

Exercice 12

UTILISATION DE L'ANALYSE EN COMPOSANTES PRINCIPALES POUR L'ANALYSE DES REGIMES PLUVIOMETRIQUES MENSUELS

Pour étudier la répartition mensuelle des précipitations sur les Alpes Maritimes, on dispose sur 27 postes pluviométriques, d'une estimation sur plus de trente ans, des précipitations moyennes mensuelles exprimées en mm (Tableau 12.1).

Tableau 12.1. Précipitations mensuelles interannuelles

Poste	X	Y	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Août	Sep	Oct	Nov	Déc
Antibes	9878	18522	77	82	73	62	47	33	12	29	67	120	108	81
Beuil	9722	19109	74	73	88	84	119	98	67	71	86	130	113	80
Breil sur Roya	10147	18949	73	86	86	83	79	66	44	60	81	139	116	101
Cannes	9732	18478	84	90	77	70	51	38	16	35	75	118	120	94
Clans	9849	19005	63	62	75	72	85	69	47	64	69	115	97	70
Colomars	9931	18750	87	81	84	74	65	44	17	40	69	129	122	88
Coursegoules	9788	18770	117	136	128	110	107	72	30	68	89	188	187	135
Escarène	10033	18849	83	83	88	78	68	52	26	41	77	125	118	98
Guillaumes	9616	19084	75	75	80	75	88	81	56	59	69	116	99	77
Levens	9925	18861	85	94	91	88	84	61	33	50	69	124	132	91
Lucéram	10184	18803	85	90	101	92	94	73	45	61	82	150	140	101
Moulinet	10079	18964	84	89	95	95	87	75	50	63	84	144	136	104
Nice	9926	18619	78	76	72	63	48	37	15	30	68	126	108	81
Peillon	10051	18776	87	82	84	71	61	44	20	40	69	138	110	87
Puget Théniers	9665	18938	77	79	83	74	84	81	43	63	67	126	101	83
Sigale	9709	18866	83	94	84	77	86	67	38	68	66	129	116	83
Sospel	10096	18909	97	100	108	85	86	55	32	49	84	154	146	107
St Auban	9537	18819	88	95	101	87	101	86	52	76	75	138	118	101
St Dalmas le Selvage	9603	19306	85	78	96	105	99	84	55	64	88	142	114	76
Auron	9674	19236	83	78	89	97	100	83	58	69	75	122	103	81
St Etienne de Tinée	9659	19272	76	71	87	91	93	81	53	64	72	121	101	75
St Jean Cap Ferrat	10019	18644	73	72	67	56	48	34	13	31	63	116	90	73
St Sauveur sur Tinée	9816	19096	63	63	81	86	91	84	63	72	83	124	101	74
St Vallier	9626	18676	103	115	112	105	93	67	28	65	89	147	132	119
Tende	10216	19121	76	73	95	89	101	70	54	63	74	121	116	89
Valdeblore	9870	19080	67	68	91	91	105	95	66	72	81	128	106	79
Vence	9854	18689	92	94	96	85	75	50	22	44	79	149	128	102

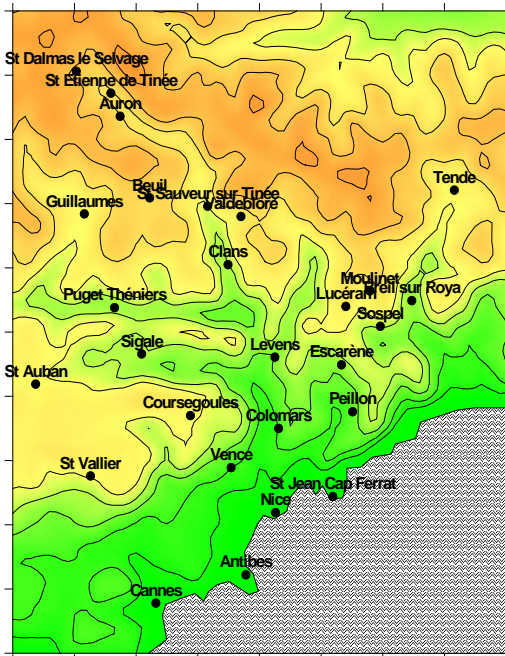


Fig.12.1. Situation des postes pluviométriques

Comme le montre la figure 12.1 ces stations sont assez uniformément réparties sur les Alpes Maritimes.

