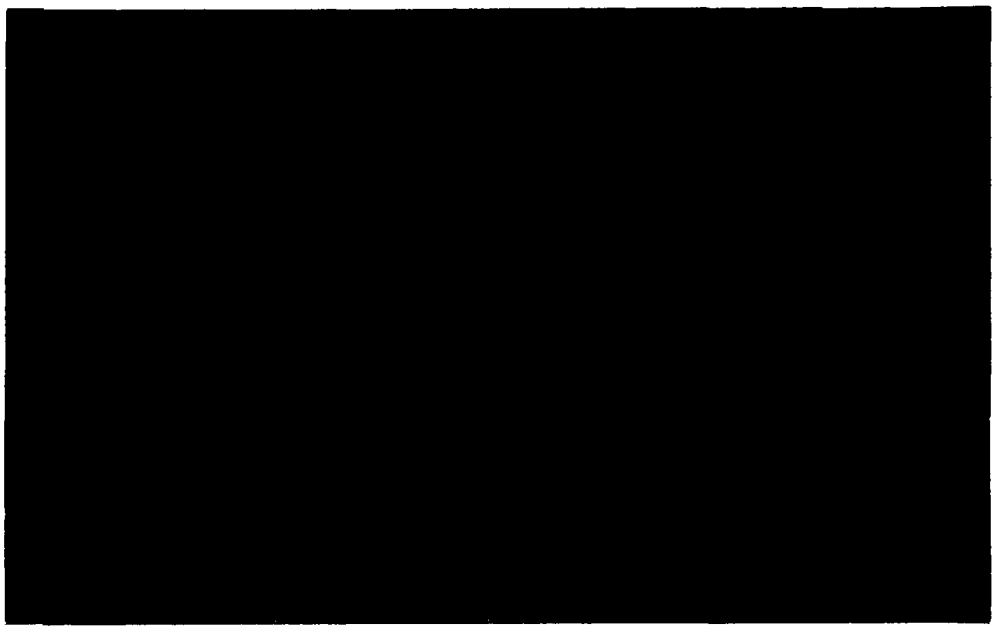


Arbeitsberichte der Schweizerischen Meteorologischen Zentralanstalt
Rapports de travail de l'Institut Suisse de Météorologie
Rapporti di lavoro dell'Istituto Svizzero di Meteorologia
Working Reports of the Swiss Meteorological Institute

Zürich

34002

No. 44

ETUDE D'UNE METHODE D'ANALYSE DISCRIMINANTE APPLIQUEE A LA PREVISION DES PRECIPITATIONS

par

J. ALTHERR, M. DUPANLOUP

Y. GANTER et E. JUNET

(groupe des prévisions objectives)

Mai 1974

Prévision 551. 509. 314:551. 509. 324. 2

Résumé :

La présente étude développe les bases d'une prévision automatique des précipitations par la méthode statistique de l'analyse discriminante multiple avec choix de variables.
Le test a été effectué pour le mois de janvier pour trois stations synoptiques suisses, en utilisant 13 prédicteurs.
Les résultats sont jugés suffisamment bons pour que cette étude soit étendue aux autres mois de l'année.

Zusammenfassung :

Die vorliegende Arbeit entwickelt die Grundlagen einer objektiven Vorhersage der Niederschläge mit Hilfe der Diskriminanz - Analyse mit Wahl der Variablen.
Getestet wurden die Resultate für den Monat Januar, wobei für drei schweizerische synoptische Stationen 13 Parameter Verwendung fanden.

Die Resultate wurden für genügend gut befunden, um eine Ausdehnung der Studien auf die übrigen Monate des Jahres als Fortsetzung vorzusehen.

Riassunto :

Il presente lavoro sviluppa le basi per una previsione obiettiva delle precipitazioni con l'aiuto dell'analisi discriminante con scelta delle variabili.
Furono presi come testi i risultati del mese di gennaio, per i quali furono utilizzati 13 parametri per 3 stazioni sinottiche svizzere.
I risultati furono trovati sufficientemente buoni, tanto da poter prevedere quale continuazione un'estensione dello studio ai rimanenti mesi dell'anno.

Summary :

This paper describes a method to forecast precipitation automatically by applying multiple discriminant analysis with chosen variables. The method was tested for the month of January for three Swiss synoptic weather stations, using thirteen parameters. The results are judged to be good enough so as to extend this investigation to include the remaining months of the year.

Remerciements

Les auteurs tiennent à dire leur gratitude à Mme A. Korten, statisticienne ; c'est la clarté de son approche théorique du problème qui a permis la réalisation de ce travail.

Introduction :

La technique développée ici et largement utilisée aujourd'hui, (par exemple Klein, 1971, Lönnqvist, 1968, Bengtsson et Moen, 1971, etc) est celle qui combine la prévision numérique (dynamique) et la prévision statistique et qu'on appelée " climatologie dynamique " (Bergeron, Chambettaz), " interprétation automatique des cartes de prévision " (Lönnqvist), ou encore "la méthode de prévision parfaite ". (" perfect prog method ", Veigas 1966).
Avec cette technique, les données du passé sont utilisées pour tirer des relations statistique entre un élément du temps (précipitations, température, ensoleillement, etc.) et certains paramètres mesurant la circulation de l'atmosphère. Ces relations sont ensuite appliquées aux cartes de prévision numérique, qui simulent la circulation réelle, pour produire une prévision automatique. Bien que les erreures dûes à la prévision numérique produisent inévitablement des erreurs correspondantes dans la prévision statistique, celle-ci s'améliore chaque fois que celle-là s'améliore. D'autre part, le récent développement des modèles numérique " humides " (Shuman et Hovermale, 1968) rendra possible dans un proche avenir l'inclusion de mesures d'humidité comme prédicteur des précipitations.
Différentes méthodes statistiques, assez proches les unes des autres, peuvent être utilisées pour la prévision automatique : la régression et l'analyse discriminante multiple par pas avec choix de variable (Klein, Miller) ou la technique de classification par type de temps (Lönnqvist). Les deux premières ont l'avantage de permettre une sélection parmi un grand nombre de prédicteurs pour n'en finalement retenir que les plus significatifs. La méthode de classification par type de temps procède plus par tâtonnement et doit se limiter à un choix généralement restreint et subjectif de prédicteurs.
Le choix de la méthode s'est donc porté sur l'analyse discriminante qui paraissait tout indiquée pour la prévision des précipitations par classes.

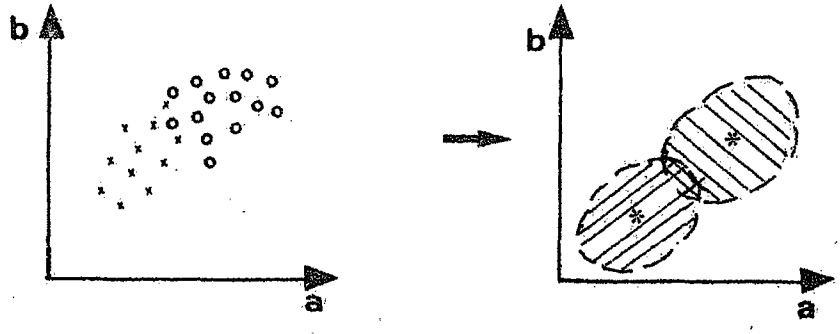
DESCRIPTION RESUMEE DE LA METHODE

Prenons un exemple simplifié à l'extrême :

- un échantillon d'observations bi-quotidiennes portant sur plusieurs années est divisé en deux groupes, le groupe RAIN et le groupe NORAIN, représentant les observations avec et sans précipitations.
- pour chacune des observations, on relève la valeur de deux autres paramètres (prédicteurs) a et b.
- le tableau suivant peut ainsi être dressé :

Observations		Mesure de a	Mesure de b	Moyennes
NORAIN	1	a'_{nr}	b'_{nr}	
	2	a''_{nr}	b''_{nr}	$\bar{a}_{nr}$, $\bar{b}_{nr}$
	3	a'''_{nr}	b'''_{nr}	
	...	...	...	
RAIN	1	a'_{r}	b'_{r}	
	2	a''_{r}	b''_{r}	$\bar{a}_{r}$, $\bar{b}_{r}$
	3	a'''_{r}	b'''_{r}	
	...	...	...	

