

MASTER : Traitement de l'Information et Exploitation des Données

Fouad Badran, Cécile Mallet, Carlos Mejia,  
Charles Sorrow, Sylvie Thiria

*Master TRIED,*

*TPA04 : Rapport*

*Sujet :*

*ACP appliquée à des variables  
climatiques liées à l'effet de serre*

*Réalisé par :*

*Exemple de Rapport*

*Année universitaire : ----/----*

## Résumé de l'énoncé :

### Objectifs :

Le travail demandé a pour objectif d'apprendre à effectuer une Analyse en Composantes Principales (ACP) et à interpréter les résultats qu'elle produit. Pour mettre en œuvre cette pratique de l'ACP, nous travaillerons sur des données environnementales, qui font partie des variables prises en compte dans les modèles géophysiques. Ces variables participent également à l'étude du phénomène de l'effet de serre.

Comme il est décrit ci-après, nous disposons de données (des moyennes mensuelles) pour plusieurs villes. Nous proposons l'étude de trois d'entre elles assez distantes les unes des autres et présentant des spécificités différentes. Pour chacune nous effectuerons 2 ACP l'une dite « saisonnière » l'autre « interannuelle ».

Le TP sera découpé en 2 parties. Dans la 1<sup>ère</sup>, après une climatologie qui supprime l'effet interannuel et permet de choisir les villes, on présentera les ACP de la ville de Reykjavik. La seconde partie sera dédiée aux ACP des villes d'Alger et de Dakar.

*Note à l'attention des étudiants : Les commentaires encadrés représentent des connaissances supplémentaires sur le sujet. Ils ne découlent pas de l'énoncé et ne sont donc pas attendus dans un rapport rédigé par un non spécialiste du sujet.*

#### **ACP : présentation**

*La représentation de données devient délicate lorsque leur dimension est supérieure à 2 ou 3. Il faut recourir à des méthodes plus avancées comme (par exemple) l'Analyse en Composantes Principales (ACP). L'ACP consiste à opérer un changement de base de façon à maximiser l'inertie de l'ensemble des données étudiées en projection sur les premiers facteurs de la nouvelle base. On cherche ainsi à ce que quelques uns seulement de ces nouveaux facteurs, qui donnent naissance à de nouvelles variables appelées composantes principales, soient suffisants à la compréhension et à l'interprétation des données. L'information est ainsi réduite mais l'interprétation est rendue délicate puisque les coordonnées des nouvelles variables sur les nouveaux facteurs, dits principaux, sont des combinaisons linéaires des axes initiaux. Nous vous renvoyons au document joint « ACPrappels » pour une présentation succincte plus formelle de l'ACP.*

## Les Données :

Les données utilisées sont issues de la base ERA-Interim du centre européen ECMWF dont on a extrait les 5 données modèles suivantes :

1) Variables de données de surface type Analyse (AN) à midi :

**t2** : Temperature at 2 meters (degC)

**tcc** : Total cloud cover (0-1)

2) Données de surface de type Forecast (FC), prévisions depuis la veille à midi à +24h (c'est-à-dire à midi du jour) :

**lsp** : Large scale precipitation (m)

**cp** : Convective precipitation (m)

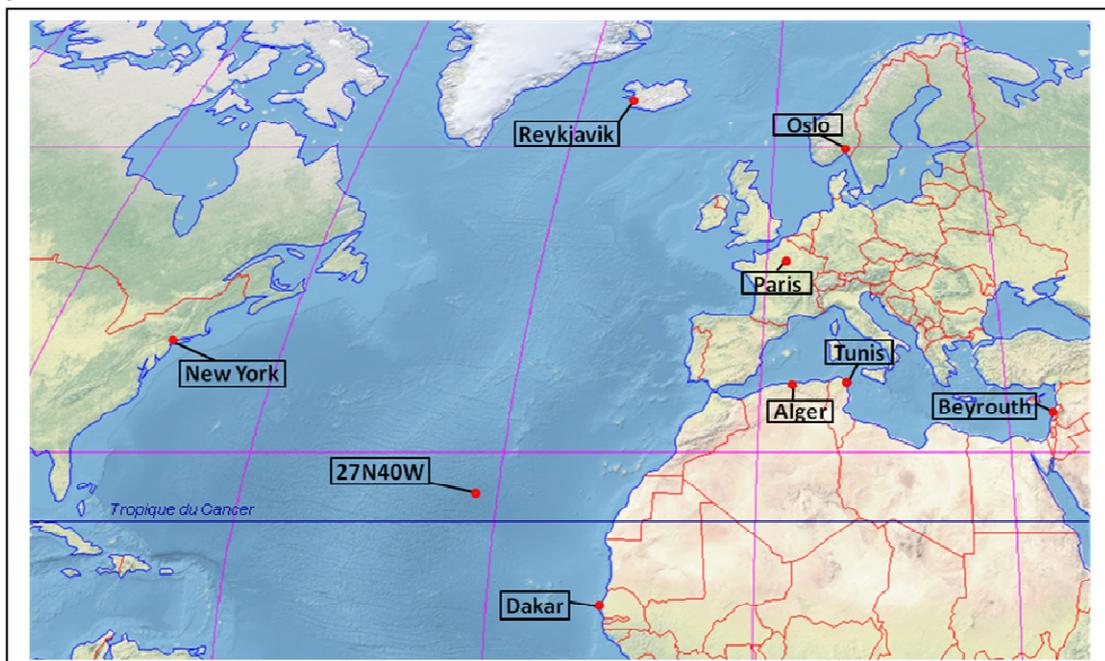
**ssr** : Surface solar radiation ((W/m<sup>2</sup>)s)

Cette liste a été complétée par des mesures de CO<sub>2</sub> fournies par la NOAA et réalisées sur le mont Mauna Loa à Hawaii :

**co<sub>2</sub>** : molfrac ppm (parties par million)

*Le climat tel qu'il est perçu par l'homme est lié à plusieurs variables qui sont la température (variable t2), la présence de nuages (variable tcc), la pluie (variables lsp et cp) ainsi que la radiation solaire (variable ssr). La pluie se divise en deux catégories : les pluies de grande échelle qui sont les pluies qui affectent habituellement les régions tempérées (lsp), et les pluies convectives qui sont les pluies tropicales et les pluies d'orage.*

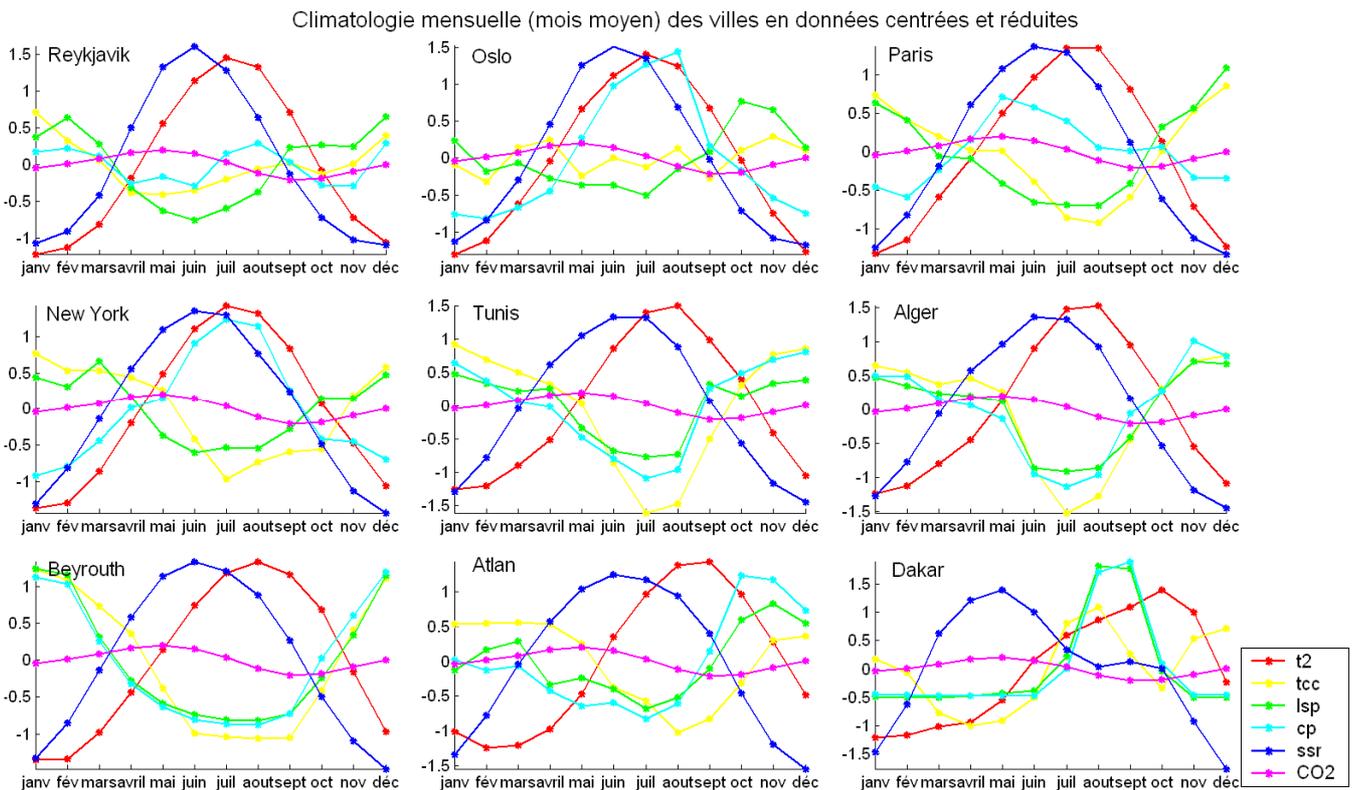
Sauf pour le CO<sub>2</sub>, les valeurs des différentes variables ont été relevées pour 8 villes et un point au milieu de l'Atlantique. Il s'agit des villes suivantes (du nord au sud) : **Reykjavik, Oslo, Paris, New York, Tunis, Alger, Beyrouth** et **Dakar**. La coordonnée du point de l'Atlantique est 27°N, 40°O. Elle vient s'intercaler entre Beyrouth et Dakar, nous l'appellerons « **Atlan** » par la suite. Nous avons extrait les données journalières de janvier 1982 à décembre 2010 soit 29 années entières, à partir desquelles nous avons calculées des moyennes mensuelles. C'est à partir de ces moyennes que nous avons réalisé le travail présenté.