On a procédé à une analyse en composantes principales sur ces 27 variables (les postes pluviométriques) et ces 12 observations (les mois).

Les résultats bruts de l'ACP sont donnés dans le tableau 12.2.

1. Seules les 5 premières composantes principales ont été calculées. Quelle serait la valeur propre maximale possible qui pourrait correspondre à la sixième composante ?

Tableau 12.2. Valeurs propres

Valeurs propres : λ_1	λ_2	λ_3	λ_4	λ_5
23,58	2,50	0,26	0,21	0,18

2. De quel ordre de grandeur pourrait être cette sixième valeur propre si on avait exprimé les pluies en cm au lieu de mm ?

3. Quels sont les pourcentages d'inertie expliqués par chacune de ces cinq composantes et quels sont les pourcentages cumulés (Tableau 12.3)? Indiquez également pourquoi ?

Tableau 12.3. Pourcentage d'inertie

Composantes :	C ₁	C ₂	C ₃	C ₄	C ₅
Pourcentage d'inertie expliquée	?	?	?	?	?
Pourcentages cumulés :	?	?	?	?	?

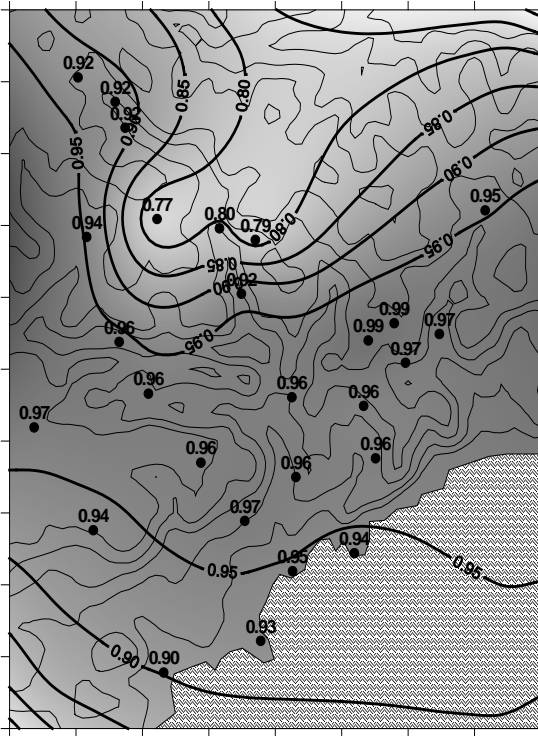
4. Combien de composantes au maximum vous paraît-il raisonnable de retenir et pourquoi ?

En tableau 12.4, sont présentées les projections des variables et des observations.

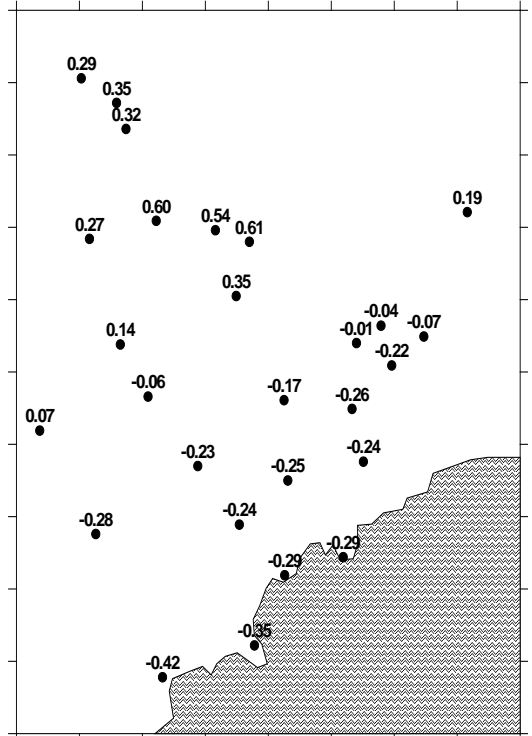
Tableau 12.4. Projection des variables et des observations