- ces valeurs peuvent être reportées sur un diagramme à 2 coordonnées ; on obtient ainsi 2 nuages de points groupés autour des moyennes de chaque classe et possédant une partie commune plus ou moins importante :



× = groupe NORAIN

o = groupe RAIN

A partir de la matrice de classification donnée ci-dessus, on calcule les matrices des sommes des carrés et des produits à l'intérieur de chaque groupe :

$$W_{nr} = \begin{Bmatrix} S^{nr}_{aa} & S^{nr}_{ab} \\ & S^{nr}_{bb} \end{Bmatrix} \quad \text{et} \quad W_r = \begin{Bmatrix} S^{r}_{aa} & S^{r}_{ab} \\ & S^{r}_{bb} \end{Bmatrix}$$

puis les matrices des variances-covariances :

$$S_{nr} = \frac{W_{nr}}{N_{nr} - 1} \qquad S_r = \frac{W_r}{N_r - 1}$$

et enfin la matrice générale de dispersion à l'intérieur des groupes

$$S = \frac{W_{nr} + W_r}{N_{nr} + N_r - 2}$$

On calcule ensuite le vecteur " différence entre les moyennes " :

$$\underline{d} = (\overline{a}_{nr} \quad \overline{b}_{nr}) - (\overline{a}_r \quad \overline{b}_r)$$

Les coefficients de la fonction discriminante sont alors tirés de la relation

$$l' = S^{-1} . d'$$

sous la forme d'un vecteur $\underline{l}$

Dans l'exemple pris ici, on obtiendra à ce stade la valeur des deux coefficients l_1 et l_2 de la fonction discriminante.

La variance à l'interérieur des groupes S étant connue, on divisera alors chacun de ces coefficients par S pour obtenir une function discriminante dont la variance à l'intérieur des groupes est unité.

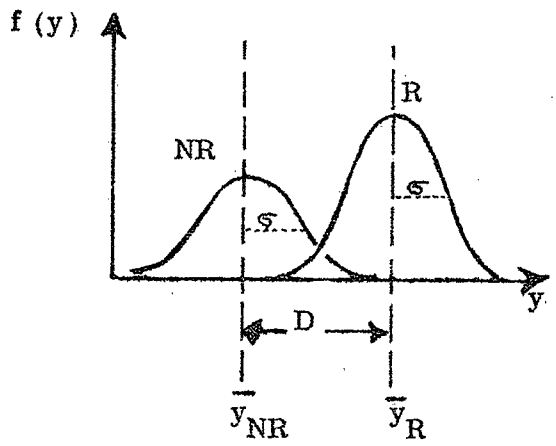
$$\underline{1} = \frac{1}{S} (l_1 \quad l_2)$$

L'analyse discriminante permet donc de calculer des coefficients l_1 et l_2 qui, appliqués aux variables a et b, fourissent une fonction discriminante

$$\boxed{y = l_1 . a + l_2 . b}$$

Cette fonction est donc une combinaison linéaire des variables, appliquée à chaque individu ; elle permet de tracer un histogramme, représentation graphique de son application aux individus du tableau précédant :

Observations		Valeur de la Fonction	Moyenne
NORAIN	1	y'_{nr}	
	2	y''_{nr}	$\bar{y}_{nr}$
	3	y'''_{nr}	
	....	....	
RAIN	1	y'_{r}	
	2	y''_{r}	$\bar{y}_{r}$
	3	y'''_{r}	
	....	....	



REMARQUES :

Après le traitement des données exposé plus haut :

- la variance à l'intérieur des groupes est unité ;
- on obtient facilement la distance D (distance de Mahalonobis) entre les moyennes des populations après standardisation des variables ;
- la fonction discriminante y est celle qui sépare le mieux les moyennes en fonction de la variabilité à l'intérieur des groupes. ELLE A MAXIMISE LE RAPPORT $\frac{D}{\sigma}$, notion importante sur laquelle nous reviendrons.

Soit maintenant (et c'est le problème posé par la prévision) un individu

inconnu dont il faut déterminer le classement dans le groupe RAIN ou dans le groupe NORAIN. On mesure les prédicteurs a et b, puis on calcule la fonction

$$y = l_1 \cdot a + l_2 \cdot b$$

Si

$$y_{\text{calculé}} - \overline{y}_{nr} < y_{\text{calculé}} - \overline{y}_{r}$$

alors l'individu inconnu sera classé dans le groupe NORAIN et réciproquement. Dans le cas où les populations des deux groupes ne sont pas égales, il est facile de calculer le seuil de discrimination qui tient compte de ce fait ; ce seuil s'exprime par la relation

$$y_{\text{limite}} = \left(\overline{y}_{nr} + \frac{1}{2} D\right) + \frac{1}{D} \ln \frac{\text{population NR}}{\text{population R}}$$

D'autre part, la probabilité d'erreur ("il y avait P% de chances que l'individu fasse partie de l'autre groupe ") est également facile à déterminer.

Le programme utilisé pour notre travail permet, bien sûr, une étude beaucoup plus approfondie du problème, en utilisant G groupes (à savoir : NORAIN ..... 0,7 à 5 mm ..... 5 à 10 mm ..... plus de 10 mm). et N prédicteurs (par exemple : précipitations dans les douze heures passées, gradients nord-sud et ouest-est, hauteurs des niveaux standard etc....).
Dans ce cas, la maximisation de $\frac{D}{\partial}$ est remplacée par la maximisation de l'expression

$$\frac{\text{somme des carrés entre les groupes}}{\text{somme des carrés à l'intérieur des groupes}}$$

ou encore, si B est la matrice des sommes des carrés et des sommes des produits entre les groupes et W la matrice des sommes des carrés et des sommes des produits à l'intérieur des groupes (B et W étant déterminées à partir de l'échantillon de développement), à maximiser

$$\frac{\underline{1}\ B\ \underline{1}'}{\underline{1}\ W\ \underline{1}'}$$

où $\underline{1} = (1_1\ 1_2 \ldots 1_n)$ représente les coefficients de la fonction discriminante.

Cette fonction aura donc la forme

$$y = l_1 \cdot a + l_2 \cdot b + \ldots + l_n \cdot n$$

Appliquée à un individu isolé, elle permettra de calculer avec quelle probabilité cet individu tombera dans le groupe NORAIN, avec quelle probabilité il tombera dans le groupe " 0.7 à 5 mm " et ainsi de suite.

On pourra ainsi prévoir le classement d'un individu inconnu, en connaissant ses probabilités d'appartenance à chacun des groupes choisis. (Un exemple pratique est donné plus loin).

Le cheminement du programme d'analyse discriminante que nous avons appliqué à notre échantillon a donc été le suivant :

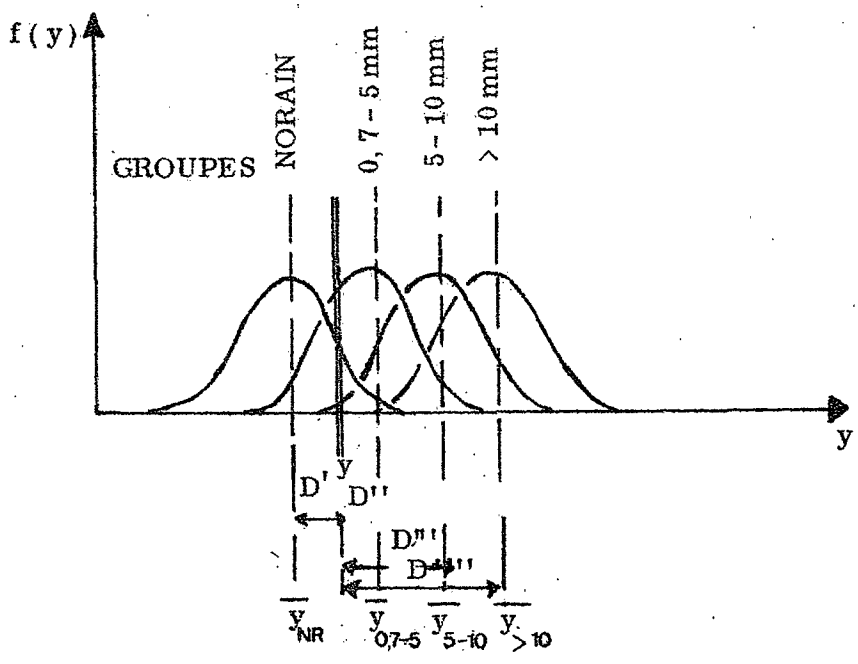
a) n prédicteurs sont " proposés " au départ ;

b) au cours de n pas successifs, un prédicteur est retenu à chaque pas en fonction de l'amélioration qu'il apporte à la séparation entre les groupes (les prédicteurs retenus en premier lieu étant ceux qui provoquent l'amélioration maximum) ;

c) dès que le choix de nouveaux prédicteurs n'apporte plus d'amélioration significative à la séparation des groupes, ces prédicteurs ne sont pas retenus ;

d) l'ordinateur calcule alors pour chaque individu de l'échantillon la valeur de la fonction discriminante y obtenue à partir des coefficients l_1, l_2, ...

e) il pourra ensuite calculer, pour chaque individu, la distance D qui sépare le point y des moyennes des groupes, ainsi que la probabilité " a posteriori " qu'a cet individu d'appartenir aux différents groupes.



APPLICATION A LA PREVISION AUTOMATIQUE

Problème : pour une station donnée, on désire connaître les probabilités de précipitation aux temps H+24, H+48 etc.

Deux possibilités se présentent, aussi intérssantes l'une que l'autre, et qui sont fonction de leurs utilisateurs potentiels.

