# Chapter 5. Data analysis or ordination, reflection illustrated with a floristic table

by Guy BOUXIN°

## Contents

Abbreviations: CA = correspondence analysis; MFA = multiple factor analysis (followed by CA, NSCA); NMDS = "non-metric multidimensional scaling"; NSCA = non-symmetrical correspondence analysis; PCA = principal component analysis. PCoA = Principal coordinate analysis.

## Introduction

       This chapter can be read alone. It explains, step by step, our final choice of the techniques best suited to an in-depth analysis of a vegetation relevé table.

---

° rue des Sorbiers 33 à B. 5101 Erpent    mail : guy.bouxin@proximus.be

In vegetation relevé tables, the lines correspond to species or environmental variables and the columns to relevés (individuals). Each cell contains either the presence or the score of a species or an environmental parameter. First of all, it is necessary to put forward a set of characteristics of vegetation data : a high diversity of possible data types (presences, abundances, cover, number of individuals, …), a high variability between sites, a mixture of common, less common and uncommon species, a mixture of dispersion patterns, and especially a high proportion of empty cells (usually much more than 50 %), each corresponding to the absence of a species and represented by a zero. All these characteristics make the data processing particularly critical.

That is the reason why the concept of data analysis (*sensu* BOUROCHE & SAPORTA, 1980) is favoured. It includes a set of very different methods of statistical analysis, that are adapted to tables. One distinguishes two processes: the automatic classification of items and variables, and the factor analysis, using the properties of the Euclidean vector space to describe the individuals and the variables (FOUCART, 1982). Only factor analysis is concerned with this paper. According to PALM (1993), the methods of factor analysis are basic tools of the analysis of the data tables that do not present any special structure. Correspondence analysis, created to treat important linguistic data, was the first link in a large set of statistical techniques that proposed to treat the statistical tables independently of any mathematical or probabilistic hypothesis on the structure of the data (HUSSON *et al*., 2009). They have essentially a descriptive goal, while condensing a table information, often composed of a high number of lines and of columns, in some graphic representations with two dimensions, accompanied by tables summing up the numerical values in order to help the user in the interpretation. The data analysis is thus multidimensional and descriptive. It produces synthetic variables, that represent several others, themselves linked together.

These techniques seem *a priori* well adapted to the analysis of vegetation relevé tables. They have been used since several tens of years in the study of the vegetation under the term "ordination". The ordination *sensu stricto* tries to put the vegetation in relation with one or more environmental gradients or axes (GOODALL, 1954). Starting with the concept that the vegetation is a continuum, it consists in placing each relevé in relation with one or more axes in such manner that their relative position brings a maximum information on its composition. The synthetic process is considered as a tool to put the floristic data in relation with the environmental factors (or other) that are likely to affect the vegetation composition (GREIG-SMITH, 1964). The ordination, following WHITTAKER (1973), is generally considered as a technique of indirect gradient analysis; to this subject the author wrote: "The gradient analysis studies the vegetation in terms of gradients (or more or less continuous variations) of environment, of specific populations and of communities. Sample arrangement according to environment gradients as base of the analysis of gradient is called ordination. The ordination is direct (location of the samples alongside one or more gradients of the environment) or indirect (the samples are arranged, by elaborated calculations, alongside axes that indicate changes in the communities). A gradient in the vegetation composition can be characterized by the quantity of "turnover" of an extremity to the other, which WHITTAKER (1967) called "beta diversity". An axis or a gradient with strong beta diversity has completely different species compositions (without any species in common) to its extremities. Inversely, an axis or gradient with weak beta diversity has species compositions rather similar to its extremities.

Factor analysis, in its statistical meaning, and ordination have thus very different conceptual bases since factor analysis does not emit any preconceived idea on the significance of the axes. The manner to exploit the results becomes strongly different. There is some confusion in the use of terms as ordination is sometimes used in place of factor analysis (WILDI, 2010 & 2013). For instance, according to GAUCH (1982) and many authors afterward, correspondence analysis (CA) revealed itself superior to principal component analysis (PCA) as an ordination technique. CA and PCA produce the arch effect and especially with PCA, the arch extremities rejoin sometimes.

Analysing the worldwide literature, von WEHRDEN et al. (2009) observed that CCA, DCA, CA, PCA or NMDS are commonly used techniques. If the arch effect is generally considered as an important default and leads to reject PCA for many vegetation data, numerous methodological discussions and the diversity of available methods has led to considerable uncertainty among researchers and students.

Current ordinations methods are reviewed and evaluated by AUSTIN (2012). Three main paradigms are discussed:
- the traditional phytosociology and the recognition and definition of associations, vegetation is assumed to be discontinuous,
- the multivariate analysis emphasizes the use of ordination and classification techniques, with three subordinate paradigms associated with several methods - CA, PCoA and NMDS,
- the emergent paradigm is the use of statistical regression methods.
-

CA is a method that uses a chi² distance as similarity measure, PCoA (or NMS) provides an ordination which is a Euclidean representation of plots based on any similarity measure seclected by the user and NMDS constructs an ordination where the distance between plots has maximum rank order agreement with the similarity measures between plots (AUSTIN, 2012).The differences between these paradigms are less than many plant ecologists assume. In another paper (AUSTIN 2013), the use of artificial data is enhanced in order to evaluate techniques; the data standardization is also an important and poorly understood step in the analysis.

One deducts that the gradient concept permanently dominates the research in vegetation ecology. When the gradient is temporal, the study becomes a succession study (AUSTIN, 1977; DEBUSSCHE et al., 1996).

One easily notes that, in most studies, only the eigenvalues of axes and the scores of the relevés or species on the ordination axes are taken into account in the result interpretation. The assistances to interpretation, such as the relative contributions of species or relevés are rarely used, probably because they are not furnished by most of the available programs. They are used by BONIN & ROUX (1978), BONIN et al. (1983) and BOUXIN (1986, 1987a & b, 1991, 1995 & 1999). The percentages of inertia associated to the axes are presented. They allow evaluating the number of axes to take into consideration. When the sum of percentages for the first two or three axes is important (without being possible to set up a norm), the multivariate analysis is considered as successful. Existing tests for the definition of the number of axes to be considered are generally not efficient and rarely used (see ESPINAR et al., 2002).