Projections des variables actives :					
STATIONS	C ₁	C ₂	C ₃	C ₄	C ₅
Antibes	0,93	-0,35	0,08	-0,02	-0,03
Beuil	0,78	0,58	-0,02	-0,08	0,07
Breil sur Roya	0,97	-0,07	0,10	-0,12	0,06
Cannes	0,90	-0,42	0,07	0,00	0,04
Clans	0,92	0,35	0,03	-0,08	0,04
Colomars	0,96	-0,25	0,02	0,04	0,00
Coursegoules	0,96	-0,23	-0,12	0,01	0,07
Escarène	0,96	-0,26	0,06	0,03	0,05
Guillaumes	0,94	0,27	-0,08	-0,09	-0,05
Levens	0,96	-0,17	-0,11	0,11	0,09
Lucéram	0,99	-0,01	0,01	0,00	0,10
Moulinet	0,99	-0,04	0,07	-0,01	0,11
Nice	0,95	-0,29	0,12	-0,04	-0,05
Peillon	0,96	-0,24	0,06	-0,03	-0,10
Puget Théniers	0,96	0,14	-0,12	-0,16	-0,12
Sigale	0,96	-0,06	-0,18	-0,09	-0,04
Sospel	0,97	-0,22	0,00	0,02	0,07
St Auban	0,97	0,07	-0,20	-0,10	-0,02
St Dalmas le Selvage	0,92	0,29	0,13	0,17	-0,15
Auron	0,92	0,32	-0,08	0,17	-0,12
St Etienne de Tinée	0,92	0,35	-0,02	0,13	-0,10
St Jean Cap Ferrat	0,94	-0,29	0,09	-0,06	-0,12
St Sauveur sur Tinée	0,80	0,54	0,22	-0,06	0,03
St Vallier	0,94	-0,29	-0,08	0,09	-0,04
Tende	0,95	0,19	-0,06	0,15	0,18
Valdeblore	0,79	0,61	0,05	-0,01	0,04
Vence	0,97	-0,24	0,03	0,02	-0,02

Projections des observations actives :					
Mois	C ₁	C ₂	C ₃	C ₄	C ₅
Jan	-0,12	-1,28	-0,39	0,54	-1,57
Fév	-0,04	-1,54	-1,18	-0,56	-0,55
Mar	0,23	-0,20	-0,33	0,84	0,15
Avr	0,01	0,27	0,56	2,38	-0,55
Mai	0,08	1,59	-1,39	0,69	0,54
Jun	-0,62	1,37	-0,53	-0,92	-0,72
Jul	-1,77	0,40	0,84	-0,28	0,69
Août	-1,09	0,22	-0,44	-0,87	-0,03
Sep	-0,33	-0,31	2,29	-0,14	0,08
Oct	2,05	0,86	0,80	-1,15	-1,19
Nov	1,37	-0,20	-0,09	0,16	1,95
Déc	0,21	-1,18	-0,14	-0,69	1,19



Carte des projections des variables sur C_1



Carte à compléter des projections des variables sur C_2

Fig.12.2. Projections des variables sur C_1

Fig.12.3. Projection des variables sur C_2

5. Au vu des valeurs prises par les projections des variables et des observations sur la première composante (valeurs et répartition spatiale) (Tableau 12.4, Fig. 12.2), quelle signification lui donneriez-vous ?

6. Tentez de cartographier les projections des variables sur la composante deux (Fig.12.3). (Courbes de 0,1 en 0,1). Quelle interprétation donneriez-vous à la seconde composante ?

7. On constate que les projections des variables sur la composante 3 présentent une certaine structure spatiale. Quelle est-elle et qu'en tireriez-vous sur la signification de cette composante (Fig.12.4) ?

8. Quel serait le coefficient de corrélation multiple entre la station de Cannes et les deux premières composantes principales ? Même question avec les trois premières composantes.

9. Si on admet que les stations du littoral sont assez bien représentées par $\rho_{C_1}=0,9$ et $\rho_{C_2}=-0,42$ (analogie avec Cannes) et la montagne par $\rho_{C_1}=0,79$ et $\rho_{C_2}=0,61$ (analogie avec Valdeblore), évaluez (en valeur centrées réduites) l'évolution mensuelle des précipitations sur le littoral et en montagne dans le tableau 12.5. Interprétez cela sur un graphique (Fig. 12.5.)