PREMIERE POSSIBILITE : CLASSIFICATION EN DEUX GROUPES (RAIN OU NORAIN)

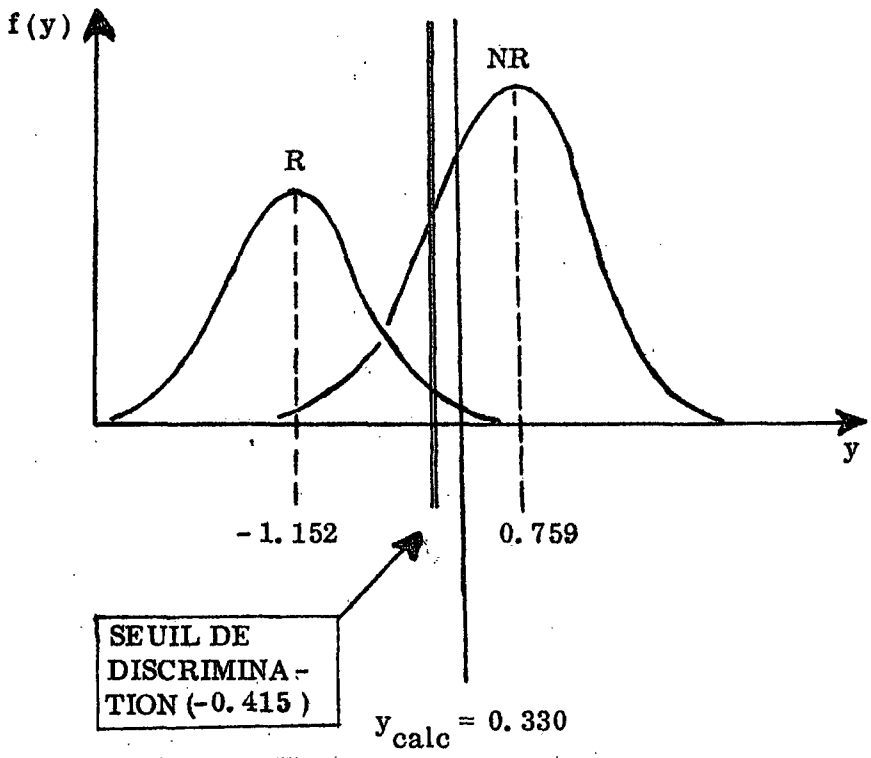
- on relève, sur les prévision numériques correspondant à la validité désirée, les valeurs des n prédicteurs choisis.
- on calcule la fonction $y = l_1 a + l_2 b + \ldots + l_n n$
- la station est alors classée dans le groupe RAIN (on prévoit de la pluie) ou dans le groupe NORAIN (on n'en prévoit pas) en fonction de la position du point y par rapport au seuil de discrimination.
- on calcule enfin le risque d'erreur au moyen des tables de la distribution normale.

Exemple :

Soit $y = 0.330$ le résultat du calcul de la fonction discriminante.

Dans l'échantillon que nous avons traité, le seuil de discrimination entre le groupe RAIN et le groupe NORAIN se trouve au point $y = -0.415$.

L'individu inconnu sera donc classé dans le groupe NORAIN, le calcul des probabilité indiquant un risque d'erreur de 7%.



DEUXIEME POSSIBILITE : CLASSIFICATION EN QUATRE GROUPES

- on relève, sur les prévisions numériques correspondant à la validité désirée, les valeurs des n prédicteurs choisis.
- on calcule la valeur d'y en fonction des coefficients de discrimination correspondant à cette classification.
- on détermine les distances D séparant l'individu à classer des moyennes des groupes (voir figure page 8).
- le calcul des probabilités permet alors de classer l'individu dans le groupe dont il a le plus de chances de faire partie, puis de calculer les probabilités qu'il avait de faire partie des trois autres groupes.

Exemple :

Soit y = 0. 066 le résultat du calcul de la fonction discriminante.

Groupe	Moyenne	D	Probabilité ponderée
NORAIN	- 0.505	- 0.571	0. 8689
0. 7 à 5 mm	1.162	1.096	0. 1133
5 à 10 mm	1. 969	1.903	0. 0127
plus de 10 mm	2.201	2.135	0. 0049

RESULTAT FINAL :

Cette méthode permettrait donc de produire des cartes de précipitations en prévision, dont la validité dépendrait de la validité des prévisions numériques ayant fourni les prédicteurs, et sur lesquelles, pour chaque station du réseau synoptique suisse, la prévision de pluie pourrait apparaître soit sous la forme simple "RAIN" ou "NORAIN", soit sous la forme d'un tableau des probabilités :

87%	11 %
NORAIN	0. 7 à 5 mm
5 à 10 mm	plus de 10 mm
1 %	< 1%

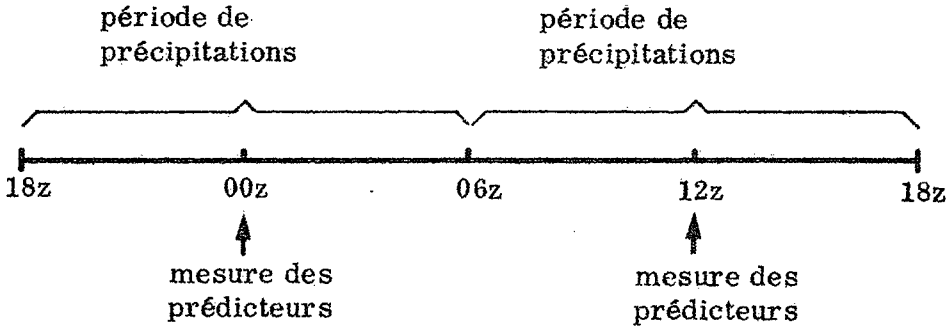
Données

Ce test a d'abord été entrepris pour rôder le programme BMDo7M sur l'ordinateur CDC 3800 de l'Université de Genève, mais nous lui avons également donné une certaine imporance quantitative afin que ses résultats puissent être immédiatement contrôlés " en routine " durant le mois de janvier 1974.
Nous avons choisi pour l'analyse 8 mois de janvier, soit de janvier 1964 à janvier 1971, à raison de deux mesures par jour. On obtient ainsi 496 cas. Les hauteurs des précipitations à Genève - Cointrin, Payerne et Zürich - Kloten ont été choisies comme prédictandes, alors que les 13 prédicteurs suivants ont été retenus :

1	RR - 12	précipitations dans les 12 heures précédentes
2	W/E	gradient ouest - est au sol
3	S/N	gradient sud - nord au sol
4	H_{500}	hauteur du niveau 500 millibars
5	H_{1000}	hauteur du niveau 1000 millibars
6	$H_{1000-500}$	topographie relative 1000 - 500 millibars
7	D_{850}	différence entre température et point de rosée à 850 millibars
8	D_{700}	différence entre température et point de rosée à 700 millibars
9	D_{500}	différence entre température et point de rosée à 500 millibars
10	Td_{850}	point de rosée à 850 millibars
11	Td_{700}	point de rosée à 700 millibars
12	DD_{500}	direction du vent à 500 millibars
13	FF_{500}	vitesse du vent à 500 millibars

Il est clair que cette liste de prédicteurs ne se veut en aucune manière exhaustive. Elle pourra être complétée lors de l'analyse définitive.

Les prédicteurs et prédictandes sont associés de la façon suivante : les mesures des prédicteurs sond prises deux fois par jour, à 00z et 12z. A ces mesures correspond une période de précipitations de 12 heures allant de 18z à 06z pour la mesure de 00z, de 06z à 18z pour celle de 12z.