New techniques are regularly presented, for instance the correspondence analysis of raw data (GREENACRE, 2010) or the fuzzy coding in constrained ordinations (GREENACRE, 2013), always in relation with the gradient paradigm. However, some multivariate analyses are almost completely ignored, such as non-symmetrical factor analysis (GIMARET-CARPENTIER, 1998) or multiple factor analyses, that simultaneously analyses several tables corresponding to the same set of relevés (ESCOFIER & PAGES, 1984, 1986, 1989).

The objective of this work is to give an example of a scientific approach leading to the in-depth analysis of the structure of a relevé table, using multivariate analyses, apart from the common paradigms. For this purpose, an already analysed vegetation (BOUXIN 1975a, b and 1983) is taken as an example. Only the analysis of the floristic tables is discussed in this article.

The question is: using multivariate analyses, is the data analysis approach, a better tool for vegetation study than the ordination approach?

## Methods

### The relevé table and the nature of data

Among the three common categories of collected data (nominal, ordinal and metric data, see WILDI, 2010 & 2013): there are many variants of floristic data: simple data of presence-absence (0-1), the frequency in quadrats or the contacts with spins (counts), the density, the BRAUN-BLANQUET abondance-dominance data (ordinal scale, 0, +, 1 to 5) often associated with a sociability coefficient (1 to 5) and sometimes converted into the middle percentage of each class (1-2.5; 2-15.0; 3-37.5; 4-62.5; 5-87.5), ignoring the lower classes r and + (TALBOT *et al.* 2010), the van der MAAREL's scale, the cover (see FEHMI, 2010) for tree canopy or dense herbaceous layer (%), the basal area (metric data) and rarely the biomass.

Some data are measured, others are estimated and WILSON (2011) considers that there are severe problems with cover-abundance scales, such as that of BRAUN-BLANQUET. Estimation of abundance "by eye", by "informed guess", etc. has too often compromised results in vegetation science. Several mathematical operations are possible with presences, densities, cover, basal area but for other data as ordinal data, some mathematical operations are illicit. For instance, in a forest relevé with trees, shrubs, herbaceous plants, mosses, what does really mean a sum of abundances-dominances with such a set of species? In CA (or DCA), the sums of species scores in the lines or columns of the table enter the $\chi^2$ distance computation and the choice of the score is primordial (STEVENS *et al*., 2011).

The primary variables are also sometimes transformed into secondary variables (ORLÓCI & KENKEL, 1985): square root, logarithmic, inverse sine, among many others which changes the form of the relation between the variables. Several transformations or standardizations are also possible (WILDI, 2010 & 2013).

A common way to change the original species scores in data analysis is to first transform the original tables into an ordinal table, by replacing the primary variables by an ordinal variable taking several entire scores or modalities (0, 1, 2, 3,...) whose number depends on the data variability. A second step leads to a simple disjunctive table, in which each variable is replaced by one or more new variables, each corresponding to one of the modalities. In each line, the presence of the modality is represented by the unity, and the absence by zero. A relevé table with only presence-absence data is the simplest form of a simple disjunctive table. Another transformation is the complete disjunctive table, in which the zeros of the ordinal variables are also considered as a modality (FOUCART, 1982).

A simple example is given with the abundances of a tree species along a stream, into a set of twenty relevés.
*Alnus glutinosa*   0+2142121220+5101022

The simple disjunctive table tables with the modalities +, 1, 2, 4 or 5 are
*Alnus glutinosa+*   01000000000010000000
*Alnus glutinosa*1   00010010100000101000
*Alnus glutinosa*2   00100101011000000011
*Alnus glutinosa*4   00001000000000000000
*Alnus glutinosa*5   00000000000001000000

For a complete disjunctive table, a line is added:
*Alnus glutinosa* absences 10000000000100010100
With this table, the sum on the columns is constant.

As many multivariate analyses are strongly influenced by the presence of rare species in a small number of relevés (see GAUTHIER, 2010 among many others), a common solution is to eliminate the rare species from an arbitrarily chosen threshold. This solution is difficult to accept and discussed in the results.

The relations between the floristic variables are not simple. They can rarely be expressed under an elegant linear or curvilinear and integrate themselves with difficulty in theoretical models. The variability of the biological data is often very large and the relations between all these variables are often very complex, multifactorial.

All these data must be integrated in a table and an adequacy between the data and the chosen mathematical model is necessary.

## Factor analyses and ordination techniques

The computation steps of principal component analysis, correspondence analysis, non-symmetrical correspondence analysis and multiple factor analysis are similar and must recalled:

1. Computation of a transformed table, with the same number of lines and columns as the original table.

2. Computation of a square covariance (*sensu lato*) matrix, whose dimension is equal to the number of species or relevés.

3. Computation of the eigenvalues and eigenvectors.

4. Computation of the coordinates of the species and the relevés on the main axes and of the contributions and associated probabilities.


The transformed tables contain:

- For principal component analysis, the data in each line are summed and centred with a unit variance. The sum in each line is thus equal to zero. The lines are summed but not the columns. Rare species have relative high scores and the empty cells receive a score that is not the same in the different lines. With the correlation coefficient computation, a great variety of variables can be handled. Theoretically, the best conditions arrive with monotonous variables that give high positive or negative coefficients but high coefficients are also observed with other variable distributions. For instance, the co-occurrence of rare species in one relevé generates high correlation coefficients.

- For correspondence analysis, both lines and columns are summed and then weighted by the margin totals. One obtains the transformed variables while multiplying the original data by the total of all the data and while dividing them by the product of the marginal sums on the corresponding lines and columns; to this quotient, one subtracts the value 1. The numerical scores of the cells of the transformed table vary therefore between -1 (absent species) and a positive value, that is so much higher than the marginal sums on corresponding lines and columns are not very important (rare species in a not very diversified relevé, for instance). The rare species are thus often overestimated. The covariance matrix is then computed. This technique is adapted to presence or frequency data (contingency tables). It produces the same eigenvalues with the transposed relevé table. The use of continuous variables is unsuitable. Considering the algorithm, it is clear that summing several kinds of variable on the columns is not recommended; making a sum concerning presences and several measurements of abundance (cover, abundance-dominance) is a non-sense. However, CA and derived techniques (such as detrended correspondence analysis), are commonly used with abundance-dominance tables.

