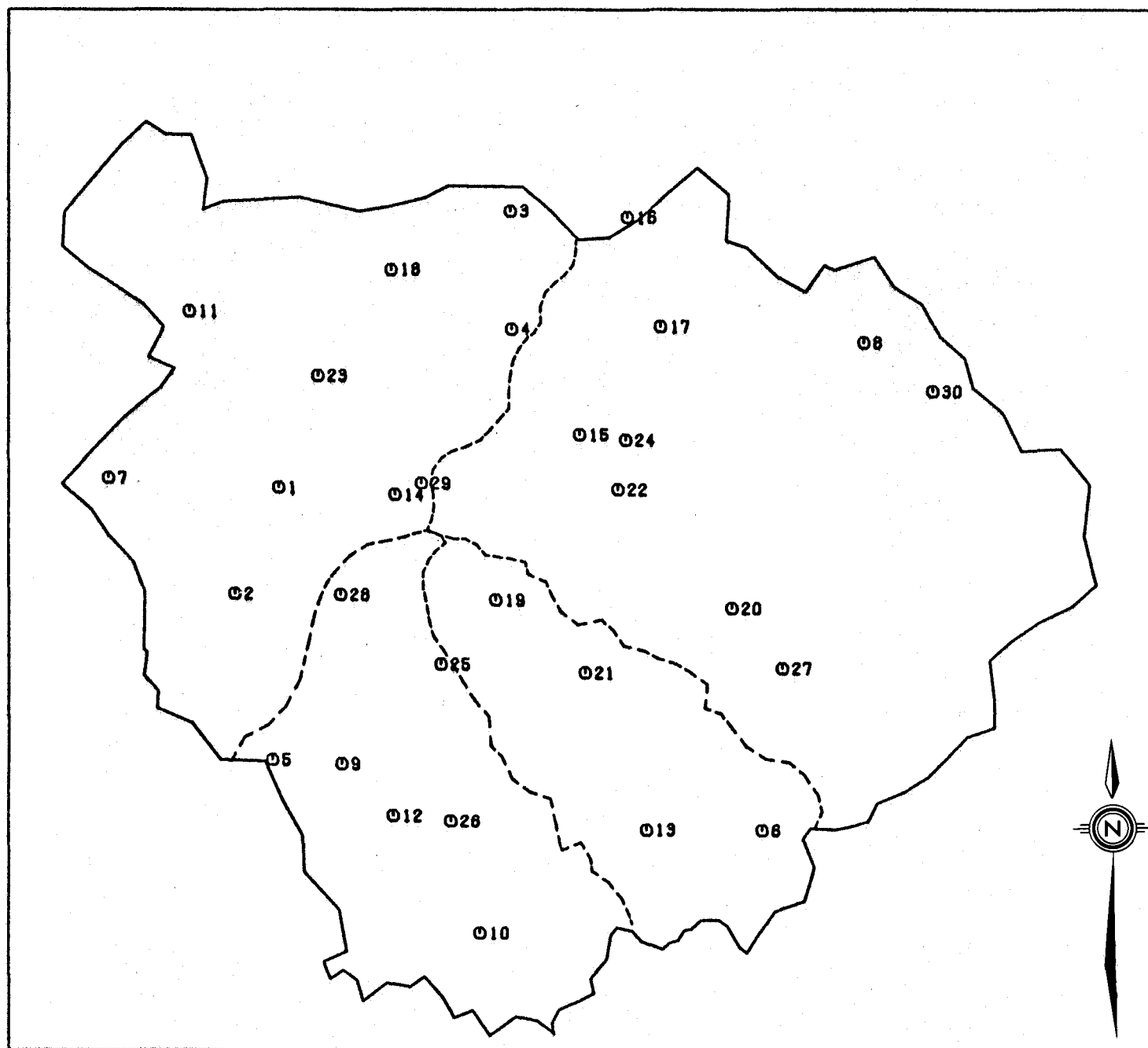


FIGURE 2.1 Emplacement sur le bassin versant de la rivière Eaton des 30 stations météorologiques.

CHAPITRE 3

CONTROLE DE LA QUALITE DE LA MESURE

Les deux méthodes proposées dans ce chapitre permettent de vérifier d'une part, avec quelle précision l'observateur effectue ses mesures et d'autre part, sa ponctualité; c'est-à-dire sa tendance ou non à cumuler les précipitations certains jours particuliers de la semaine afin d'éviter le dérangement occasionné par le relevé de la mesure.

3.1 FREQUENCE DES EVENEMENTS SELON LES JOURS DE LA SEMAINE

Les événements météorologiques ne sont aucunement liés aux jours de la semaine; ils peuvent tout aussi bien se produire un lundi ou un jeudi, par exemple. L'observateur, pour sa part, a une façon de vivre qui est liée à certains jours de la semaine. Prend-il la peine d'aller lire les instruments, les jours de congés? S'il y a pluie, il laisse peut-être s'accumuler l'eau et fait une seule lecture totale. La méthode est d'autant plus précise que la série de mesures est longue, à la condition que l'observateur soit resté le même et que la localisation de la station soit demeurée inchangée.

3.1.1 Description de la méthode

Il suffit de cumuler les valeurs mesurées d'un phénomène (pluie, neige, température) pour chacun des jours de la

semaine. Statistiquement la distribution du nombre d'événements ainsi que les quantités mesurées suivent une distribution uniforme.

3.1.2 Analyse des résultats

Les résultats présentés au tableau 3.1 montrent que sur le bassin étudié, les observateurs météorologiques semblent consciencieux pour ce qui a trait à la régularité des prises de mesures. Toutefois, il faut noter que seulement 14 stations ont plus de 3 ans d'observations.

La fréquence des pluies est en règle générale faible le mercredi et maximum le jeudi; si l'on cumule les nombres de cas observés pour ces deux journées, le mercredi totalise 1550 cas alors que le jeudi en a 1902, soit un accroissement de 23%. Les quantités de pluie sont également minimales le mercredi avec une moyenne de 0.29 pouce par jour de pluie alors que le samedi enregistre en moyenne 0.34 pouce, soit une augmentation de 17%. Ces constatations n'ont aucun support physique et il est difficile de trouver l'explication au niveau des observateurs puisque le phénomène est général sur tout le bassin. Il faut donc admettre que sur le bassin de la rivière Eaton, le mercredi soit un jour légèrement privilégié durant la période retenue.

TABLEAU 3.1 Variabilité des pluies selon les jours de la semaine

STATION	PLUIE MOYENNE EN POUCE, NOMBRE DE CAS							PERIODE
	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche	
7026COE	.33 27	.39 29	.25 24	.32 31	.33 30	.44 28	.24 25	1973 à 1975
702GGR7	.13 9	.21 8	.28 6	.24 10	.31 9	.26 13	.17 7	1973
702HACH	.36 93	.34 82	.34 80	.34 102	.34 95	.44 88	.35 83	1965 à 1975
702HSHA	.35 30	.41 28	.28 24	.36 33	.38 34	.42 29	.28 27	1973 à 1975
702HBHG	.35 28	.52 27	.25 27	.35 30	.31 31	.40 27	.24 23	1973 à 1975
702K3OE	.40 89	.33 89	.32 80	.34 104	.38 100	.36 92	.29 84	1965 à 1975
7020785	.39 95	.32 91	.28 84	.35 106	.32 104	.35 97	.29 94	1965 à 1975
7021460	.36 20	.39 19	.29 23	.44 31	.25 29	.44 21	.31 21	1973 à 1975
7021718	.31 28	.52 28	.20 26	.31 32	.32 34	.39 26	.22 22	1973 à 1975
7021719	.42 84	.29 86	.33 76	.33 90	.41 91	.41 76	.32 73	1965 à 1975
7021918	.35 101	.30 102	.29 90	.35 107	.32 111	.37 105	.27 103	1964 à 1975
7022288	.33 29	.51 26	.26 22	.37 25	.37 30	.39 24	.15 19	1973 à 1975
7022306	.38 87	.31 87	.30 76	.28 101	.34 94	.34 90	.26 87	1966 à 1975
7022374	.37 28	.56 24	.30 22	.39 31	.50 27	.41 21	.53 18	1973 à 1975
7023312	.40 89	.33 88	.30 82	.36 101	.36 104	.40 89	.31 89	1965 à 1975
7023313	.35 20	.31 19	.31 18	.40 20	.31 20	.69 16	.44 16	1973 - 1974
7024263	.50 85	.29 85	.29 77	.34 96	.34 101	.36 89	.29 88	1966 à 1975
7024270	.36 25	.41 29	.22 28	.37 31	.31 25	.34 24	.28 25	1973 à 1975
7024623	.42 87	.36 85	.31 80	.36 95	.35 100	.38 84	.28 84	1965 à 1975
7024624	.39 99	.31 96	.26 94	.33 109	.37 103	.35 98	.26 96	1965 à 1975
7024625	.25 8	.50 9	.17 9	.17 9	.35 12	.31 9	.11 8	1973
7025425	.39 20	.42 28	.28 23	.37 23	.37 26	.35 26	.28 25	1973 à 1975
7026005	.36 26	.47 24	.22 25	.37 30	.42 32	.38 21	.22 22	1973 à 1975
7026345	.58 4	.36 4	.16 4	.42 3	.38 6	.40 4	.21 4	1973
7027372	.40 101	.31 101	.28 90	.30 109	.31 109	.38 102	.25 89	1964 à 1975
7027550	.46 25	.52 23	.26 23	.39 28	.39 23	.37 22	.27 22	1973 à 1975
7027800	.41 162	.46 144	.37 127	.35 167	.36 170	.41 158	.41 150	1952 à 1975
702802	.33 129	.35 117	.28 108	.30 133	.32 131	.31 124	.26 118	1961 à 1975
7028906	.37 104	.30 96	.27 102	.33 115	.33 109	.33 104	.27 104	1965 à 1975

3.2 OCCURRENCE DES CHIFFRES

Même si un appareil de mesure est susceptible d'être lu avec une précision donnée, parfois un observateur, pour différentes raisons, le lira avec une autre précision. Au Canada, la quantité d'eau qui tombe sous forme de pluie peut être lue au centième de pouce. La courbe théorique de distribution des fréquences d'observation des lames de pluie variant entre 0 et 9 centièmes de pouce (nombres à 1 chiffre) montre une décroissance exponentielle. Dans les cas de nombres à 2 chiffres, la courbe de distribution de fréquence du chiffre des dizaines est aussi une exponentielle décroissante alors que la distribution de fréquence du chiffre des centièmes a une distribution théorique uniforme. Un raisonnement similaire s'applique aux nombres à 3 ou 4 chiffres.

3.2.1 Description de la méthode

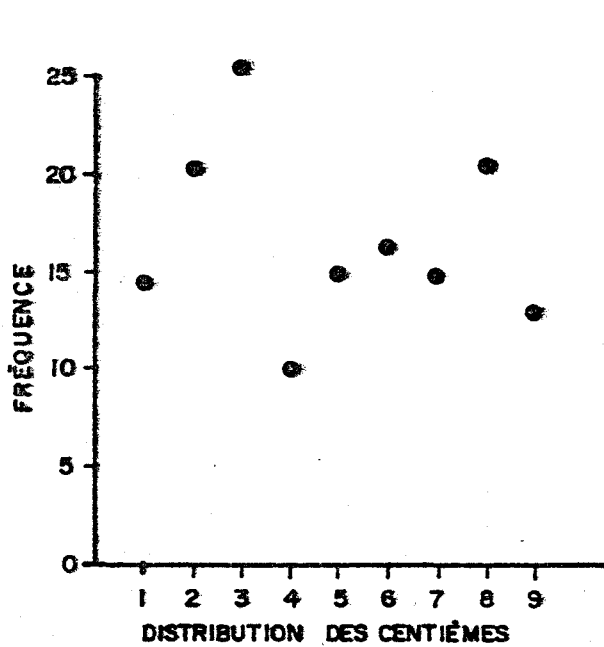
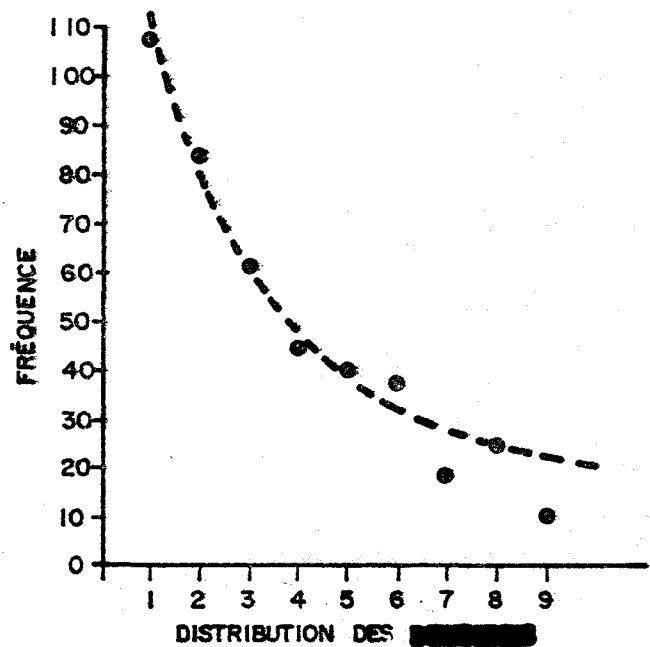
Dans le but de clarifier la procédure, il est pratique de raisonner sur un exemple. La précipitation est mesurée au centième de pouce. La méthode isole les cas où il est tombé entre 0 et 9 centièmes de pouce d'eau et analyse la fréquence d'occurrence de chacun des chiffres (0, 1, 2, ..., 9). De même, les nombres à 2 chiffres sont isolés soit de 10 à 99 centièmes de pouce et on analyse alternativement la fréquence d'occurrence des dixièmes et des centièmes.

TABLEAU 3.2 Occurrence des chiffres.

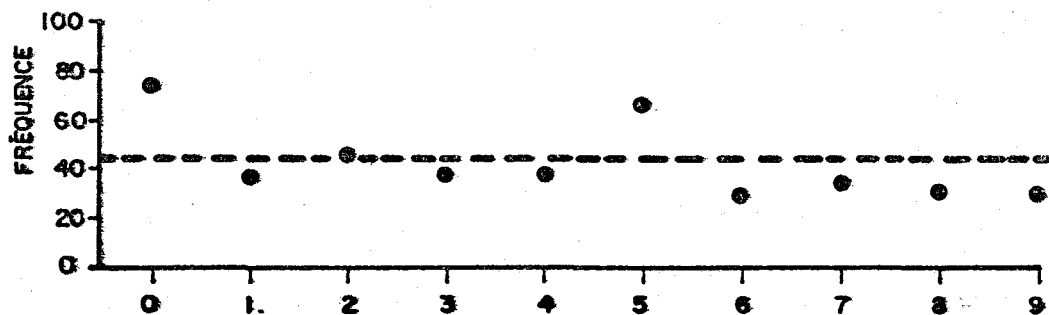
STATION		TENDANCE A L'ARRONDISSEMENT					REMARQUE PARTICULIERE
Numéro	Nom	Très fort	Fort	Moyenne	Faible	Très faible	
7021918	Cookshire					X	.01 pouce n'est pas souvent observé
7022306	Eaton Second Branch					X	
7023312	Island Brook					X	
7024263	Lawrence			X			Décroissance exponentielle pas bien marquée
7024623	Maple Leaf		X				Décroissance exponentielle pas bien marquée
7024624	Maple Leaf East					X	
7027372	Saint-Isidore				X		
7027800	Sawyerville			X			Très peu d'observations de .01 à .04 pouce
7027802	Sawyerville Nord					X	
7028906	West Ditton					X	
702HACH	Sherman Brook North		X				.39 pouce jamais observé
702K302	Eaton River North			X			
7020785	Birchton					X	
7021719	Clifton River South				X		

TABLEAU 3.3 Occurrence des chiffres (station Maple Leaf).

NOMBRES	0	1	2	3	4	5	6	7	8	9	
1 chiffre											
1	863	14	21	26	10	15	17	15	21	13	0.00 à 0.09
2 chiffres											
1	0	104	84	62	45	40	38	19	24	10	0.10 à 0.99
2	75	39	46	39	39	64	30	35	32	31	
3 chiffres											
1	0	27	4	0	0	0	0	0	0	0	1.00 à 9.99
2	6	6	9	3	2	0	1	1	2	1	
3	5	2	4	3	0	6	0	6	1	4	

3.1 ~~TABLEAU DE 0.00 à 0.00 pouce~~3.2 ~~TABLEAU DE 0.10 à 0.00 pouce~~

NOTE : Les deux distributions de fréquences ci-haut devraient être des exponentielles décroissantes alors que la distribution ci-bas devrait être uniforme.



3.3 DISTRIBUTION DE CHIFFRE DES CENTIÈMES POUR UNE POUCE DE 0.10 à 0.00 pouce

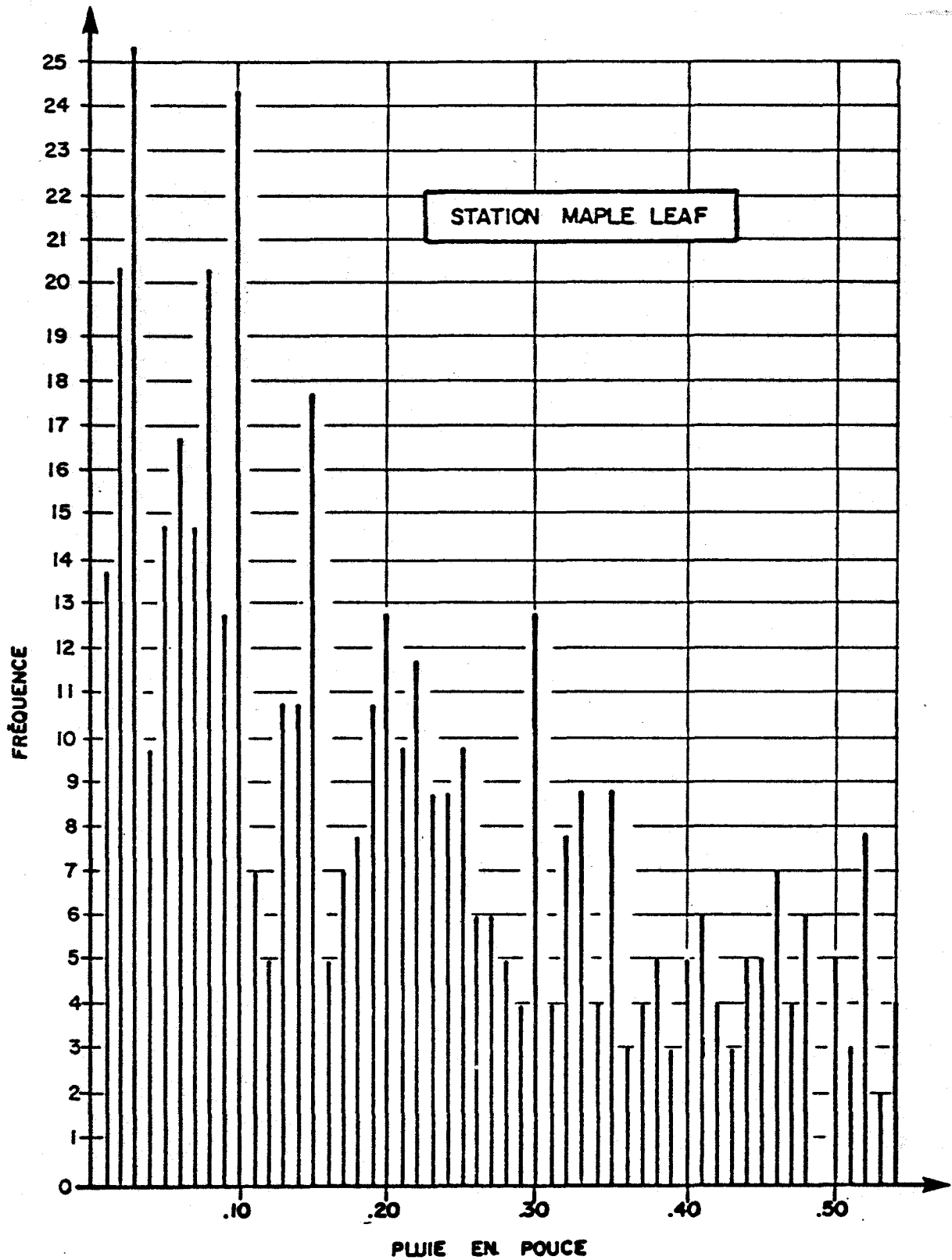


FIGURE 3.4 Fréquence d'observation des lames de pluies classées pour la station "Maple Leaf".

Une seconde partie du programme trace l'hystogramme des fréquences d'occurrence des quantités décroissantes de la précipitation.

Ces tests très simples sont révélateurs des tendances à l'arrondissement des lectures effectuées par certains observateurs.

3.2.2 Analyse des résultats

L'ensemble des stations pluviométriques est en général très bon en regard de l'occurrence des chiffres. Le tableau 3.2 résume les résultats obtenus aux 14 stations où nous possédons de longues séries d'observations puisqu'avec de courtes séries, la méthode n'est pas précise. Nous incluons, à titre d'exemple, le tableau 3.3 et les figures correspondantes (3.1 à 3.3) de même que la figure 3.4 se rapportant à la station "Maple Leaf".

Il ressort de ce test que la station "Maple Leaf" serait plus critiquable que les autres stations du bassin, principalement en ce qui a trait à la décroissance des précipitations de .01 à .09. De plus, on observe à la figure 3.4 que les observations sont arrondies aux "0" ou au "5".

La station "Sawyerville" n'enregistre presque jamais de quantités de pluie inférieures à 0.04 pouce, ce qui est anormal.

CHAPITRE 4

ETUDE DE LA VARIATION SPATIALE DES PRECIPITATIONS
PAR LA METHODE DES COMPOSANTES PRINCIPALES

L'étude du comportement des stations les unes par rapport aux autres va nous permettre de caractériser le réseau actuel et d'en faire ressortir, s'il y a lieu, les redondances. Cette étude est réalisée en recourant à la méthode des composantes principales, décrites en annexe A.

4.1 DONNEES UTILISEES

Pour appliquer la méthode des composantes principales, qui utilise les corrélations entre les stations prises deux à deux, il est nécessaire de s'assurer une période d'observation concomitante suffisamment longue pour que les coefficients de corrélation soient significatifs.

Dans notre cas on a choisi d'utiliser les 14 stations pour lesquelles les données sont disponibles sur support informatique pour au moins 10 ans (tableau 2.1).

On a choisi d'étudier le réseau pour les précipitations de 10 jours, d'autres périodes pourraient être utilisées (voir CONCLUSION, page 56). Pour travailler sur des données homogènes, on a fait l'étude de quatre périodes, soit l'hiver, le printemps, l'été et l'automne. Chaque saison nous donne 8 ou 9 observations, de 10 jours distribuées de la façon suivante:

<u>SAISONS</u>	<u>NOMBRE DE PERIODES</u>	<u>JOUR DU DEBUT</u>
Hiver	8	1 ^{er}
Printemps	8	75 ^e
Eté	9	150 ^e
Automne	9	240 ^e

Notons qu'il existe une superpositions de 5 jours au passage hiver-printemps et printemps-été.