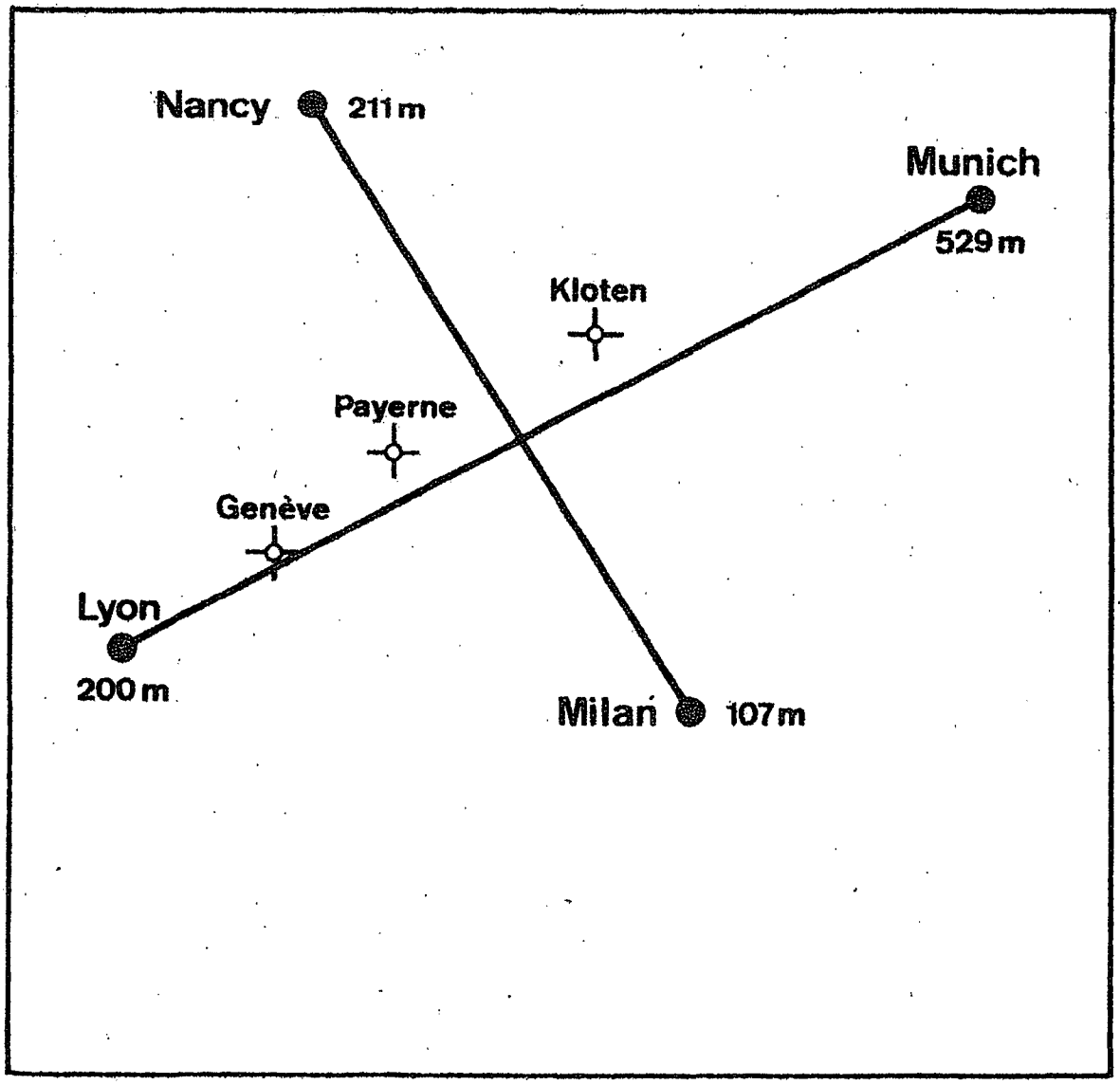


Les hauteurs de précipitations RR (prédictande) et RR - 12 ont été relevées sur les synops de Genève - Cointrin, Payerne (Berne pour l'année 1964), et Zürich - Kloten. Les précipitations dans les 12 heures précédentes ont été classées en 2 groupes seulement, à savoir d'une part, pas de pluie ou pluie inférieure à 1 mm, d'autre part, pluie supérieure à 1 mm. Les gradients W/E et S/N ont été relevés sur les Météos suisses de 00z et 12z et calculés d'après la formule indiquée sur la figure 1. Les autres prédicteurs ont été tirés des sondages de Payerne (00z et 12z).
L'analyse discriminante a été effectuée en 2 et 4 classes, de la manière suivante :

Division en 2 classes : 1er groupe NORAIN = RR inférieurs à 1 mm (0 - 0.6 mm)
2e groupe RAIN = RR supérieures à 1 mm (0.7 - et plus)

Division en 4 classes : 1er groupe = RR 0 - 0.6 mm
2e groupe = RR 0.7 - 4.9 mm
3e groupe = RR 5.0 - 9.9 mm
4e groupe = RR plus de 10 mm

Une seconde classification a été testée, en prenant comme limite entre RAIN et NORAIN RR= 0.3 mm et en mettant dans le groupe NORAIN les cas de neige de stratus. L'analyse discriminante donne pour ce découpage des résultats légèrement inférieurs à la division précédente.



Gradient S/N donné par $PPP_{(Milan)} - PPP_{(Nancy)}$

E/W " " $PPP_{(Munich)} - PPP_{(Lyon)}$

GENEVE - COINTRIN (moyennes et déviations standards)

NUMBER OF VARIABLES 13

NUMBER OF GROUPS 2

NUMBER OF CASES IN EACH GROUP 372 124

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP NORAIN	RAIN	
1	0.13978	0.59677	0.25403
2	-3.43817	44.49194	8.54435
3	6.46237	-20.53226	-0.28629
4	553.29570	545.41129	551.32460
5	181.59140	97.36290	160.53427
6	5352.02151	5357.15323	5353.30444
7	45.23656	13.13710	37.21169
8	84.06183	18.71774	67.72581
9	86.22581	22.44355	70.28024
10	-63.73656	-25.42742	-54.15927
11	-168.61559	-109.22581	-153.76815
12	21.19355	25.52419	22.27621
13	31.95161	42.63484	34.61492

STANDARD DEVIATIONS

VARIABLE	GROUP NORAIN	RAIN
1	0.34723	0.49254
2	66.20237	71.04945
3	41.91406	60.38541
4	13.21759	10.19007
5	75.59582	80.21124
6	93.56837	67.87004
7	49.74520	22.65447
8	78.91924	48.56821
9	72.06663	33.41104
10	52.63494	35.52491
11	74.52444	62.75981
12	12.09866	6.32258
13	17.97281	20.66528

Les variables sont données dans les unités suivantes :

W/E	en dixièmes de millibar
S/N	en dixièmes de millibar
H_{500}	en décamètres
H_{1000}	en mètres
$E_{1000-500}$	en mètres
$D_{850,700,500}$	en dixièmes de degré
$Td_{850,700}$	en dixièmes de degré
DD_{500}	en dizaines de degré
FF_{500}	en noeuds

NUMBER OF VARIABLES 13

NUMBER OF GROUPS 4

NUMBER OF CASES IN EACH GROUP 372 75 31 18

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP 0-0,6MM	0,7-5MM	5-10MM	PL10MM	
1	0.13978	0.54667	0.67742	0.66667	0.25403
2	-3.43817	27.36000	59.93548	89.27778	8.54435
3	6.46237	-18.33333	-34.29032	-6.00000	-0.28629
4	553.29570	546.20000	543.61290	545.22222	551.32460
5	181.59140	110.65667	82.38710	67.72222	160.53427
6	5352.02151	5351.20000	5355.67742	5384.50000	5353.30444
7	45.23656	15.57333	9.54839	9.16667	37.21169
8	84.06183	19.12000	24.45161	7.16667	67.72581
9	86.22581	28.61333	17.03226	6.05556	70.28024
10	-63.73656	-30.28000	-21.58065	-11.83333	-54.15927
11	-168.61559	-111.89333	-116.54839	-85.50000	-153.76815
12	21.19355	25.21333	26.54839	25.05556	22.27621
13	31.95161	39.38667	46.70968	48.94444	34.61492

STANDARD DEVIATIONS

VARIABLE	GROUP 0-0,6MM	0,7-5MM	5-10MM	PL10MM
1	0.34723	0.53117	0.47519	0.48537
2	66.20237	70.59336	70.57475	46.77833
3	41.91406	50.49253	80.37379	57.40158
4	13.21759	10.66061	9.61137	9.21352
5	75.59582	78.37804	75.36298	87.56296
6	93.56837	69.50948	66.80438	58.89797
7	49.74520	25.94953	14.50480	18.56071
8	78.91924	45.62448	65.53210	15.02253
9	72.06663	38.94448	22.88156	6.71672
10	52.63494	37.73116	29.90627	31.97839
11	74.52444	61.35796	75.03636	38.06689
12	12.09866	7.62516	3.10740	4.06524
13	17.97281	20.73963	21.97755	15.68991

PAYERNE (moyennes et déviations standards)

NUMBER OF VARIABLES 13

NUMBER OF GROUPS 2

NUMBER OF CASES IN EACH GROUP 387 109

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP NORAIN	RAIN	
1	0.13695	0.53211	0.22379
2	0.33333	37.69725	8.54435
3	7.29716	-27.21101	-0.28629
4	552.87855	545.80734	551.32460
5	175.70801	106.66055	160.53427
6	5353.75969	5351.68807	5353.30444
7	44.24548	12.23853	37.21169
8	81.56589	18.58716	67.72581
9	83.14729	24.59633	70.28024
10	-61.59432	-27.76147	-54.15927
11	-164.58140	-115.37615	-153.76815
12	21.34109	25.59633	22.27621
13	31.99483	43.91743	34.61492

STANDARD DEVIATIONS

VARIABLE	GROUP NORAIN	RAIN
1	0.34424	0.50127
2	66.89618	75.42712
3	43.11338	56.86958
4	13.20654	10.44959
5	80.52747	78.33702
6	91.59803	73.16694
7	49.36225	21.37795
8	78.86883	49.02403
9	71.12914	42.81758
10	52.14759	39.96756
11	76.32434	62.24753
12	11.82197	7.12341
13	18.53410	18.81643

Les variables sont données dans les unités suivantes :

W/E	en dixièmes de millibar
S/N	en dixièmes de millibar
H_{500}	en décamètres
H_{1000}	en mètres
$E_{1000-500}$	en mètres
$D_{850, 700, 500}$	en dixièmes de degré
$Td_{850, 700}$	en dixièmes de degré
DD_{500}	en dizaines degré
FF_{500}	en noeuds

