

# SCILAB à l'École nationale des ponts et chaussées

<http://cermics.enpc.fr/scilab>

Statistiques de précipitations  
sur l'archipel de Polynésie

Michel DE LARA

(avec le soutien du Ministère de l'Outre-Mer)

15 mars 2005 (dernière date de mise à jour)

## Table des matières

1	Statistique descriptive	2
2	Analyse en composantes principales	8

Le fichier `precipitations.txt` comporte des données de précipitations sur plusieurs îles de Polynésie, Atuona, Faaa, Here, Rapa, Tubuai, Takaroa, Omoa, Paea, Papara, Papeeno. Par années, mois et jours, les précipitations sont données en mm de pluie (-99 correspond à une donnée manquante).

On récupérera le fichier `precipitations.txt`. On transformera les données en matrice comme suit.

```
stacksize(10000000);  
// taille mémoire  
M=fscanfMat('precipitations.txt');  
// convertit un tableau de données en matrice  
Stations =list("Atuona", "Faaa", "Here", "Rapa", "Tubuai", "Takaroa", ...  
    "Omoa", "Paea1", "Papara1", "Papeeno1");  
// liste des stations de mesure  
stations=size(Stations);  
// nombre des stations de mesure
```

De nombreuses données manquantes sont repérées par le nombre -99.

# 1 Statistique descriptive

**Question 1** Identifier les stations sans données manquantes. Tracer la courbe des précipitations.

```
ZU=bool2s(M==-99);
// matrice de 1 (si élément = -99) et de 0 (sinon)
[lhs,rhs]=maxi(ZU($:-1:1,4:$),"r");
// rhs contient les indices, comptés depuis la fin, à partir desquels
// -99 n'apparaît plus sur la colonne correspondante
// sauf si -99 n'apparaît jamais dans une colonne !!!

printf('indices des stations sans donnees manquantes :\n');
SDM=find(rhs==1)

for i=SDM
xset("window",i); xbasec();
plot2d(M(:,i+3));
xtitle('Precipitations pour la station'+string(Stations(i)),...
'jours','precipitations (mm)');
end

xdel(SDM);
```

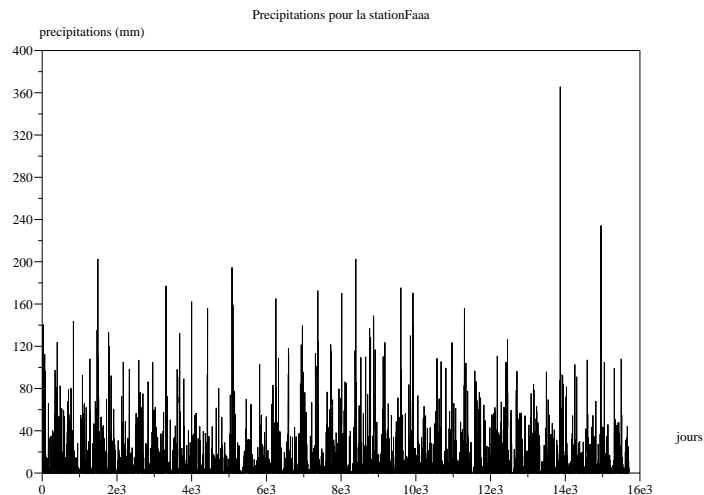


FIG. 1 – Précipitations à Faaa (1961-2003)