4.2 APPLICATIONS DE LA METHODE DES COMPOSANTES PRINCIPALES

Les données observées aux 14 stations sont regroupées dans la matrice X de dimension $14 \times n$ ($n = \text{nombre d'années} \times 8$ ou 9 observations selon la saison). La matrice y des observations centrées réduites s'obtient ensuite par:

$$y_{ij} = \frac{x_{ij} - \bar{X}_j}{\sigma_j}$$

où

x_{ij} = variable observée

$\bar{X}_j$ = moyenne de la $j^{\text{ième}}$ station

σ_j = écart type de la $j^{\text{ième}}$ station

De la matrice y nous obtenons facilement la matrice des coefficients de corrélation:

$$r_{jk} = \frac{1}{n} \sum_{i=1}^n y_{ij} y_{ik}$$

où

r_{jk} est le coefficient de corrélation entre les stations j et k .

y_{ik} les observations centrées réduites

Les coefficients c_{jk} , qui sont appelés cosinus directeurs, sont les coordonnées des vecteurs propres de la matrice des coefficients de corrélation et sont obtenus en diagonalisant cette matrice à l'aide de la sous-routine EIGEN de IBM. Les valeurs propres λ_j également obtenues par la sous-routine EIGEN, sont les variances des nouvelles variables et sont calculées par ordre décroissant. Finalement la corrélation entre la station j et la composante principale d'ordre k nous est donné par:

$$r_{jk} = c_{jk} \sqrt{\lambda_k}$$

L'étude de ce coefficient de corrélation nous renseigne sur le comportement des stations.

4.3 ANALYSE DES RESULTATS

4.3.1 Période "HIVER"

Une première analyse en composantes principales des précipitations cumulées sur 10 jours a été effectuée pour la saison "hiver". Les résultats en sont résumés au tableau 4.1.

Ce tableau nous montre que la première composante explique plus de 75% de la variance et qu'avec les trois premières composantes on explique 84.5% de la variance totale des observations, ce qui est un indice d'homogénéité et de redondance du réseau étudié (14 stations) pour les précipitations de 10 jours.

Si l'on compare les coefficients de corrélation des stations avec la première composante principale on voit que l'influence de cette composante n'est pas la même pour toutes les stations. Par exemple, la variance expliquée par la première composante est de .910 pour la station 7027802 et de seulement .619 pour la station 7020785.

Les valeurs des coefficients de corrélation ou de variance expliquée varient d'une composante à l'autre et dépendent des observations. Si le coefficient de corrélation entre 2 stations est égal à 1 (sans que les stations aient nécessairement une même moyenne ou un même écart type) les coefficients de corrélation entre les

TABLEAU 4.1 Corrélation entre composantes et station;
saison "HIVER".

IDENTIFICATION DE LA STATION	NO.	COMPOSANTES			COEFFICIENT DE CORRELATION MULTIPLE
		No. 1	No. 2	No. 3	
702HACH	1	.834	-.314	-.310	.890
702K30E	2	.835	.393	-.062	.856
7020785	3	.787	-.453	.215	.871
7021719	4	.796	.431	-.145	.841
7021918	5	.890	-.191	-.194	.867
7022306	6	.801	.005	.451	.847
7023312	7	.912	-.050	-.015	.834
7024263	8	.846	.039	-.240	.775
7024623	9	.843	.140	.132	.748
7024624	10	.912	.093	.110	.860
7027372	11	.873	.110	.185	.809
7027800	12	.902	-.158	-.041	.841
7027802	13	.954	-.022	-.042	.913
7028906	14	.936	-.012	-.013	.877
Valeurs propres		10.544	.751	.535	
Variance expliquée		75.3%	5.4%	3.8%	
Variance totale expliquée		75.3%	80.7%	84.5%	

composantes et ces deux stations seront égaux.

Comme on observe rarement des corrélations parfaites, même pour deux stations météorologiques voisines, par suite d'erreurs de mesure et de conditions micro-climatiques locales, on doit rechercher les groupes homogènes en associant les stations de comportement semblable sur l'ensemble des premières composantes.

Le nombre de composantes à considérer ne suit pas de règle rigide. En général, on tient compte de la variance expliquée par la composante et de la variance totale expliquée. Dans ce cas-ci on considère les trois premières composantes.

Il existe deux méthodes simples de présenter le tableau 4.1 sous forme de figure de manière à déterminer les groupes de stations dont le comportement est semblable. La première est la représentation dans l'espace à trois dimensions des corrélations entre les stations et les trois premières composantes. La deuxième méthode, qui est moins employée mais plus discriminante et qui contient plus d'informations à notre point de vue, est une projection, dans le plan des composantes 2 et 3, des corrélations avec les autres composantes. Contrairement à la première méthode cette deuxième méthode peut présenter plus de trois composantes simultanément.

Les résultats pour la saison "hiver" sont présentés et discutés avec les deux formes de graphiques pour qu'on puisse se

rendre compte des avantages de l'une sur l'autre.

La figure 4.1 est la présentation en trois dimensions du tableau 4.1. Les axes A, B et C représentent respectivement les composantes 1, 2 et 3. L'origine des segments de droite qui représentent chacun une station a pour coordonnées les corrélations de la station avec les composantes deux et trois, tandis que la longueur du segment de droite représente la corrélation de la station avec la première composante. Au vu de la figure 4.1 les stations qu'on peut grouper sont:

- A) 9, 10 et 11
- B) 7, 13 et 14
- C) 4 et 2

La figure 4.2 explique comment on met en graphique les données du tableau 4.1 selon la seconde méthode. Comme dans le cas précédent, on détermine tout d'abord le point A d'un segment de droite $\vec{AB}$ en prenant comme coordonnées sur les axes R2 et R3 (plan des composantes 2 et 3) les valeurs respectives des coefficients de corrélation entre la station et les composantes deux et trois. La longueur du segment $\vec{AB}$ est égale à la valeur prise par le coefficient de corrélation entre la station et la première composante principale. Ce segment $\vec{AB}$ est construit dans le plan des axes R2 et R3 perpendiculairement au segment $\vec{OA}$ (orthogonalité des composantes principales). Enfin, si par convention, la corrélation est

1 = 702HACH	8 = 7024263
2 = 702K30E	9 = 7024623
3 = 7020785	10 = 7024624
4 = 7021719	11 = 7027372
5 = 7021918	12 = 7027800
6 = 7022306	13 = 7027802
7 = 7023312	14 = 7028906

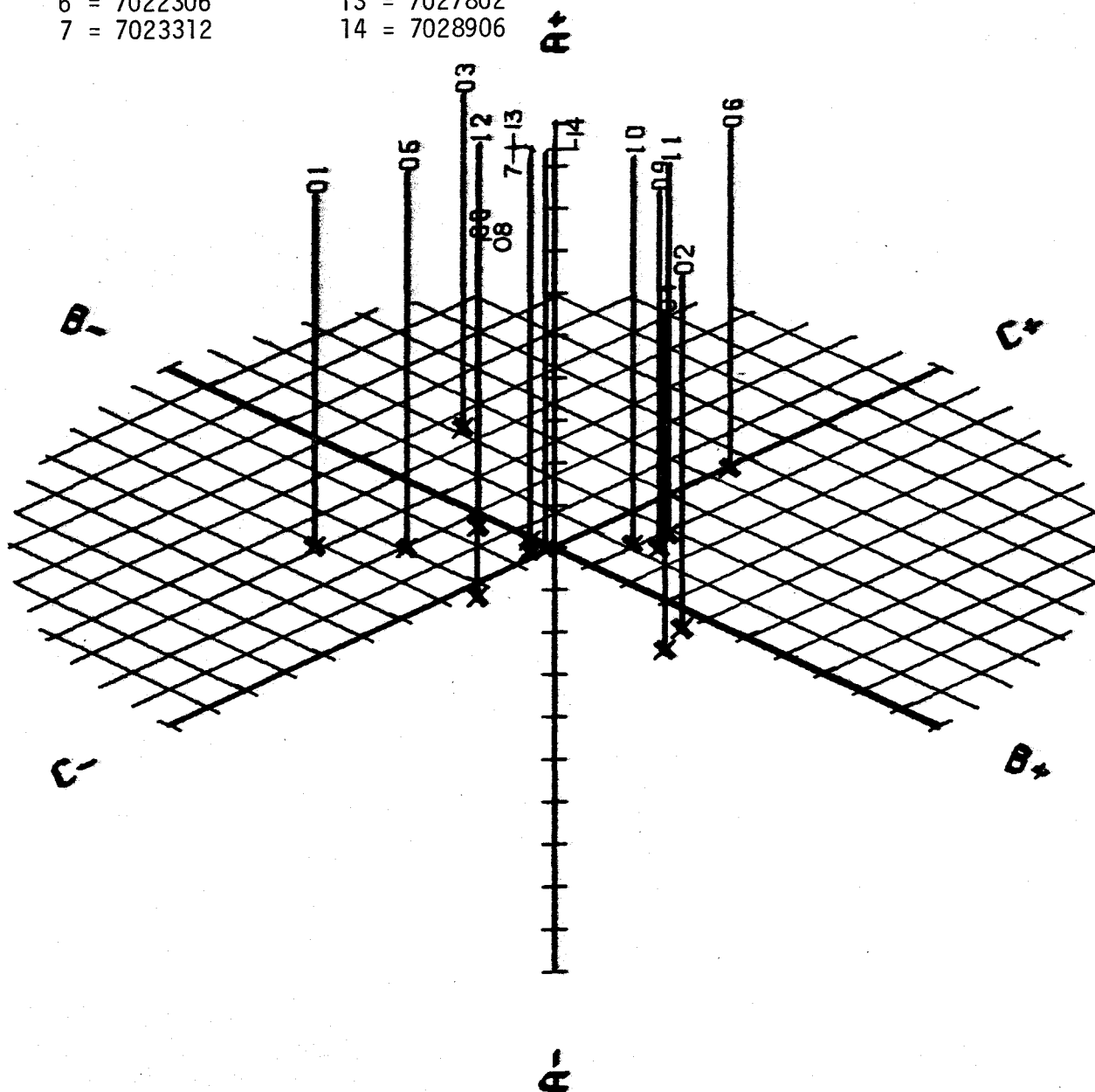
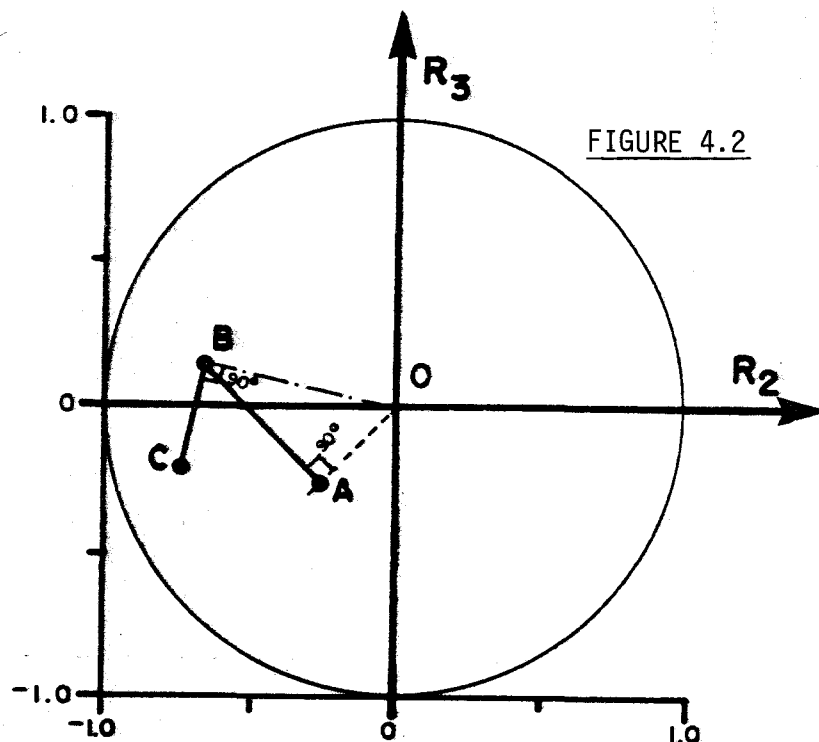


FIGURE 4.1 Représentation des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours mesurées à chaque station pour la période "HIVER".

positive, on trace le segment $\overrightarrow{AB}$ dans le sens des aiguilles d'une montre et dans le sens contraire si elle est négative. Le segment $\overrightarrow{OB}$ est le coefficient de corrélation multiple entre la station et les trois premières composantes. On peut tracer ainsi toutes les composantes. Cependant compte tenu du peu de variance expliquée par les dernières composantes on trace rarement plus de 6 composantes.



La figure 4.3 permet de trouver les groupes de stations par la proximité des segments et leur pente. Elle nous montre également l'importance des premières composantes et le nombre minimum utile pour la reconstitution des observations à chaque station.

1 = 702HACH	8 = 7024263
2 = 702K30E	9 = 7024623
3 = 7020785	10 = 7024624
4 = 7021719	11 = 7027372
5 = 7021918	12 = 7027800
6 = 7022306	13 = 7027802
7 = 7023312	14 = 7028906

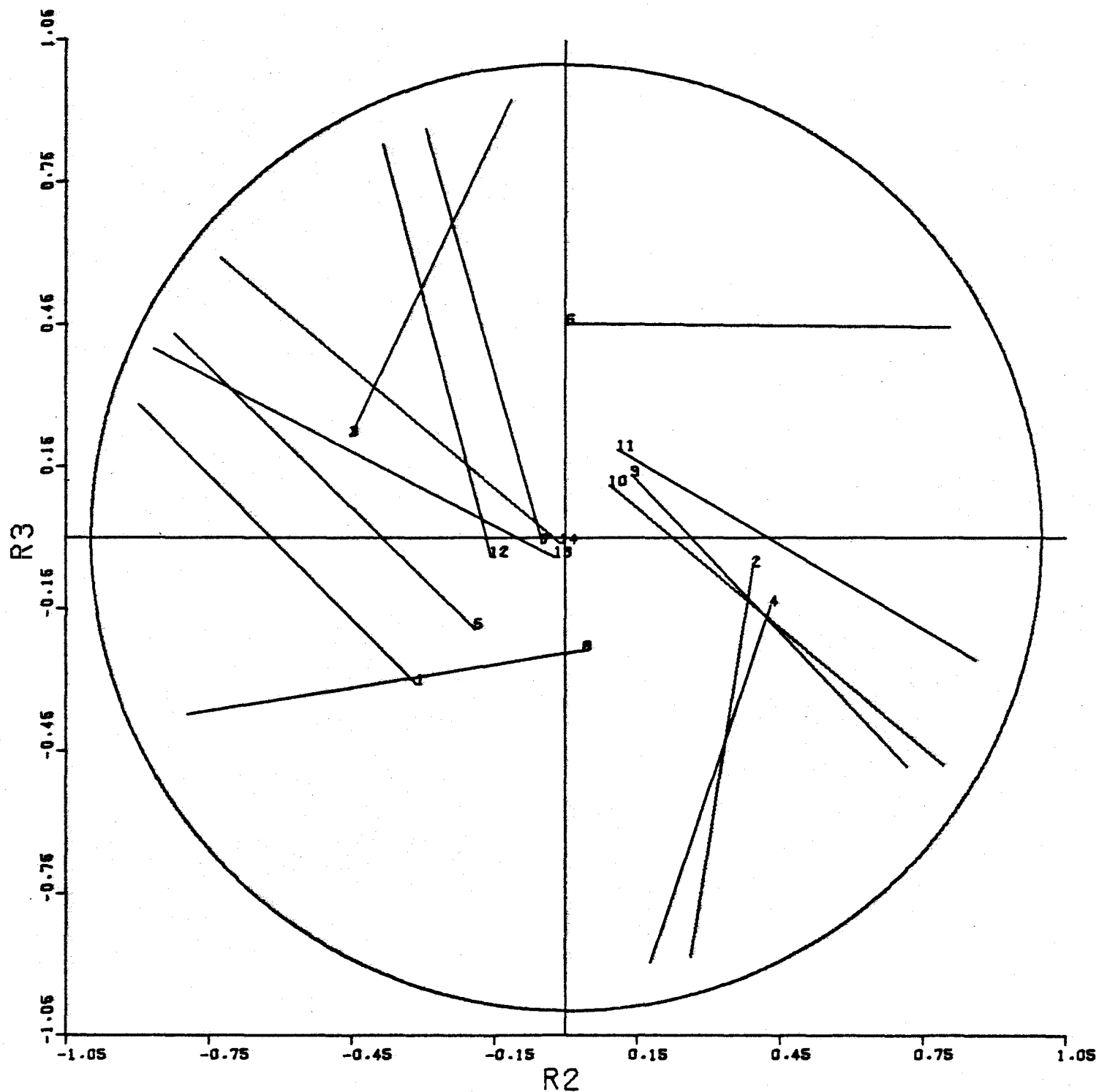


FIGURE 4.3 Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "HIVER".

Rappelons en effet que l'on peut reconstituer les observations pour chaque station à l'aide des K premières composantes. Cette reconstitution des observations sera d'autant meilleure que le coefficient de corrélation multiple s'approchera de 1, ce qui nous est indiqué sur le graphique par la proximité de l'extrémité de la ligne brisée relative à la station avec le cercle qui correspond à une corrélation de 1.

L'étude de la figure 4.3 nous montre qu'avec 3 composantes les observations de toutes les stations peuvent être bien reproduites (proximité de la fin de chaque segment avec le cercle). On groupe comme précédemment les stations 9, 10 et 11. Les stations 7, 12, 13 et 14 se groupent ainsi que 2 et 4. Les stations 1 et 5 peuvent à la limite se grouper (la pente du premier segment tient compte simultanément des composantes 2 et 3).

Les groupes de stations ont été identifiés sur les graphiques par des points, ou segments de droite, voisins indiquent des contributions identiques sur le calcul des composantes.

On peut, si on le désire, faire ressortir des groupes en faisant subir une rotation aux premiers axes principaux.

Pour les précipitations d'hiver, par exemple si l'on fait cette rotation (on utilise la méthode varimax*) avec les

*Statistical Package for the social sciences - McGraw-Hill, 1975.

3 premiers axes principaux, la variance expliquée sur l'ensemble demeure 84.5% mais est distribuée différemment sur les 3 premiers axes. La rotation étant faite on détermine les groupes en faisant ressortir pour chaque station l'axe de contribution maximum. Le tableau 4.2 montre les valeurs obtenues après rotation et les stations soulignées sur une même colonne forment un groupe. On a les groupes suivants:

- A) 1, 5, 7, 8, 12, 13 et 14
- B) 2, 4, 9 et 10
- C) 3, 6 et 11

La rotation des axes est faite après avoir imposé le nombre d'axe que l'on désire garder. Dans ce cas-ci on a imposé la rotation en ne conservant que les trois premiers axes. Il est évident que si l'on impose par exemple la rotation avec 4 axes, quelques stations à la limite d'un groupe peuvent former un autre groupe. Par exemple la contribution de la station no. 11 sur l'axe 3 n'est peut-être pas significativement différente de sa contribution à l'axe 2 et pourrait à la rigueur faire partie du groupe de l'axe 2 ou d'un quatrième groupe.

Afin de porter un jugement sur la composition des groupes il est intéressant de tracer dans le plan des axes 2 et 3 des ellipses de proximité des coefficients de corrélation pour un

TABLEAU 4.2 Corrélation entre composantes et station
après rotation; saison "HIVER".

IDENTIFICATION DE LA STATION	NO.	COMPOSANTES		
		No. 1	No. 2	No. 3
702HACH	1	<u>.856</u>	.319	.235
702K30E	2	.303	<u>.819</u>	.305
7020785	3	.639	.081	<u>.675</u>
7021719	4	.299	<u>.841</u>	.209
7021918	5	<u>.756</u>	.426	.338
7022306	6	.250	.398	<u>.791</u>
7023312	7	<u>.592</u>	.512	.470
7024263	8	<u>.615</u>	.585	.231
7024623	9	.359	<u>.590</u>	.519
7024624	10	.443	<u>.602</u>	.549
7027372	11	.368	.574	<u>.586</u>
7027800	12	<u>.664</u>	.428	.464
7027802	13	<u>.615</u>	.565	.464
7028906	14	<u>.583</u>	.557	.477

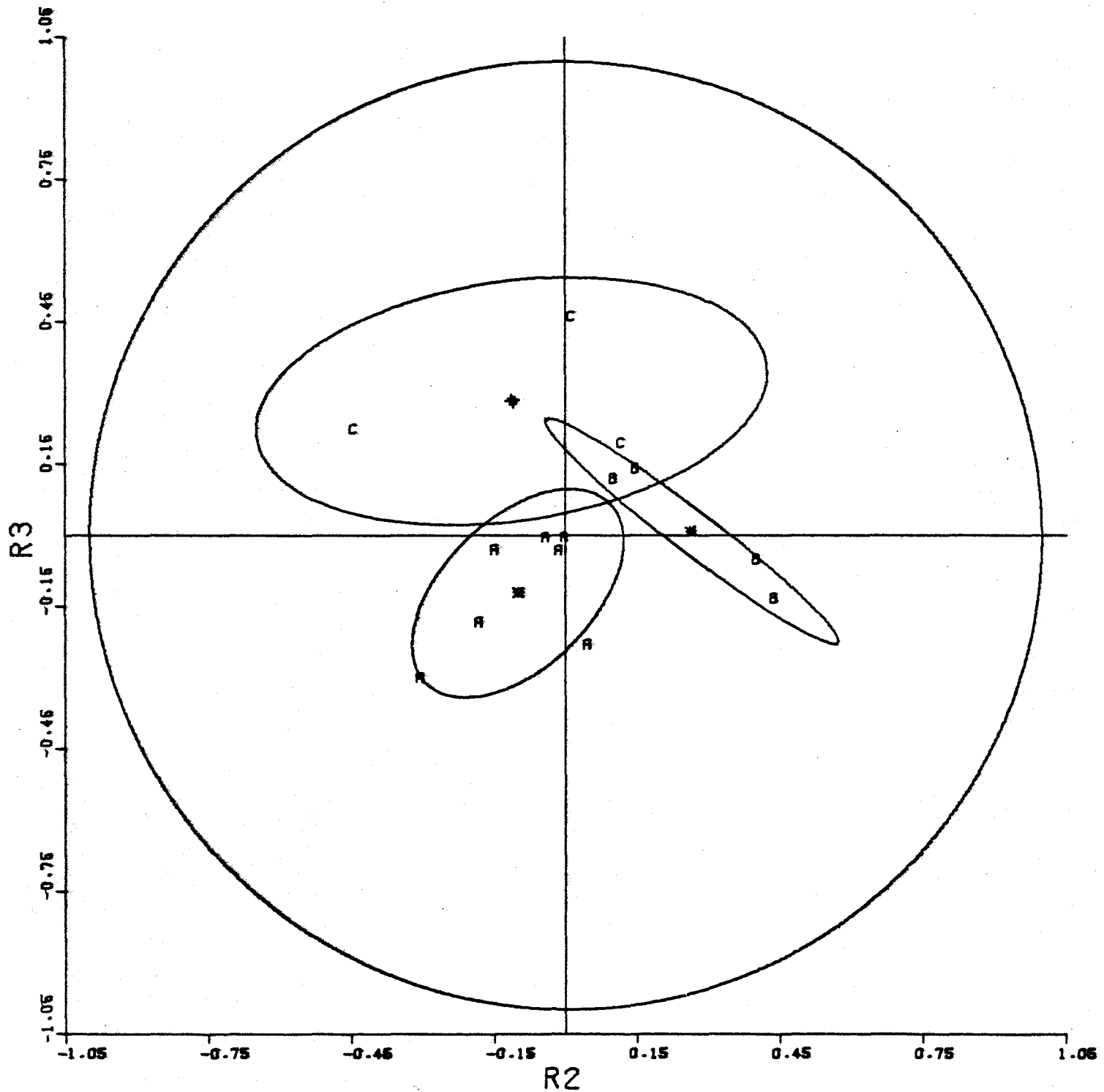


FIGURE 4.4 Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2^e et 3^e composantes principales pour les groupes établis par la méthode varimax; saison "HIVER".

niveau de probabilité donné. Ces ellipses d'égale probabilité sont calculées à l'aide d'une loi normale à deux variables, en prenant pour acquis que telle ou telle station fait partie d'un groupe (au moins 3 stations). Si une station est à l'intérieur de l'ellipse on ne peut rejeter l'hypothèse qu'elle fait partie du groupe selon le niveau de probabilité correspondant au tracé de l'ellipse. Inversement, si une station est à l'extérieur de l'ellipse on ne peut accepter l'hypothèse qu'elle fait partie du groupe toujours à un niveau de probabilité donné qui est de 80% dans les graphiques de cette étude.