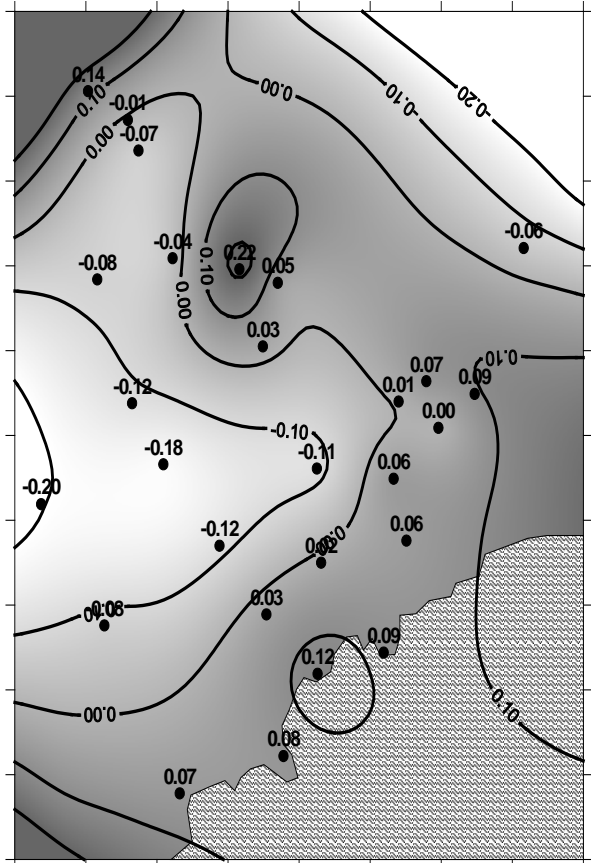


Fig.12. 4. Projections des variables sur la composante 3

Tableau 12.5. Evolution mensuelle des précipitations en fonction des composantes C_1 et C_2

	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Août	Sep	Oct	Nov	Déc
C_1	-0,12	-0,04	0,23	0,01	0,08	-0,62	-1,77	-1,09	-0,33	2,05	1,37	0,21
C_2	-1,28	-1,54	-0,20	0,27	1,59	1,37	0,40	0,22	-0,31	0,86	-0,20	-1,18
Littoral												
Montagne												

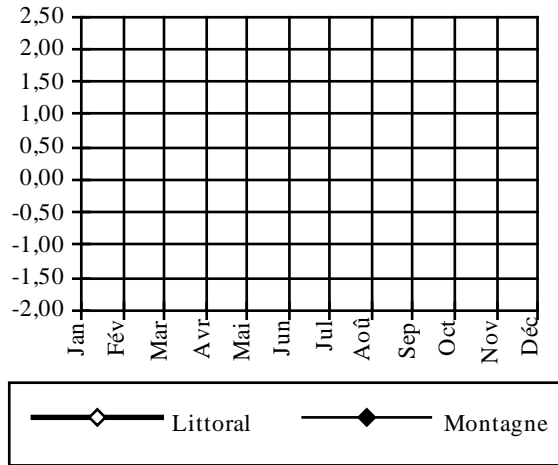


Fig.12.5. Evolution mensuelle des précipitations sur le littoral et en montagne

Corrigé :

1. On sait que la somme des valeurs propres est telle que $\sum_{i=1}^{nv} \lambda_i = nv$ et donc

dans notre cas $\sum_{i=1}^{27} \lambda_i = 27$

La somme des valeurs propres des composantes 1 à 5 est $26,73 = (23,58 + 2,50 + 0,26 + 0,21 + 0,18)$;

La somme des valeurs propres des composantes 6 à 27 est donc $0,27 (=27 - 26,73)$.

Si toutes ces valeurs étaient égales elles le seraient à $0,0123 (= 0,27 / 22)$. On sait donc que λ_6 doit être inférieur ou égal à $0,18$ et à $0,27$ et supérieur ou égal à $0,0123$ on aura donc forcément : $0,0123 \leq \lambda_6 \leq 0,18$.

2. Dans une Analyse en Composante Principale, on travaille sur la matrice des coefficients de corrélation et les différents résultats sont absolument indépendants des systèmes d'unité choisis.