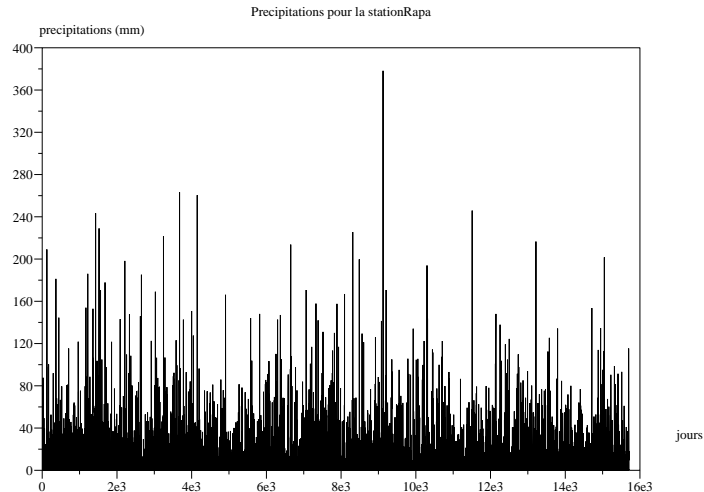


FIG. 2 – *Précipitations à Rapa (1961-2003)*

**Question 2** Identifier, pour chaque station, la première date à partir de laquelle il n'y a pas de donnée manquante après. Tracer, pour chaque station, la courbe des précipitations à partir de cette date.

```
ZU=bool2s(M== -99);
[lhs,rhs]=maxi(ZU($:-1:1,4:$),"r");
// rhs contient les indices, comptés depuis la fin, à partir desquels
// -99 n'apparait plus sur la colonne correspondante
// sauf si -99 n'apparait jamais dans une colonne !!!
// dates-rhs : pour chaque station, indice de date à partir de laquelle
// il n'y a pas de donnée manquante

for i=1:stations
if rhs(i)==1 then rhs(i)=dates, end
xset("window",i); xbas();
plot2d(dates-rhs(i)+2:dates,M(dates-rhs(i)+2:dates,i+3));
xtitle('Precipitations pour la station'+string(Stations(i)),...
'jours','precipitations (mm)');
end

xdel(1:stations);
```