L'interprétation des ellipses peut être la suivante: la redondance est d'autant plus forte pour un groupe que les axes de l'ellipse sont de faible dimension et de longueur sensiblement équivalente.

Une ellipse allongée signifie que le phénomène ne peut être bien représenté par une seule station. Des ellipses qui se chevauchent, d'axes confondus ou légèrement décalés signifient que l'information pluviométrique est comparable et redondante sur les groupes correspondants.

La figure 5.4 nous montre les ellipses de proximité tracées en supposant les groupes de stations tels que définis au tableau 4.2. Sur ce graphique les lettres indiquent la position

des stations, les lettres A indiquent les stations du premier groupe, B les stations du second groupe, etc... Pour déterminer le numéro de station correspondant à une lettre il faut comparer avec la figure 4.3 ou l'on voit au début de chaque segment à la même position dans le plan, le numéro de station.

Le groupe de l'ellipse A peut être estimé redondant; cependant la station 8 ne peut être introduite dans le groupe avec un niveau de probabilité de 80%. Sur le tableau 4.2 on voit d'ailleurs que la contribution de la station 8 est de .615 au premier groupe et de .585 au second ce qui nous laisse supposer que cette station peut être considérée à la limite des deux groupes. Il est intéressant de refaire l'ellipse du groupe A (figure 4.5) excluant la station 8, on obtient une ellipse moins large.

L'ellipse du groupe B (figure 4.4) qui a une forme très allongée indique que la précipitation de ce groupe ne peut être bien représentée par une seule station.

L'ellipse du groupe C est très étendue et on peut conclure que ces trois stations sont nécessaires si on les garde comme un groupe. Cependant la station 11 est très proche de l'ellipse B, et en considérant le tableau 4.2, on voit que la contribution de cette station est de .586 au groupe C et de .574 au groupe B.

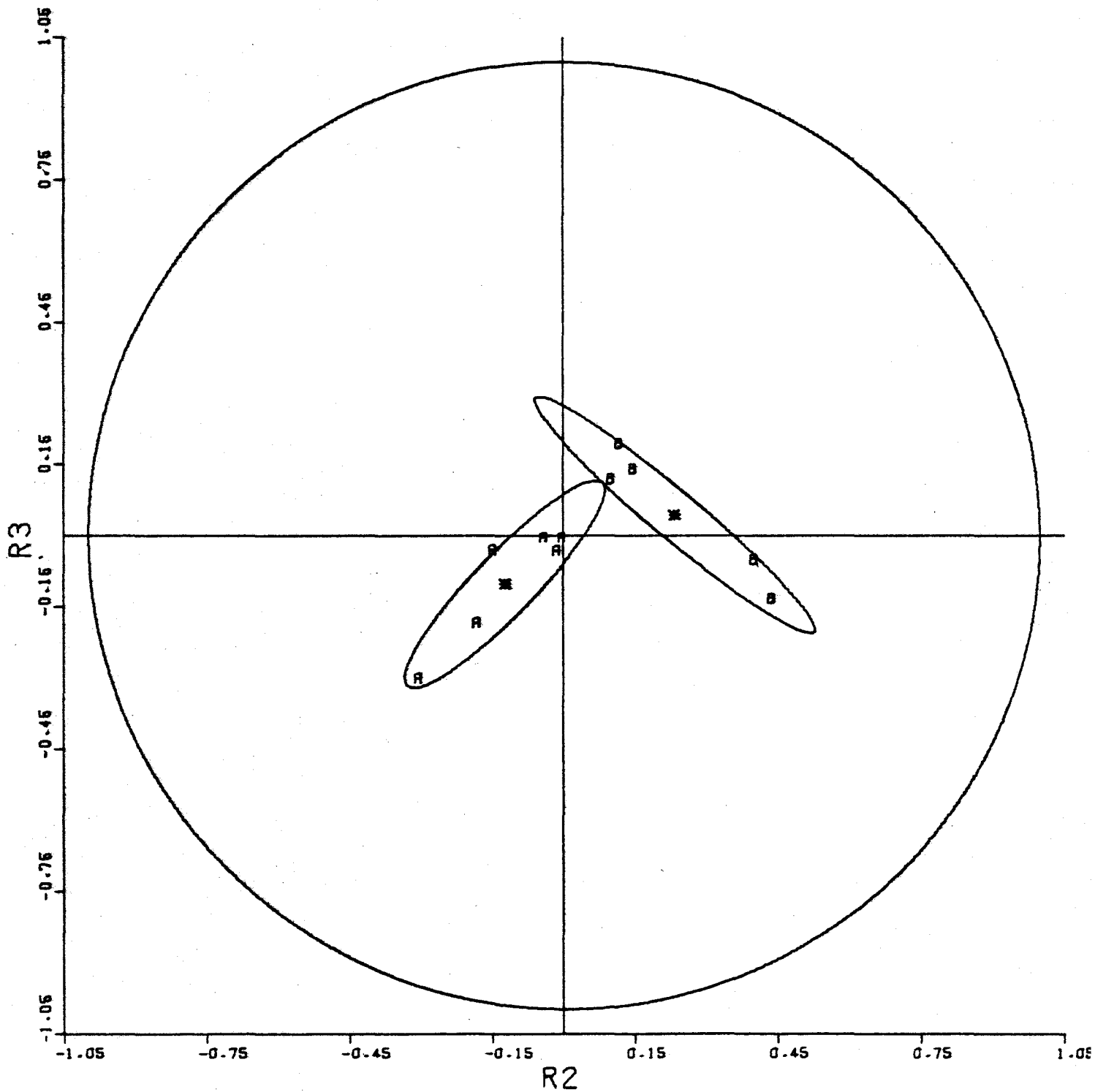


FIGURE 4.5 Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2^e et 3^e composantes principales en changeant les groupes établis par la méthode varimax; saison "HIVER".

TABLEAU 4.3 Corrélation entre composantes et station;
saison "PRINTEMPS".

IDENTIFICATION DE LA STATION	NO.	COMPOSANTES			COEFFICIENT DE CORRELATION MULTIPLE
		No. 1	No. 2	No. 3	
702HACH	1	.887	.304	.260	.948
702K30E	2	.945	-.116	.184	.941
7020785	3	.937	-.000	-.186	.914
7021719	4	.907	-.327	-.022	.931
7021918	5	.948	.146	-.028	.921
7022306	6	.940	-.162	.177	.942
7023312	7	.945	.072	.039	.901
7024263	8	.945	-.008	-.121	.908
7024623	9	.950	.043	-.068	.910
7024624	10	.968	.063	-.015	.943
7027372	11	.952	-.170	.111	.947
7027800	12	.928	.037	-.209	.907
7027802	13	.971	.051	-.123	.961
7028906	14	.968	.069	.017	.942
Valeurs propres		12.244	.310	.262	
Variance expliquée		88.9%	2.2%	1.9%	
Variance totale expliquée		88.9%	91.1%	93.0%	

1 = 702HACH	8 = 7024263
2 = 702K30E	9 = 7024623
3 = 7020785	10 = 7024624
4 = 7021719	11 = 7027372
5 = 7021918	12 = 7027800
6 = 7022306	13 = 7027802
7 = 7023312	14 = 7028906

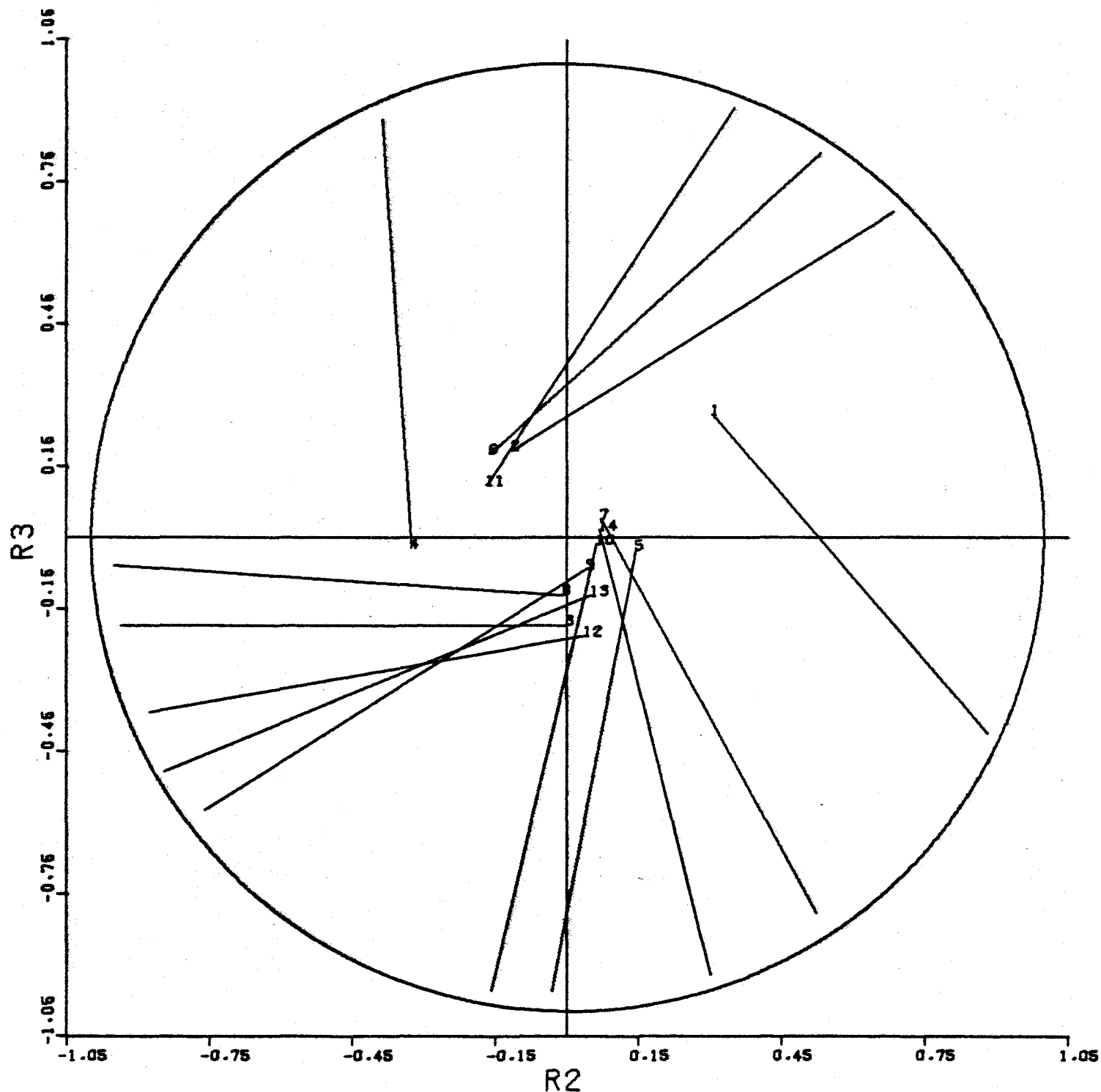


FIGURE 4.6 Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "PRINTEMPS".

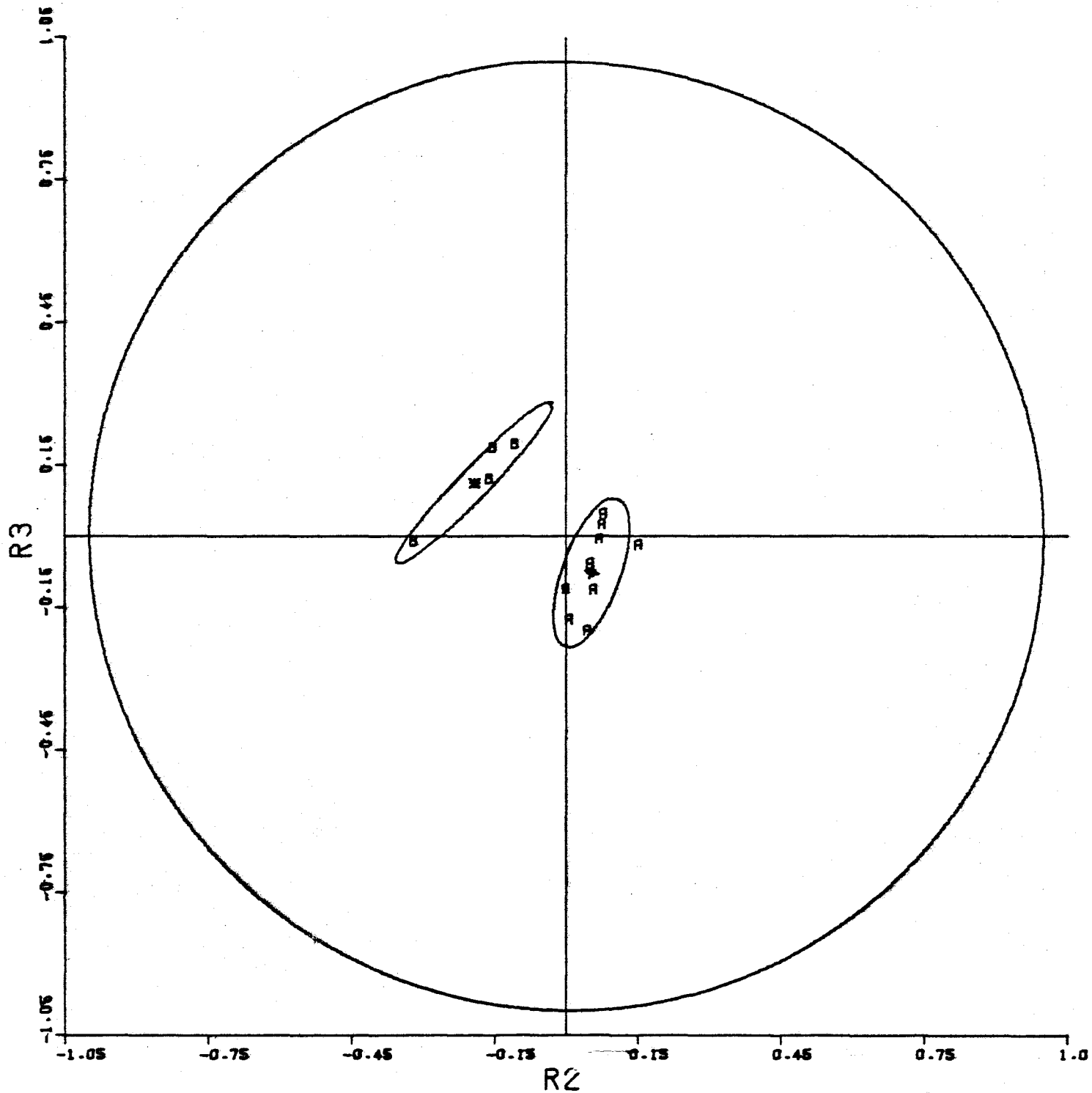


FIGURE 4.7 Ellipses de proximités à 80% établies d'après les coefficients de corrélation entre les 2^e et 3^e composantes principales pour les groupes établis par la méthode varimax; saison "PRINTEMPS".

La figure 4.7 qui montre les ellipses tracées en supposant les groupes de stations tels que définis ci-dessus, nous indique que la station 5 ne peut, avec une probabilité de 80%, être considérée comme faisant partie du groupe A. La nouvelle ellipse tracée en excluant cette station aurait une forme plus allongée, ce qui indique la nécessité de garder plus d'une station pour bien représenter le phénomène, ce qui est également le cas du groupe B.

4.3.3 Période "ETE"

Le tableau 4.4 présente les résultats pour la saison d'été. L'ensemble des trois premières composantes explique 90.7% de la variance dont 80.8% pour la première. La figure 4.8 qui est la schématisation du tableau 4.4, fait ressortir les stations ayant le même comportement.

La rotation des trois premiers axes a donné les groupes suivants:

- A) 2, 4, 6, 10 et 11
- B) 7, 9, 12, 13 et 14
- C) 1, 3, 5 et 8

que la figure 4.9 présente avec les ellipses de proximité. Comme précédemment l'ellipse A, de forme allongée nous indique que le phénomène ne peut être bien représenté avec une seule station.

TABLEAU 4.4 Corrélation entre composantes et station;
saison "ETE".

IDENTIFICATION DE LA STATION	NO.	COMPOSANTES			COEFFICIENT DE CORRELATION MULTIPLE
		No. 1	No. 2	No. 3	
702HACH	1	.786	.488	.954	.860
702K30E	2	.925	-.275	.065	.937
7020785	3	.859	.165	.371	.903
7021719	4	.846	-.396	.187	.909
7021918	5	.856	.329	.275	.916
7022306	6	.903	-.353	.137	.960
7023312	7	.919	.121	-.248	.922
7024263	8	.908	.272	-.066	.904
7024623	9	.920	-.068	-.240	.909
7024624	10	.917	-.196	-.121	.894
7027372	11	.912	-.191	.039	.870
7027800	12	.950	.074	-.092	.916
7027802	13	.942	.097	-.158	.922
7028906	14	.921	-.000	-.127	.865
Valeurs propres		11.313	.916	.465	
Variance expliquée		80.8%	6.5%	3.3%	
Variance totale expliquée		80.8%	87.4%	90.7%	

1 = 702HACH	8 = 7024263
2 = 702K30E	9 = 7024623
3 = 7020785	10 = 7024624
4 = 7021719	11 = 7027372
5 = 7021918	12 = 7027800
6 = 7022306	13 = 7027802
7 = 7023312	14 = 7028906

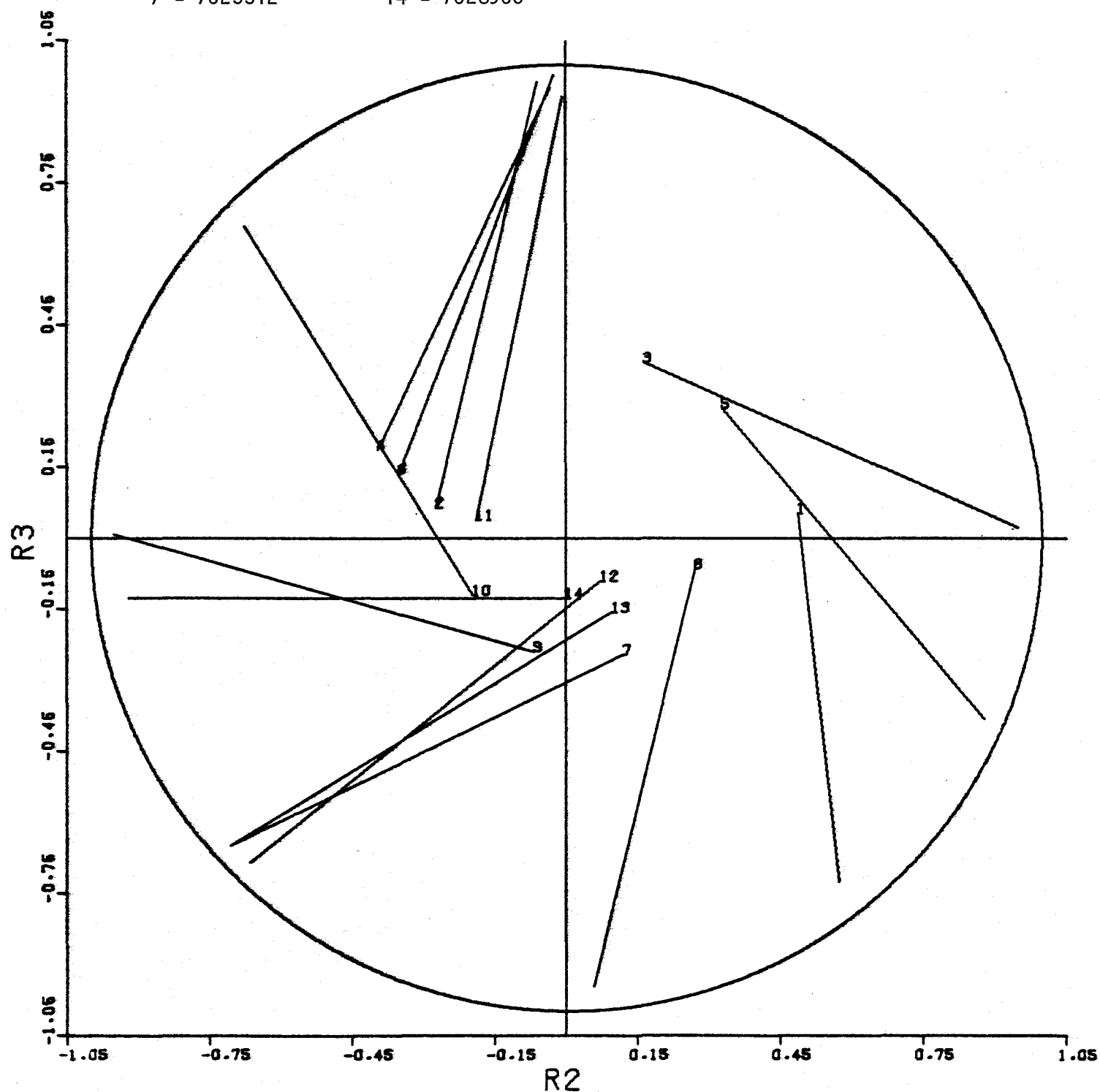


FIGURE 4.8 Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "ETE".