NUMBER OF VARIABLES 13

NUMBER OF GROUPS 4

NUMBER OF CASES IN EACH GROUP 387 68 36 5

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP NORAIN	.7-4.9	5-9.9	PL10	
1	0.13695	0.48529	0.58333	0.80000	0.22379
2	0.33333	29.32353	49.44444	67.00000	8.54435
3	7.29716	-24.82353	-34.08333	-10.20000	-0.28629
4	552.87855	545.47059	546.80556	543.20000	551.32460
5	175.70801	109.61765	112.80556	22.20000	160.53427
6	5353.75969	5344.64706	5356.91667	5409.80000	5353.30444
7	44.24548	13.50000	10.19444	9.80000	37.21169
8	81.56589	22.77941	13.11111	1.00000	67.72581
9	83.14729	31.25000	14.83333	4.40000	70.28024
10	-61.59432	-31.20588	-24.44444	-4.80000	-54.15927
11	-164.58140	-124.22059	-106.05556	-62.20000	-153.76815
12	21.34109	25.73529	25.19444	26.60000	22.27621
13	31.99483	40.86765	46.55556	66.40000	34.61492

STANDARD DEVIATIONS

VARIABLE	GROUP NORAIN	.7-4.9	5-9.9	PL10
1	0.34424	0.50350	0.50000	0.44721
2	66.89618	74.69809	77.34799	65.51336
3	43.11338	51.53667	64.11613	77.32852
4	13.20654	9.88210	11.78575	8.81476
5	80.52747	67.23850	88.59049	109.02156
6	91.59803	76.94246	66.57987	35.80084
7	49.36225	17.64873	27.40158	21.91347
8	78.86883	53.15420	43.56327	2.23607
9	71.12914	50.07140	24.74729	5.17687
10	52.14759	36.71128	44.99489	43.37280
11	76.32434	61.25012	63.94102	18.95257
12	11.82197	7.74523	6.35079	3.04959
13	18.53410	17.07307	20.37428	13.86723

ZURICH - KLOTEN (moyennes et déviations standards)

NUMBER OF VARIABLES 13

NUMBER OF GROUPS 2

NUMBER OF CASES IN EACH GROUP 376 120

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP NORAIN	RAIN	
1	0.14628	0.55000	0.24395
2	2.07713	28.80833	8.54435
3	10.52660	-34.16667	-0.28629
4	552.92287	546.31667	551.32460
5	175.61436	113.28333	160.53427
6	5354.26330	5350.30000	5353.30444
7	45.58777	10.96667	37.21169
8	83.98138	16.79167	67.72581
9	83.70213	28.22500	70.28024
10	-62.73404	-27.29167	-54.15927
11	-166.60904	-113.53333	-153.76815
12	20.88298	26.64167	22.27621
13	31.99468	42.82500	34.61492

STANDARD DEVIATIONS

VARIABLE	GROUP NORAIN	RAIN
1	0.35385	0.49938
2	69.23288	70.89515
3	41.34130	53.87697
4	13.12146	11.16491
5	80.14911	82.44837
6	91.61931	74.91918
7	49.47449	20.51641
8	79.64784	41.84101
9	71.84232	43.67594
10	52.90992	36.14615
11	77.59604	54.91031
12	11.84037	6.72234
13	18.58432	18.63838

Les variables sont données dans les unités suivantes :

W/E	en dixièmes de millibar
S/N	en dixièmes de millibar
H_{500}	en décamètres
H_{1000}	en mètres
$E_{1000-500}$	en mètres
$D_{850, 700, 500}$	en dixièmes de degré
$Td_{850, 700}$	en dixièmes de degré
DD_{500}	en dizaines de degré
FF_{500}	en noeuds

NUMBER OF VARIABLES 13

NUMBER OF GROUPS 4

NUMBER OF CASES IN EACH GROUP 376 81 31 8

MEANS (THE LAST COLUMN CONTAINS THE GRAND MEANS OVER THE GROUPS USED IN THE ANALYSIS)

VARIABLE	GROUP NORAIN	.7-4.9	5-9.9	PL10	
1	0.14628	0.48148	0.67742	0.75000	0.24395
2	2.07713	25.03704	34.83871	43.62500	8.54435
3	10.52660	-29.92593	-38.32258	-61.00000	-0.28629
4	552.92287	546.08642	546.58065	547.62500	551.32460
5	175.61436	116.50617	109.35484	95.87500	160.53427
6	5354.26330	5344.72839	5357.09677	5380.37500	5353.30444
7	45.58777	11.62963	11.83871	0.87500	37.21169
8	83.98138	18.67901	15.87097	1.25000	67.72581
9	83.70213	32.27160	22.64516	8.87500	70.28024
10	-62.73404	-30.50617	-23.90323	-7.87500	-54.15927
11	-166.60904	-119.28395	-108.00000	-76.75000	-153.76815
12	20.88298	27.02469	26.48387	23.37500	22.27621
13	31.99468	40.51852	46.01290	51.50000	34.61492

STANDARD DEVIATIONS

VARIABLE	GROUP NORAIN	.7-4.9	5-9.9	PL10
1	0.35385	0.50277	0.47519	0.46291
2	69.23288	73.42963	64.32423	74.25042
3	41.34130	50.45364	54.67564	79.59182
4	13.12146	11.12340	11.36089	12.21167
5	80.14911	76.72779	87.37450	122.58226
6	91.61931	81.35032	57.96571	61.55123
7	49.47449	16.69539	29.93671	2.10017
8	79.64784	43.21280	43.45629	2.43487
9	71.84232	45.97799	41.65137	8.88719
10	52.90992	34.59014	40.52683	29.94966
11	77.59604	56.05360	53.02641	35.12325
12	11.84037	6.64826	6.20146	9.17975
13	18.68432	17.55855	19.61237	22.85982

Moyennes et déviations standards des prédicteurs

Elles sont calculées en premier lieu par le programme et figurent dans les tableaux ci contre. Vu le manque de temps, il n'a pas été possible de transformer certains prédicteurs dont la distribution ne s'avère pas absolument normale et qui ont été utilisés tels quels dans l'analyse. Si l'on compare les moyennes et les déviations standards des différents groupes pour les 3 stations Genève, Payerne et Zürich, on remarque d'und façon générale, en fonction de ce qui a été dit plus haut (maximisation du rapport $\frac{A}{\sigma}$), que la variabilité à l'intérieur des groupes étant viruellement la même pour les trois stations, ce sont les écarts entre les moyennes qui apporteront la discrimination.

Division en 2 groupes :

A Genève, on note une composante est pour le groupe NORAIN (-3. 4), alors qu'a Payerne (0. 33) et à Zürich (2. 07) existe une faible composante ouest, confirmant le fait que les situations d'est sont plus favorables au temps sec dans l'ouest que dans l'est du Plateau (fin du barrage du nord). Dans le groupe RAIN, l'inverse se produit, la composante ouest étant la plus élevée à Genève, la plus faible à Zürich.

Quant au gradient S/N, la composante nord la plus élevée se trouve dans le groupe RAIN de Zürich, soulignant bien l'effet de barrage. La composante sud est également plus élevée qu'à Payerne et à Genève dans le groupe NORAIN, confirmant cette fois l'effet du foehn.

La hauteur des niveaux 500 et 1000 millibars et la topographie relative 1000 - 500 millibars sont pratiquement identiques pour les trois stations. On notera cependant que H_{1000} à Zürich reste plus élevé dans le groupe RAIN qu'à Genève et Payerne et inférieur pour le groupe NORAIN. Cette différence peut également être imputée aux situation de barrage du nord, qui provoquent des précipitations dans l'est de la Suisse, alors même que la pression au sol est relativement élevée.

La topographie relative 1000 - 500 millibars est semblable pour les groupes RAIN et NORAIN et nous verrons que ce paramètre n'est jamais retenu comme significatif pour discriminer entre

ces deux groupes précipitations.

Les mesures de l'humidité (D et Td), ainsi que la vitesse du vent à 500 millibars ne présentent pas de différences significatives entre les trois stations. Quant aux moyennes des direction du vent à 500 millibars, on ne peut en tirer de conclusions valables pour l'instant, ce paramètre étant précisément l'un de ceux qu'il faudra transformer dans l'analyse définitive.

Division en 4 groupes :

Dans l'analyse en 4 groupes, on note les mêmes tendances que pour 2 groupes, c'est-à-dire, pour le gradient W/E, le renforcement de la composante ouest à mesure que la hauteur des précipitations augmente (avec la progression la plus nette pour Genève). Avec le gradient S/N, même augmentation des précipitations à Zürich pour la composante nord, alors qu'à Payerne et Genève, celle-ci diminue entre les groupes 5. 0 - 9. 9 mm et plus de 10 mm, indiquant par là que les fortes précipitations se produisent avec une composante sud-ouest à ces 2 stations (plutôt nord-ouest à Zürich).

La hauteur du niveau 500 millibars donne des valeurs irrégulières : relativement bas pour les fortes précipitations à Payerne, alors qu'il reste élevé pour ce groupe aux 2 autres stations.