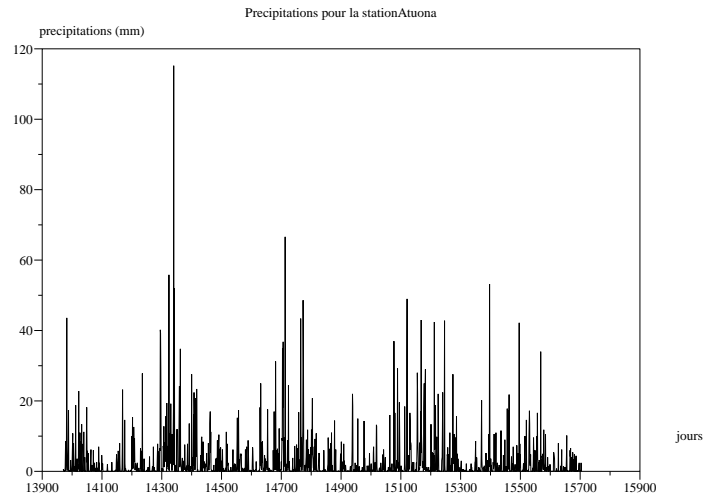


FIG. 3 – *Précipitations à Atuona*

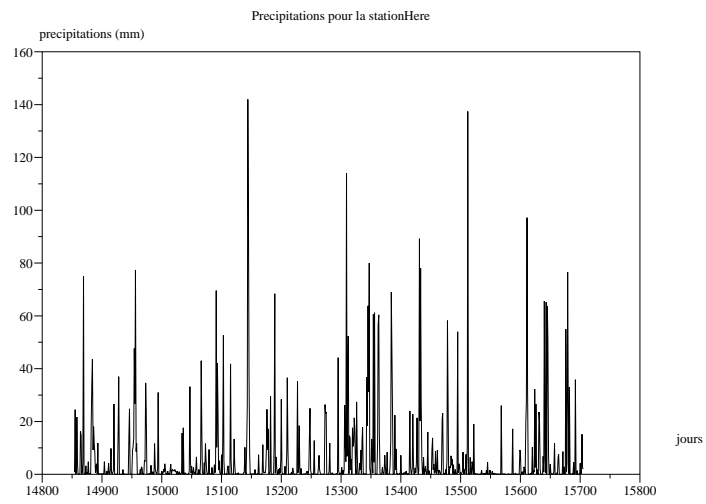


FIG. 4 – *Précipitations à Here*

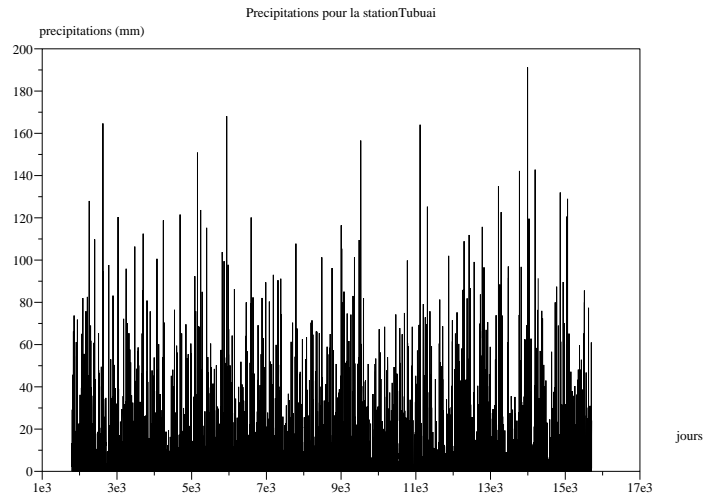


FIG. 5 – *Précipitations à Tubuai*

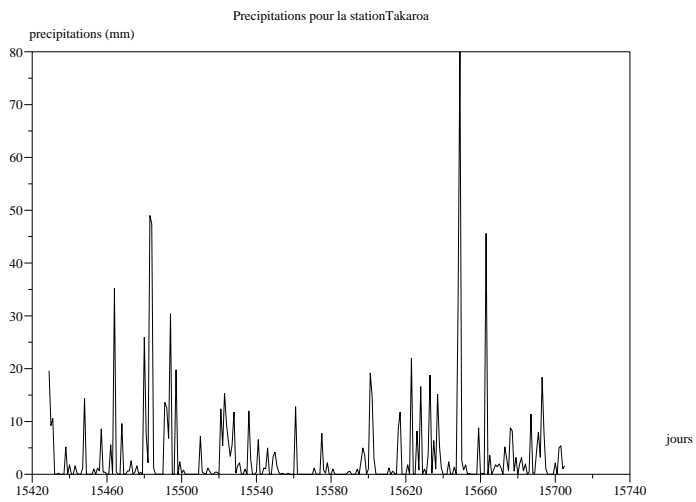


FIG. 6 – *Précipitations à Takarua*

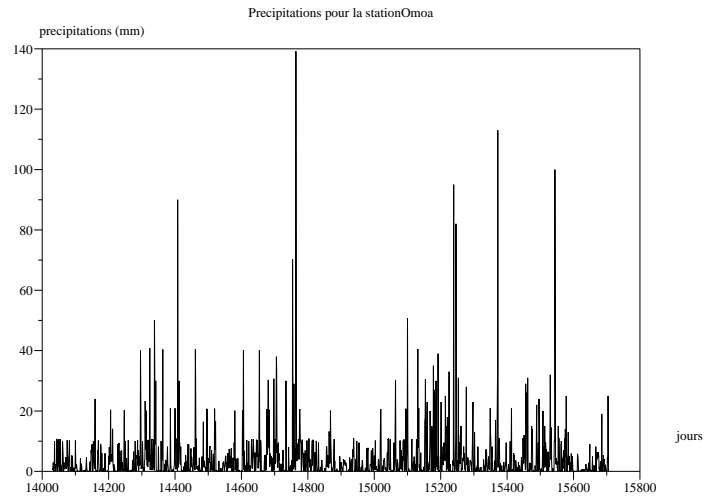


FIG. 7 – *Précipitations à Omoa*

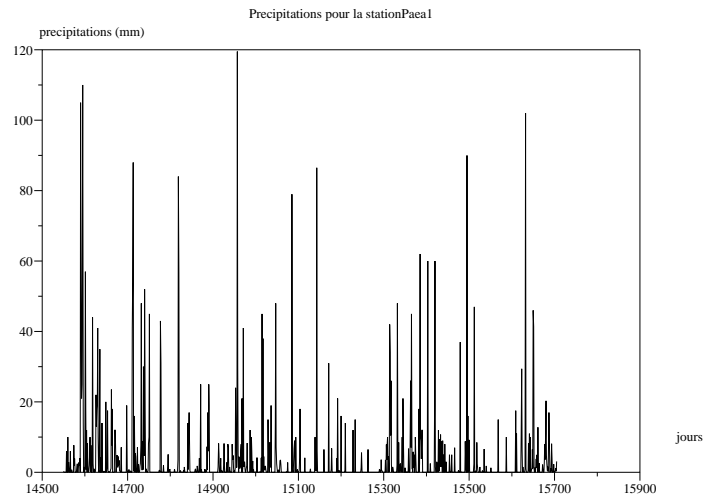


FIG. 8 – *Précipitations à Paea1*

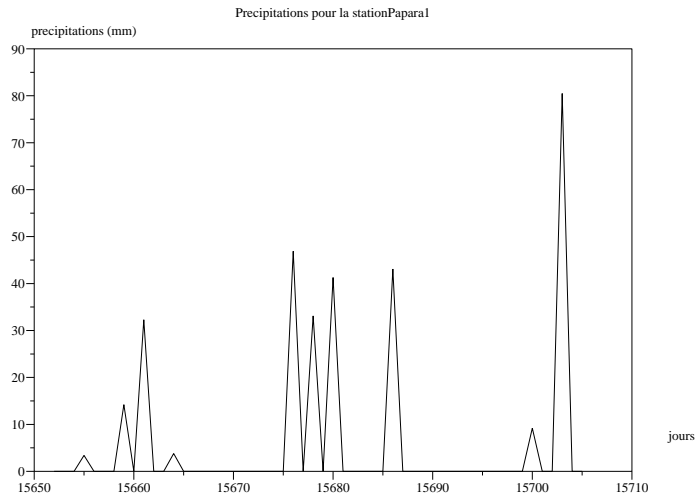


FIG. 9 – *Précipitations à Papeara1*

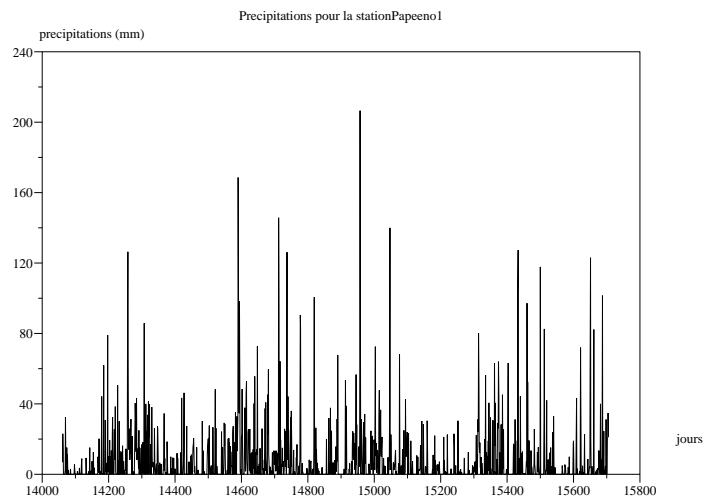


FIG. 10 – *Précipitations à Papeeno1*

## 2 Analyse en composantes principales

On veut traiter ces données par analyse en composantes principales (ACP). On récupérera le fichier ACP.sci.

**Question 3** *Extraire du tableau les lignes n'ayant aucune donnée manquante. Ceci correspondant aux jours pour lesquels toutes les stations ont une mesure de précipitation. Calculer la matrice des corrélations pour ces données. Représenter les valeurs propres par ordre décroissant. Tracer projections du nuage et cercles des corrélations sur les premiers plans factoriels.*

```