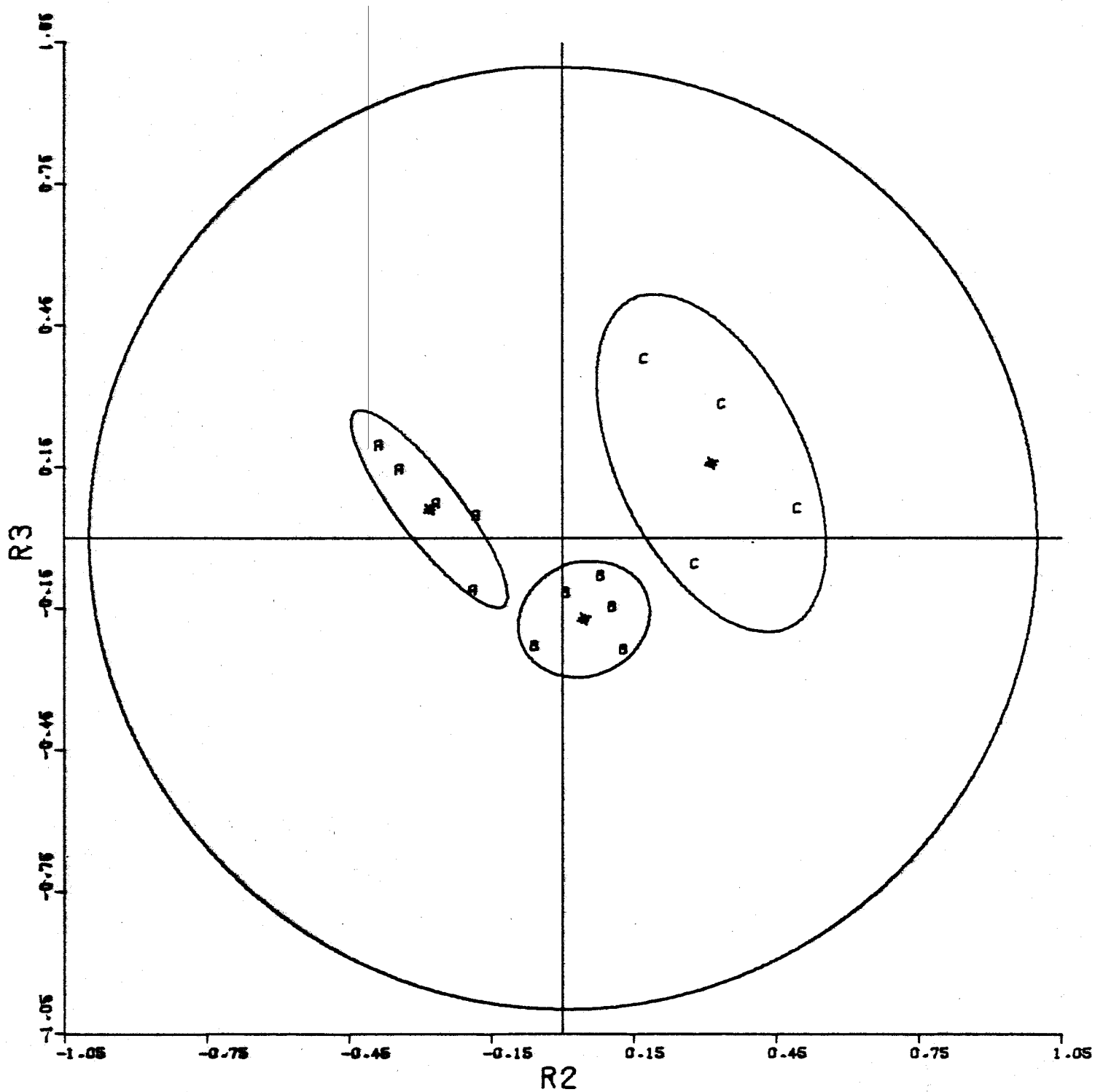


FIGURE 4.9 Ellipses des proximités à 80% établies d'après les coefficients de corrélation entre les 2^e et 3^e composantes principales pour les groupes établis par la méthode varimax; saison "ETE".

Par contre, les stations du groupe B peuvent être considérées comme plus homogènes. La décision de garder une seule station de ce groupe reste cependant délicate.

4.3.4 Période "AUTOMNE"

Le tableau 4.5 présente les résultats de l'analyse des précipitations de la période "automne".

La variance expliquée par les trois premières composantes est de 94.2%. Et comme précédemment la figure 4.10 qui synthétise les résultats, nous montre les stations qui se groupent, la station 1 peut être considérée comme particulière.

Les groupes obtenus par rotation sont les suivants:

- A) 2, 4, 6 et 11
- B) 3, 5, 7, 8, 9, 10, 12 et 13
- C) 1 et 14

La figure 4.11 montre les ellipses tracées avec ces groupes de stations. La station 5 ne peut avec certitude être considérée comme faisant partie du groupe B. Ce groupe demeure quand même nombreux, soit 7 stations pour une ellipse relativement petite, ce qui indique qu'il y a redondance.

TABLEAU 4.5 Corrélation entre composantes et station;
saison "AUTOMNE".

IDENTIFICATION DE LA STATION	NO.	COMPOSANTES			COEFFICIENT DE CORRELATION MULTIPLE
		No. 1	No. 2	No. 3	
702HACH	1	.857	.436	.213	.972
702K30E	2	.944	-.171	.204	.963
7020785	3	.943	-.033	-.200	.931
7021719	4	.916	-.211	.159	.909
7021918	5	.958	.114	-.111	.945
7022306	6	.943	-.165	.185	.951
7023312	7	.964	.071	-.107	.946
7024263	8	.950	-.007	-.000	.903
7024623	9	.954	.020	-.141	.930
7024624	10	.969	.042	.036	.943
7027372	11	.959	-.108	.003	.932
7027800	12	.970	-.043	-.112	.956
7027802	13	.966	-.036	-.176	.965
7028906	14	.958	.122	.081	.941
Valeurs propres		12.564	.344	.283	
Variance expliquée		89.7%	2.5%	2.0%	
Variance totale expliquée		89.7%	92.2%	94.2%	

1 = 702HACH	8 = 7024263
2 = 702K30E	9 = 7024623
3 = 7020785	10 = 7024624
4 = 7021719	11 = 7027372
5 = 7021918	12 = 7027800
6 = 7022306	13 = 7027802
7 = 7023312	14 = 7028906

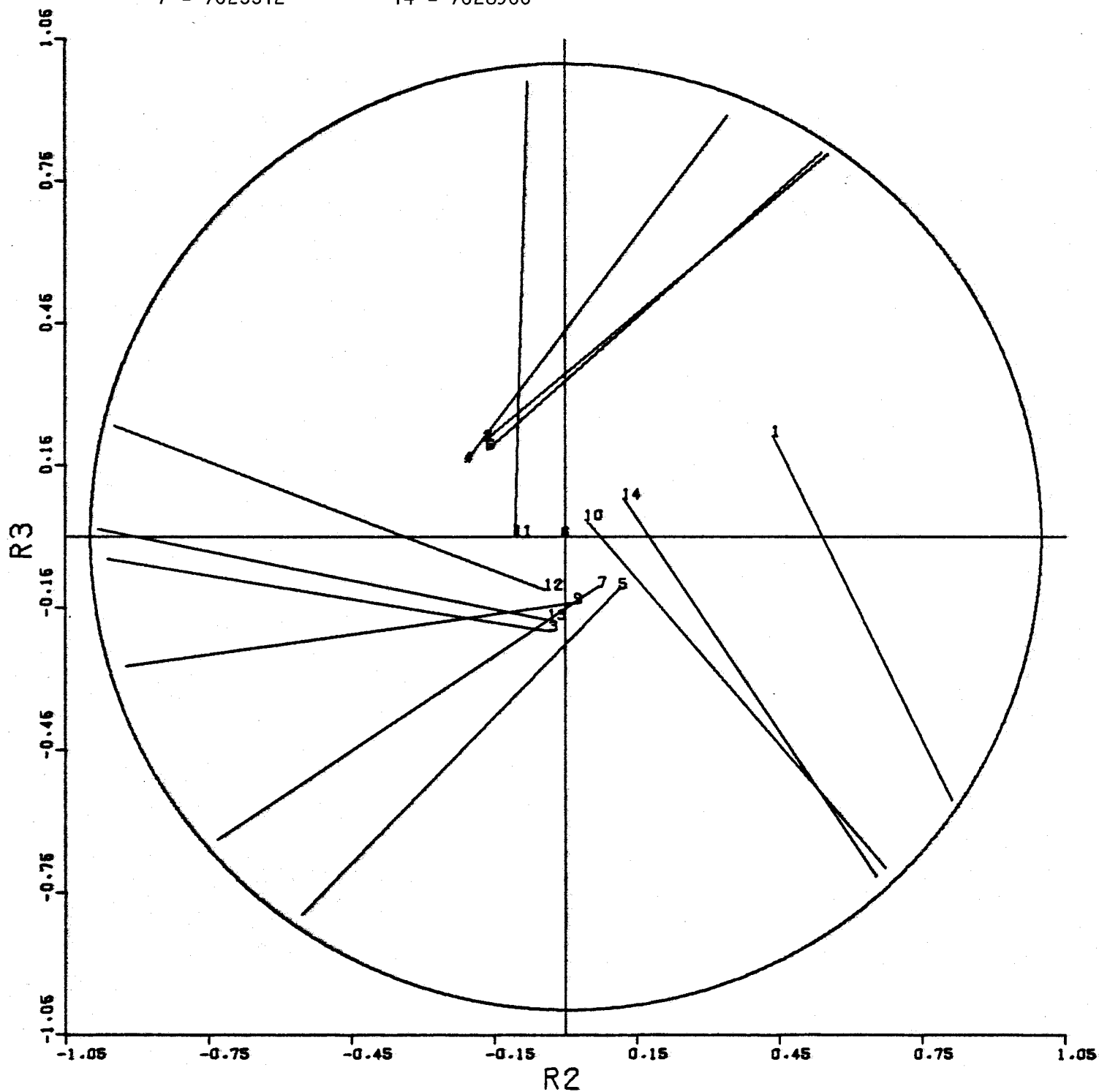


FIGURE 4.10 Représentation dans le plan des projections des coefficients de corrélation entre les stations et les trois premiers axes principaux pour les précipitations de 10 jours pour la période "AUTOMNE".

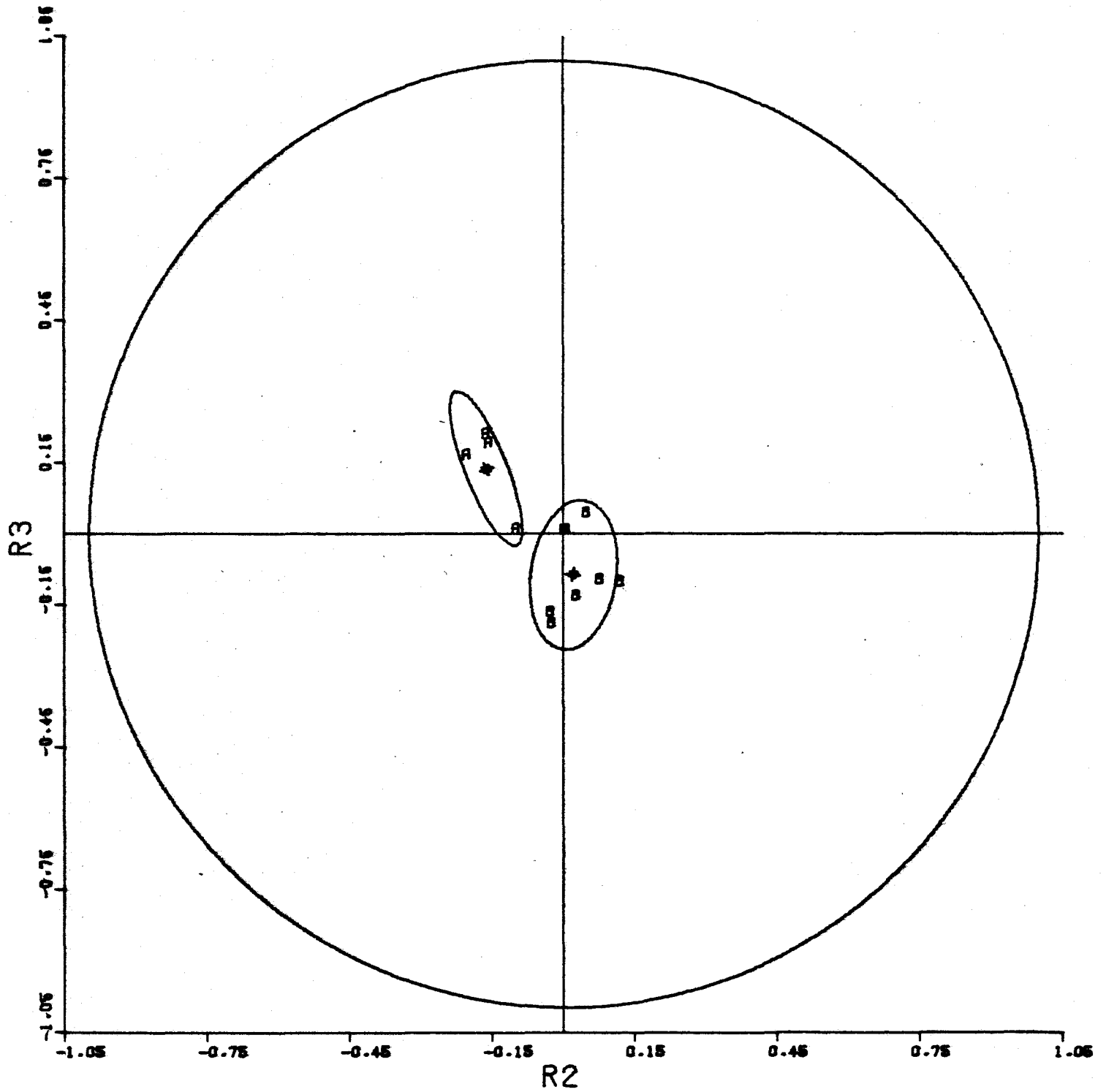


FIGURE 4.11 Ellipses des proximités à 80% établies d'après les coefficients de corrélation entre les 2^e et 3^e composantes principales pour les groupes établis par la méthode varimax; saison "AUTOMNE".

4.4 CONCLUSION

A chaque station on a trouvé par la méthode des composantes principales des groupes de stations de comportement semblable ce qui nous a permis de voir la redondance du réseau étudié (14 stations) et de faire ressortir les stations particulières, saison par saison sans s'occuper de la position géographique de ces stations sur le bassin versant. Faire la synthèse des saisons en tenant compte de la position géographique des groupes est évidemment nécessaire pour conclure. La figure 4.12 nous montre pour chaque saison la position géographique des groupes tels que déterminés après rotations des axes principaux par la méthode varimax, sans tenir compte des restrictions qui sont apparues en étudiant les ellipses de proximité de chaque groupe. L'étude de la figure 4.12 nous montre que les groupes divisent le bassin en 3 régions du nord au sud (sauf pour la période "hiver" dont les régions sont très irrégulières). Le groupe du centre occupe une superficie importante du bassin versant. Le groupe du nord varie beaucoup d'une saison à l'autre.

De l'évolution saisonnière on déduit que les stations 12, 13 et 7 sont toujours ensemble et que sauf pour l'été les stations 8 et 5 sont toujours avec celles-ci. La station 9 est également groupée avec les stations 12, 13 et 7, sauf pour la saison hiver. Les stations 2 et 4 sont toujours ensemble et les

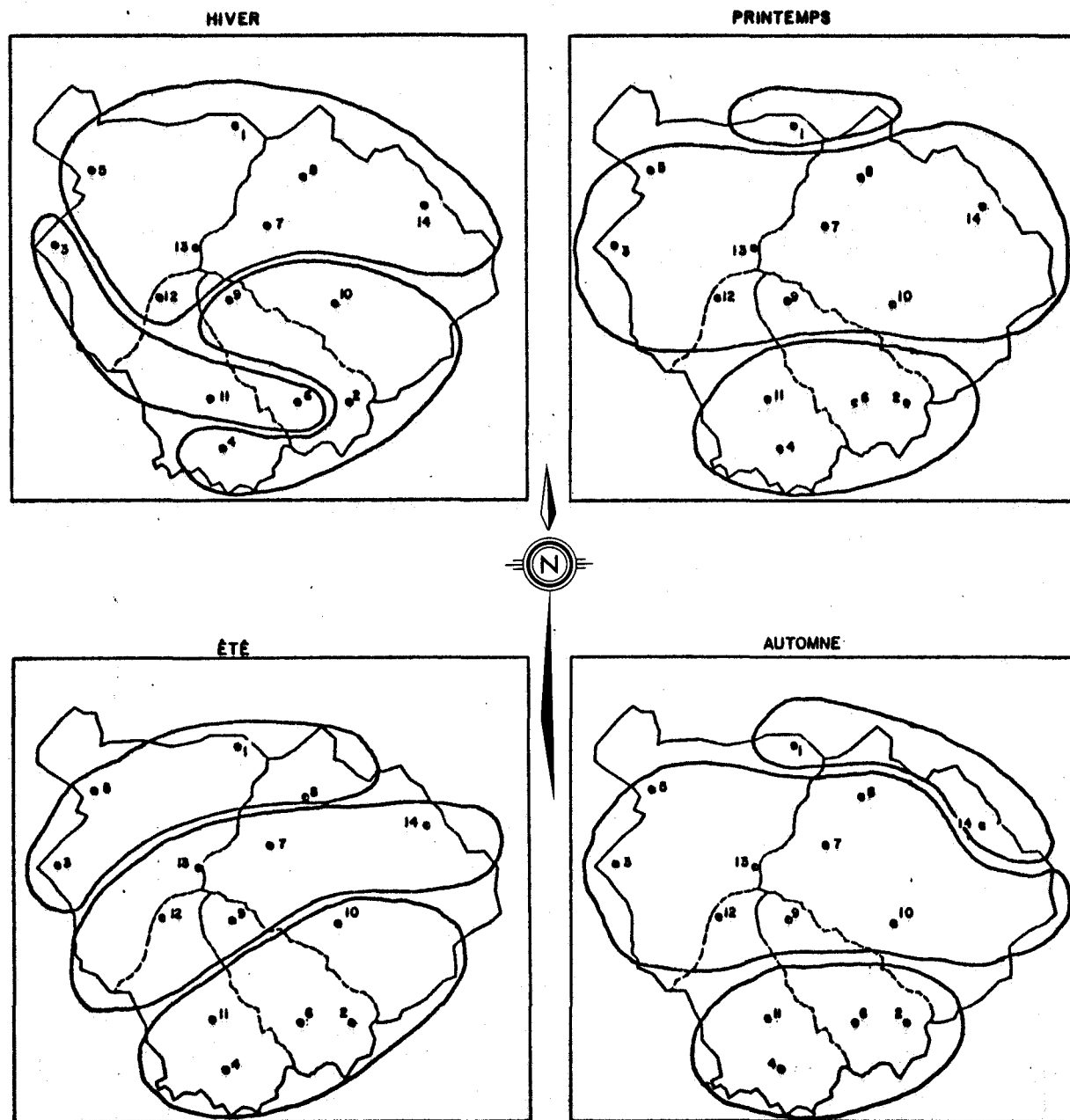


FIGURE 4.12 Répartition géographique des groupes de stations pour chacune des saisons.

stations 2, 4, 6 et 11 sont groupées pour toutes les saisons sauf l'hiver.

La station 1 qui demeure seule ou qui se groupe différemment d'une saison à l'autre peut être considérée comme hétérogène. Finalement l'étude de l'ensemble nous montre que pour les trois saisons de précipitations liquides soit le printemps, l'été et l'automne, les régions sont semblables. La saison d'hiver dont les précipitations sont exclusivement solides montre des régions très différentes des autres saisons et de plus, les régions sont très irrégulières. L'imprécision de la mesure des précipitations solides est probablement la cause d'une bonne partie de ces irrégularités.

On pourrait raffiner l'étude du réseau en faisant les mêmes calculs avec d'autre pas de temps. Il serait intéressant par exemple d'étudier le réseau à l'échelle journalière ou encore à l'échelle d'un orage. Il serait également intéressant de reprendre cette étude avec l'ensemble des trente stations disponibles sur le bassin versant.

CHAPITRE 5

CONTROLE DE SERIES CHRONOLOGIQUES DANS UNE REGION

La méthode dite des "doubles masses" ou "doubles cumuls" utilisée pour contrôler l'homogénéité temporelle de données climato-
logiques consiste à porter en ordonnées le cumul dans l'ordre chronologique des valeurs d'une variable en fonction du cumul, dans le même ordre chronologique, d'une variable représentative du même phénomène pour la région, par exemple la moyenne des stations voisines pondérées parfois en fonction de leur éloignement. Si ces variables sont assez bien corrélées, ($R > 0.7$), et homogènes, la courbe obtenue est voisine d'une droite; au contraire lorsqu'une erreur systématique a eu lieu pendant assez longtemps, on observe une "cassure" qui permet de détecter l'hétérogénéité.

5.1 DESCRIPTION DE LA METHODE "BOIS"

Le bref rappel de la méthode bien connue des "doubles masses" qui vient d'être fait, va servir pour présenter la méthode de contrôle de l'homogénéité des données (Bois, 1971).

La description se divise en 2 parties: la première traite de la formation de la station témoin et la seconde de la méthode proprement dite.

5.1.1 Recherche d'une station témoin

La station témoin est la résultante d'une combinaison linéaire des (n) stations représentant au mieux l'échantillon; ce qui revient à projeter chaque observation des (n) stations sur l'axe principal d'inertie ou première composante principale (voir annexe A). Ceci est obtenu en calculant les vecteurs propres de la matrice de corrélation inter-stations puis en faisant le produit scalaire de chaque vecteur observation (les variables étant centrées réduites) par le vecteur propre ayant la plus grande valeur propre, c'est-à-dire que si $x_i(j)$, $\overline{x_i}$ et σ_i sont respectivement les valeurs de la variable "i" pour le pas de temps "j", $\overline{x_i}$ sa moyenne et σ_i son écart type, la station témoin aura pour le pas de temps "j" une valeur:

$$y(j) = a_1 \frac{x_1(j) - \overline{x_1}}{\sigma_1} + a_2 \frac{x_2(j) - \overline{x_2}}{\sigma_2} + \dots + a_n \frac{x_n(j) - \overline{x_n}}{\sigma_n}$$

où

$a_1, a_2 \dots a_n$ sont proportionnels aux composantes du 1^{er} vecteur propre.

Généralités sur les composantes principales

On a "n" observations d'une variable à "p" stations. A cause de la dépendance qui existe entre les stations, on

va rechercher une transformation linéaire de façon à obtenir des nouvelles stations "z" ou composantes principales, qui soient indépendantes en probabilité.

Cette opération revient à effectuer un changement d'axes et chaque composante principale peut s'exprimer comme une combinaison linéaire des observations.

De plus, cette transformation présente l'avantage que le premier axe (ou première composante (z)) est choisi en rendant minimum la distance entre tous les points d'observation et celui-ci, au sens des moindres carrés. Cette composante (z_1) est donc celle qui explique le maximum de variance.

On voit l'avantage d'utiliser la technique des composantes principales pour former la station (fictive) témoin. En effet, si dans le groupe de stations de base une station présente des anomalies qui contrastent avec l'homogénéité de l'ensemble des stations, cette station sera relativement peu corrélée avec ses voisines et de ce fait verra sa projection sur le premier axe (ou première composante z_1) être minimum. Le vecteur formé des projections sur le premier axe est surtout constitué par les stations homogènes, ce qui ne peut pas être déterminé a priori. Et inversement, la station hétérogène donnera un axe singulier après transformation.

5.1.2 Cumul des résidus

Le vecteur témoin et la station étudiée sont désormais considérés comme deux stations normales y et x ; d'effectif " n "; de moyenne $\bar{y}$ et $\bar{x}$; d'écart type σ_y et σ_x .