Pour H_{1000}, la baisse est progressive (sauf Payerne : légère hausse pour le groupe 5. 0 - 9. 9 mm) à mesure que les précipitations s'élèvent.

La topographie relative 1000 - 500 millibars reste uniforme pour les trois premiers groupes, alors qu'elle augmente pour le quatrième (plus de 10 mm), indiquant bien que les fortes précipitations se produisent en association avec un afflux marqué d'air chaud (du moins en janvier).

Les différences entre la température et le point de rosée tendent - ce qui est logique - vers la saturation et les températures du point de rosée augmentent en fonction des précipitations.

Quant à la vitesse du vent à ce au niveau 500 mb, elle augmente elle-aussi en fonction de la hauteur des précipitations, ce qui avait déjà été mis en lumière par Courvoisier (1970).

Choix des prédicteurs

Le programme sélectionne un par un les prédicteurs qui contribuent de façon significative à la séparation des groupes. A chaque pas, n'est inclu que le prédicteur qui ajoute la plus grande quantité à la mesure de la dispersion (discrimination) entre les groupes. On teste alors ce nouveau prédicteur conditionnellement sur les variables déjà inclues. De même, une variable inclue précédemment peut être éliminée, si elle cesse d'être significative par l'introduction d'une nouvelle variable.
A chaque pas, le programme produit également une matrice de classification dans les différents groupes en fonction des n prédicteurs déjà introduits. On peut, à partir de ces matrices, calculer le nombre de succès obtenus (en %), si la prévision des précipitations avait été faite à l'aide des prédicteurs introduits.

Résumé des matrices de classification

GENEVE - COINTRIN (division en 2 classes)

Pas	Prédicteurs choisis	Succès (%)
1	RR - 12	= 79% **
2	RR - 12 + D_{500}	= 81% **
3	RR - 12 + D_{500} + W/E	= 79% **
4	RR - 12 + D_{500} + W/E + H_{1000}	= 82% **
5	RR - 12 + D_{500} + W/E + H_{1000} + S/N	= 82% **
6	RR - 12 + D_{500} + W/E + H_{1000} + S/N + D_{850}	= 83% **
7	RR - 12 + D_{500} + W/E + H_{1000} + S/N + D_{850} + FF_{500}	= 83% *
8	RR - 12 + D_{500} + W/E + H_{1000} + S/N + D_{850} + FF_{500} + DD_{500}	= 82% *

autres prédicteurs non - significatifs
niveau de signification : * = 0.05, ** = 0.01

PAYERNE (division en 2 classes)

pas	Prédicteurs choisis	Succès (%)
1	RR - 12	= 79%**
2	RR - 12 + D_{500}	= 79%**
3	RR - 12 + D_{500} + H_{1000}	= 78%**
4	RR - 12 + D_{500} + H_{1000} + S/N	= 78%**
5	RR - 12 + D_{500} + H_{1000} + S/N + W/E	= 80%**
6	RR - 12 + D_{500} + H_{1000} + S/N + W/E + FF_{500}	= 82%**
7	RR - 12 + D_{500} + H_{1000} + S/N + W/E + FF_{500} + D_{850}	= 82%**
8	RR - 12 + D_{500} + H_{1000} + S/N + W/E + FF_{500} + D_{850} + DD_{500}	= 82%*

autres prédicteurs non - significatifs

niveau de signification : *= 0. 05, **= 0. 01

ZURICH - KLOTEN (division en 2 classes)

pas	Prédicteurs choisis	Succès (%)
1	RR - 12	= 78%**
2	RR - 12 + D_{500}	= 79%**
3	RR - 12 + D_{500} + S/N	= 80%**
4	RR - 12 + D_{500} + S/N + W/E	= 80%**
5	RR - 12 + D_{500} + S/N + W/E + D_{850}	= 82%**
6	RR - 12 + D_{500} + S/N + W/E + D_{850} + H_{1000}	= 82%*

autres prédicteurs non - significatifs

niveau de signification : *= 0. 05, **= 0. 01

GENEVE - COINTRIN (division en 4 classes)

Pas	Prédicteurs choisis	Succès
1	RR - 12	= 69%**
2	RR - 12 + W/E	= 68%**
3	RR - 12 + W/E + D_{500}	= 67%**
4	RR - 12 + W/E + D_{500} + S/N	= 71%**
5	RR - 12 + W/E + D_{500} + S/N + H_{1000}	= 70%**
6	RR - 12 + W/E + D_{500} + S/N + H_{100} + FF_{500}	= 69%**

autres prédicteurs non - significatifs

niveau de signification : * = 0. 05, ** = 0. 01

PAYERNE (division en 4 classes)

Pas	Prédicteurs choisis	Succès
1	RR - 12	= 68%**
2	RR - 12 + D_{500}	= 68%**
3	RR - 12 + D_{500} + H_{1000}	= 68%**
4	RR - 12 + D_{500} + H_{1000} + FF_{500}	= 70%**
5	RR - 12 + D_{500} + H_{1000} + FF_{500} + S/N	= 70%**
6	RR - 12 + D_{500} + H_{1000} + FF_{500} + S/N + W/E	= 71%**

autres prédicteurs non - significatifs

niveau de signification : * = 0. 05, ** = 0. 01

ZURICH - KLOTEN (division en 4 classes)

Pas	Prédicteurs choisis	Succès
1	RR - 12	= 66%**
2	RR - 12 + D_{500}	= 67%**
3	RR - 12 + D_{500} + S/N	= 67%**
4	RR - 12 + D_{500} + S/N + W/E	= 68%**
5	RR - 12 + D_{500} + S/N + W/E + D_{850}	= 67%**

autres prédicteurs non significatifs

niveau de signification : * = 0. 05, ** = 0. 01

D'une façon générale, les mêmes prédicteurs sont sélectionnés pour les trois stations, mais dans un ordre légèrement différent. A Zürich seulement, la vitesse et la direction du vent à 500 millibars ne sont jamais choisis comme paramètres significatifs. Cela peut paraître étrange à première vue. En fait, par un phénomène de redondance inévitable en météorologie, l'information contenue dans certains prédicteurs est déjà exprimée en grande partie par d'autres. On ne peut donc pas déduire que les prédicteurs rejetés à un certain niveau de signification sont sans valeur aucune.

Le fait que RR - 12 soit choisi en premier aux trois stations pour 2 et 4 groupes indique bien la valeur de critère de persistance pour la prévision des précipitations.

En ce qui concerne les gradients W/E et S/N, nous avons déjà remarqué que celui-ci est plus significatifs à Zürich et à Payerne, alors que c'est l'inverse qui se produit à Genève. Ceci démontre l'importance des situations de barrage (du nord ou du sud) pour l'incidence des précipitations dans l'est de la Suisse. Cet effet s'atténue pour Payerne, mais y est encore sensible, alors qu'il est plus négligeable pour Genève.

En comparant les mesures de l'humidité, on voit que les différences entre la température et le point de rosée sont systématiquement choisies avant les point de rosée (le programme indique également un ordre de choix pour les prédicteurs non-significatifs). La mesure la plus significative de l'humidité est D_{500} et peut paraître un niveau un peu élevé. Il est remarquable que D_{700} soit moins significatif que les autres niveaux. Notons en passant qu'il faut être prudent à l'égard des valeurs de l'humidité relevées sur les sondages, celle-ci étant souvent une conséquence des précipitations et non pas sa cause.

La seule hauteur de niveau significative est H_{1000} qui est une mesure de la valeur de la pression au sol. On peut s'étonner que H_{500} ne soit pas retenu. Cela provient du fait que son information est en grande partie déjà contenue dans H_{1000}.

Pour le vent à 500 millibars, même remarque ci-dessus : son information est déjà contenue dans les gradients S/N et W/E.

On note aussi que la vitesse du vent à ce niveau est plus signifi - cative que sa direction (cf. Courvoisier 1970).

Suite aux remarques faites dans l'introduction, il est essentiel de se limiter à un certain nombre de prédicteurs. En effet, chaque prédicteur devant être lui - même prévu par le modèle numérique, les erreurs de prévision de chaque prédicteur se cumulant, celles - ci vont fausser d'autant la prévision finale des précipitations. C'est la raison pour laquelle il faut procéder à un choix plus restreint, même en éliminant des variables qui fournissent un apport signifi - catif à la discrimination.

Le problème peut être simplifié pour des raisons pratiques. En effet, les prévisions numériques de l'humidité n'étant pas encore à notre disposition, il semble aléatoire de tenir compte de ce paramètre pour le moment. D'autre part, il semble très difficile d'obtenir en prévision avec une exactitude suffisante - même à l'avenir - les valeurs de la pression aux points de la croix pour le calcul des gradients S/N et W/E. On pourrait donc procéder à un choix des prédicteurs en se basant sur nos possibilités actuelles en prévisions numériques, quitte à modifier la fonction discriminante au fur et à mesure des différentes améliorations apportées aux prévisions numériques.