# 1ère partie : Climatologie mensuelle et ACP de Reykjavik

## 1) Climatologies mensuelles

Afin d'avoir un vue d'ensemble des variables étudiées nous avons, pour chaque ville, tracées leurs courbes de climatologie mensuelle en données centrées et réduites sur les 29 années. On précise que chaque point de ces courbes représente la moyenne de chaque mois sur toute la période. Il s'agit donc ici d'une représentation saisonnière des données dont les variations interannuelles sont neutralisées par la moyenne.



L'observation de ces figures nous permet d'effectuer plusieurs commentaires sur les différents régimes qui se dégagent:

- Tout d'abord, concernant le CO<sub>2</sub>, on rappelle qu'il s'agit d'une variable mesurée en un seul lieu géographique qui est donc commune à toutes les villes. On obtient une même courbe pour toutes les villes, dont on voit qu'elle est plutôt dé-corrélée des autres variables quelque soit la ville considérée.

*On remarque une variation saisonnière de petite amplitude du CO<sub>2</sub>. Sa concentration est maximum en mai, et minimum en septembre. Le minimum de septembre est dû à l'assimilation chlorophyllienne dont l'effet cumulé dans l'hémisphère nord est maximum à la fin de l'été.*

- On observe que les courbes de ssr et t2 sont en forme de « cloche convexe » et qu'elles varient d'une manière similaire ; t2 présente un petit décalage vers la droite (par rapport à ssr), ce qui signifierait assez logiquement que la température suit la radiation solaire avec un certain temps de réaction. Par

la suite, d'une façon un peu abusive nous utiliserons le terme d' « ensoleillement » à la place du terme de « radiation solaire ».

*Ce retard correspond à l'inertie thermique du système terre, atmosphère, océan par rapport au forçage solaire.*

- On voit assez vite que Dakar montre un régime bien différent des autres villes. Il apparaît que dans le nord, ssr joue un rôle prépondérant pour les valeurs de température, alors que plus au sud (Dakar en particulier), ssr ne serait plus le principal et seul élément, d'autres phénomènes pourraient intervenir dans la définition des valeurs de température.

- Les autres variables, tcc, lsp et cp (que l'on va associer à une notion de « mauvais temps ») sont souvent dé-corrélées ou anti-corrélées à ssr et t2 (que nous associons à l'idée de « beau temps ») ; la ville d'Alger étant un cas typique de cette opposition.

- Si l'on se penche sur les variables que nous avons associées au « mauvais temps » (tcc, lsp, cp), on remarque que tcc et lsp sont visuellement plus proches entre elles. La variable cp, présente un comportement un peu différent puisque si elle suit souvent tcc et lsp, dans certains cas, (Oslo, New York, et un peu moins Paris) elle se rapproche plutôt des courbes convexes de ssr et t2.

*Les courbes plates représentent des phénomènes qui n'ont pas de variation saisonnière. Par exemple, à Reykjavik, il pleut tout le temps, plus ou moins de la même façon.*

## 2) Les ACP

Il serait fastidieux de procéder à l'ACP de chacune des villes. Dans le cadre de cet exercice, nous nous sommes limités aux 3 villes de Reykjavik, Alger et Dakar. Pour chacune d'elle, nous avons réalisé 2 ACP.

- La 1<sup>ère</sup> porte sur les données moyennes mensuelles soit 348 « individus-mois » (12 mois x 29 années) pour les 6 variables (t2, tcc, lsp, cp, ssr, CO<sub>2</sub>). Il s'agit donc d'une approche saisonnière et qui conserve la variabilité interannuelle. Nous l'évoquerons sous le nom d'ACP saisonnière.
- La seconde ACP sera effectuée avec les données de moyennes annuelles, soit 29 « individus-années », avec les mêmes 6 variables. Il s'agit donc d'une étude simplement interannuelle.

. Dans les 2 cas, les données ont été centrées et réduites pour faire l'ACP de sorte que toutes les variables aient le même impact dans l'analyse.

. Pour chaque étude d'ACP, nous présenterons :

- Pour l'étude saisonnière les courbes des climatologies (représent du §1).  
Pour l'étude interannuelle une courbe qui représente la moyenne de chaque année pour chaque variable.
- L'histogramme des pourcentages d'inertie des valeurs propres avec l'indication des pourcentages cumulés (courbe rouge).
- Le cercle des corrélations (CC).

- Le nuage des individus (NI), chaque individu étant représenté par 2 triangles orientés l'un selon l'abscisse et l'autre selon l'ordonnée et dont les tailles sont proportionnelles à la qualité de représentation sur l'axe correspondant. Pour avoir un point de comparaison, nous avons représenté, en bas et à gauche de la figure, 3 triangles noirs dont les tailles correspondent respectivement à des qualités de représentation égales à 1.0 , 0.5 et 0.1 (la qualité de représentation étant un pourcentage). Pour une bonne interprétation des axes, il faudra toujours bien tenir compte de cette qualité. Bien entendu, le résultat d'une ACP est en partie qualitative, elle ne permet pas de statuer sur tous les individus. L'interprétation du nuage s'appuiera sur les individus les mieux représentés et sur le pourcentage d'inertie que l'on regarde. Il y a continuité dans la projection, ce qui amènerait à concevoir une graduation dans l'interprétation des individus.

La couleur des triangles est déterminée par une échelle de couleur qui est associée, selon le cas, soit au mois, soit à l'année (ce qui est indiqué par la mention **Ech.col** sur la figure). De plus les points ont été labélisés par leur mois dans le cas de l'étude saisonnière, et par leur année pour l'étude interannuelle.

. En général (donc sauf indication contraire), les axes 1 et 2 des composantes principales sont respectivement portés par l'abscisse et l'ordonnée des cercles de corrélation ; il en est bien sûr de même pour les nuages des individus. Sauf cas particulier, l'interprétation ne portera que sur ces 2 premiers axes.

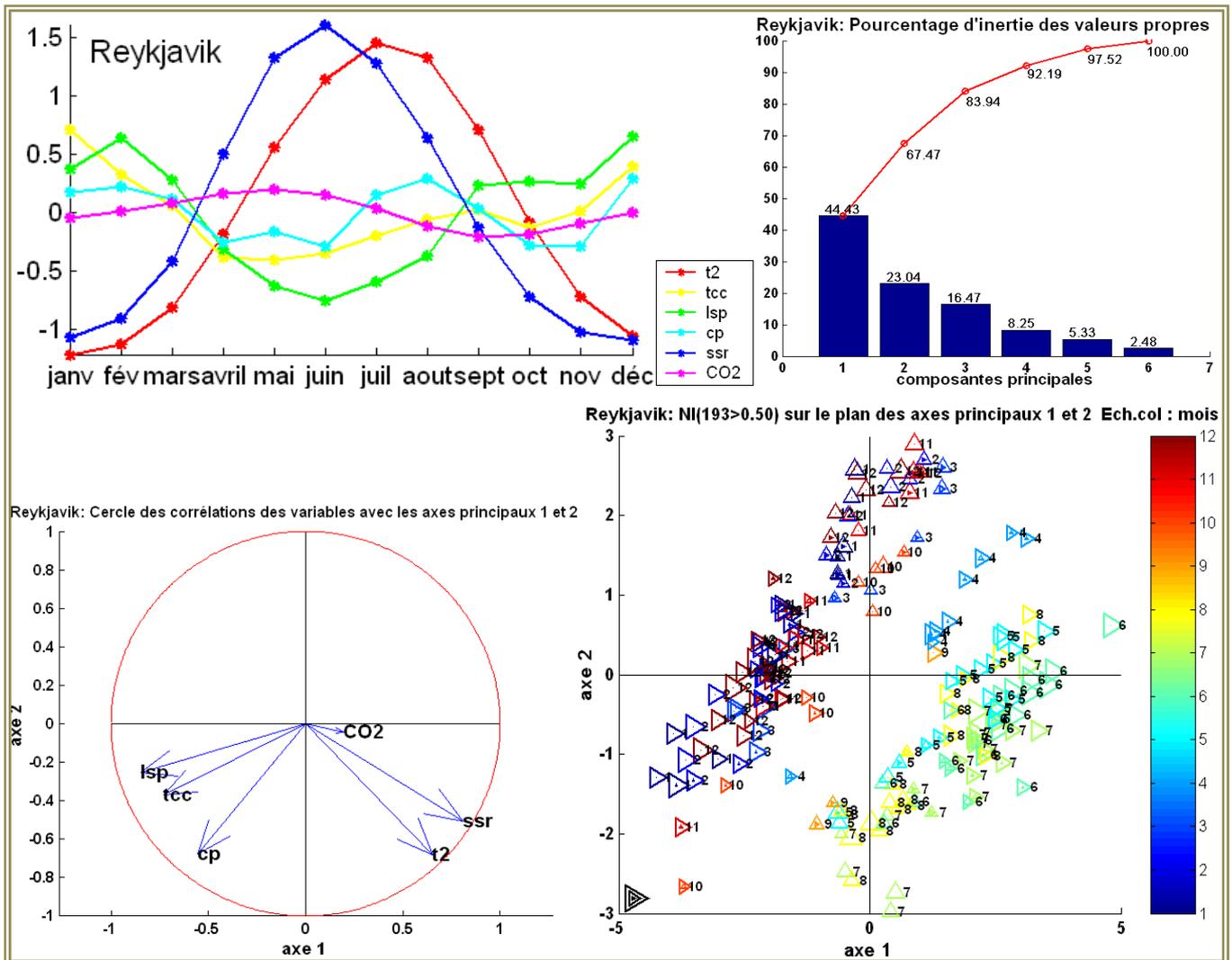
. Pour éviter que les nuages des individus des études saisonnières ne soient trop surchargés et donc illisibles, ils ont été présentés en ne retenant, parmi les 348 individus, que ceux dont la qualité de représentation est supérieure à 0,5. Le nombre  $n$  de points utilisés pour représenter le nuage est indiqué sur la figure par la mention : NI( $n>0.5$ ).