Par une méthode régressive on établit l'équation qui lie la station étudiée à la station témoin ainsi que le coefficient de corrélation " r ".

$$x_i^* = ay_i + b$$

où:

x_i^* : la valeur calculée à partir de la station témoin

y_i : la station témoin

a et b : les constantes de régression

et

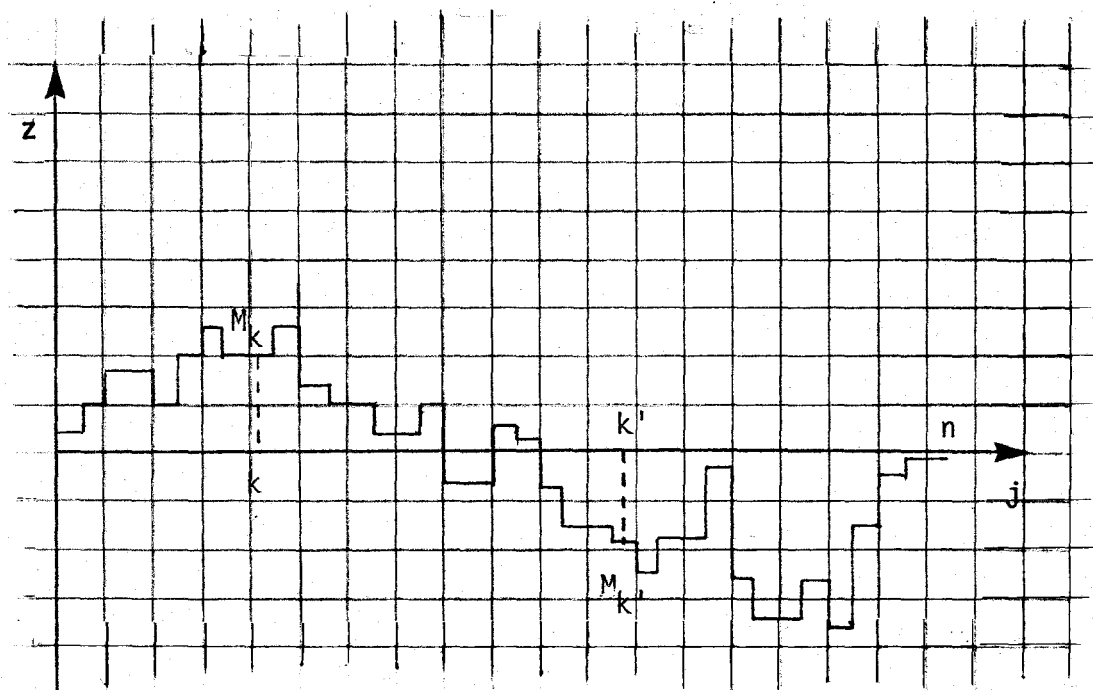
$$\mathcal{E}_i = x_i - x_i^*$$

$\mathcal{E}_i$ = l'erreur résiduelle

avec $\overline{\mathcal{E}} = 0$

$$\sigma_{\mathcal{E}} = \sigma_x \sqrt{1 - r^2}$$

Posons $z_j = \sum_{k=1}^{k=j} \mathcal{E}_k$. Si l'on porte en graphique z_j en fonction de " j ", on obtient une ligne brisée partant de $(0, 0)$ et aboutissant à $(n, 0)$ car $z_n = n\overline{\mathcal{E}} = 0$



Les résidus sont indépendants entre eux mais ne le sont pas dans leur ensemble (somme nulle). Si l'on fait abstraction de cette dernière contrainte, l'écart selon l'axe "z" entre deux points M_k et $M_{k'}$ est la somme de $k' - k$ variables normales centrées, d'écart type $\sigma_{kk'}^2 = (k' - k) \sigma_x^2 (1 - r^2)$ c'est-à-dire que la distance verticale séparant deux points représentant deux observations distantes de "m" dans la série chronologique est une variable normale centrée d'écart type

$$\sigma_m = \sigma_x \sqrt{1 - r^2} \sqrt{m}$$

Sachant que les courbes passent par $(0, 0)$ et $(n, 0)$, on en déduit, si l'on conserve l'hypothèse d'indépendance

des résidus, que le chemin aléatoire est limité par 2 paraboles de sommets respectifs $(0, 0)$ et $(n, 0)$ et d'équations:

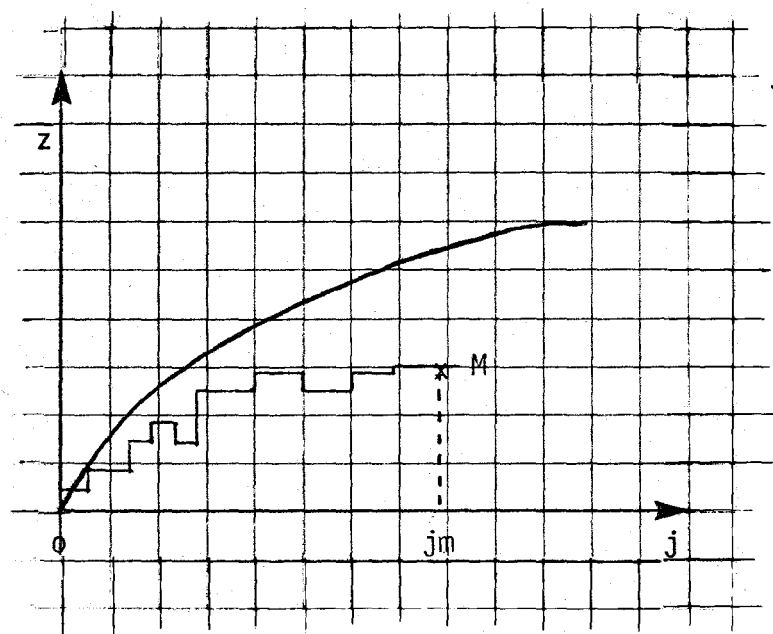
$$z^2 = \sigma_x^2 \cdot (1 - r^2) \cdot t^2(p) \cdot i$$

$$z^2 = \sigma_x^2 \cdot (1 - r^2) \cdot t^2(p) \cdot (n - i)$$

où:

$t(p)$ est la valeur de la variable centrée réduite normale correspondant à la probabilité "p" de ne pas sortir de ce domaine.

En effet, si l'on considère la distance verticale entre 0 et M, d'abscisse j_m , elle est la somme des j_m variables normales centrées d'écart type $\sigma_x \sqrt{1 - r^2}$.



Donc pour une probabilité donnée "p", M_{jm} est la valeur d'une variable normale centrée d'écart type: $\sigma_x \sqrt{1 - r^2} \sqrt{jm}$

$$M_{jm} = t(p) \cdot \sigma_{M_{jm}} = t(p) \cdot \sigma_x \sqrt{1 - r^2} \cdot \sqrt{jm}$$

Si l'on tient compte de la contrainte qui stipule que la somme des résidus est nulle, donc que la ligne brisée part du point (0, 0) et s'arrête au point (n, 0) la courbe de probabilité "p" devient une ellipse et la valeur M_{jm} s'écrit:

$$M_{jm} = t(p) \cdot \sigma_{M_{jm}} = t(p) \cdot \sigma_x \sqrt{1 - r^2} \cdot \sqrt{\frac{jm(n - jm)}{n - 1}}$$

Nous prendrons pour "p" les niveaux:

$$p = 95\% \text{ soit } t(p) = 1.97$$

$$p = 99\% \text{ soit } t(p) = 2.57$$

5.1.3 Comparaison avec la méthode classique des "Doubles masses"

Après de nombreux essais comparatifs des deux méthodes, Bois (1971) conclut que lorsque l'erreur est détectable par la méthode des doubles masses, l'analyse du cumul des résidus

est simple et donne rapidement des résultats, mais de façon beaucoup plus claire (exemple: double masse et cumul des résidus, figures 5.1 et 5.2).

5.2 ANALYSE DES RESULTATS

La méthode du cumul des résidus a été appliquée au 14 stations dont les observations sont disponibles sous forme informatique pour la période 1966 à 1975 inclusivement.

Ainsi que la figure 5.3 permet de le constater, deux cas peuvent se présenter: le tracé obtenu ne montre aucun caractère marqué, ce qui signifie que les résidus sont aléatoires c'est l'exemple du tracé relatif à la station Eaton Sc B. Le tracé de la station Maple Leaf Est montre, à partir du 1^{er} juillet 1971, une tendance dans le temps des résidus, qui sont tous négatifs, jusqu'au 20 juillet 1973.

De plus la méthode du cumul des résidus permet de quantifier de manière probabiliste la confiance à accorder aux données successives. Puisque les résidus sont indépendants théoriquement, il est toujours possible de tracer les ellipses de probabilité 5% et 1% à partir d'une date quelconque, par exemple à partir du 1^{er} juillet 1971. L'indépendance des résidus implique que nous aurions

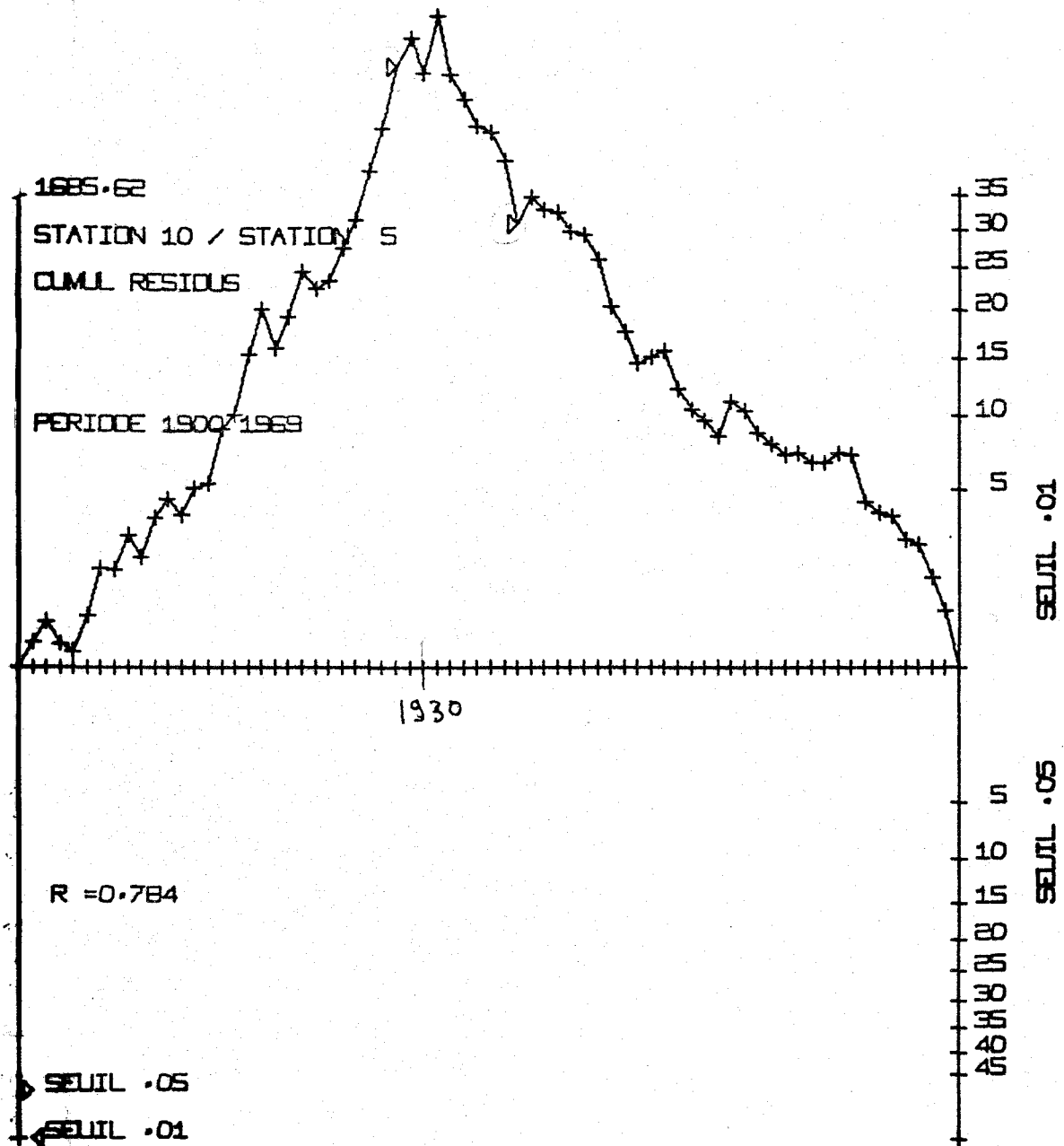


FIGURE 5.1 Sensibilité de la méthode (Multiplication par 1,1 de 1900 à 1930, $\bar{X} = 1000$, $\sigma = 160$).

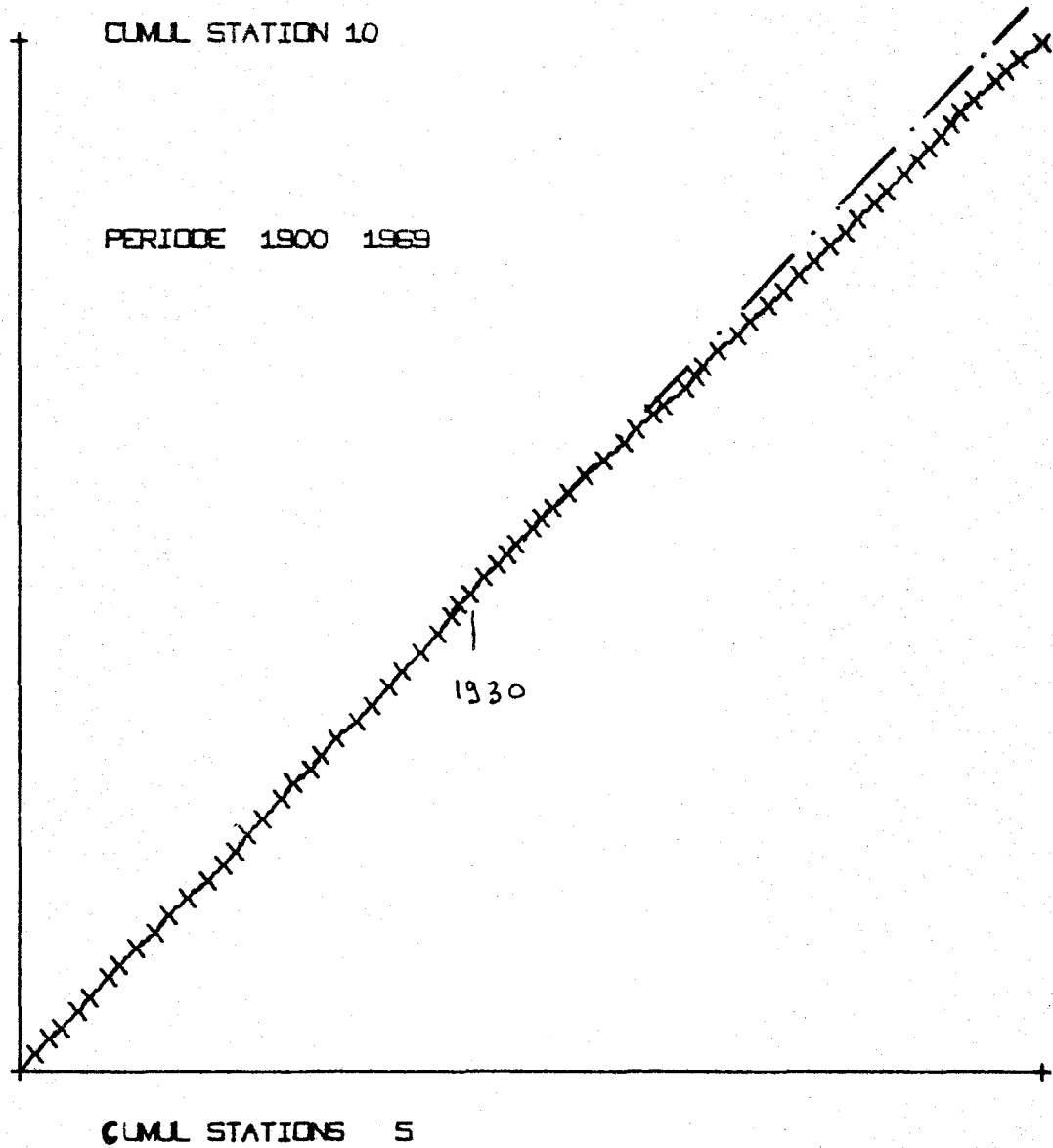


FIGURE 5.2 Double masse correspondant à la figure 5.1
 (Multiplication par 1,1 de 1900 à 1930,
 $\bar{X} = 1000$, $\sigma = 160$).

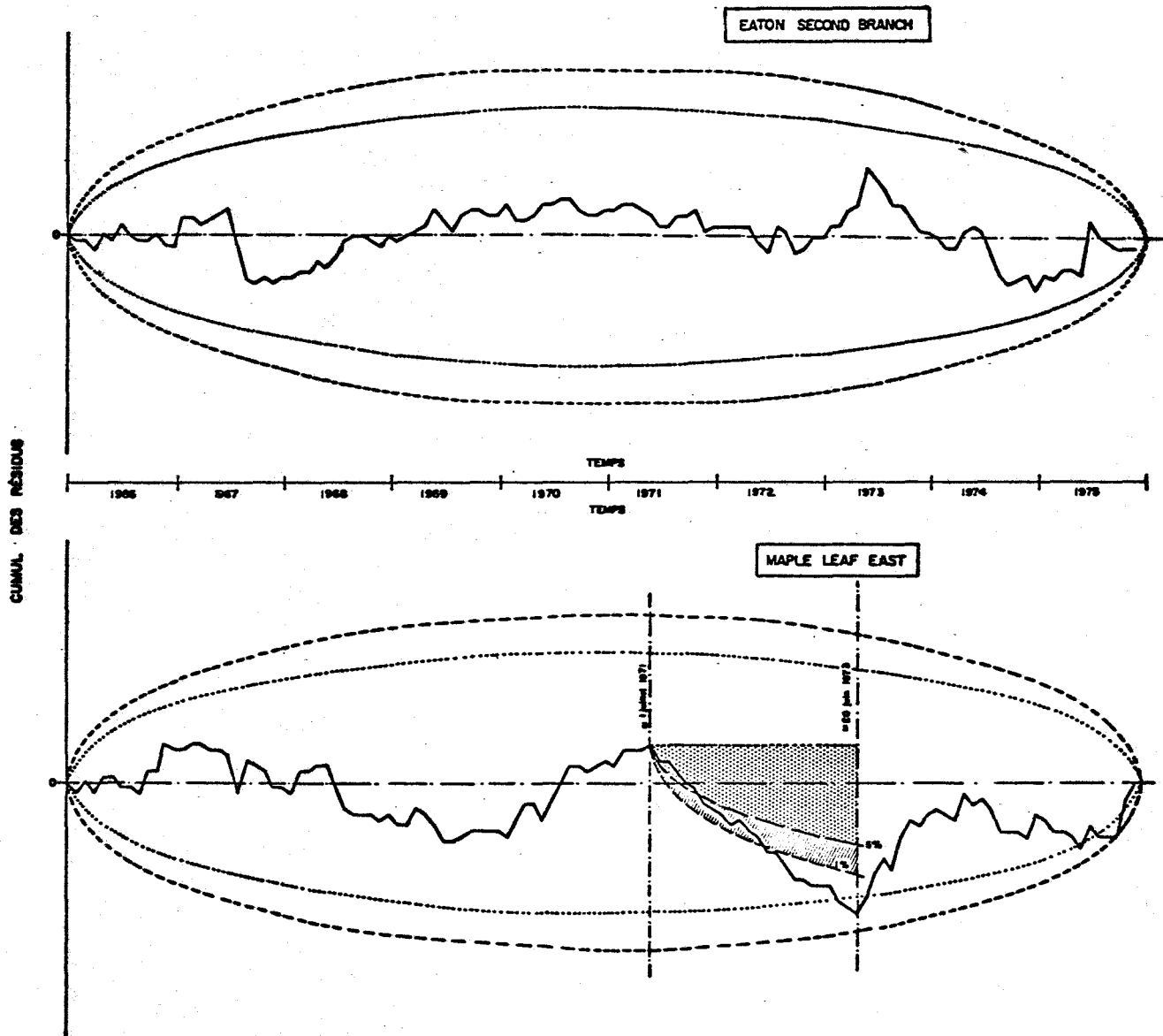


FIGURE 5.3 Présentation en fonction du temps du cumul des résidus.

eu les 2 mêmes ellipses si les données de janvier 1966 à juin 1971 avaient été relevées après 1975. On en conclut donc que les ellipses relatives au 1^{er} juillet 1971 auraient même excentricité, et par conséquent on peut translater directement les arcs d'ellipses commençant en janvier 1966 pour vérifier les résidus postérieurs à juillet 1971.

C'est ce qui est représenté en grisé sur la figure 5.3 (B). Ainsi donc le comportement des résidus entre le 1^{er} juillet 1971 et **le 20 juillet 1973 a une probabilité inférieure à 1%, ce qui rend la station suspecte sur cette période.**

Le tableau 5.1 résume la critique des 14 stations analysées de la même manière. Pour éviter au maximum de classer les stations de manière subjective, deux personnes différentes ont qualifié les **stations de leur côté, et sont arrivées aux mêmes conclusions.**

La figure 5.4 montre les graphiques de Cookshire et Island Brock qui présentent des "anomalies" en début de période.

TABLEAU 5.1 Résultats de l'analyse des 14 stations pour le cumul des résidus.

STATION	NUMERO FEDERAL	REMARQUES
Shermann BN	702HACH	légère hétérogénéité
Eaton Rv N	702K30E	légère hétérogénéité
Birchton	7020785	légère hétérogénéité
Clifton Rv S	7021719	normale
Cookshire	7021918	hétérogène en début de période (figure 5.4)
Eaton Sc B	7022306	normale (figure 5.3)
Is. Brook	7023312	hétérogène avant 70 (figure 5.4)
Lawrence	7024263	légère hétérogénéité
Maple Leaf	7024623	Juillet 1967 douteux
Maple Leaf E	7024624	hétérogène (figure 5.3)
St-Isidore	7027372	légère hétérogénéité
Sawyerville	7027800	normale
Sawyerville	7027802	normale
W. Ditton	7028906	normale

En résumé, on retiendra que 4 stations sont "bonnes":

- Eaton Sc B
- Sawyerville
- Sawyerville N
- W. Ditton

et que 3 stations sont douteuses:

- Cookshire
- Is. Brook
- Maple Leaf E.

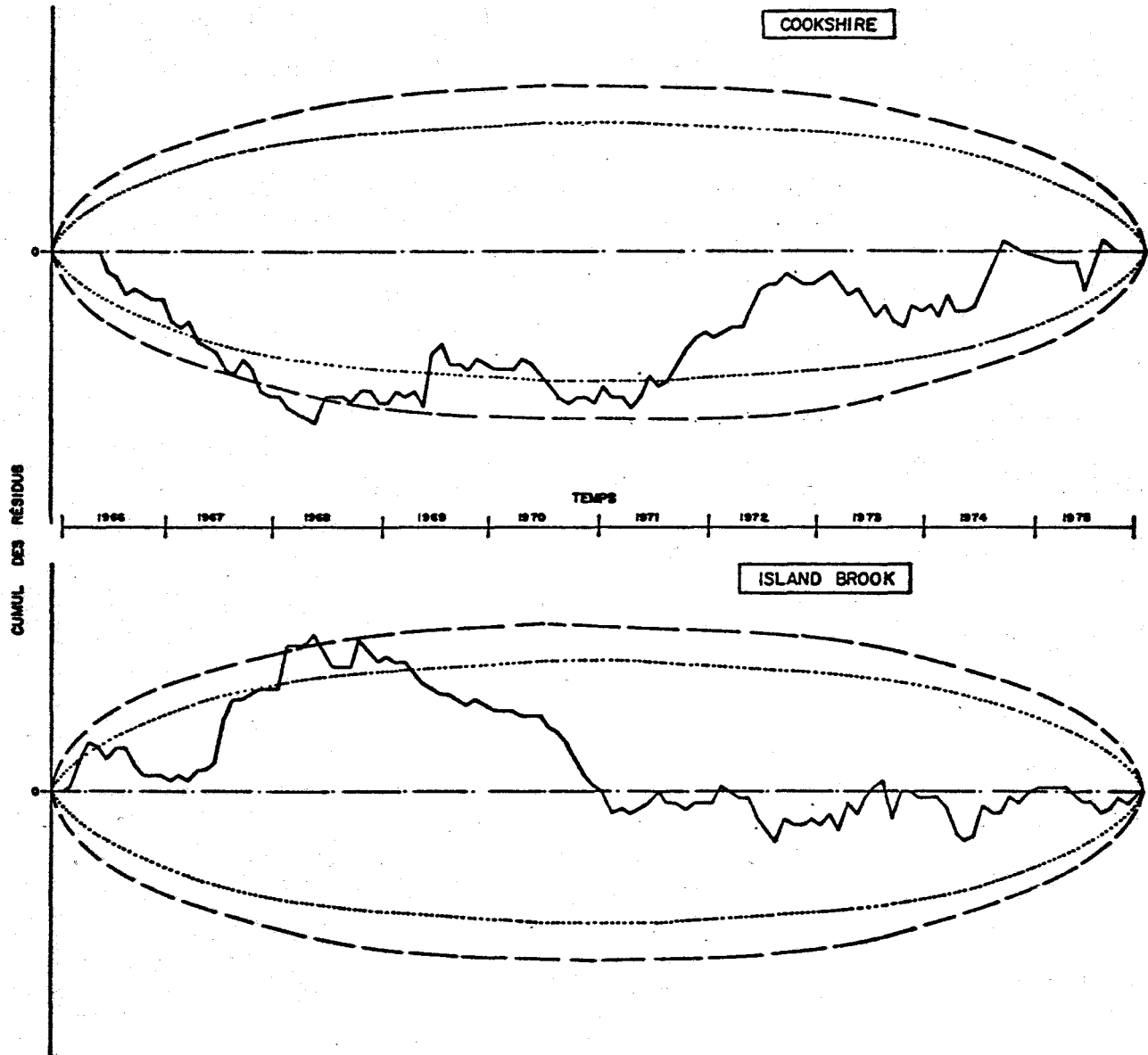


FIGURE 5.4 Présentation en fonction du temps du cumul des résidus.