3. Le pourcentage d'inertie expliqué par la $i^{\text{ème}}$ composante dépend de la valeur propre λ_i et du nombre de variable nv : $\% = \lambda_i / nv$. Les résultats sont dans le tableau 12.6.

Tableau 12.6. Calcul des composantes

Composantes :	C ₁	C ₂	C ₃	C ₄	C ₅
Pourcentage d'inertie expliquée	87,32	9,25	0,97	0,78	0,67
Pourcentages cumulés :	87,32	96,57	97,55	98,32	98,99

4. Une règle empirique consiste à ne retenir que les composantes correspondant à des valeurs propres supérieures à 1. Par ailleurs, on constate également une chute importante des valeurs propres entre la deuxième et la troisième composante. Il paraît donc raisonnable de ne travailler que sur les deux premières composantes.

5. On constate que toutes les variables sont corrélées fortement et positivement avec la première composante. Il est évident que C₁ représente la répartition moyenne des précipitations dans l'année. Dans les Alpes Maritimes, en général, les précipitations se répartissent mensuellement comme l'indique le graphique (Fig.12. 6) construit à partir des projections des observations sur la première composante.

De Décembre à Mai les précipitations sont sensiblement constantes. Succèdent ensuite une saison sèche de Juin à Septembre (principalement Juillet et Août) puis deux mois très humides (Octobre et Novembre).

D'un point de vue spatial, les plus fortes corrélations avec C₁ se rencontrent dans la zone centrale du domaine d'étude. Ceci était prévisible puisque C₁ représente l'évolution moyenne spatiale des précipitations. Par contre en allant au nord ou au sud le coefficient de corrélation avec C₁ diminue, il est donc prévisible qu'apparaîtra une différenciation nord-sud dans les composantes d'ordre supérieur.

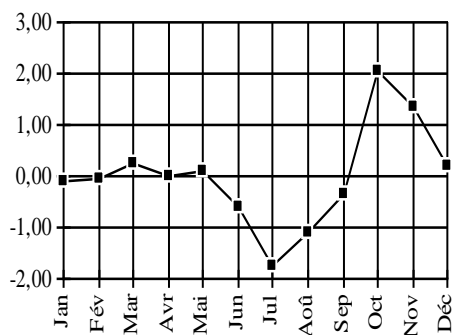


Fig.12.6. Projections des observations sur C₁

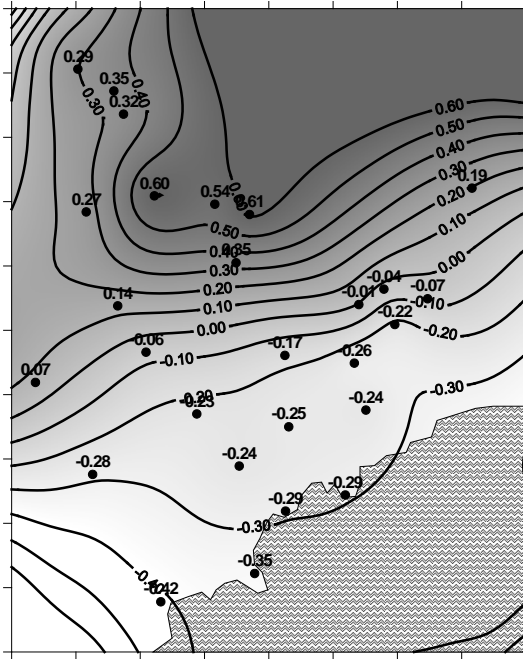


Fig.12.7. Répartition spatiale des projections sur C_2

7. Manifestement la troisième composante oppose l'est à l'ouest du département ou mieux encore le Massif du Cheiron au reste du département. Bien que la valeur propre soit nettement inférieure à un, cette composante représente certainement un phénomène physique dans la répartition saisonnière des précipitations. La particularité essentielle concerne le mois de Septembre (valeur forte de C_3). Ainsi sur le Cheiron (corrélé négativement avec C_3) les pluies de Septembre sont plus faibles que la normale (Fig.12.9).

6. La seconde composante oppose manifestement la zone côtière à l'intérieur du pays. Sur la côte les coefficients de corrélation avec la seconde composante sont négatifs, alors qu'ils sont positifs à l'intérieur du pays (Fig.12.7).