- For non-symmetrical correspondence analysis, the transformed variables is obtained while dividing the original data by the corresponding column total and while subtracting from this result the quotient between the sum on the corresponding line and the general total of the table. In the transformed table, the sums on the columns equal zero and the totals on the lines are so higher than the species are more frequent or abundant. As for PCA, the empty cells receive a score that is not the same in the different lines. The covariance matrix is then computed. The restriction about the data type is similar as for CA. With such an algorithm, the weigth of the species is all the more important as the sum on the lines is high. The weigth of rare species is not overestimated as in CA. The computation of the eigenvalues and eigenvectors is similar for the three analyses.


The multivariate analyses are most often used as descriptive techniques. The statistical inference is rarely taken into account. Several techniques facilitate the passage to the statistical inference: these are the randomization, the bootstrap and MONTE CARLO methods (MANLY, 1997). Some techniques were developed in order to evaluate the number of significant axes (KNOX & PEET, 1989; REED *et al.*, 1993; PILLAR, 1999).


I developed an automatic algorithm with randomization (BOUXIN, 1999) applied to principal component and to correspondence analyses; it refers to the relative contributions of the axes, columns and species. The test is easily constructed on relative contributions as there are always positive; they are important components in the interpretation of the results as they directly enter in the calculation of the species or relevés scores on axes. At each permutation step, all

the lines of the table (here the species) are permutated separately. It is a long procedure, as a PCA, a CA or a NSCA is calculated for each permutated table; the relative contributions of the axes, the species and the relevés are compared with the same parameters of the original table. If the proportion of relative contributions of the permutated tables superior to the corresponding relative contribution of the original table is lower than 0.05, then the observed relative contribution is significant. The number of permutations is fixed with the number of species (for PCA) or the number of relevés (for CA or NSCA); it varies between 1000 for large tables and 100,000 (the ideal number) for small tables.

The multiple factor analysis is a method of simultaneous analysis of several tables. The MFA is therefore a generalization of the techniques of coupling tables. It is one of the most practical methods among those of simultaneous ordination of *K* tables (DRAY *et al.*, 2007). With this technique, the *K* tables must have the same set of individuals (relevés) in common, each of them corresponding to a different group of variables (species or environmental parameters). The MFA makes uniform the incidence of each table in the simultaneous analysis; it consists in putting the first eigenvalues (of PCA, for instance) of each table at the same scale. Three main steps are necessary to perform a multiple factor analysis. Firstly, a PCA, a CA or a NSCA of the complete relevés table (obtained in putting together all the tables corresponding to the same set of relevés) is calculated. Secondly, the same analysis is applied to every elementary table. Thirdly, a new analysis is applied on the complete table, but in introducing the first eigenvalues of the separate tables as a weighting parameter (exactly the square root of that first eigenvalue). It implies that no column in the separate tables be empty. That condition makes MFA impossible with some tables.

This technique was rarely used (see BOUXIN, 2011 & 2013; DEVINEAU, 2001; HAURY & BAGLINIÈRE, 1990). This kind of analysis is adapted to a lot of vegetation studies, for instance for pluristratal vegetations. The common softwares produce the coordinates of the species and the relevés along the first axes of ordination, but without any assistance to the interpretation. But the computation of relative contributions of relevés or species for each axe and the probability linked to that contribution is useful. So, relevés and species may be ordered in function of their relative contributions, that have priority in the interpretation. Any test is calculated for multiple factor analyses as the permutation tests can be applied to the separate tables. My software also furnishes the transformed table, the distribution of the transformed scores, the covariance matrix, the distribution of the covariance coefficients into 20 classes and the eigenvectors.

The principal coordinate analysis (PCoA) is a generalization of principal component analysis. It was proposed by GOWER (1966). In contrast to the PCA, that calculates the distance between points of a space (relevés or species) using the Euclidean distance, PCoA allows use of other distance measures likely to better reflect the difference between relevés, for example. The distance between two statements may be the number of species present in a relevé or another but not in the two (PIELOU, 1977), or the coefficient of BRAY & CURTIS or even a chord distance.

The non-metric multidimensional scaling is a novel technique, proposed by KRUSKAL (1964a, b). It is explained by ORLòCI & KENKEL (1985) and McCUNE & MEFFORD (1999) but it calls a few remarks:
- it also starts with the calculation of a distance matrix; as for PCoA, the choice of the distance coefficient is an important step (ROOT & NELSON, 2011).

- the first dimension in a two-dimensional space is not the same as the first dimension in three dimensions or a single space. For a given number of dimensions, the solution for a particular axis is unique.

- the axis numbers are arbitrary, which makes the percentage of variance associated with an axis does not necessarily form a descending series with the growing number of axes.

- the appropriate number of axes should be determined by the evolution of final stress depending on the number of axes. You must choose a number of axes beyond which stress reduction is small.

The solution is influenced by the initial order of relevés in the original table. If you choose the "random" option, several analyses provide different results. There are many variants of this technique (PRENTICE, 1977 and 1980; MINCHIN, 1987). It is thus a completely different technique from CA and PCA as there is no eigenvector computation.

## Data

The scientific approach is illustrated with a floristic table, from the Akagera national park; it comprises 238 floristic variables and 80 relevés. The savannah data and main results were presented in BOUXIN (1975a, b and 1983). Eighty relevés were placed inside a transect of 2400 x 200 m between Ihéma lake and Kionja hill. The relevés of 25 x 10 m were located along forty 200 m long traverses. Along each traverse, two relevés were randomly selected among eight possibilities. The measure of abundance of species changes with growth and life form. Whenever possible, the circumference at breast height was measured, and the data transformed into basal area. The abundance of the adult trees of each species in a relevé is given by the amount of the basal areas of all the measurable individuals.