L'ensemble de ces représentations sont regroupés dans un encadré suivi des commentaires qui s'y rapportent.

## 2.1 ) REYKJAVIK

### • Cas saisonnier (données moyennes mensuelles)

En utilisant les deux premiers axes, on regarde ici 67,47% de l'inertie du nuage de point ce qui en fait un plan assez représentatif.



Constatant tout d'abord que le CO<sub>2</sub> est mal représenté (i.e. proche du centre du cercle des corrélations) son interprétation n'est pas opportune ici.

◆ L'axe 1 du cercle des corrélations de Reykjavik montre que les variables ssp et t2 sont corrélées, comme on pouvait s'y attendre grâce aux tracés des climatologies (mensuelles) ; de même lsp avec tcc. On voit aussi que cp est presque dé-corrélés avec t2. A droite de l'axe 1 on retrouve les périodes ensoleillé, chaudes et sèches (par opposition) tandis qu'à gauche se situent les mois pluvieux, couverts et froids.

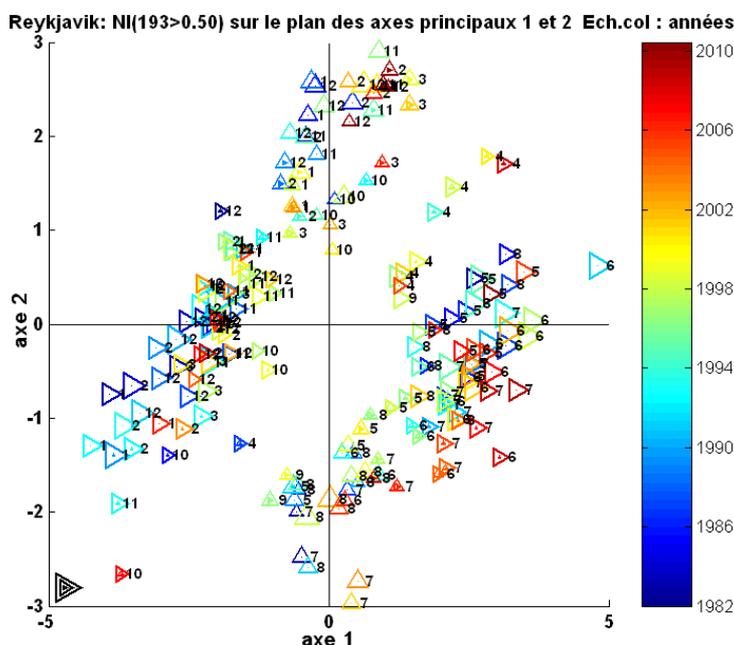
Les variables t2 et cp sont celles qui apparaissent les plus corrélées à l'axe 2. La partie inférieure de cet axe doit correspondre à des périodes chaudes et pluvieuses. On peut dire que la partie supérieure de l'axe 2, pourrait concerner alors des périodes plutôt sèches et froides. On le vérifie sur la climatologie.

◆ Le Nuage des données mensuelles s'étirent selon une droite perpendiculaire à t2 qui sépare les mois froids d'hiver des mois chaud d'été, le mois d'avril faisant office de frontière. En se référant au cercle des corrélations, on peut grossièrement dresser le tableau suivant :

<u>Positionnement</u>	<u>mois</u>	<u>caractérisation des individus</u>
à droite de l'axe 1 :	avril, <b>mai, juin, juillet août</b>	ensoleillé, chaud et sec
à gauche de l'axe 1 :	oct, <b>nov, déc, janv, fév</b>	pluvieux, couvert et froid
en haut de l'axe 2 :	<b>nov, déc, janv, févr</b> , mars	sec et froid
en bas de l'axe 2 :	<b>mai, juin, juillet août</b> , septembre	chaud et pluvieux

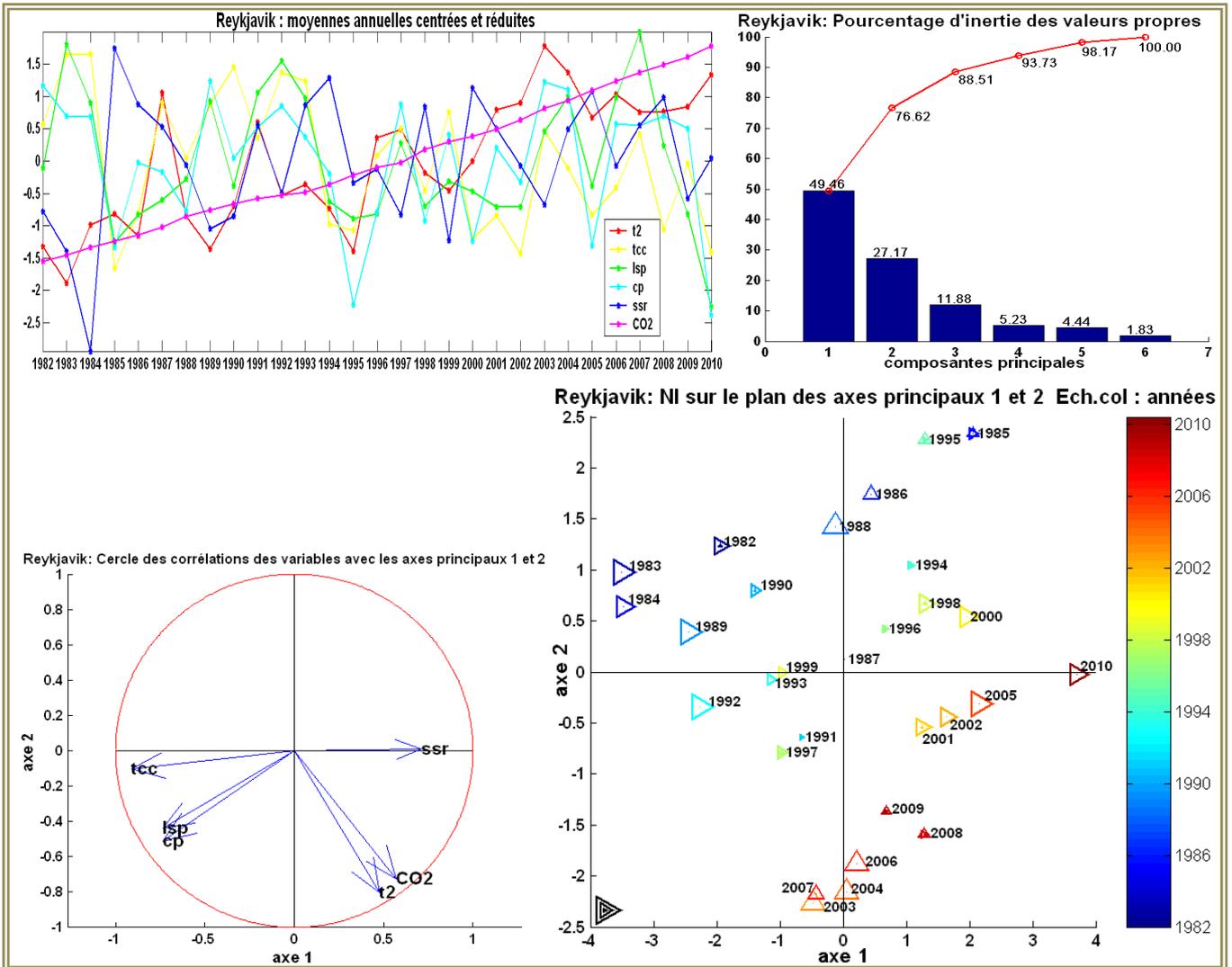
Beaucoup de mois (ceux mis en gras) se retrouvent sur deux types de temps dont la température est cependant la même (chaud ou froid). C'est la disposition par rapport à la variable t2 qui permet de comprendre cette cohérence par rapport à l'opposition chaud/froid. Par exemple, décembre peut être froid et pluvieux ou froid et sec, mais il n'est jamais chaud. La position d'un mois donné à deux types de temps différents provient de la variabilité interannuelle. Certains mois, comme par exemple le mois de novembre, se trouvent être sec une année et pluvieux une autre.