CHAPITRE 6

ERREURS D'ESTIMATION ASSOCIEES A L'INTERPOLATION DES PRECIPITATIONS

Les données météorologiques disponibles sur un bassin versant étant ponctuelles en raison de la nature discrète de tout réseau de mesures, l'hydrologue doit, pour répondre à ses besoins, interpoler ces données de façon à estimer le phénomène en tout point du territoire.

Les méthodes les plus utilisées par les hydrologues pour estimer les précipitations, en un point quelconque pour une surface donnée, à partir des points de mesure sont:

- méthode des polygones de Thiessen;
- méthode des iso-valeurs

Cette seconde méthode, qui est en somme une interpolation linéaire, est plus précise dans bien des cas que l'estimé par les polygones de Thiessen, mais tout comme la première, elle ne tient pas compte à la fois des particularités des points de mesure et de la structure spatiale du phénomène. La méthode d'interpolation optimale (décrite en annexe B) que nous nous proposons d'utiliser, est une méthode probabiliste qui tient compte de la structure du phénomène à étudier et ainsi permet non seulement d'estimer la valeur ponctuelle d'un phénomène ou sa valeur moyenne sur une surface donnée à l'intérieur de la région retenue mais en outre de calculer l'erreur d'estimation dans chaque cas.

6.1 DONNEES UTILISEES

Pour appliquer la méthode d'interpolation optimale, on doit connaître la structure dans l'espace du phénomène à interpoler. Comme précédemment on a choisi d'étudier les précipitations de 10 jours. Pour travailler sur des données homogènes, l'étude porte sur les quatre saisons telles que définies au chapitre 4. Pour les calculs de la fonction de structure on utilise 13 stations soit les mêmes qu'au chapitre 4, sauf Sherman Brook North (702HACH) que ce chapitre a fait ressortir comme hétérogène.

6.2 CALCUL DE LA FONCTION DE STRUCTURE

La fonction de structure (F.S.) est déterminée en calculant pour tous les couples de stations la valeur:

$$F.S. = \frac{\sum_{i=1}^m (f'_{ij} - f'_{ik})^2}{m} \quad (6-1)$$

où

f'_{ij} est la $i^{\text{ième}}$ déviation de la station j
 et f'_{ik} est la $i^{\text{ième}}$ déviation de la station k
 f' étant la déviation de l'observation par rapport
 à la moyenne des observations de la station
 m est le nombre d'observations

Les valeurs obtenues de l'équation (6-1) sont utilisées pour définir la fonction de structure du phénomène étudié en fonction de la distance qui sépare chaque couple de stations. En pratique l'axe des distances est subdivisé en classes et on calcule à l'intérieur de chaque classe la valeur moyenne de la fonction de structure. L'ajustement d'une courbe, de la forme $f(x) = a - b e^{-x/c}$, passant par ces valeurs en fonction de la distance nous permet par la suite d'interpoler la fonction de structure (ou fonction covariance) entre les stations et le point à interpoler à l'aide des distances de ce point à chacune des stations. De plus en extrapolant la courbe à une distance zéro on peut estimer l'erreur minimale ponctuelle à laquelle on peut s'attendre. Cette erreur est due à la fois à l'imprécision des mesures et à l'influence du micro-climat.

6.3 INTERPOLATION OPTIMALE

Pour estimer les précipitations en un point on doit déterminer le poids à donner à chacune des stations. Ces poids sont déterminés en minimisant l'erreur quadratique moyenne, ce qui nous donne un système à N équations à résoudre lorsqu'on a N stations. Soit à résoudre:

$$\sum_{j=1}^N \sigma_{ij}^2 p_j^2 = \sigma_{oi}^2 \quad (i = 1, 2 \dots N) \quad (6-2)$$

où

σ_{ij}^2 sont les covariances entre les stations i et j

σ_{oi} sont les covariances entre le point à extrapoler et les stations

p_j sont les poids à donner à chacune des stations

Ce système est régulier et admet une seule solution qui donne les N poids p_j qui minimisent la variance. Les valeurs σ_{ij}^2 et σ_{oi}^2 sont prises sur la fonction de structure en fonction des distances inter-stations et des distances entre le point à estimer et les stations.

Connaissant les poids, on peut calculer la variance de l'estimation par:

$$E_o^2 = \sum_{i=1}^N p_i \sigma_{oi}^2 \quad (6-3)$$

où

E_o^2 = variance de l'estimé au point o

σ_{oi}^2 = variance entre le point et chacune des stations, prise sur la fonction de structure en fonction de la distance

L'estimé du phénomène au point nous est donné par:

$$V_o = \sum_{i=1}^N p_i f_i + f_o \quad (6-4)$$

où

V_0 = valeur au point zéro

p_i = poids de chacune des stations

f'_i = déviation de l'observation de la station i
par rapport à la moyenne des observations de
cette station

$\overline{f_0}$ = norme ou moyenne au point 0

Si l'on ne connaît pas la norme à tous les points, on peut faire les calculs directement avec les observations (cf. annexe B; problème de la norme) en imposant une contrainte supplémentaire ($\sum p_i = 1$) pour que l'estimé soit sans biais. C'est cette méthode que nous utilisons. Le système d'équations 6-2 devient un système à $N + 1$ équations, le $N + 1^{\text{ième}}$ paramètre est le multiplicateur de Lagrange. En pratique, pour tracer les iso-valeurs ou les iso-erreurs on fait les calculs sur un grand nombre de points répartis également sur la région. Pour le bassin versant de la rivière Eaton, on a calculé les erreurs aux intersections d'une grille de 1 km par 1 km. Le but de cette étude étant de faire ressortir l'erreur d'estimation on n'a pas tracé les iso-valeurs. On peut cependant dire que les iso-valeurs calculées par cette méthode sont plus précises que par les méthodes citées précédemment parce qu'elle tient compte de la structure spatiale du phénomène et de l'erreur de mesure et de micro-climat.

6.4 ANALYSE DES RÉSULTATS

6.4.1 Période "HIVER"

La fonction de structure pour la saison "HIVER" est présentée sur la figure 6.1. Connaissant la fonction de structure dans l'espace des précipitations de 10 jours, on peut à l'aide de l'interpolation optimale, déterminer les cartes d'iso-erreurs d'estimation pour un réseau quelconque, en supposant que la fonction de structure est valable pour les stations de la région.

Pour montrer les possibilités de la méthode, on a tracé les cartes d'iso-écart type d'estimation pour le réseau complet, c'est-à-dire 30 stations, et pour un réseau de 5 stations bien distribué sur le bassin versant. Les cinq stations utilisées sont Clifton River South, Cookshire, Maple Leaf Est, Sawyerville North et West Ditton. Ces stations correspondent aux numéros 10, 11, 20, 29 et 30 de la figure 2.1.

La figure 6.2 montre les résultats pour un réseau de 30 stations. La figure 6.3 montre les résultats pour un réseau de 5 stations. La ligne pointillée sur ces figures indique le contour du bassin versant de la rivière Eaton. L'étude de ces deux figures montre que l'écart type d'estimation sur le bassin versant a une valeur minimum de 0.26 pour le réseau de 30 stations et de 0.28 pour le réseau de 5 stations. La valeur maximum est dans le

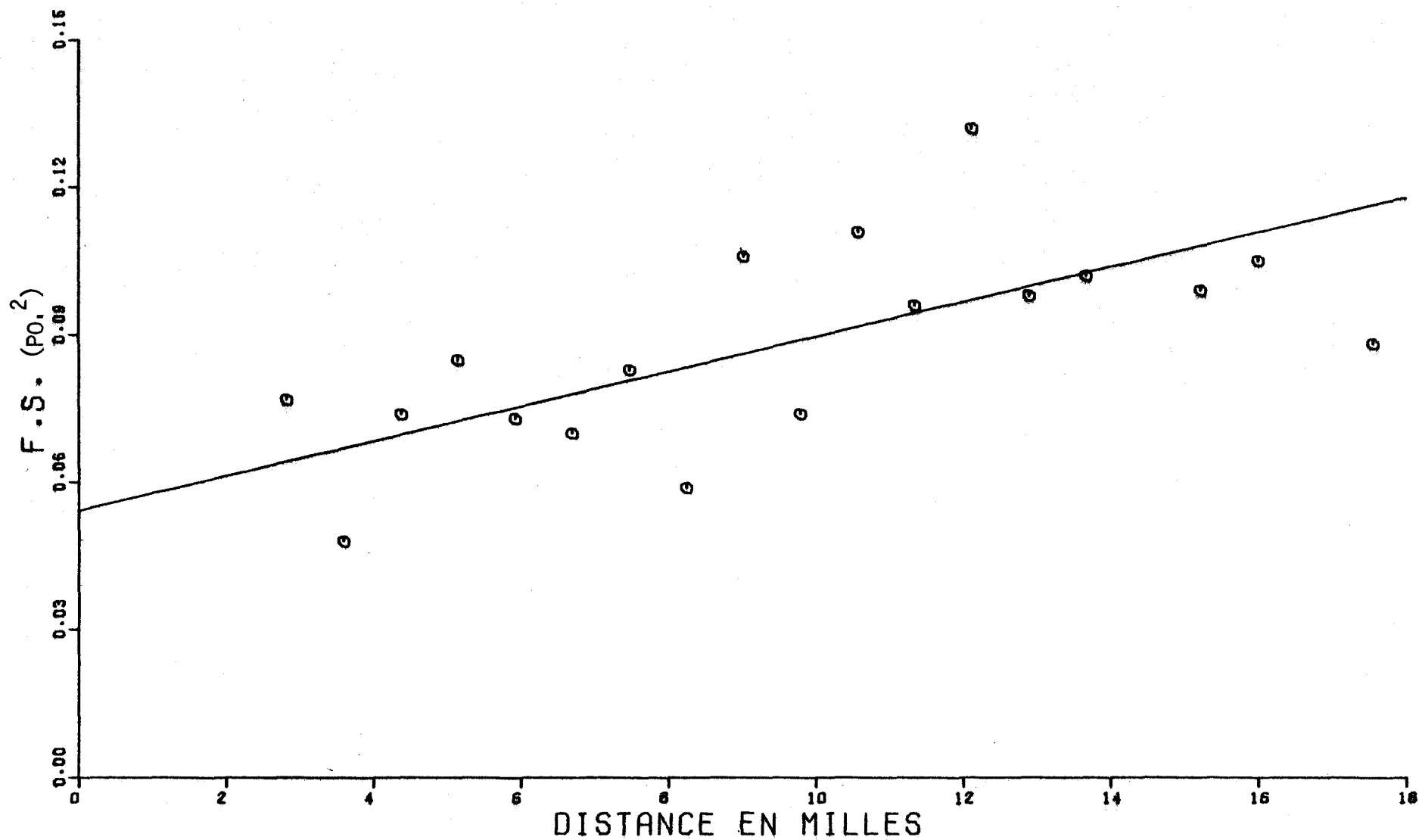


FIGURE 6.1 Fonction de structure dans l'espace des précipitations de 10 jours; saison "HIVER".

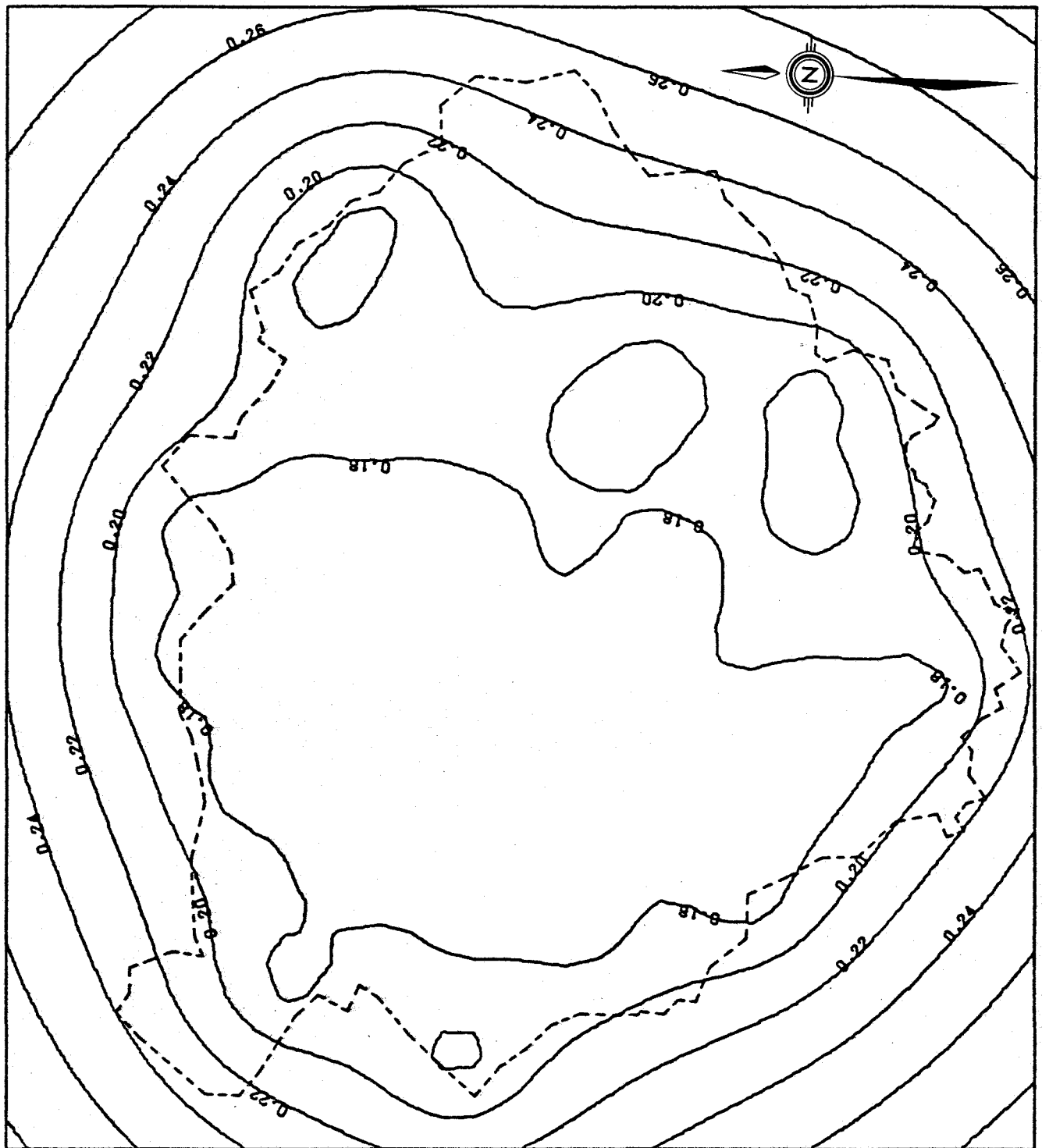


FIGURE 6.2 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 30 stations; saison "HIVER".

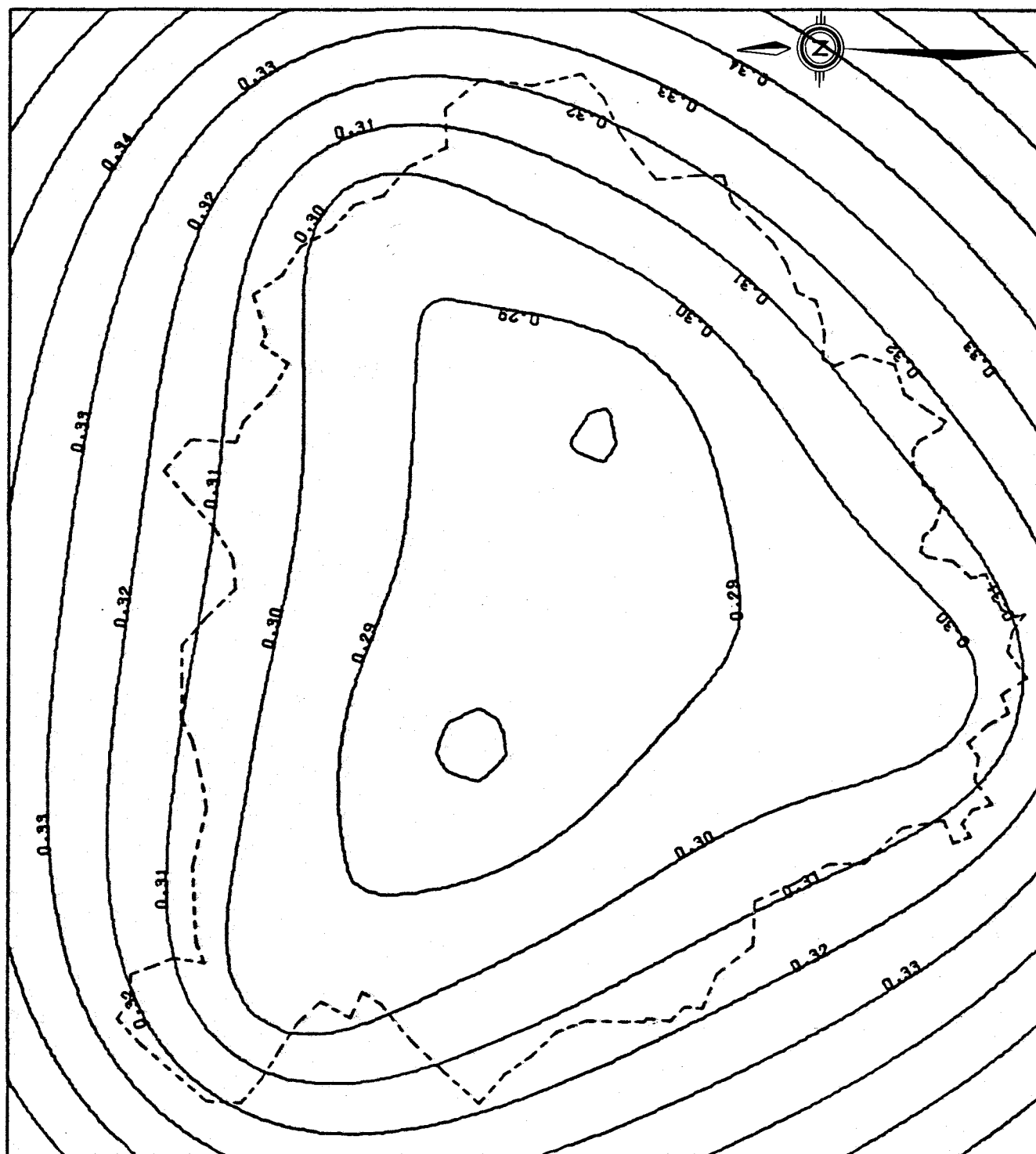


TABLEAU 6.3 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 5 stations; saison "HIVER".

premier cas de 0.30 et de 0.32 dans le second. Ces chiffres nous montrent que l'écart type d'estimation ponctuel diminue de seulement 0.02 si l'on ajoute 25 stations. Il faut noter ici que la prise en compte d'un grand nombre de stations pour l'interpolation ne diminue pas l'écart type d'estimation provenant de l'erreur de mesure ou de micro-climat mais diminue seulement la partie de l'erreur rattachée à l'interpolation en fonction des distances aux stations. On voit sur la courbe de fonction de structure (figure 6.1) que la variance d'estimation provenant des erreurs de mesure ou de micro-variation est d'environ .056 (ordonnée à l'origine) ceci implique que l'écart type d'estimation ponctuel ne peut être inférieur à 0.24 ($\sqrt{.056}$) même avec une infinité de stations ce qui explique le peu de gains obtenus en considérant 30 stations. En hydrologie, il est nécessaire de considérer non seulement l'erreur ponctuelle mais également l'erreur de l'estimation pour une surface donnée qui varie plus avec le nombre de station, comme on le verra plus loin. Notons enfin que l'augmentation régulière de l'erreur d'estimation à l'extérieur du bassin est due au fait que le réseau étudié ne comprend pas de station à l'extérieur du bassin.

6.4.2 Période "PRINTEMPS"

La fonction de structure pour la période "PRINTEMPS" nous est montrée à la figure 6.4. On voit que l'ordonnée à l'origine est plus faible que pour la période "HIVER" et que la pente de la

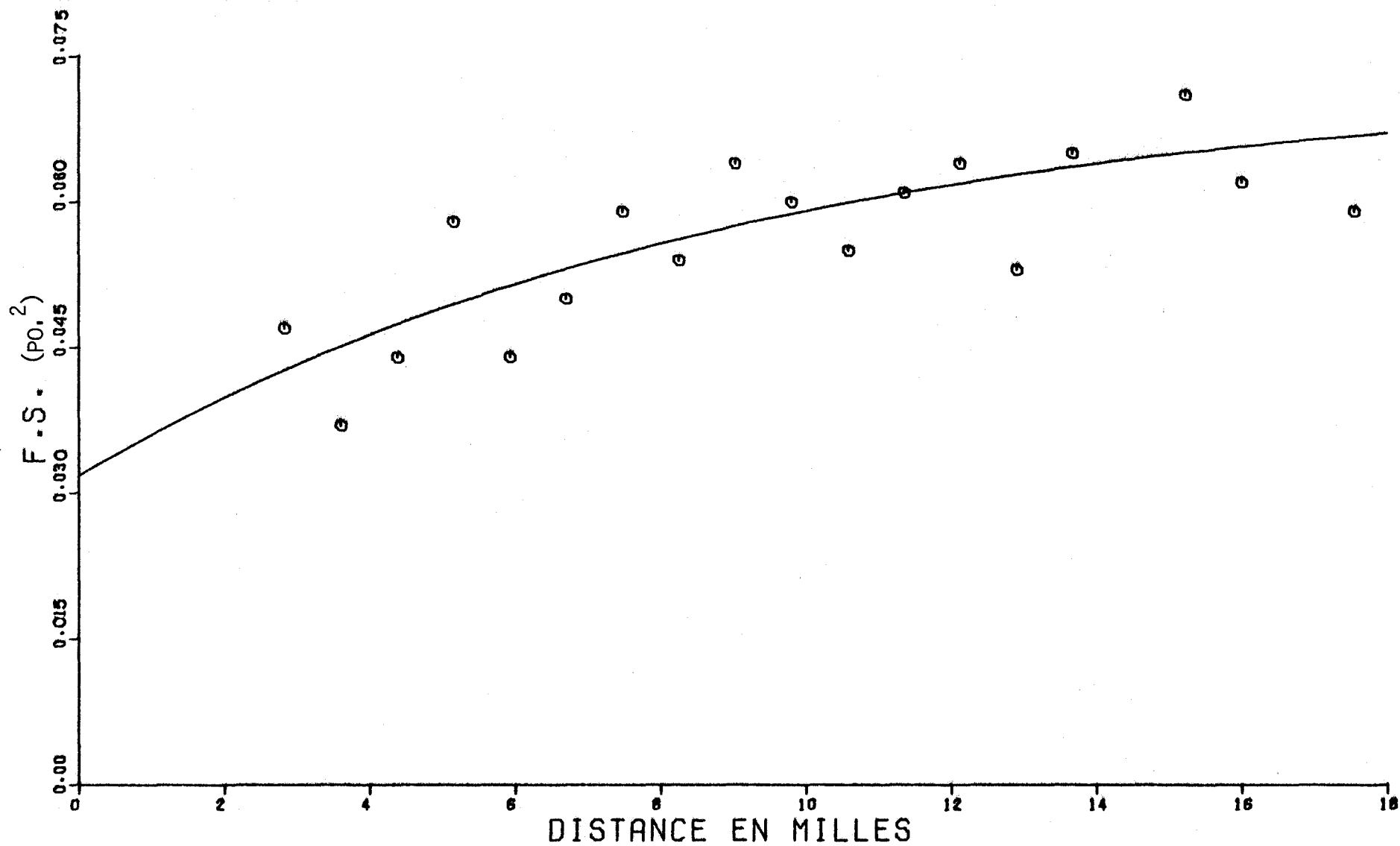


FIGURE 6.4 Fonction de structure dans l'espace des précipitations de 10 jours; saison "PRINTEMPS".

courbe, est plus forte pour des faibles distances, ce qui est interprété comme une diminution des liens entre les stations. Les figures 6.5 et 6.6 présentent les courbes d'iso-écart type d'estimation pour la saison "PRINTEMPS" avec 30 et 5 stations. La figure 6.6 fait très bien ressortir la position des 5 stations qui sont situées à l'intérieur des courbes d'iso-écart type 0.23. Pour les mêmes raisons que précédemment l'erreur augmente peu avec 5 stations.