 For all the young trees and shrubs less than 1.5 m high, the abundance is expressed by the number of individuals. The abundance of every tree and shrub is given by two variables: the first one for adult individuals and the second one for small ones. As many herbaceous species have a tussock form, the abundance was assessed in each relevé with 125 points regularly spaced along five lines running parallell to the short side of the rectangle at equal distance; every sampling point is materialized by the point of a pin; for uncommon species found between the lines of points, the presence is noted.

In the relevé tables, the floristic variables of trees with a measurable basal area are followed by A and the youngest by J.

To make the interpretation of the various results more explicit, we carried out a classification of the relevés with a search for characteristic species, based on the multiple factorial analysis of a 113 file, following the techniques described in the chapter on the relevé classification.

While leaving the lake Ihéma to the hill Kionja, the following communities have been encountered:

- Wooded and dense, narrow savannah, with large *Acacia polyacantha* subsp. *campylacantha* on humus soil, along the lake, relevés A, B and C.

- Wooded savannah with *Acacia hockii* and *Themeda triandra* >1, with *Solenostemon thyrsiflorus*, *Plectranthus caninus*, relevés 1, 2, 3, 4.5, 6, 7, 8, 28, 30, 32, 37, 42, 46, 48, 50, 51, 52, 54

- Xerophilous thickets with *Acacia brevispica*, *Teclea nobilis, Haplocoelum gallense, Grewia trichocarpa*, etc., relevés 9, 10, 11, 20, 29, 31, 38, 44.

- Shrub savannah with *Acacia brevispica, Boscia angustifolia, Markhamia obtusifoliaJ, Commelina erecta* subsp. *livingstonii, Blepharis maderaspatensis, Plectranthus pauciflorus*, relevés 12, 13, 14, 15, 16, 17, 18, 19, 21, 23, 24, 39.

- Wooded savannah with *Acacia gerrardii, Acacia siberana, Eustachys paspaloides* >1, *Indigofera colutea, Panicum atrosanguineum*, relevés 22, 25, 26, 27, 58, 60, 61, 63.

- Herbaceous savannah with *Heteropogon contortus, Microchloa kundtii, Sporobolus stapfianus, Vernonia miombicola*, relevés 33, 35, 36, 40, 41, 43, 45, 47.

- Fairly dense wooded savannah with *Acacia senegal* (mainly at the foot of the hill), *Pennisetum polystachion, Achyranthes aspera, Panicum maximum*>1, *Albuca kirkii*, relevés 34, 59, 62, 64, 65, 66, 67, 68, 69, 70.

- Herbaceous savannah with *Hyparrhenia dissoluta, Hibiscus cannabinus, Commelina purpurea*, relevés 49, 53, 55, 56, 57.

- an open grassy or wooded savannah on rocky and shallow soil, dominated by the grass *Loudetia simplex*, relevés D, E, F, G, H, I, K, L.


The dispersion pattern has also been described for each species and environmental or pedological parameter (BOUXIN, 1983).


# Results interpretation

The interpretation of the results is based on:

- The distribution of correlation coefficients between floristic variables, covariances or distances between surveys. A very strong asymmetry of these parameters has a strong impact on the calculation of the eigenvalues and means that the statistical technique is not consistent with the data in the statement table.

- The importance of the eigenvalues arranged in descending order.

- The coordinates of the species and relevés or relevés alone on the axes of the multivariate analysis, which leads to the representation of floristic and mesological variables in plans. Only the first two axes are illustrated. The complete results are given in the annex.

- The relative contributions of species or readings on the axes.

- The probabilities associated with the eigenvalues or the relative contributions of floristic variables or surveys, which reveal the characteristic species or not, the simple companion species, in the phytosociological sense of the term.


In   analyses, the programs provide the coordinates of relevés on axes and the eigenvalues. In a non-metric multi-dimension analysis, only the coordinates of relevés are provided.


# Analysis of the original floristic table

The original table presents 238 species. There is a very high proportion of absences as 81,017% as the table cells are empty. A small number of relevés is characterized by the cooccurrence of uncommon species. With the original data, PCA

is the only possible technique. CA and the derived techniques are absolutely not suited to these data as well as any technique based on a calculation of distance or similarity.

In the PCA transformed table, the reduced centred scores corresponding to the species absence vary between -2 and scores that are close to zero. The correlation matrix shows a great asymetry (figure 1) in the distribution of the coefficients, many more on the negative side of the distribution. There is a very high proportion (82.250%) of coefficients between -0.168 and 0.123. The first five eigenvalues rapidly decrease and their sum corresponds to 26.419 % of the matrix trace and then, their decrease is slow and regular. The first twelve ones are significative (Figure 1).
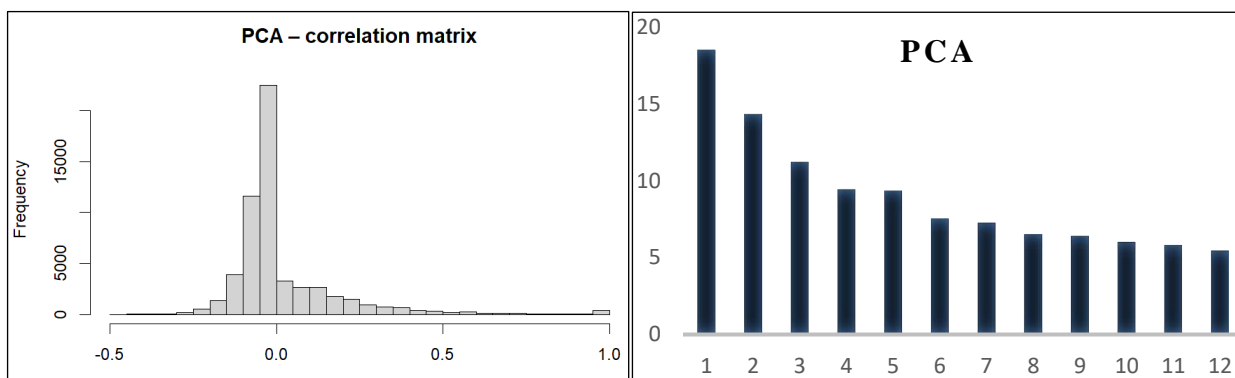


Figure 1. Distribution of correlation coefficints into 40 classes and the first twelve eigenvalues.