◆ Nous avons également produit, ci-dessous, le nuage des individus en utilisant les années comme échelle de couleur. On doit convenir que cela ne nous a pas apporté d'élément intéressant pour notre réflexion puisque qu'on retrouve de part et d'autre des axes des années anciennes et récentes. Peut être pouvons nous constater une certaine dominante des années anciennes à gauche et plus récentes à droite, ce qui traduirait l'augmentation des températures. Ceci étant, cette apparence n'est peut être pas assez franche pour conclure ici à ce réchauffement. Mais cela montre tout de même que les variations interannuelles sont cachées par les variations saisonnières.



●) Cas Interannuel (données moyennes annuelles)

Avec 76,62% de d'inertie pour les deux premiers axes, on a une idée assez représentative du nuage des données.



◆ Alors que les courbes de climatologie mensuelle sont plus ou moins interprétables visuellement, Il n'en est plus de même avec les courbes interannuelles qui apparaissent plus mélangées. A priori, les moyennes annuelles des variables présentent une grande variabilité qui ne s'interprète pas au premier coup d'œil. Seule la courbe du CO<sub>2</sub>, avec sa progression régulière se démarque.

◆ Le cercle des corrélations exhibe plusieurs types de dépendance:

- les variables lsp et cp d'une part puis t2 et CO<sub>2</sub> d'autre part sont fortement corrélées. La corrélation de t2 et CO<sub>2</sub> est en accord avec les hypothèses du réchauffement climatique qui dit que l'augmentation de CO<sub>2</sub> provoque un effet de serre (et donc une augmentation des températures).

- les variables tcc et ssr sont assez logiquement anti corrélées car on comprend bien que globalement, plus la couverture nuageuse est importante (tcc) moins il y a d'ensoleillement (ssr) et inversement. Ce sont principalement ces deux variables qui permettent une compréhension de l'axe 1 : nuageux, pluvieux et un peu frais pour la gauche de l'axe et ensoleillé, un peu chaud et sec sur la partie positive de l'axe.

L'axe 2 a sa partie inférieure associée à un temps chaud et un peu pluvieux. Par opposition on déduit que la partie supérieure de cet axe correspond à un temps froid et sec.

◆ Si l'on regarde le nuage des individus à la lumière du cercle des corrélations, on peut dresser, pour certains individus (i.e. les années), ceux qui sont les mieux représentés, le tableau suivant :

<u>Positionnement</u>	<u>mois</u>	<u>caractérisation des individus</u>
à droite de l'axe 1 :	2010, 2005, 2000,2001, 2002	ensoleillé, un peu chaud et sec
à gauche de l'axe 1 :	1983, 1984, 1989, 1992,1982	nuageux, pluvieux et un peu frais
en haut de l'axe 2 :	1985, 1995, 1986, 1988	froid et sec
en bas de l'axe 2 :	2003, 2004, 2007, 2006	chaud et un peu pluvieux

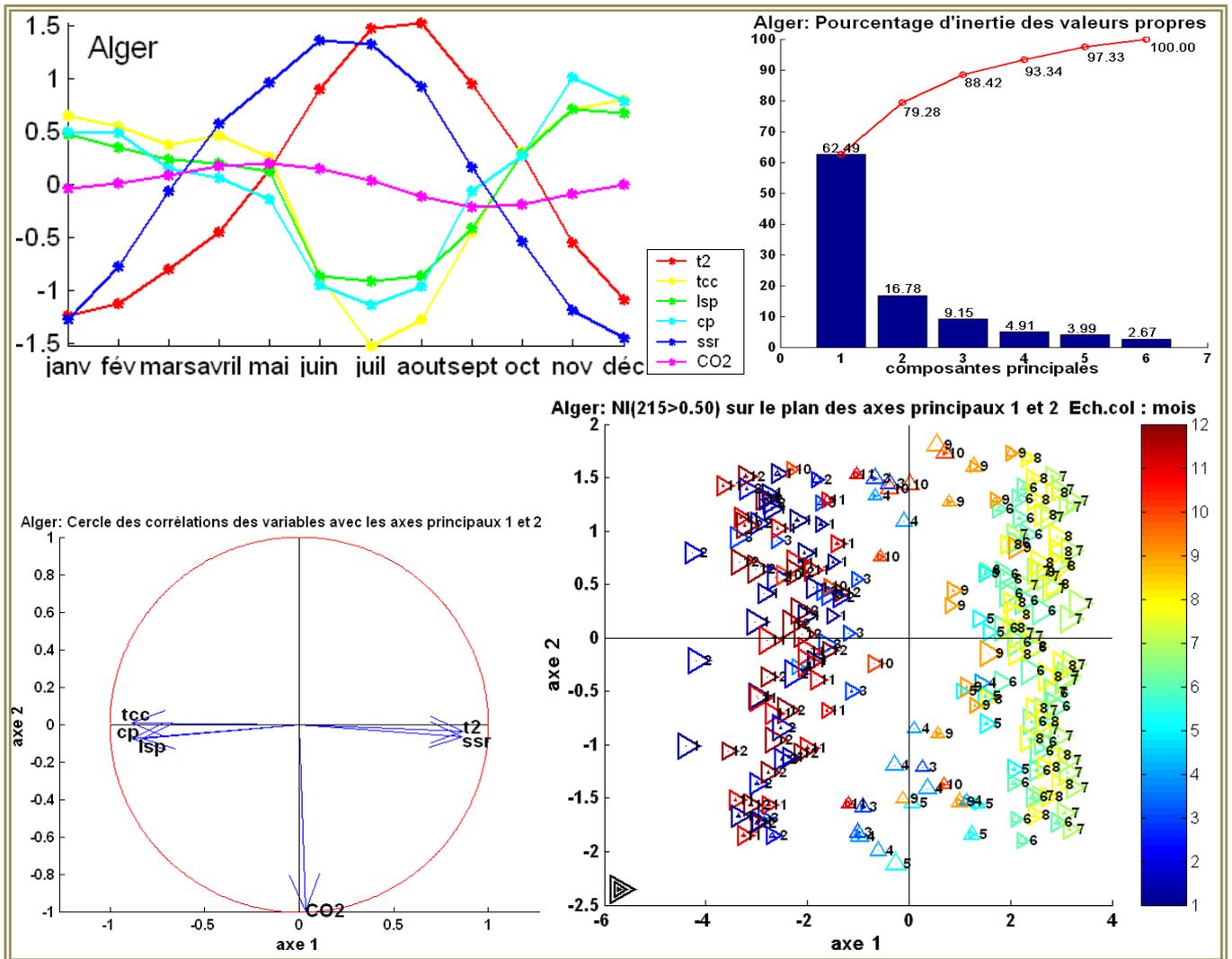
Le tableau ci dessus montre clairement que les années marquantes (qui s'éloignent le plus de la climatologie mensuelle calculée sur les 29 années) opposent des années avant l'an 2000 et des années après. On peut supposer que le temps a évolué d'un type moins chaud et plus pluvieux (pour les années de début de période à partir de 1982 à un type plus chaud et moins pluvieux pour les années de fin de période (à partir des années 2000 environ). Cette information obtenue à l'aide de l'ACP peut maintenant être plus facilement observée sur la figure des courbes où l'on voit effectivement que, dans la partie gauche des courbes, les variables dites de « mauvais temps » (tcc, lsp, cp) sont plutôt au dessus de celles de « beau temps » (t2, ssr) et que sur la partie droite, c'est plutôt l'inverse, le point de croisement se situant environ au alentour de l'année 1996. On remarque effectivement que l'année 1996 se situe assez proche de l'origine du plan qui correspond à l'année moyenne

## 2<sup>ème</sup> partie : ACP d'Alger et de Dakar

Se reporter à la première partie pour le descriptif des ACP

### 1) ALGER

#### • Cas saisonnier (données moyennes mensuelles)



Par rapport aux autres villes, la première composante principale d'Alger est celle qui recueille le plus d'inertie. Le plan formé par ces 2 composantes totalise 79.28% de l'inertie, ce qui est suffisamment représentatif pour entreprendre une interprétation valide.

◆ Alger est la ville où l'opposition entre variables de « beau temps » (t2, ssr) et de « mauvais temps » (tcc, lsp, cp) est le plus clairement dessinée sur la figure des courbes de climatologie mensuelle. Les pluies présentent une forte variabilité saisonnière. Elles sont importantes l'hiver et nulle l'été. De même, il y a une forte variabilité de la radiation solaire et des températures, qui sont fortes en été et faibles en hiver). Cette franche opposition se retrouve sur le cercle des corrélations où ces deux groupes de variable qui sont d'ailleurs bien représentées sont clairement anti-corrélés sur l'axe 1. On remarque dans ce cas l'émergence du CO<sub>2</sub> sur l'axe 2, lui même très bien représenté. On remarque de plus que, dans cette approche saisonnière, ce dernier apparaît tout à fait dé-corrélé des autres

variables<sup>1</sup>. Bien que visible à partir des courbes climatologiques, cela est bien mieux mis en évidence sur le cercle.

La gauche de l'axe 1 représente donc un temps couvert, pluvieux et froid, le coté positif de l'axe correspond à des périodes (mois) ensoleillées, chaudes et sèches. L'axe 2 met en opposition les mois de plus forte concentration en CO<sub>2</sub> pour la partie inférieure de l'axe aux mois de plus faible concentration pour la partie positive de l'axe.