getf("ACP.sci");
// chargement du fichier ACP.sci

// PROBLEME DES DONNES MANQUANTES
z=mini(M,"c");
// minimum par ligne
ind=find(z>=0);
// indices correspondants
A=M(ind,4:$);
// matrice extraite correspondant aux jours pour lesquels
// il n'y a pas de donnée manquante pour aucune station

mat_cor=correlation(A);
// matrice des corrélations

valp=val_prop(mat_cor);
// valeurs propres

xset("window",0);
xbasc();plot2d2(1:sum(ones(valp)),valp/sum(valp));
xtitle("Eboulement des valeurs propres (normalisées)")

Coord=acp_indiv(A);
i=1;j=2;
xset("window",1);
nuage(Coord,i,j);
xset("window",2);
cercle(Coord,i,j);
xtitle("cercle des correlations : plan 1-2");

i=1;j=3;
xset("window",3);
nuage(Coord,i,j);
```

```
xset("window",4);
cercle(Coord,i,j);
xtitle("cercle des correlations : plan 1-3");

i=2;j=3;
xset("window",5);
nuage(Coord,i,j);
xset("window",6);
cercle(Coord,i,j);
xtitle("cercle des correlations : plan 2-3");

xset("window",7);
i=1;j=2;
qualite_indiv(A,i,j);

xdel(0:7)
```