The results of the PCA are illustrated in Table 1 with the first three axes, with the first eight floristic variables and the first eight readings. Appear: the coordinates of the variables and readings on the axes, the eigenvalues or relative contributions and the probabilities associated with the eigenvalues or relative contributions. For all our personal programs, the results are presented in the same way.

| | coord1 | Cr% 1 | P 1 | coord2 | Cr% 2 | P 2 | coord3 | Cr% 3 | P 3 |
|---|---|---|---|---|---|---|---|---|---|
| Vp.Vp% and P | 18.51314944 | 7.811455461 | 0 | 14.32723806 | 6.045248126 | 0 | 11.23227507 | 4.73935657 | 0 |
| *Acacia senegal* A | -0.19571259 | 0.038303417 | 19.2 | 0.10248238 | 0.010502638 | 51.5 | -0.02378593 | 0.000565771 | 89 |
| *Acacia senegal* J | -0.18995608 | 0.036083311 | 15.8 | 0.021164268 | 0.000447926 | 86.7 | -0.03629284 | 0.00131717 | 81 |
| *Acacia hockii* A | -0.20201625 | 0.040810566 | 19.3 | 0.122969538 | 0.015121507 | 45.2 | 0.077731764 | 0.006042227 | 64 |
| *Acacia hockii* J | -0.23943011 | 0.057326777 | 12.7 | 0.0874593 | 0.007649129 | 55 | 0.100553848 | 0.010111076 | 49 |
| *Acacia gerrardii* A | -0.03428807 | 0.001175672 | 74.2 | 0.048950513 | 0.002396153 | 68.5 | 0.090421979 | 0.008176134 | 49 |
| *Acacia gerrardii* J | 0.048073944 | 0.002311104 | 62.9 | 0.075515768 | 0.005702631 | 58.9 | 0.126620787 | 0.016032824 | 39 |
| *Acacia polyacantha* A | -0.04223656 | 0.001783927 | 67 | -0.04340878 | 0.001884322 | 72.7 | -0.19747609 | 0.038996805 | 17 |
| *Acacia sieberana* A | -0.00872359 | 7.61E-05 | 93.1 | 0.032843132 | 0.001078671 | 79.7 | -0.04418022 | 0.001951892 | 74 |
| ∘∘∘∘∘∘∘∘ | coord1 | Cr% 1 | P 1 | coord2 | Cr% 2 | P 2 | coord3 | Cr% 3 | P 3 |
| A | -1.29898443 | 0.003645468 | 54.3 | -1.04010962 | 0.002337242 | 66.1 | -3.9808609 | 0.034237293 | 16 |
| B | -0.74014334 | 0.002097109 | 67.6 | -0.70639926 | 0.001910248 | 68.8 | -3.57801123 | 0.049008647 | 9.5 |
| X1 | -3.0228084 | 0.023309792 | 17.7 | 1.177807502 | 0.00353888 | 60.5 | 0.317372797 | 0.000256955 | 90 |
| X2 | -3.48515046 | 0.026031501 | 13.8 | 0.956005392 | 0.00195874 | 68.9 | -1.22671268 | 0.00322509 | 63 |
| X3 | -2.43930017 | 0.038617455 | 10.7 | 1.469432141 | 0.014013681 | 30.6 | 1.685308675 | 0.018433681 | 26 |
| C | -0.97157414 | 0.003333479 | 57.9 | -1.17653786 | 0.004888298 | 51.5 | -3.5116866 | 0.043548919 | 11 |
| X4 | -1.39749233 | 0.007006308 | 41.6 | 2.332630088 | 0.019520103 | 22.9 | 1.971395356 | 0.013942409 | 32 |
| X5 | -2.41115364 | 0.015005441 | 25.8 | 1.836433007 | 0.008704607 | 41.4 | 0.849086922 | 0.001860815 | 71 |

Table 1. Partial results of the principal component analysis of the raw Akagera file, for the first eight species and the first eight relevés. Coord: coordinate on the axes or eigenvalues, Vp, Vp%: % inertia of the eigenvalues or relative contribution on the axes of the species or relevés. P: probabilities associated with eigenvalues or relative contributions.