Le programm permettant de forcer le choix des prédicteurs dans un ordre quelconque, on a pris RR - 12, DD_{500}, FF_{500}, H_{1000} et H_{500} qui sont actuellement à notre disposition. On obtient les pourcentages de succès suivant :

GENEVE - COINTRIN (division en 2 groupes) : $RR - 12 + FF_{500} + DD_{500} = 78\%$

PAYERNE " " " : $RR - 12 + FF_{500} + DD_{500} = 78\%$

ZURICH - KLOTEN " " " : $RR - 12 + FF_{500} + DD_{500} = 76\%$

Si l'on ajoute encore H_{1000} et H_{500}, le pourcentage de réussite reste identique pour Genève et Payerne, alors qu'il monte également à 78% pour Zürich.

GENEVE - COINTRIN (division en 4 groupes) : $RR-12+FF_{500}+DD_{500}$ = 68%

PAYERNE " " " : $RR-12+FF_{500}+DD_{500}$ = 68%

ZURICH - KLOTEN " " " : $RR-12+FF_{500}+DD_{500}$ = 66%

L'adjonction de H_{1000} et H_{500} n'apporte pas d'amélioration supplémentaire.

On voit donc que le taux de réussite est à peine inférieur à celui vu précédemment.

Comme nous l'avons noté en page 17, la présentation des vents au niveau 500 mb sous la forme DDD/FF est mathématiquement incorrecte, les moyennes sur les directions n'ayant aucun sens. Nous avons donc modifié la présentation de ce prédicteur en décomposant le vecteur " vent au niveau 500 mb " en ses deux composantes S/N et W/E ; la notion de moyenne devient à ce moment logique, et l'analyse a fourni les résultats suivants pour Genève :

	Moyennes des groupes			
	NORAIN	0.7 - 4.9 mm	5 - 9.9 mm	10 mm
Vecteur S/N	13. 63	4. 50	- 1. 63	-10. 33
Vecteur W/E	-08. 79	-27. 11	-38. 66	-36. 66

Le vent moyen à 500 mb est donc NNW pour le groupe NORAIN, puis, au fur et à mesure de l'augmentation des quantités de précipitations, il s'oriente progressivement à l'W puis à l'WSW pour le dernier groupe.

Pour Genève, les pourcentages de succès sont alors les suivants :

Division en 2 classes :

1	RR_{-12}	79%
2	$RR_{-12}+W/E_{500}$	78%
3	$RR_{-12}+W/E_{500}+S/N_{500}$	79%
4	$RR_{-12}+W/E_{500}+S/N_{500}+H_{500}$	80%

Division en 4 classes :

1	RR_{-12}	67%
2	$RR_{-12}+W/E_{500}$	68%
3	$RR_{-12}+W/E_{500}+S/N_{500}$	69%
4	$RR_{-12}+W/E_{500}+S/N_{500}+H_{500}$	68%

QUELQUES REMARQUES POUR CONCLURE

- comme nous l'avons dit dans l'introduction, cette première étape de notre travail avait surtout pour but de tester le programme d'analyse discriminante. Toutfois, les résultats obtenus sont déjà parfaitement utilisables dans la pratique.
- la souplesse du programme utilisé permettra d'améliorer encore ces " scores ", en particulier en transformant quelque peu la présentation de certains prédicteurs afin que leur répartition soit plus proche de la normale, en introduisant d'autres prédicteurs (par exemple, le mouvement vertical), ou en combinant entre eux plusieurs prédicteurs (pour obtenir, par exemple, l'humidité moyenne de plusieurs niveaux).
- nous proposons donc le programme suivant pour la suite des opérations :

a) perforation du reste des données ;

b) passage du programme BMD 07 M mois par mois, afin d'établir la possibilité - très vraisemblable - d' un groupage des résultats par saisons et l'éventuelle apparition d'un ordre de sélection différent des prédicteurs suivant l'époque de l'année ;

c) test " en routine " des résultats obtenus à partir de l'échantillon de développement complet. Ceci sous - entend une poursuite rapide de la suite de notre travail, afin d'être en mesure de commencer ces tests dans le courant de l'année 1974 ;

d) perfectionnement de la méthode par l'introduction éventuelle des nouveaux prédicteurs cités plus haut.

LE GROUPE DE TRAVAIL DE GENEVE

J. D. ALTHERR — M. DUPANLOUP
Y. GANTER — E. JUNET

Février 1974

APPENDICE

EXEMPLES DE PRESENTATION DES RESULTATS A LA SORTIE DE L'ORDINATEUR :

Classification en deux groupes RAIN et NORAIN

1) Mode de sélection des 4 premiers prédicteurs.

2) Tableau récapitulatif.

3) Représentation graphique.

EXTRAIT DE L'APPLICATION DU PROGRAMME BMD 07 M A LA CLASSIFICATION EN DEUX GROUPES: MODE DE SELECTION DES 4 PREMIERS PREDICTEURS.

```
SUBPROBLEM
F-LEVEL FOR INCLUSION   3.9000
F-LEVEL FOR DELETION    3.7500
TOLERANCE LEVEL         0.0001
CONTROL VALUES      1111111111111

*****************************************************************************************

STEP NUMBER         0

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1  494

  1 128.6644          3  30.4273          5 111.9456          7  48.2446          9  90.5467         11  63.6724         13  30.2492
  2  46.9722          4  36.8092          6   0.3172          8  75.4244         10  56.9912         12  14.5487

          FUNCTION
         NORAIN         RAIN
VARIABLE

CONSTANT
         -0.00000       -0.00000

           NUMBER OF CASES CLASSIFIED INTO GROUP -
             NORAIN RAIN
GROUP
NORAIN         0    372
RAIN           0    124

*****************************************************************************************

STEP NUMBER         1
VARIABLE ENTERED    1

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1  494

  1 128.6644

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1  493

  2  52.6597          4  11.4169          6   2.7536          8  37.6474         10  31.2250         12   5.5205
  3   9.2418          5  55.6255          7  17.2962          9  68.1369         11  47.9645         13  10.2916

U-STATISTIC           0.79336      DEGREES OF FREEDOM    1    1  494
APPROXIMATE F       128.66438      DEGREES OF FREEDOM    1  494.00

F MATRIX - DEGREES OF FREEDOM    1  494

             GROUP
             NORAIN
GROUP
RAIN        128.66438

          FUNCTION
         NORAIN         RAIN
VARIABLE
  1       0.92603        3.95342

CONSTANT
         -0.06472       -1.17965

           NUMBER OF CASES CLASSIFIED INTO GROUP -
             NORAIN RAIN
GROUP
```

```
NORAIN      326      52
RAIN         50      74

*******************************************************************************************

STEP NUMBER        2
VARIABLE ENTERED   9

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1  493

  1 104.7283          9  68.1359

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1  492

  2  34.7554          4   5.7151          6   2.4585          8   5.6323         11  12.5109         13   9.6211
  3   4.3207          5  33.3578          7   9.9558         10  14.7889         12   0.4238

U-STATISTIC          0.69733      DEGREES OF FREEDOM    2    1  494
APPROXIMATE F      107.14369      DEGREES OF FREEDOM    2  493.00

F MATRIX - DEGREES OF FREEDOM    2  493

            GROUP
            NORAIN
GROUP
RAIN       107.14368

           FUNCTION
          NORAIN          RAIN
VARIABLE
  1       0.99942         3.97423
  9       0.02076         0.08588

CONSTANT
         -0.96503        -1.25182

           NUMBER OF CASES CLASSIFIED INTO GROUP -
             NORAIN RAIN
GROUP
NORAIN       327     45
RAIN          50     74

*******************************************************************************************

STEP NUMBER        3
VARIABLE ENTERED   2

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1  492

  1 112.3253          2  34.7554          9  49.6963

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1  491

  3  16.1180          5  17.5492          7  17.2602         10   1.9645         12   0.0306
  4  15.5045          6   5.2714          8   7.4026         11   2.5103         13  11.6187

U-STATISTIC          0.65184      DEGREES OF FREEDOM    3    1  494
APPROXIMATE F       87.90494      DEGREES OF FREEDOM    3  492.00

F MATRIX - DEGREES OF FREEDOM    3  492

            GROUP
```

```
          NORAIN
GROUP
RAIN        87.90494

          FUNCTION
         NORAIN         RAIN
VARIABLE
  1       1.04469        4.23570
  2       0.00231        0.01334
  9       0.02196        0.00757

CONSTANT
         -0.97684       -1.64552

          NUMBER OF CASES CLASSIFIED INTO GROUP -
            NORAIN RAIN
GROUP
NORAIN       312     70
RAIN          32     92
```