6.4.3 Période "ETE"

Si l'on étudie la fonction de structure pour la période "ETE" présentée à la figure 6.7 on voit que la pente de la courbe est plus importante que pour les autres saisons. Ceci implique que l'erreur d'interpolation va augmenter plus rapidement avec la distance. Cette augmentation était prévisible compte tenu de la fréquence des orages localisés pour la saison été. Les figures 6.8 et 6.9 présentent les erreurs d'interpolation pour les réseaux de 30 et 5 stations. Pour un réseau donné, l'erreur (de mesure et de micro-climat, reste toujours à peu près la même en "ETE" et en "HIVER", tandis que l'écart type de l'estimé est plus grand pour la saison "ETE" ce qui est expliqué par la différence de pente des deux fonctions de structure.

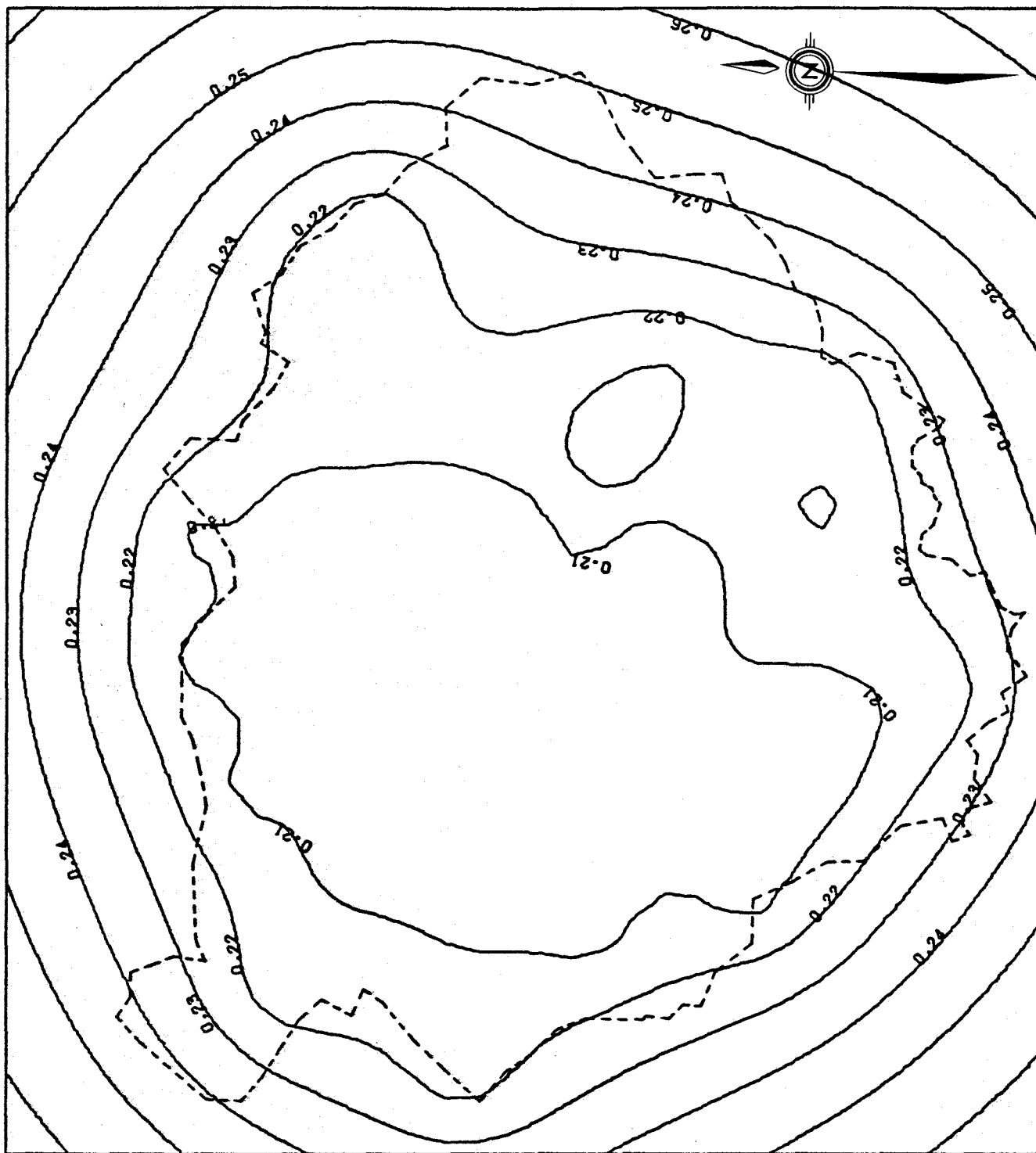


FIGURE 6.5 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 30 stations; saison "PRINTEMPS".

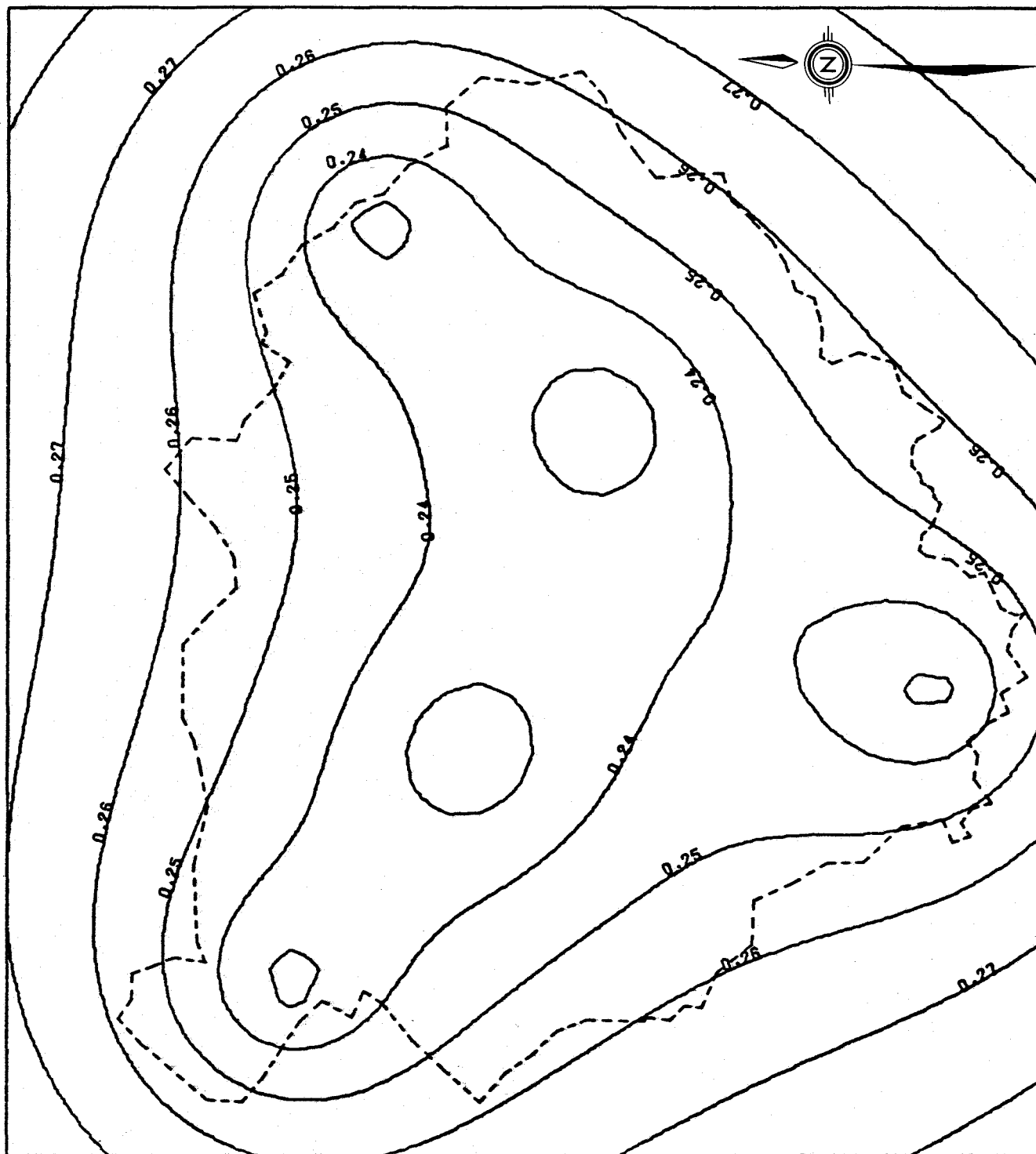


FIGURE 6.6 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 5 stations; saison "PRINTEMPS".

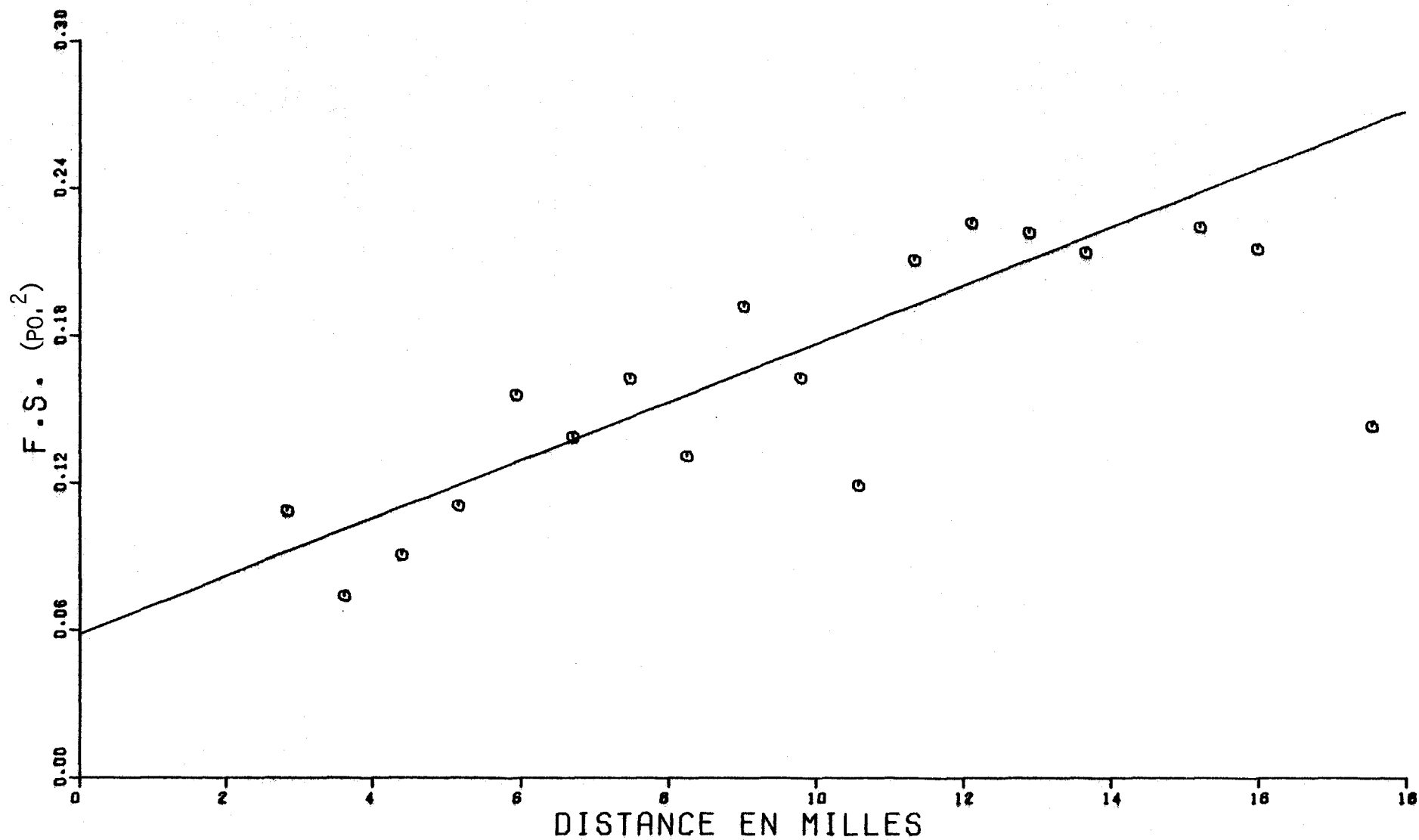


FIGURE 6.7 Fonction de structure dans l'espace des précipitations de 10 jours; saison "ETE".

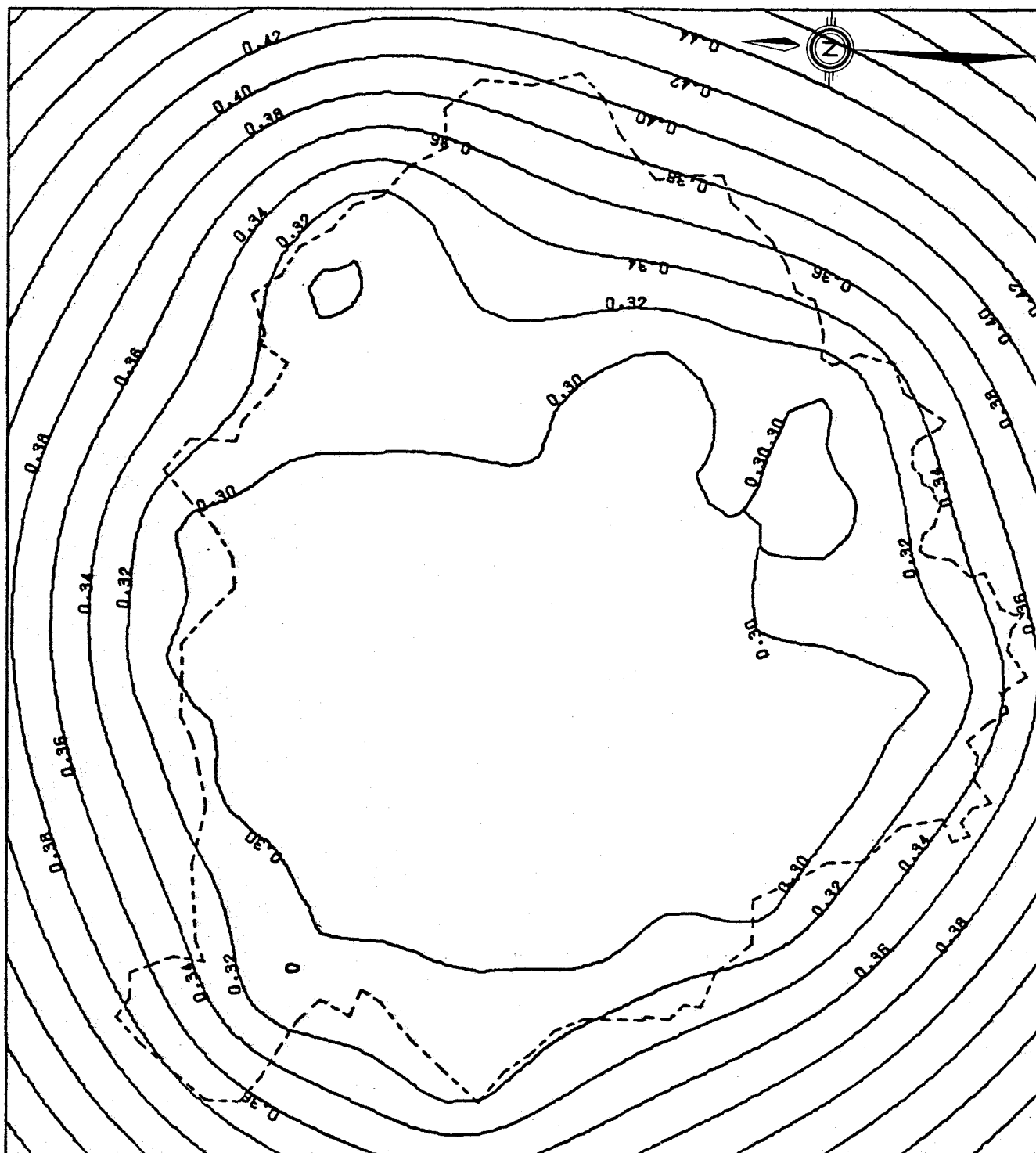


FIGURE 6.8 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 30 stations; saison "ETE".

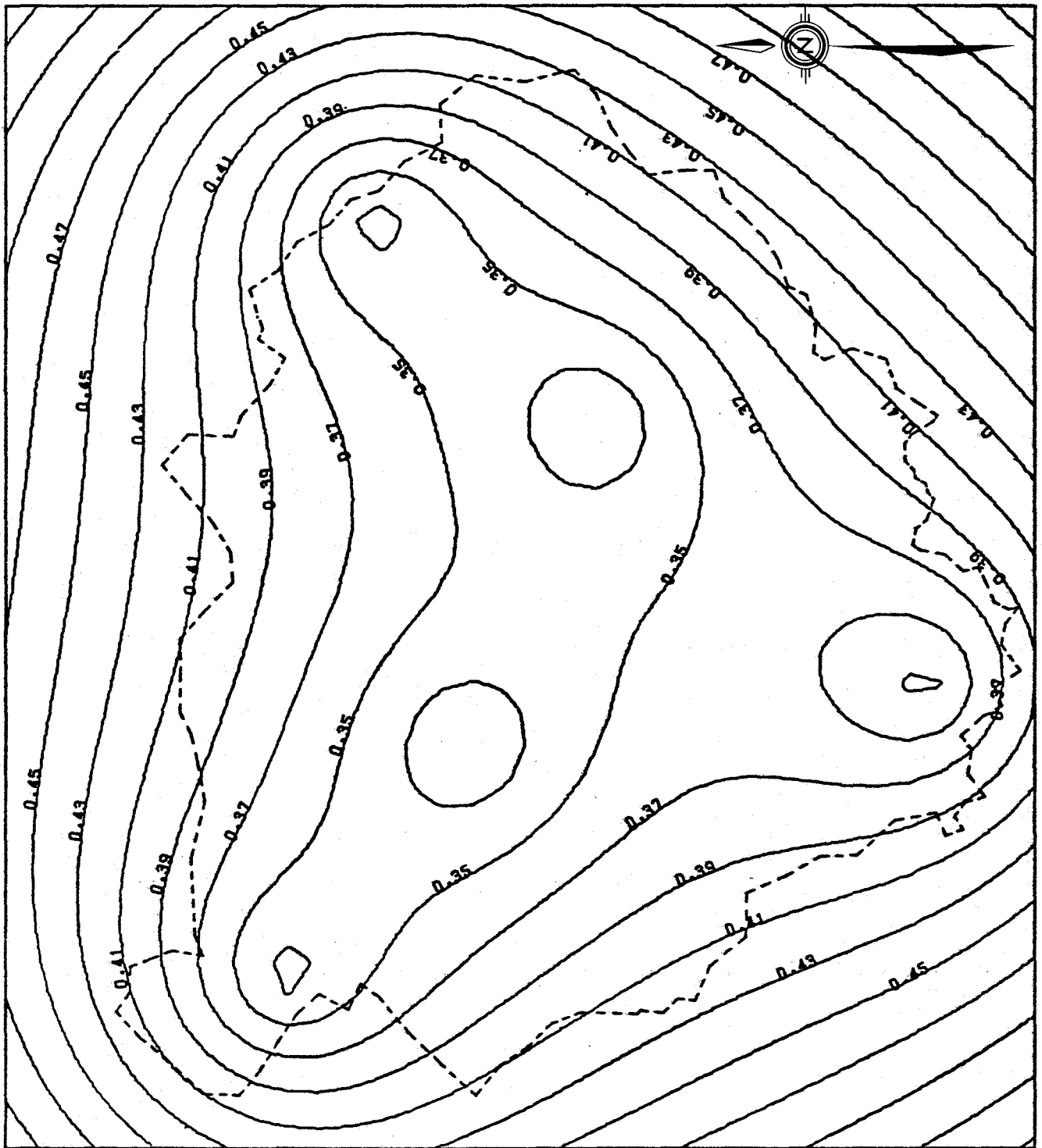


FIGURE 6.9 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 5 stations; saison "ETE".

6.4.4 Période "AUTOMNE"

Les résultats pour cette saison sont présentés sur les figures 6.10, 6.11 et 6.12. On remarque que la fonction de structure a pour ordonnée à l'origine une valeur faible, ce qui peut être expliqué par le peu d'orages locaux observés durant cette saison. On observe sur les figures 6.11 et 6.12 que l'écart type d'estimation est plus faible que pour les autres saisons.

6.5 ERREUR SUR DES SURFACES

Jusqu'à maintenant nous n'avons discuté que de l'évaluation des erreurs ponctuelles. Il est bien évident que l'intégration de ces erreurs sur une surface donnée ne correspond pas à l'erreur de l'estimé sur cette surface. Les études hydrologiques, plus particulièrement les modèles paramétriques ou statistiques pluie-débit, nécessitent la connaissance de ces erreurs.

Kagan (1966) a développé une méthode simple pour estimer l'erreur commise sur l'évaluation de la précipitation moyenne sur une superficie donnée.