| **Species** | **coord1** | **Cr% 1** | **P 1** | **Relevés** | **coord1** | **Cr% 1** | **P 1** |
|---|---|---|---|---|---|---|---|
| *Acacia brevispica* J | 0.762585664 | 0.581536895 | 0 | X9 | 24.53609198 | 0.577846353 | 0.7 |
| *Jasminum schimperi* | 0.751526263 | 0.564791724 | 0 | X11 | 13.97306655 | 0.402345314 | 0.7 |
| *Panicum brevifolium* | 0.729652577 | 0.532392883 | 0 | X10 | 10.60380036 | 0.25775096 | 1.3 |
| *Tricalysia ruandensis* J | 0.721874407 | 0.52110266 | 0.4 | X18 | 5.466898187 | 0.154947798 | 1.3 |
| *Pupalia lappacea* | 0.699407098 | 0.489170288 | 0 | X14 | 6.437972862 | 0.141537766 | 2.1 |
| *Cyanotis lanata* | 0.694475938 | 0.482296829 | 0 | X13 | 7.266526484 | 0.14098438 | 2.2 |
| *Psilotrichum elliotii* | 0.676476692 | 0.457620715 | 1.4 | X20 | 4.778346309 | 0.105484267 | 3.7 |
| *Vernonia brachycalyx* | 0.674331474 | 0.454722936 | 0 | X66 | -3.3395508 | 0.095853114 | 2.8 |
| *Canthium lactescens* A | 0.670569496 | 0.449663449 | 2.4 | X15 | 5.134898167 | 0.088941845 | 3.2 |
| *Allophylus rubifolius* J | 0.66609473 | 0.443682189 | 1.7 | X50 | -3.50971725 | 0.0884272 | 3.3 |
| **Species** | **coord2** | **Cr% 2** | **P 2** | **Relevés** | **coord2** | **Cr% 2** | **P 2** |
| *Pappea capensis* J | -0.82045405 | 0.673144852 | 0.2 | F | -26.8941049 | 0.699190758 | 0 |
| *Canthium vulgare* A | -0.79990289 | 0.639844633 | 0.3 | K | -8.78060644 | 0.132258256 | 3 |
| *Olea europaea var. africana* A | -0.79939663 | 0.639034965 | 0.4 | H | -7.68915933 | 0.129159165 | 3 |
| *Pappea capensis* A | -0.79939663 | 0.639034965 | 0.2 | E | -6.27699606 | 0.087575895 | 3 |
| *Allophyllus africanus* J | -0.79939663 | 0.639034965 | 0.1 | G | -6.38895057 | 0.068398766 | 8 |
| *Indigofera dendroides* | -0.79939663 | 0.639034965 | 0.2 | I | -4.09766745 | 0.055891745 | # |
| *Ziziphus mucronata* A | -0.79327214 | 0.629280688 | 0.3 | X9 | 6.940871032 | 0.046241248 | 8 |
| *Rubiaceae* | -0.73151367 | 0.535112247 | 0.1 | D | -2.74345102 | 0.032652776 | # |
| *Haplocoelum gallaense* A | -0.73105462 | 0.534440854 | 0.4 | X16 | 2.401468383 | 0.02612025 | # |
| *Loudetia simplex* | -0.70341041 | 0.494786206 | 0 | X11 | 3.409192179 | 0.023950714 | # |
| **Species** | **coord3** | **Cr% 3** | **P 3** | **Relevés** | **coord3** | **Cr% 3** | **P 3** |
| *Kalanchoe integra* | -0.62626438 | 0.392207079 | 0.6 | K | -13.5072317 | 0.312972711 | 1 |
| *Rhoicissus tridentata* | -0.60106497 | 0.361279101 | 0.4 | G | -11.6531707 | 0.227550308 | 1.2 |
| *Tricalysia ruandensis* A | -0.55830165 | 0.311700729 | 0.2 | F | 12.14007294 | 0.142470215 | 2.4 |
| *Parinari curatellifolia* A | -0.55220311 | 0.304928273 | 0.7 | I | -6.46402548 | 0.139085024 | 2.9 |
| *Dalbergia nitidula* J | -0.55025083 | 0.30277598 | 0.9 | X9 | -10.7225196 | 0.110355833 | 2.9 |
| *Allophylus rubifolius* J | -0.49667153 | 0.246682609 | 1.3 | H | -6.68669056 | 0.097676471 | 5.2 |
| *Commelina bracteosa* | -0.48241912 | 0.232728212 | 2.2 | X17 | 5.230411148 | 0.080876426 | 5.5 |
| *Parinari curatellifolia* J | -0.47558396 | 0.226180102 | 1.5 | X20 | 3.450101567 | 0.054991581 | 9.4 |
| *Lannea schimperi* A | -0.45343907 | 0.205606993 | 2.7 | X12 | 4.192425155 | 0.054787445 | 8.2 |
| *Ficus glumosa* A | -0.45343907 | 0.205606993 | 2.2 | B | -3.57801123 | 0.049008647 | 9.5 |
| **Species** | **coord 4** | **Cr% 4** | **P 4** | **Relevés** | **coord 4** | **Cr% 4** | **P 4** |
| *Bidens pilosa* | -0.56355414 | 0.317593273 | 0.1 | X7 | -10.7347085 | 0.234722375 | 0.9 |
| *Achyranthes aspera var. sicula* | -0.47454213 | 0.225190234 | 1.3 | A | -9.62302007 | 0.200063655 | 1.4 |
| *Panicum atrosanguineum* | 0.452427304 | 0.204690465 | 0.3 | X33 | 7.815465399 | 0.189212812 | 1.4 |
| *Heliotropium strigosum* | 0.438973219 | 0.192697487 | 0.5 | X36 | 5.62058449 | 0.169181598 | 1.6 |
| *Indigofera colutea* | 0.431514461 | 0.18620473 | 0.8 | B | -5.95482477 | 0.135746078 | 2.2 |
| *Hibiscus cannabinus* | 0.422760864 | 0.178726748 | 0.6 | X32 | 3.887219281 | 0.092894753 | 4.4 |
| *Justicia matammensis* | -0.41650734 | 0.173478361 | 2.5 | X44 | 2.615658056 | 0.076035514 | 5.7 |
| *Themeda triandra* | 0.410099591 | 0.168181675 | 1.2 | X47 | 4.422359603 | 0.07571376 | 6.1 |
| *Panicum maximum* | -0.39848937 | 0.15879378 | 0.7 | X48 | 3.460811498 | 0.075362557 | 7.2 |
| *Pavonia patens* | -0.39314735 | 0.154564839 | 3.2 | X9 | 8.315577021 | 0.066372225 | 7.3 |
| *Species* | coord1 | Cr% 1 | P 1 | Relevés | coord1 | Cr% 1 | P 1 |
| *Acacia brevispica* J | 0.762585664 | 0.581536895 | 0 | X9 | 24.53609198 | 0.577846353 | 0.7 |
| *Jasminum schimperi* | 0.751526263 | 0.564791724 | 0 | X11 | 13.97306655 | 0.402345314 | 0.7 |
| *Panicum brevifolium* | 0.729652577 | 0.532392883 | 0 | X10 | 10.60380036 | 0.25775096 | 1.3 |
| *Tricalysia ruandensis* J | 0.721874407 | 0.52110266 | 0.4 | X18 | 5.466898187 | 0.154947798 | 1.3 |
| *Pupalia lappacea* | 0.699407098 | 0.489170288 | 0 | X14 | 6.437972862 | 0.141537766 | 2.1 |
| *Cyanotis lanata* | 0.694475938 | 0.482296829 | 0 | X13 | 7.266526484 | 0.14098438 | 2.2 |
| *Psilotrichum elliotii* | 0.676476692 | 0.457620715 | 1.4 | X20 | 4.778346309 | 0.105484267 | 3.7 |

| *Vernonia brachycalyx* | 0.674331474 | 0.454722936 | 0 | X66 | -3.3395508 | 0.095853114 | 2.8 |
| *Canthium lactescens* A | 0.670569496 | 0.449663449 | 2.4 | X15 | 5.134898167 | 0.088941845 | 3.2 |
| *Allophylus rubifolius* J | 0.66609473 | 0.443682189 | 1.7 | X50 | -3.50971725 | 0.0884272 | 3.3 |

Table 2. First ten species and ten first relevés statements arranged in descending order of relative contributions. Axes 1 to 4.