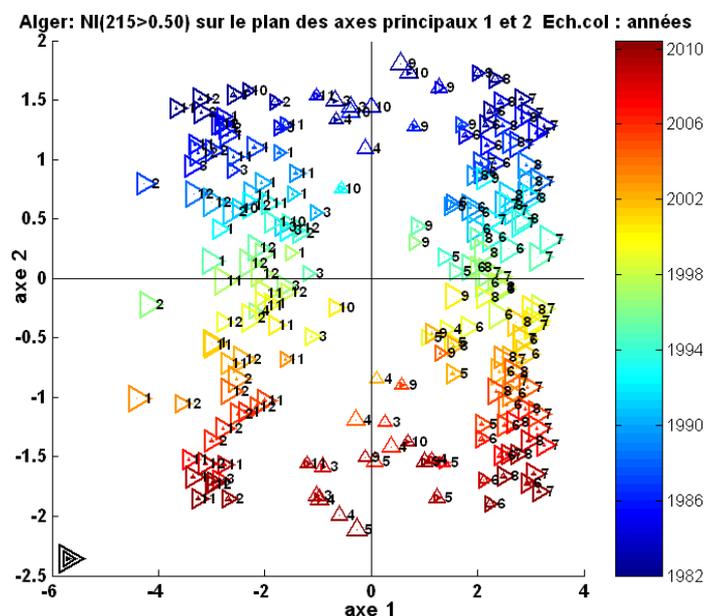
◆ Suivant ces caractéristiques, et à l'aide du nuage des individus, nous pouvons maintenant dégager une classification des mois en se fiant à ceux qui ont les plus fortes qualités de représentation :

<u>Positionnement</u>	<u>mois</u>	<u>caractérisation des individus</u>
à droite de l'axe 1 :	juin, juil, août, mai, sept	ensoleillé, chaud et sec
à gauche de l'axe 1 :	nov, déc, janv, fév,	couvert, pluvieux et froid
en haut de l'axe 2 :	sept, oct, mars, avril	faible concentration (de CO <sub>2</sub> )
en bas de l'axe 2 :	sept, oct, , mars, avril, mai	forte concentration (de CO <sub>2</sub> )

On peut constater que l'axe 1 permet de bien séparer les mois d'été (juin, juillet, août), ensoleillés, chauds et secs (à droite de l'axe) de ceux d'automne/hiver (novembre, décembre, janvier, février) couverts, pluvieux et froids (à gauche de l'axe).

Concernant l'axe 2, celui-ci recueille des mêmes mois de part et d'autre de l'axe 2 (septembre, octobre, mars, avril). Il ne permet donc pas d'opérer une séparation sur le critère des périodes mensuelles utilisé qui du coup n'apparaît pas pertinent ici pour la variable du CO<sub>2</sub>.

◆ L'utilisation d'une échelle de couleur basée sur l'année, comme c'est le cas pour la figure du nuage des individus présentée ci-dessous, nous permet de comprendre la position des points sur l'axe 2 en fonction des données que l'on a utilisées.

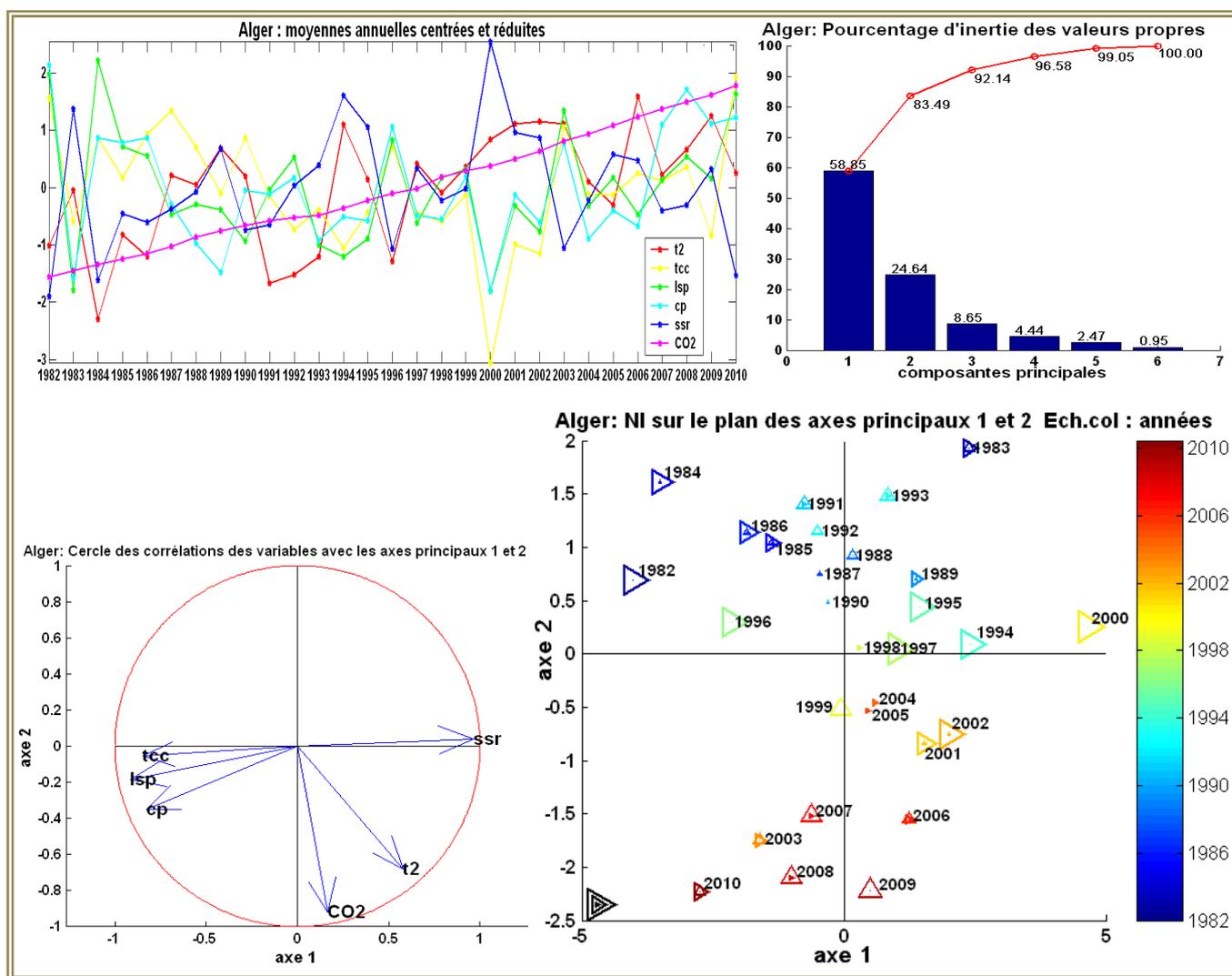


Cette fois, c'est pour l'axe 1 que le critère de l'année, utilisé comme échelle de couleur, n'est pas

<sup>1</sup> En fait, on a vu (TP sur l'effet de serre (indicateurs bidimensionnels)) qu'il y avait une corrélation négative entre la concentration de CO<sub>2</sub> et la température qui pouvait être mise en évidence à condition d'opérer un décalage temporelle de quelques mois ce qui n'a pas été fait ici.

pertinent. Par contre, ce critère nous permet ici, une compréhension de l'axe 2. On voit en effet que celui-ci reflète l'évolution temporelle qui va des années les plus anciennes en haut de l'axe, aux années les plus récentes dans la partie inférieure de l'axe. Puisque c'est dans cette direction qu'est orientée la variable du CO<sub>2</sub> sur le cercle des corrélations, cela montre que la concentration de CO<sub>2</sub> a régulièrement augmentée au cours des trois dernières décennies étudiées. Remarquons que, bien que travaillant dans un cadre saisonnier, c'est une évolution interannuelle qui ressort, sur l'axe 2 associé au CO<sub>2</sub>. D'ailleurs, en se reportant aux courbes interannuelles (ci-après), l'augmentation régulière du CO<sub>2</sub> apparaît clairement. Cette tendance interannuelle cache alors les variations saisonnières du CO<sub>2</sub> qui ne ressortent donc pas sur cette ACP.

●) Cas Interannuel (données moyennes annuelles)



Cette ACP est remarquable du fait que les deux premiers axes cumulent 83,49% d'inertie, ce qui donne à l'interprétation une validité un peu supérieure aux autres analyses déjà effectuées.

◆ Le cas d'Alger présente des similitudes avec celui de Reykjavik. Sur le cercle des corrélations, d'une part, on retrouve un axe 1 qui oppose l'ensoleillement (ssr) et dans une moindre mesure la température (t2) dans sa partie droite, à une couverture nuageuse (tcc) et des précipitations (lsp, cp) dans sa partie gauche. D'autre part, on retrouve aussi une association de l'axe 2 avec la température

(t2) et le CO<sub>2</sub>, ces 2 variables étant relativement corrélées entre-elle, mais un peu moins qu'elles le sont à Reykjavik.

◆ Le nuage des individus nous permet de préciser les années qui sont le plus concernées par ces caractérisations. Ainsi on trouve que les années 2000, 1994, 1995, 1997, 2001 et 2002, qui sont les mieux représentées à la droite de l'axe 1 ont du être ensoleillée, chaude et sèches, les années 1982, 1984, 1985, 1986, 1996 ainsi que 2010, positionnées sur la gauche de l'axe 1, devaient être nuageuses, pluvieuses et plutôt fraîches, les années 2009, 2008 et 2007 bien représentées sur la partie inférieure de l'axe 2 furent des années chaudes. Pour le coté positif de cet axe il n'y a pas d'individu qui y soit réellement bien représenté. Les années les mieux représentées, 1991 1992 et 1993, ont dues être plus froides (ce que l'on peut vérifier sur le graphe des courbes des moyennes annuelles).