La base de la méthode proposée par Kagan est la fonction de corrélation $\rho(d)$ qui dépend de la distance entre les stations. La fonction $\rho(d)$ sera représentée par la forme exponentielle

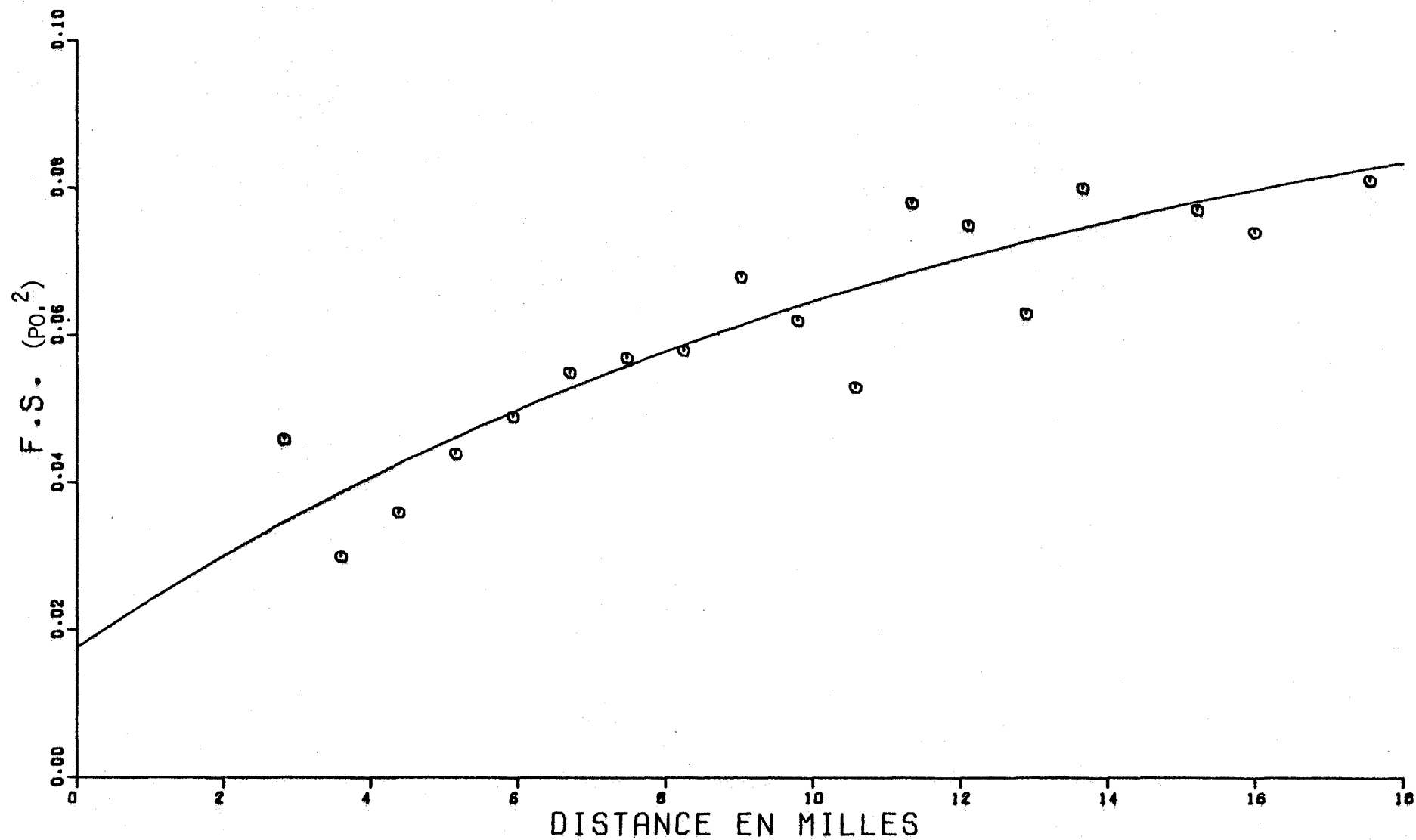


FIGURE 6.10 Fonction de structure dans l'espace des précipitations de 10 jours; saison "AUTOMNE".

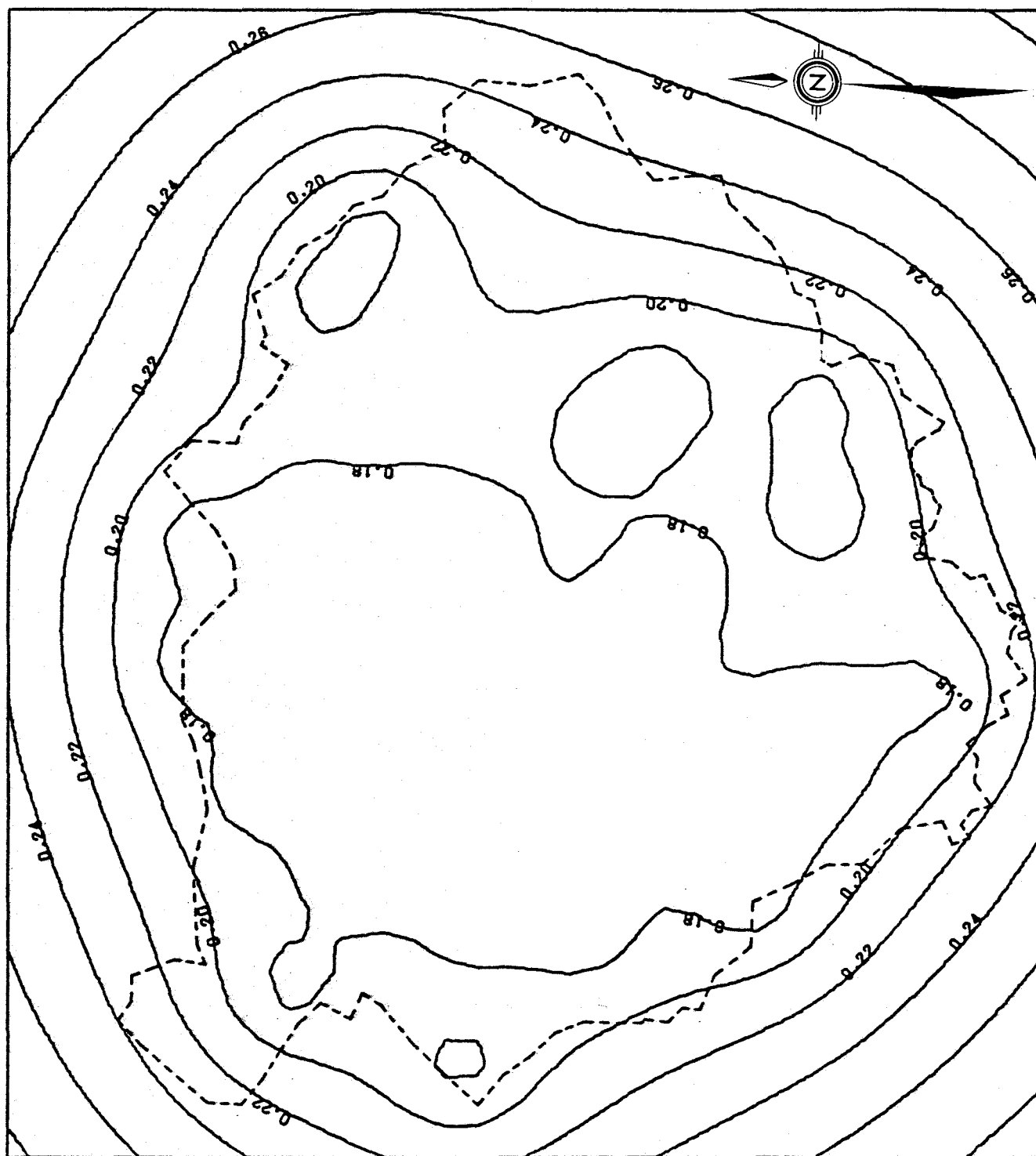


FIGURE 6.11 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 30 stations; saison "AUTOMNE".

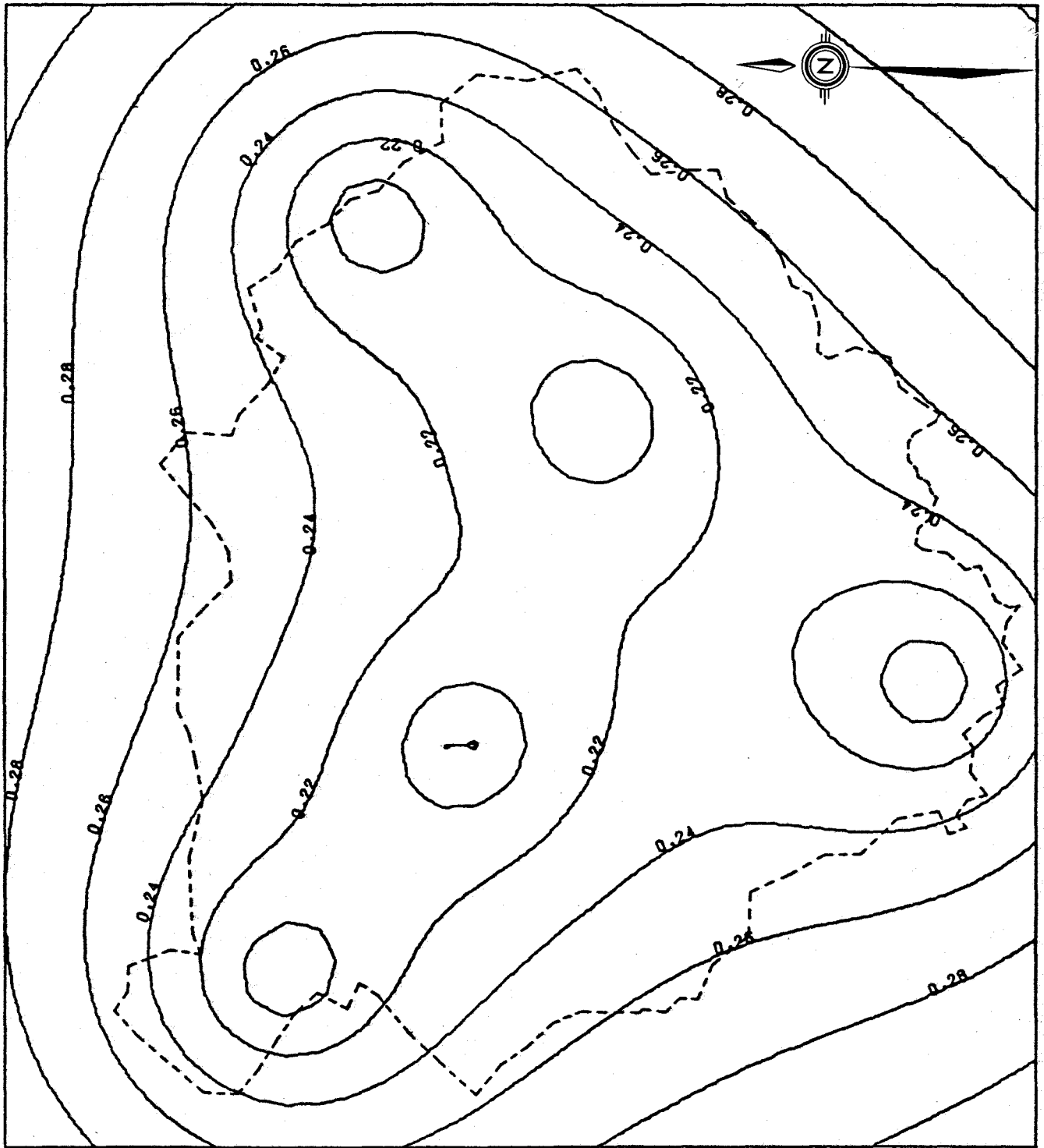


FIGURE 6.12 Ecart type (po.) d'estimation des précipitations de 10 jours.
Réseau de 5 stations; saison "AUTOMNE".

sui vante:

$$\rho(d) = \rho(o) e^{-d/d_o} \quad (6-5)$$

où:

$\rho(o)$ est la corrélation correspondant à une distance zéro

d_o = le "rayon de corrélation" ou la distance pour laquelle la corrélation est égale à $\rho(o)/e$

Théoriquement $\rho(o)$ doit être égal à l'unité mais en pratique c'est rarement le cas. Ceci est dû aux erreurs de mesures des précipitations et au micro-climat tout comme on l'a vu précédemment pour la fonction de structure. On peut d'ailleurs estimer l'écart type de ces erreurs par:

$$\sigma = \sqrt{(1 - \rho(o))} \sigma_h^2 \quad (6-6)$$

σ = écart type des erreurs de mesure

σ_h^2 = variance des précipitations en un point

Kagan a montré que la variance de l'erreur de la précipitation moyenne sur une surface avec une station centrale est donnée par:

$$V = \sigma_h^2 [1 - \rho(o)] + 0.23 \sigma_h^2 \frac{\sqrt{s}}{d_o} \quad (6-7)$$

Le premier terme est attribué à l'erreur de mesure et le second terme est attribué à la variance des champs de précipitations.

Pour une surface S avec n station distribuée uniformément la variance de l'erreur de la précipitation moyenne est donnée par

$$V_n = \frac{\sigma_h^2}{n} \left[1 - \rho(o) + \frac{0.23 \sqrt{S}}{d_o \sqrt{n}} \right] \quad (6-8)$$

L'erreur relative est alors définie par:

$$Z = \frac{\sqrt{V_n}}{\bar{h}} = C_v \sqrt{\frac{1 - \rho(o) + 0.23 \frac{\sqrt{S}}{d_o \sqrt{n}}}{n}} \quad (6-9)$$

où

Z = est l'erreur relative

$C_v = \sigma_h / \bar{h}$

$\bar{h}$ = précipitation moyenne sur S

Si $\rho(o)$ et d_o sont connus, on peut calculer le nombre de stations requises pour obtenir une erreur spécifiée Z , ou inversement évaluer Z pour un nombre n de stations.

Le réseau météorologique de la Eaton n'est pas distribué uniformément; on peut cependant estimer l'erreur sachant que les chiffres obtenus correspondent à un réseau dont les stations sont uniformément distribuées; ce vers quoi on doit tendre.

6.5.1 Analyse des résultats

La fonction de corrélation est déterminé en calculant pour tous les couples de stations, les coefficients de corrélations et la distance séparant les stations. Comme pour la fonction de structure, l'axe des distances est subdivisé en classes et on calcule à l'intérieur de chaque classe la corrélation moyenne. On ajuste alors la courbe exponentielle décrite par l'équation 6-5.

La figure 6.13 montre ces ajustements. Le tableau 6.1 donne les caractéristiques des précipitations de 10 jours du bassin versant de la rivière Eaton pour les quatre saisons.

L'étude du tableau 6.1 montre que d_0 est de 57 milles pour la saison été et de plus de 200 milles pour le printemps et l'automne. La corrélation pour une distance nulle varie de 0.833 pour l'hiver à .955 pour l'automne. Ces chiffres ne sont pas incompatibles avec les valeurs citées dans la littérature. Ainsi Haragan (1976) trouve des valeurs de d_0 de 110 à 220 milles pour les précipitations mensuelles des mois de mai à septembre au Texas. Gushchina (1967) trouve 140 milles pour des précipitations de 10 jours en été dans la région de Valdaï et de 200 milles dans une autre région de Russie.

Les valeurs entre parenthèses du tableau 6.1 donnent l'écart type d'estimation de $\rho(0)$ et de d_0 . L'écart type sur d_0 est

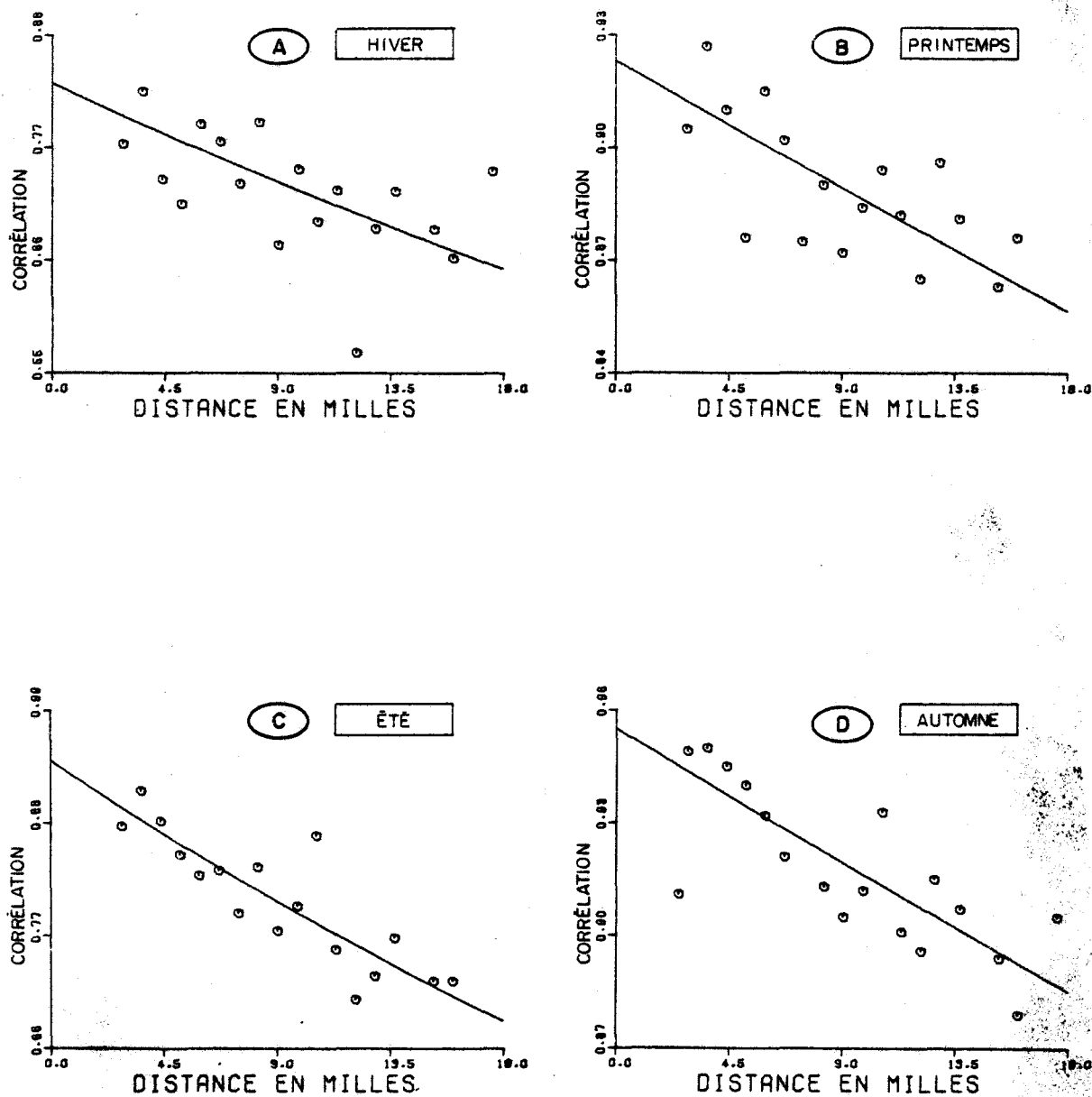


FIGURE 6.13 Exponentielle ajustée aux coefficients de corrélation des précipitations de 10 jours en fonction de la distance.

TABLEAU 6.1 Caractéristique des précipitations de 10 jours.
Bassin versant de la rivière Eaton.

SAISON	MOYENNE (POUCE)	σ (POUCE)	$\hat{c}_v$	$\rho(o)$	d_o (MILLES)
Hiver	1.02	.55	.54	.833(.026)*	73 (20)*
Printemps	.88	.70	.80	.923(.009)*	240 (65)*
Eté	1.47	.90	.61	.941(.02)*	57 (8.2)*
Automne	1.25	.78	.62	.955(.006)*	237 (56)*

* Les valeurs entre parenthèses sont les écarts type de l'estimé du paramètre.

TABLEAU 6.2 Erreur relative en % des précipitations de 10 jours en fonction du nombre de postes météorologiques.

NOMBRE DE STATIONS	HIVER	PRINTEMPS	ETE	AUTOMNE
30	4.13	4.12	2.95	2.47
20	5.09	5.07	3.69	3.05
10	7.29	7.23	5.42	4.37
5	10.50	10.35	8.06	6.31
2	17.16	16.74	13.90	10.35
1	25.13	24.27	21.35	15.22

important mais si l'on considère l'équation 6-9 on voit que Z est plus sensible à l'erreur sur $\rho(0)$ qu'à celle sur d_0 .

A l'aide de ces données et de la formule 6-9 on peut estimer l'erreur relative des précipitations de 10 jours calculées avec un nombre variable de stations; sachant que le bassin versant de la rivière Eaton est de 248 milles carrés et en supposant la répartition uniforme des stations. Les valeurs du tableau 6.2 nous montrent que l'erreur moyenne des précipitations sur une surface varie beaucoup plus avec le nombre de stations que l'erreur ponctuelle. La comparaison des résultats des saisons "été" et "automne" nous montre que même avec des "rayons de corrélation" très différents l'erreur relative varie peu lorsque le nombre de stations est assez grand, du moins pour une surface de dimension faible par rapport à ces "rayons de corrélation".

6.6 CONCLUSION

Ce chapitre a fait ressortir une grande variation des erreurs de mesure ou de micro-climat d'une saison à l'autre. Ainsi l'erreur est $.167 \left((1 - \rho(0)) \right)$ pour l'hiver et de .045 pour l'automne. Ces résultats sont facilement explicables si l'on considère la difficulté de mesurer avec précision les précipitations solides. Le "rayon de corrélation", qui est un indice de la portée d'une

station, est très faible pour l'été, ce qui s'explique par des orages localisés. La faible valeur de d_0 pour la saison hiver est plus difficile à expliquer physiquement; on peut cependant supposer que l'erreur de mesure a également une influence sur ce paramètre.

Finalement ce chapitre a fait ressortir l'importance que prend l'erreur de mesure et le micro-climat sur l'estimé des précipitations ponctuelles ou des précipitations moyennes sur une surface.

BIBLIOGRAPHIE

- BOIS, P. (1971).
Une méthode de contrôle de séries chronologiques utilisées
en climatologie et en hydrologie. Laboratoire de mécanique
des fluides, Université de Grenoble, section hydrologie, mai,
49 pages.
- GUSHCHINA, M.V., KAGAN, R.L. and POLISHCHUK, A.I. (1967).
Accuracy in determining the mean precipitation depth over an
area. Hydrology: Selected papers, Issue No. 6: 585-596.
- HARAGAN, D.R. (1976).
Spatial variation of precipitation on the Texas high plains.
Water resources bulletin, 12(6): 1191-1204, december.
- KAGAN, R.L. (1966).
An evaluation of the representativeness of precipitation data.
Works of the Main Geophysical Observatory, USSR, Vol. 191.

L'ellipse de la figure 4.5 nous montre qu'on peut avec un niveau de possibilité à 80% considérer que la station 11 fait partie du groupe B.

4.3.2 Période "PRINTEMPS"

Le tableau 4.3 présente les résultats de l'analyse en composantes principales des précipitations de 10 jours pour la période du printemps.

La première composante explique 88.9% de la variance et les composantes 2 et 3 expliquent respectivement 2.2% et 1.9% soit un total de 93%. La figure 4.6 nous montre que la corrélation multiple entre chaque station et les trois premières composantes s'approche de 1. Si l'on compare les figure 4.3 et 4.6 on voit que les coefficients de corrélation multiples sont plus près de 1 pour la station du "printemps" ce qui indique une plus grande corrélation spatiale qu'en hiver. Les stations qu'on peut grouper facilement sont 2, 4, 6 et 11 ainsi que 3, 4, 5, 7, 8, 9, 10, 12, 13 et 14.

Comme précédemment on a fait la rotation des trois premiers axes et les groupes obtenus sont:

- A) 3, 5, 7, 8, 9, 10, 12, 13 et 14
- B) 2, 4, 6 et 11
- C) 1