On axis 1 (table 2 and figure 1), there are three thicket relevés that have the highest cos² and coordinates. They are associated with a large number of species located mainly in these three sites. Some species have their maximum abundance in thickets, others are exclusive (sometimes with only one presence) or preferential to such sites.

On axis 2, only one relevé, on rocky and steep ground, has a high cos²; many species are abundant or present in only this and similar neighboring relevés.

Axis 3 is redundant and axis 4 reveals several relevés (positive coordinates) close to Lake Ihéma with species characterized by their abundance (such as *Acacia polyacantha* subsp. *campylacantha*), presence or frequency and open savannah sites (negative scores). With the following axes, only local variations are extracted.
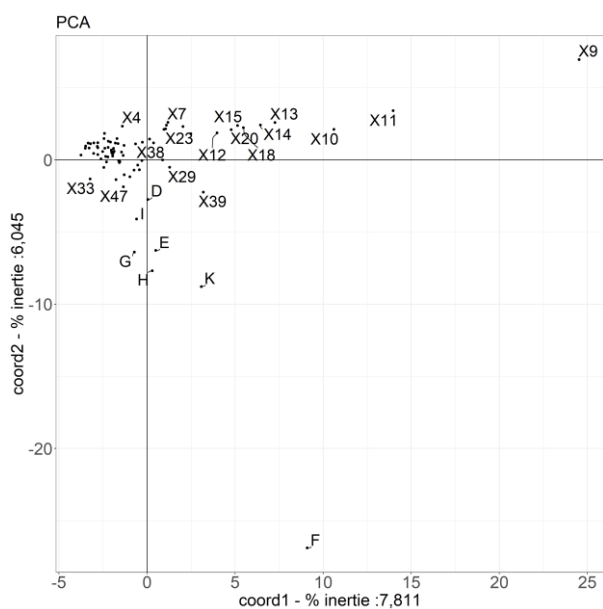


Figure 2. Principal component analysis of the raw file Akagera. Coordinates of relevés on axes 1 and 2.
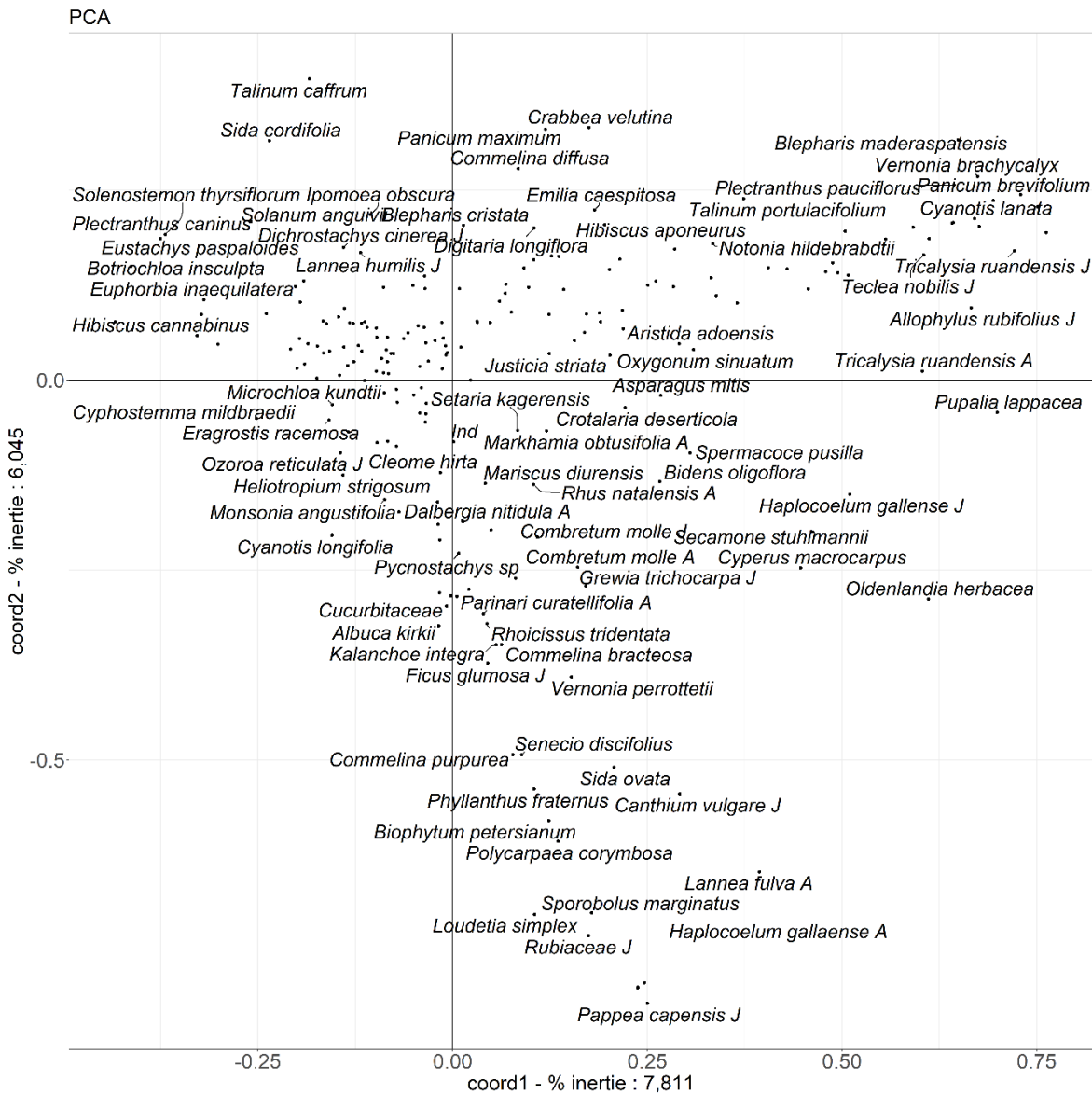
PCA



Figure 3. Principal component analysis of the raw file Akagera. Coordinates of species on axes 1 and 2.