On remarque de grands écarts à la moyenne (origine du plan), positifs ou négatifs, lorsque l'on regarde de part et d'autre de l'axe 2.

L'ensemble de ces remarques peuvent être synthétisées par le tableau suivant:

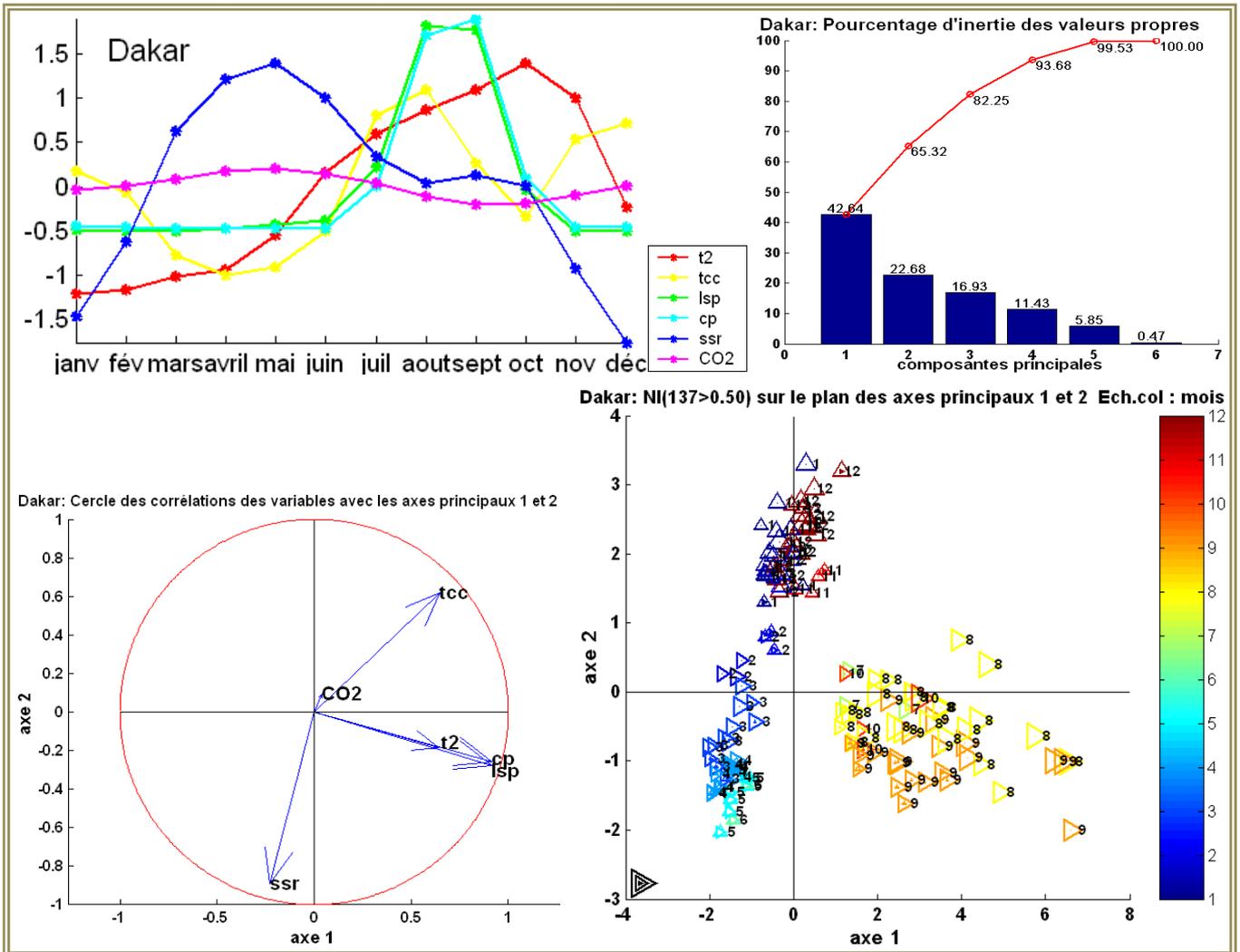
<u>Positionnement</u>	<u>années</u>	<u>caractérisation des individus</u>
à droite de l'axe 1 :	2000 1994, 1995, 1997, 2001, 2002	ensoleillé et chaud et sec
à gauche de l'axe 1 :	1982, 1984, 1985, 1986, 1996, 2010	nuageux, pluvieux et frais
en haut de l'axe 2 :	1991, 1992, 1993	froid
en bas de l'axe 2 :	2009, 2008, 2007	chaud

En fait, le nuage des individus représenté à partir d'une échelle de couleur associée à l'année reflète bien l'augmentation des températures, puisqu'en particulier, l'axe 2, qui y est associé en partie, correspond à la série chronologique qui va des années les plus anciennes en bleu aux plus récentes en rouge foncé.

Remarquons de plus que pour les deux cas d'étude (saisonnier et interannuelle), le CO<sub>2</sub> est représenté sur l'axe 2, mais l'étude interannuelle fait ressortir le lien entre l'augmentation du CO<sub>2</sub> et celle de la température (ce qui n'est pas le cas du saisonnier).

## 2) DAKAR

### ● Cas saisonnier (données moyennes mensuelles)



◆ Les courbes de climatologie mensuelle de Dakar montrent que cette ville suit une évolution saisonnière bien différente des autres villes. La courbe de ssr et t2 sont très fortement décalées, celles des précipitations (cp et lsp) sont quasiment confondues.

*En effet il pleut pendant une courte période de temps (Juin Juillet Aout), période liée au déplacement de l'ITCZ (Inter Tropical Continental Zone) qui atteint sa latitude maximale durant cette période. L'ITCZ est une zone de convergence pluvieuse qui se déplace entre l'équateur et le tropique du cancer.*

On remarque que la courbe de couverture nuageuse (tcc) change plusieurs fois de sens.

*Cela est dû à la présence de plusieurs types de nuages : nuages dus à la condensation de la vapeur d'eau en été, nuages liés aux aérosols et aux poussières sahariennes pendant les autres mois.*

◆ Cela se retrouve sur le cercle des corrélations, Dakar se démarque ici des analyses effectuées les autres villes. Ceci est particulièrement visible pour la dé-corrélation de t2 et ssr, mais aussi entre la variable tcc d'une part, et les variables lsp et cp d'autre part. Ces deux dernières sont par ailleurs très fortement corrélées entre elles. On retrouve donc aussi une dé-corrélation de cp et lsp d'avec ssr. On garde néanmoins et toujours assez logiquement une bonne anti-corrélation entre l'ensoleillement (ssr)

et la couverture nuageuse (tcc). L'opposition entre ce que nous avons identifié, de « beau » et de « mauvais » temps (à propos des autres villes) n'apparaît désormais plus aussi clairement. La partie droite de l'axe 1 correspondrait plutôt à des périodes chaudes et humides (que l'on pourrait associer à la « mousson africaine »). Par opposition, la partie gauche de l'axe 1 dénoterait des mois plutôt froids et secs. Le rôle de l'axe 2, quant à lui, serait d'opposer des mois ensoleillés pour ceux qui se portent sur sa partie inférieure, à des mois couverts pour ceux dont les coordonnées sont positives sur cet axe.

*A Dakar, on constate que ssr n'est plus la seule variable qui provoque l'augmentation de température. A cette latitude, l'ITCZ entraine une météorologie perturbée car c'est un phénomène qui intervient fortement pour modifier les conditions météorologiques.*

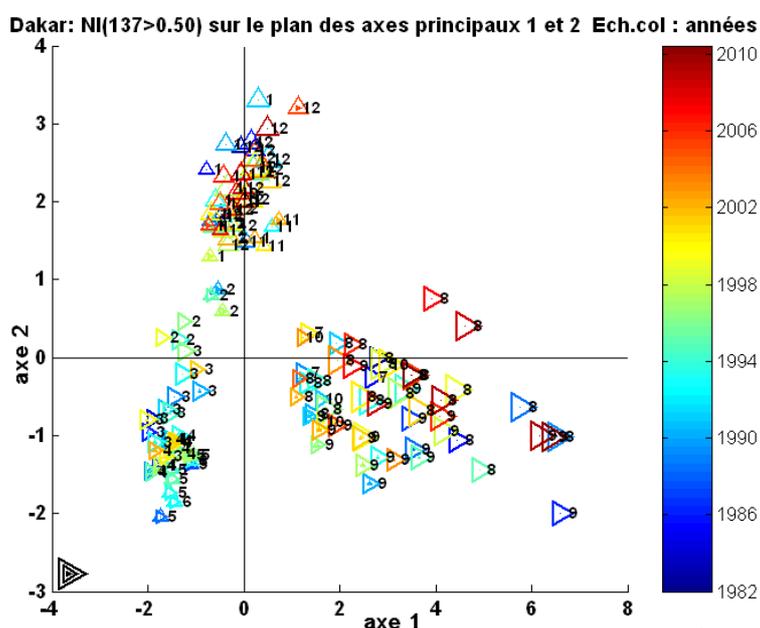
◆ A la suite de cette première étude, nous pouvons classer les individus-mois du nuage des individus sur lequel, pour une fois, on voit que trois groupes se sont assez bien formés.

On peut tenter une description synthétique de l'ensemble des mois de l'année, sachant que certains mois sont assez délicats à caractériser.