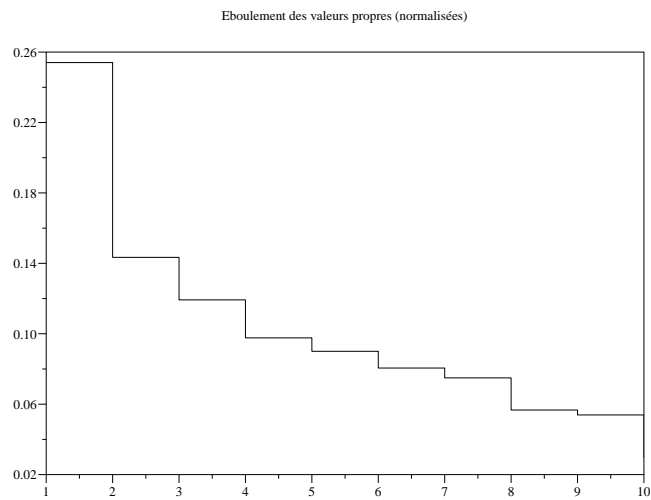


FIG. 11 – *Éboulement des valeurs propres (normalisées)*

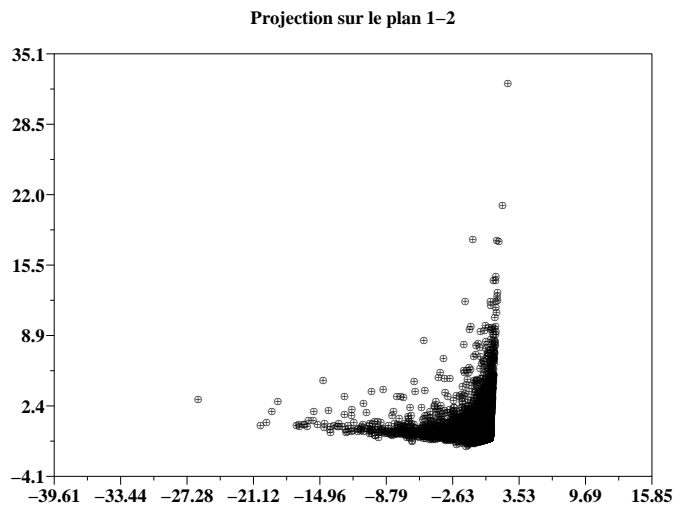


FIG. 12 – *Projection du nuage sur le plan factoriel 1-2*

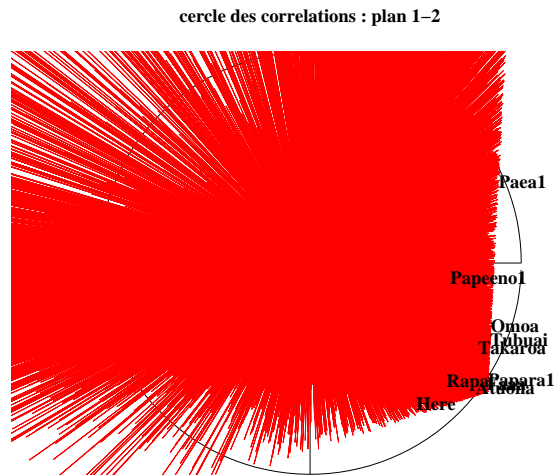


FIG. 13 – *Cercle des corrélations : plan factoriel 1-2*

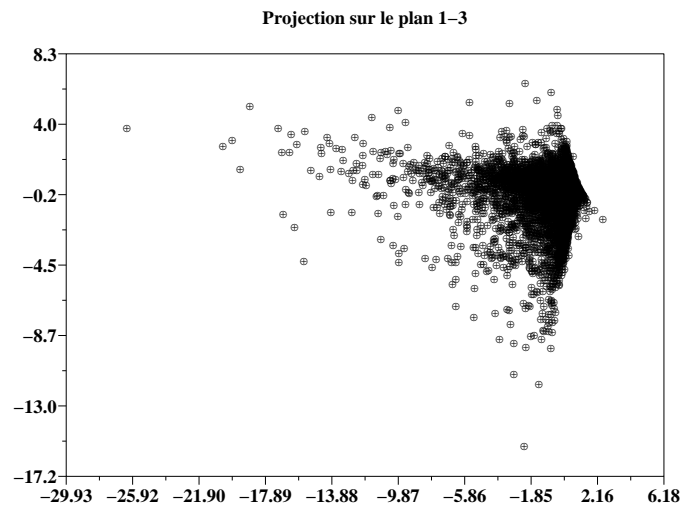