Cette seconde composante oppose les mois d'hiver (Décembre à Février) à ceux du printemps et de l'automne (Mai-Juin et Octobre) comme le montre la figure 12.8.

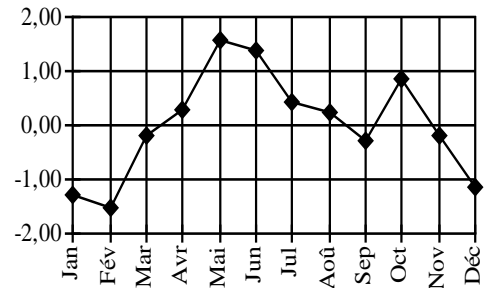


Fig.12.8. Chronique mensuelle de C_2

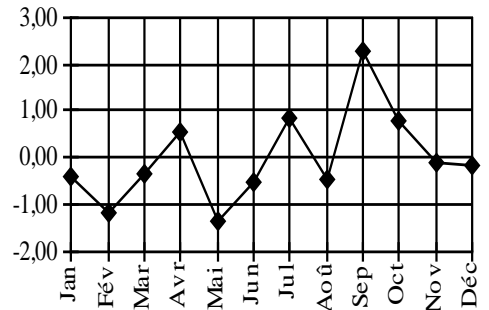


Fig.12.9. Chronique mensuelle de C_3

8. Les composantes étant indépendantes entre elles, le coefficient de corrélation multiple R entre une variable X et les k premières composantes principales s'exprime ainsi en fonction des projections $\rho_x C_j$ de x sur la j^{ème} composante :

$$R = \sqrt{\sum_{j=1}^k \rho_x^2 C_j} \quad (12.1)$$

Cannes avec les deux premières composantes et Cannes avec les trois premières composantes :

$$R = \sqrt{0,90^2 + 0,42^2} = 0,9938$$

$$R = \sqrt{0,90^2 + 0,42^2 + 0,07^2} = 0,9962$$

9. En appelant C_{1k} , la projection du k^{ème} mois sur la composante 1 et C_{2k} la projection de ce mois sur la seconde composante (Tableau 12.7), la valeur centrée réduite de la variable X pour le mois k est :

$$X_k = \rho_x C_1 * C_{1k} + \rho_x C_2 * C_{2k} \quad (12.2)$$

Tableau 12.7. Chronique synthétique des précipitations mensuelles

Mois	Jan	Fév	Mar	Avr	Mai	Jun	Jul	Août	Sep	Oct	Nov	Déc
C ₁	-0,12	-0,04	0,23	0,01	0,08	-0,62	-1,77	-1,09	-0,33	2,05	1,37	0,21
C ₂	-1,28	-1,54	-0,20	0,27	1,59	1,37	0,40	0,22	-0,31	0,86	-0,20	-1,18
Littoral	0,43	0,61	0,29	-0,10	-0,59	-1,13	-1,76	-1,07	-0,16	1,48	1,32	0,68
Montagne	-0,87	-0,97	0,06	0,17	1,04	0,35	-1,15	-0,72	-0,45	2,15	0,96	-0,56

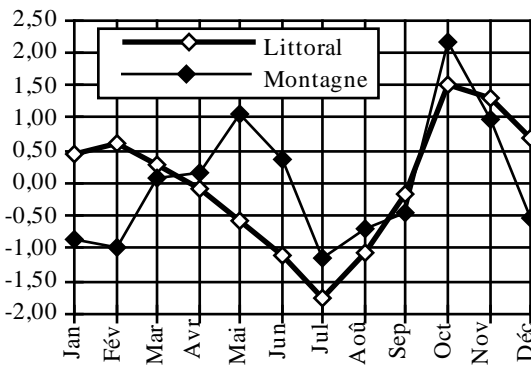


Fig.12.10. Chronique des précipitations

En figure 12.10, on voit clairement deux régimes pluviométriques.

Sur le littoral il n'y a qu'une baisse des précipitations d'Octobre à Juillet suivie d'une hausse d'Août à Octobre.

Pour la montagne on a deux saisons sèches de Décembre à Février et de Juillet à Septembre et deux saisons pluvieuses (Automne et Printemps).

&&&&&