<u>Positionnement</u>	<u>mois</u>	<u>caractérisation des individus</u>
à droite de l'axe 1 :	août, sept	chaud et humide
à gauche de l'axe 1 :	<b>fév</b> , mars, avril	froid et sec
en haut de l'axe 2 :	janv, <b>fév</b> , nov, déc	couvert
en bas de l'axe 2 :	mai, juin	ensoleillé

Remarquons que l'axe 2 n'agrège pas d'individu réellement bien représenté dans sa partie inférieure, ce qui signifierait que des mois qui seraient juste ensoleillés (sans chaleur ni précipitation) doivent être peu fréquents ; il pourrait plutôt s'agir des mois de mai et juin. Il reste les 3 groupes apparents associés à leur caractérisation : « chaud et humide », « froid et sec », « couvert ». Il est à noter que le mois de février qui est dans une position intermédiaire entre un type de temps « couvert » et « froid et sec », doit jouer un rôle de transition entre ces 2 types de temps. Le mois de juillet apparaît peu sur le plan principal puisque nous n'avons retenu que les individus qui avaient une qualité de représentation supérieure à 0,5.

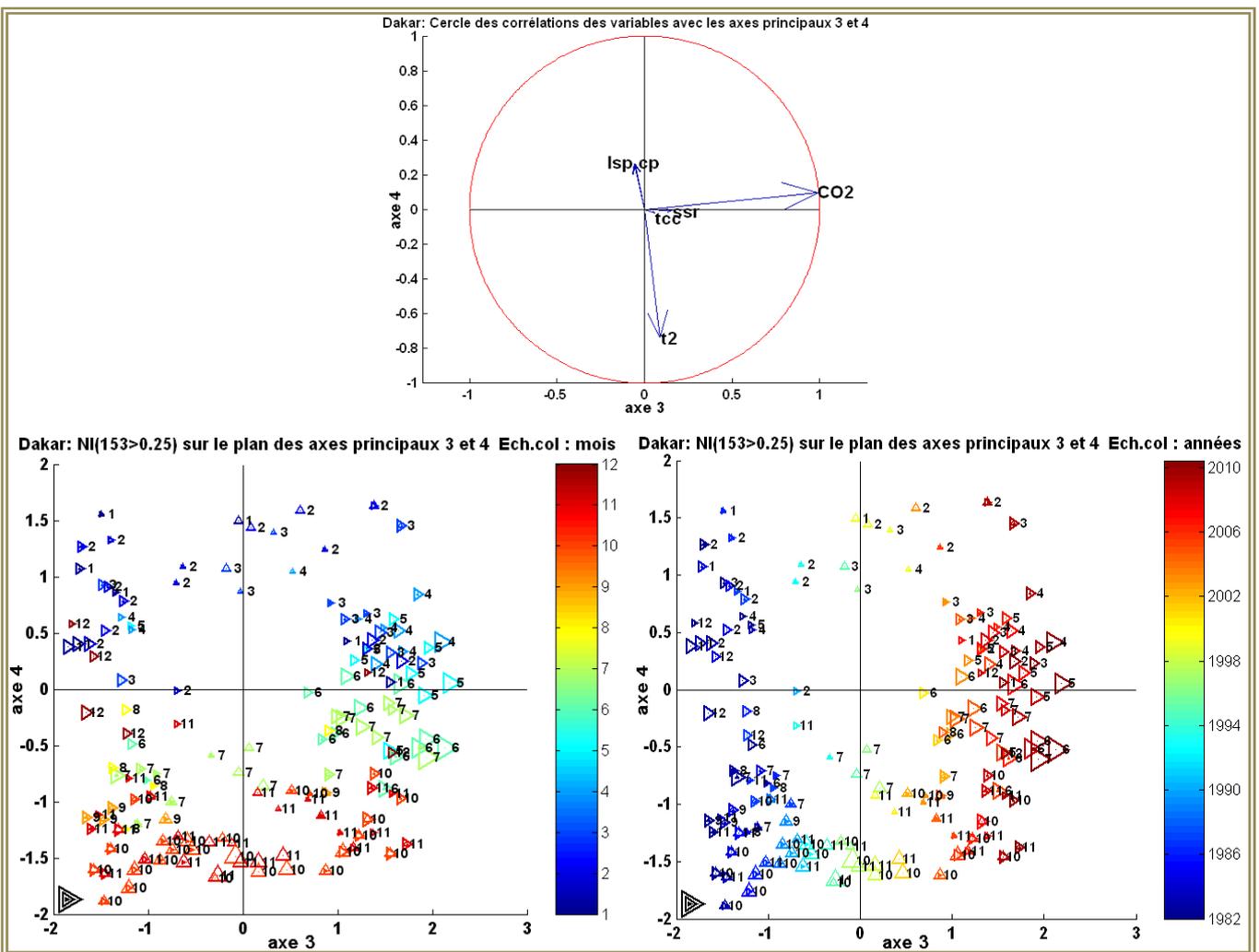
◆ Nous produisons également, ci-dessous, le nuage des individus en utilisant les années comme échelle de couleur. On voit, que des années anciennes comme récentes apparaissent sur les parties positives des 2 axes signifiant ainsi que le critère de l'année n'est pas pertinent pour distinguer les individus.



► Etude complémentaire du plan des axes 3 et 4 :

Nous n'avons jusqu'à présent qu'étudié les projections sur les deux premiers axes. Il peut être intéressant de regarder les projections sur les autres axes. Dakar présente une inertie non négligeable sur les axes 3 et 4 (supérieure à 28%). Avec les climatologies mensuelles par ville (§1) on avait vu que les courbes des moyennes mensuelles de Dakar présentaient plus de décalage par rapport aux autres villes et en particulier pour  $ssr$  et  $t_2$ . De ce fait, on comprend que l'ACP a utilisé plus d'axe pour représenter les phénomènes significatifs. Nous allons donc étudier de plus près le plan formé par les axes 3 et 4 de Dakar.

Les figures des nuages des individus ci-dessous, ne reprennent que ceux dont la qualité de représentation est supérieure à 0,25. Ceci car, la qualité globale de représentation est en général plus faible sur les plans de moindre inertie ; il y aurait eu trop peu d'individus représentés si on avait conservé le seuil de 0,50.



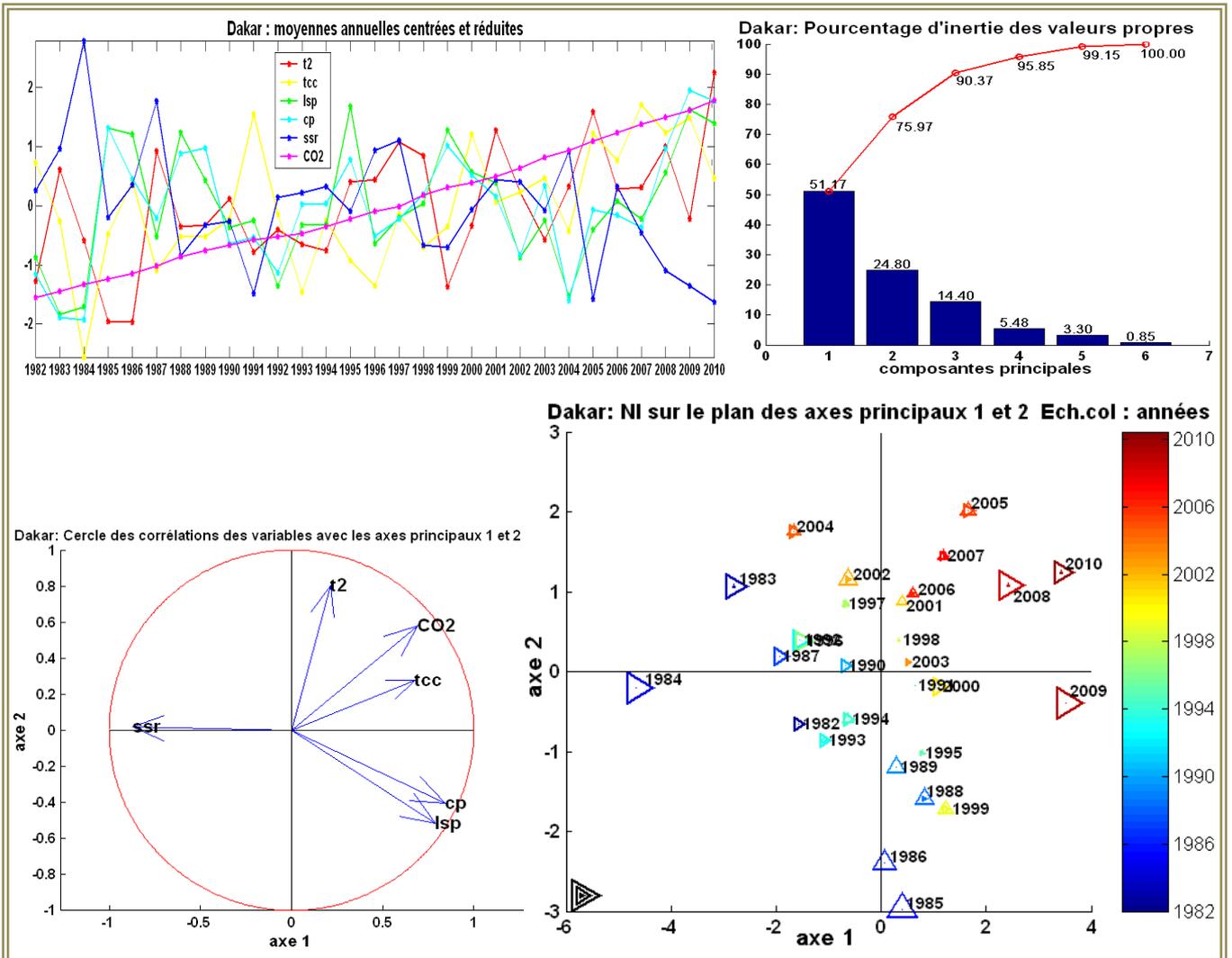
◆ Sur le cercle des corrélations, l'axe 3 est très représentatif de la concentration de CO<sub>2</sub>. L'axe 4 pour sa part est plutôt représentatif de la température t<sub>2</sub>. Sur ce plan factoriel et pour cette approche saisonnière, ces 2 variables apparaissent dé-corrélées. On ne peut rien dire des autres variables trop proches du centre du cercle.