FIG. 14 – *Projection du nuage sur le plan factoriel 1-3*

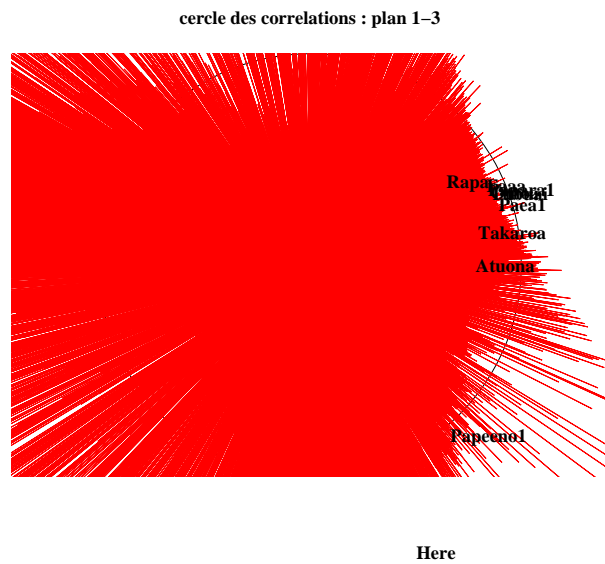


FIG. 15 – *Cercle des corrélations : plan factoriel 1-3*

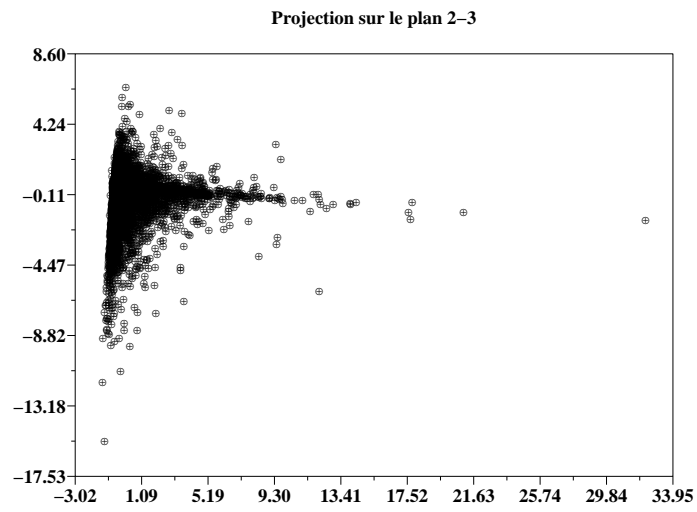


FIG. 16 – *Projection du nuage sur le plan factoriel 2-3*

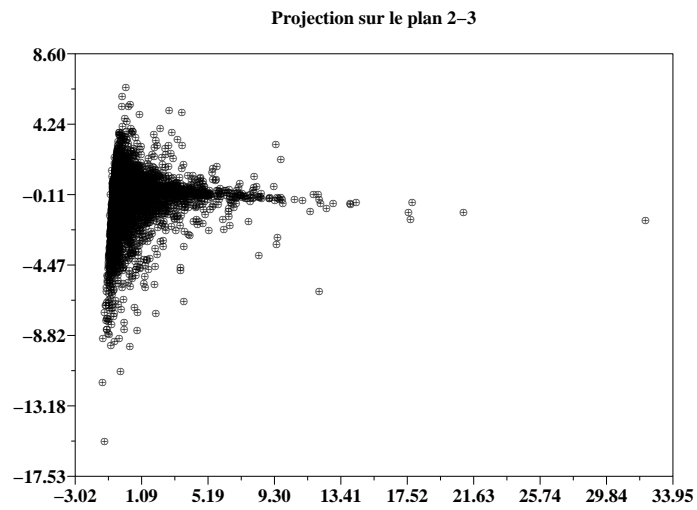


FIG. 17 – *Cercle des corrélations : plan factoriel 2-3*