****

```
STEP NUMBER        4
VARIABLE ENTERED   5

VARIABLES INCLUDED AND F TO REMOVE - DEGREES OF FREEDOM    1  491

  1  75.4362         2  19.0023         5  17.6492         9  38.2519

VARIABLES NOT INCLUDED AND F TO ENTER - DEGREES OF FREEDOM    1  490

  3  16.9285         6   0.4896         8   3.1928        11   2.9686        13  13.2719
  4   0.5688         7  10.3987        10   2.7735        12   0.5539

U-STATISTIC          0.62845     DEGREES OF FREEDOM    4    1  494
APPROXIMATE F       72.57203     DEGREES OF FREEDOM    4  491.00

F MATRIX - DEGREES OF FREEDOM    4  491

            GROUP
            NORAIN
GROUP
RAIN        72.57203

          FUNCTION
         NORAIN         RAIN
VARIABLE
  1       2.64237        5.47824
  2       0.01253        0.02128
  5       0.03418        0.02658
  9       0.01573        0.01343

CONSTANT
         -3.94474       -3.44065

          NUMBER OF CASES CLASSIFIED INTO GROUP -
            NORAIN RAIN
GROUP
NORAIN       318     64
RAIN          26     98
```

****

```
STEP NUMBER        5
```

SUMMARY TABLE

STEP NUMBER	VARIABLE ENTERED	VARIABLE REMOVED	F VALUE TO ENTER OR REMOVE	NUMBER OF VARIABLES INCLUDED	U-STATISTIC
1	1		128.6644	1	0.7934
2	9		68.1369	2	0.6970
3	2		34.7354	3	0.6510
4	5		17.6492	4	0.6284
5	3		15.9285	5	0.6075
6	7		7.4975	6	0.5983
7	13		6.2053	7	0.5908
8	12		5.0375	8	0.5847

EIGENVALUES

0.71016	0.00030	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

CUMULATIVE PROPORTION OF TOTAL DISPERSION

1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

CANONICAL CORRELATIONS

0.64441	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

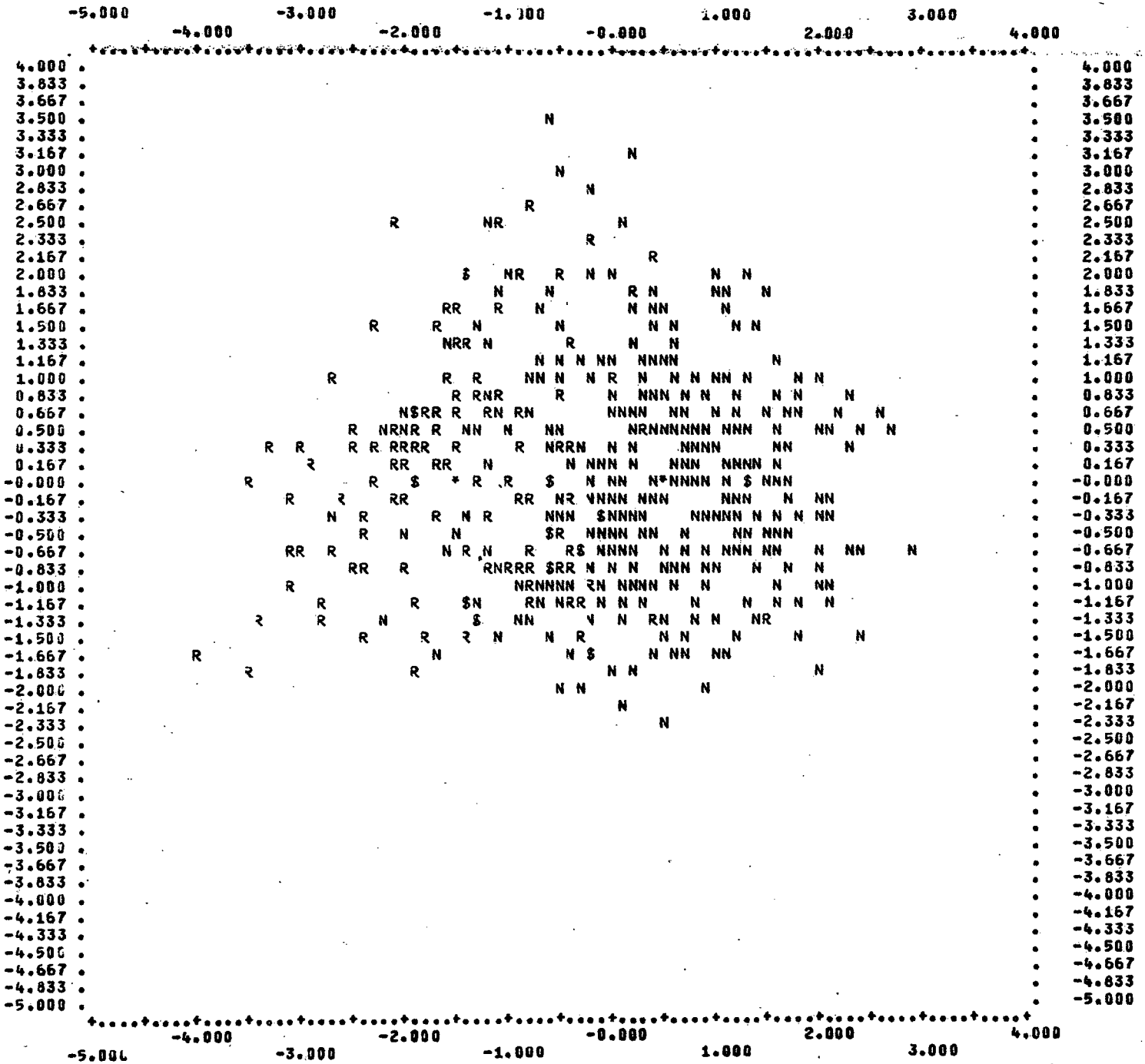
COEFFICIENTS FOR CANONICAL VARIABLE -

ORIGINAL VARIABLE	1	2	3	4	5	6	7	8
1	-1.22129	0.58591	0.66381	0.99544	-0.99290	-1.67177	0.09637	0.72300
2	-0.00689	-0.01354	-0.00352	0.00636	0.00073	0.00000	0.00215	0.00046
3	0.00582	0.00000	0.02339	0.00000	0.00000	0.00000	0.00046	-0.00136
5	0.00391	0.00000	0.00005	0.01410	0.00000	0.00000	0.00031	-0.00092
7	0.00454	0.00000	-0.00258	-0.00728	-0.02254	0.00000	-0.00256	-0.00101
9	0.00589	-0.00597	-0.00337	-0.00328	0.00411	-0.01198	-0.00233	-0.00126
12	0.01562	0.00000	0.01938	0.01154	0.00000	0.00000	-0.09560	-0.00188
13	-0.01069	0.00000	0.01091	-0.00613	-0.00399	0.00000	0.00702	-0.05359

GROUP	CANONICAL VARIABLES EVALUATED AT GROUP MEANS							
1	0.48556	0.00000	0.00000	-0.00000	0.00000	-0.00000	0.00000	0.00000
2	-1.45667	-0.00000	-0.00000	-0.00000	-0.00000	0.00000	0.00000	0.00000

CHECK ON FINAL U-STATISTIC 0.58474

CLASSIFICATION EN DEUX GROUPES: TABLEAU RECAPITULATIF.



CLASSIFICATION EN DEUX GROUPES: REPRESENTATION A DEUX DIMENSIONS DE L'APPLICATION DE LA FONCTION DISCRIMINANTE AUX PREDICTEURS CHOISIS.

(N = groupe NORAIN
R = groupe RAIN)

BIBLIOGRAPHIE

[1] BENTSSON AND MOEN, An operational system for numerical weather prediction, WMO no 283, Geneva 1971.

[2] BERGERON, T., Richtlinien einer dynamischen Klimatologie, Met. Zeitschrift 47, 1930.

[3] CHAMBETTAZ, A., La prévision en Suisse par la climatologie dynamique, Cointrin 1969.

[4] COURVOISIER, H. W., Die quantitative Niederschlagsprognose winterlicher zykloner Witterungslagen auf der Alpennordseite der Schweiz, Veröffentlichung der MZA Nr. 15, Zürich 1970.

[5] DUPANLOUP, KORTEN, CHAMBETTAZ, Prevision des brouillards à l'Aéroport de Genève - Cointrin. Rapport de travail de l'ISM no 27, Genève 1971.

[6] GANTER, Y., La prévision des orages à l'aide d'indices d'instabilité dans la partie suisse du nord des Alpes. Rapport de travail de l'ISM no 31, Genève 1971.

[7] KIRCHHOFER, W., Abgrenzung von Wetterlagen im zentralen Alpenraum. Veröffentlichung MZA Nr. 23, Zürich 1971.

[8] KLEIN, H. W., Computer prediction of precipitation for 108 cities in the United States. NOAA Technical Memorandum, Siver Spring 1971.

[9] LOENNQVIST, O., Experiments with automatic interpretation of meteorological forecasts charts. Meteorological and Hydrological Institute, Stockholm, 1968.

[10] MILLER, R. G., Statistical prediction by discriminant analysis, Meteorological Monographs, vol 4, no 23, 1962.

[11] SHUMAN AND HOVERMALE, An operational Six - Layer primitive equation Model. Journal of Applied Meteorology, vol 7, no 4, 1968.

[12] VEIGAS, K. W., The development of a statistical - physical Hurricane Prediction Model. Final Report Hartford, Conn, 1966.