◆ Pour la température  $t_2$ , c'est l'échelle de couleur par mois qui permet une interprétation de l'axe 4 du nuage des individus. Le haut de l'axe correspond à des mois froids (janvier, février, mars, avril) ; vers le centre de l'axe on rencontre plutôt des mois de température intermédiaire (mai, juin, juillet, décembre) ; les mois les plus chauds (août, septembre, octobre, novembre) se trouvent en bas de l'axe conformément à la position de la variable  $t_2$  sur le cercle des corrélations. En fait cela est en accord avec ce que l'on peut lire directement sur la courbe de moyenne  $t_2$  (en rouge) sur la figure des climatologies mensuelles.

◆ Le nuage des individus dont l'échelle de couleur est l'année, montre une concentration de  $\text{CO}_2$  qui augmente avec l'échelle du temps. En effet, la partie gauche de l'axe 3, qui au vu du cercle des corrélations correspond à une plus faible concentration en  $\text{CO}_2$ , porte les coordonnées des années les plus anciennes, puis, en allant progressivement vers la droite de l'axe, qui représente des concentrations de  $\text{CO}_2$  plus élevés, on rencontre les années plus récentes. Bien que travaillant avec des données mensuelles (étude saisonnière), c'est un phénomène interannuel qui apparaît ici. On peut comprendre que la l'augmentation du  $\text{CO}_2$  est identique en été et en hiver, ainsi, quelque soit la saison, on observe la progression du  $\text{CO}_2$  avec les années. L'échelle de couleur annuelle retrouve une pertinence pour l'étude de ce plan (alors qu'elle n'en avait pas pour le plan des axes 1 et 2 étudié précédemment).

● **Cas Interannuel (données moyennes annuelles)**

Le pourcentage d'inertie expliquée du plan 1-2 (75,97%) est suffisant pour en entreprendre valablement une étude.



◆ Mis à part la progression régulière de la variable  $CO_2$ , les courbes des moyennes annuelles ne permettent pas aisément de dégager d'information synthétique de façon évidente. L'utilisation du cercle des corrélations, par contre, nous renseigne de plusieurs choses.

Sur l'axe 1, la variable d'ensoleillement  $ssr$  (à gauche (de l'axe)) s'oppose à toutes aux variables ( $cp$ ,  $lsp$ ,  $tcc$  (il s'agit des variables dites de « mauvais temps » qui caractérisent une séquence couverte et pluvieuse)). Cette variable  $ssr$  est plutôt dé-corrélée avec la variable  $t2$ . Même si elle est un peu moins marquée que pour d'autre ville, on retrouve la proximité entre la concentration de  $CO_2$  et la température  $t2$ . Ainsi la partie positive de l'axe 2 doit correspondre à des individus, c'est-à-dire, des années globalement plus chaudes. Par opposition, la partie inférieure de l'axe 2 devrait correspondre à des années plus fraîches.

◆ On peut à l'aide du cercle des corrélations, interpréter les années du nuage des individus en ne retenant que les plus représentatifs. Sur l'axe 1 on remarque qu'à Dakar, les années 1984 et 1983 ont dues être bien ensoleillées (plus qu'à l'habitude), les années 2009, 2010 et 2008, ont connu des couvertures nuageuse plus importantes et ont du recevoir plus de précipitation. La lecture de l'axe 2

nous permet de voir que les années 1985 et 1986 ont du être, peut être, plus fraîches qu'à l'habitude. Concernant la partie supérieure de l'axe 2, nous ne remarquons pas d'année très représentative, signifiant par là qu'aucune année n'est apparue particulièrement plus chaude pour ces dernières décennies à Dakar.

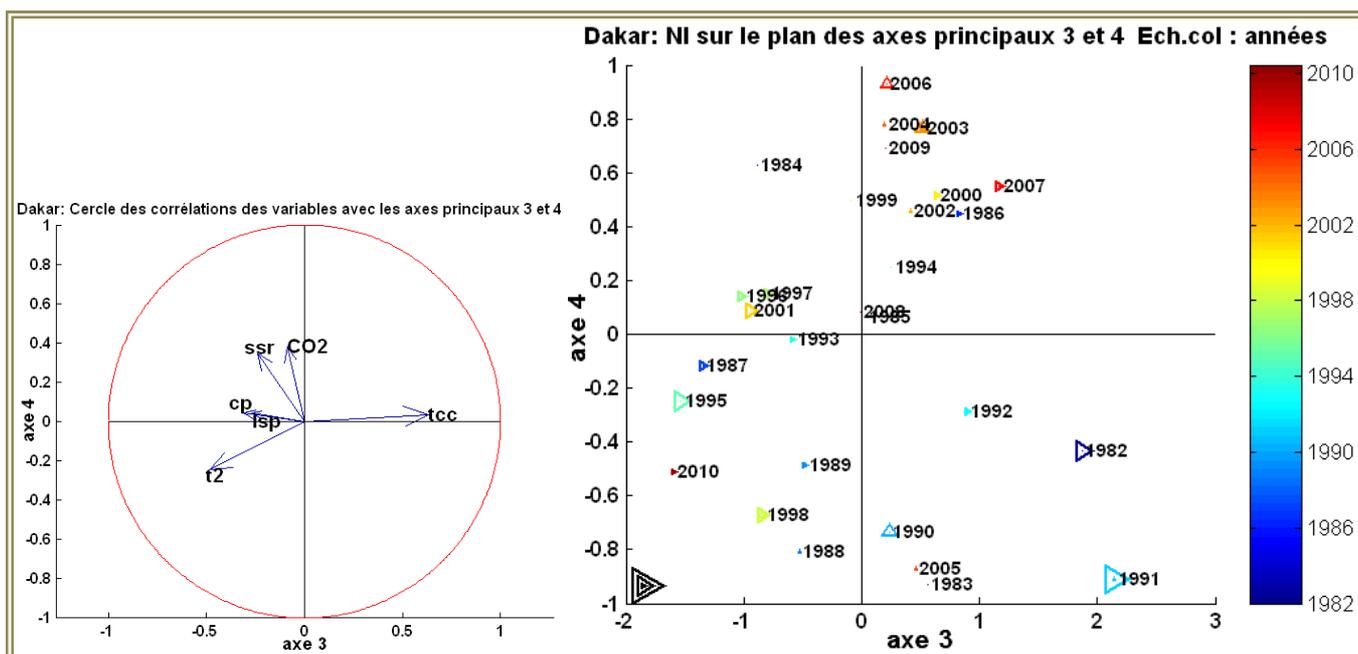
Dans le tableau suivant, on reprend de façon synthétique ce qui à été dit :

<u>Positionnement</u>	<u>mois</u>	<u>caractérisation des individus</u>
à droite de l'axe 1 :	2009, 2010, 2008	couvert pluvieux
à gauche de l'axe 1 :	1984, 1983 (1987,1992, 1996)	ensoleillé
en haut de l'axe 2 :	(2002,2005)	chaud
en bas de l'axe 2 :	1985, 1986 (1989, 1988)	fraîches

Sur un autre thème que celui du type de temps, il est intéressant de remarquer que l'orientation du vecteur CO<sub>2</sub> correspond tout à fait à la façon dont sont disposées les années des plus anciennes aux plus récentes, et cela en accord avec la courbe des moyennes. Une fois de plus on peut observer l'augmentation de la concentration de CO<sub>2</sub> sur les trente années de l'étude.

► Etude complémentaire du plan des axes 3 et 4 :

Parmi les plans formés par les axes 3 et 4 des l'ACP interannuelles, c'est celui de Dakar, qui avec près de 20% de l'inertie en recueille le plus. Il mérite d'être examiné.



On voit sur ce plan (3-4) que l'analyse s'est focalisée sur quelques années pour lesquelles les relations entre tcc et t2 sont différentes de ce que l'on a constaté précédemment pour Dakar. Seules quelques années qui ont une bonne qualité de représentation semblent permettre d'expliquer la variance de ce plan : il s'agit des années 1991, 1982, 1995. Remarquons que ces années étaient mal représentées sur le plan (1-2), ceci indique qu'elles forment un sous groupe caractéristique dans le nuage. L'axe 3 oppose des années ayant une couverture nuageuse plus élevée et une température plus faible que la moyenne à des années pour lesquelles l'opposition entre ces deux variables est inverse.

Les autres variables (lsp, cp, ssr et CO<sub>2</sub>), ne semblent pas assez bien représentées sur le cercle des corrélations pour que nous cherchions à en tirer un enseignement suffisamment fiable (même si on peut éventuellement, en s'appuyant sur l'axe 4, y retrouver un indice de l'augmentation du CO<sub>2</sub>).