This analysis mainly highlights localized species and is of little synthetic interest. The main plant groups and frequent species are fairly well overwritten by the analysis.

Is there a way to improve the results? One might be tempted to eliminate the rare species, as it is often the case, but this approach is hardly acceptable in phytosociological terms. First of all, should we eliminate species with 1, 2, 3 presences or more? It is facing an arbitrary choice. In addition, it often happens that some relevés are characterized by a set of localized species, so infrequent in the data table. Remove rare species returned to amputate some relevés of their characteristic species. In addition, in dense forest or savannah vegetation, rare species sometimes show a high biomass and have therefore of great importance in the functioning of the ecosystem. A second step should be the elimination of the species that have no structured dispersion pattern in the table (BOUXIN, 1983) but then, many common species are eliminated from the table and it is no longer recommended.

We now explore other adaptations: the data transformation and the choice of multivariate analyses that are compatible with such data.

## First data transformation

Before starting any new analysis, the next step is to transform the original table and to replace all raw abundances by an ordinal variable ranging from 1 to 5, it is recommended by PODANI (2006).

| Variables | The five coefficients and the class limits | | | | |
|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** |
| Basal areas in dm² | < 5 | 5-9.9 | 10-14.9 | 15-19.9 | ≥ 20 |
| Number of contacts with tufts | 1-25 | 26-50 | 51-75 | 76-100 | > 100 |
| Number of individuals | 1-20 | 21-40 | 41-60 | 61-80 | > 100 |

Table 3. Transformation of original abundance species scores into ordinal scores.

In the relevé table, there are several possible lines for each species; the name alone means that the coefficient is 1, for higher abundances, the names are followed by 2, 3, 4 or 5 (in addition to A or J for woody species). With an ordinal scale, several analyzes are possible: PCA, PCoA and NMDS. CA is not really recommended because this technique is originally adapted to presence or contingency tables. However, CA (or DCA) is commonly used with ordinal scales. NSCA is not adapted to any abundance data, since only abundant species are highlighted.

## Principal component analysis of ordinal data

The correlation coefficients are a little more spread out than in the previous analysis and the first twelve eigenvalues are significant (figure 4).
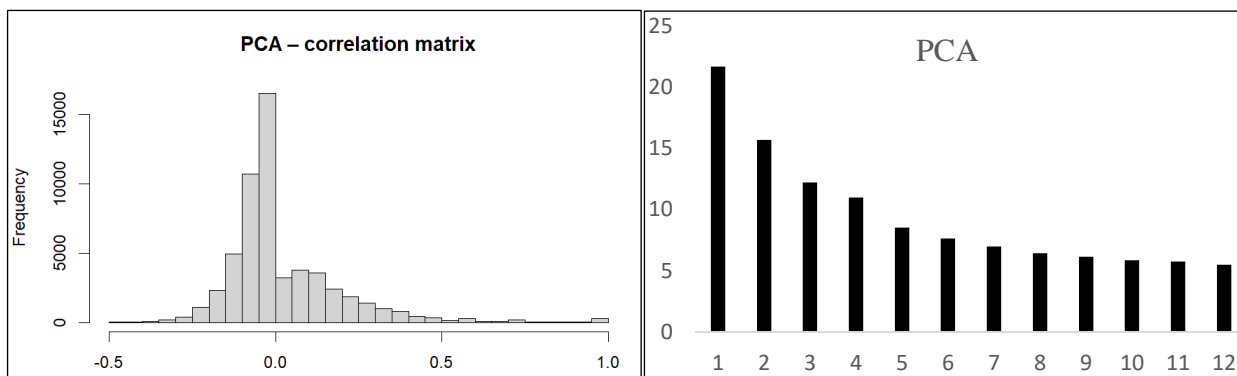


Figure 4. Distribution of the correlation coefficients into the 40 classes and the first twelve eigenvalues.

By arranging on axis 1, the species and relevés in descending order of relative contribution (table 4), we find, as in the analysis of the raw data, species and records of xerophilous thickets.

Axis 2 also corresponds to the species and sites of grassy savannahs on shallow soil with large blocks or rock slabs. Axis 3 is redundant.

Axis 4 mixes gallery forest and savannah surveys on surface soil.

| Species | coord1 | Cr%1 | P1 | Relevés | coord1 | Cr%1 | P1 |
|---|---|---|---|---|---|---|---|
| *Acacia brevispica* J | 0.851106579 | 0.72438241 | 0 | X9 | 20.1124359 | 0.4694869 | 0.48 |
| *Sanseviera cylindrica* | 0.778397524 | 0.60590271 | 0 | X11 | 13.7051696 | 0.41105965 | 0.54 |
| *Cyanotis lanata* | 0.737597943 | 0.54405072 | 0 | X10 | 12.279254 | 0.36155667 | 0.76 |
| *Pupalia lappacea* | 0.722954864 | 0.52266373 | 0 | X18 | 9.1552484 | 0.30348263 | 0.89 |
| *Panicum brevifolium* | 0.719676534 | 0.51793431 | 0 | X15 | 6.9404216 | 0.25915897 | 1.26 |
| *Plectranthus pauciflorus* | 0.718605418 | 0.51639375 | 0 | X14 | 8.38343057 | 0.25271443 | 1.21 |
| *Vernonia brachycalyx* | 0.710244468 | 0.5044472 | 0 | X20 | 8.17870591 | 0.22304754 | 1.55 |
| *Blepharis maderaspatensis* | 0.701139934 | 0.49159721 | 0 | X13 | 9.18001198 | 0.20831384 | 1.3 |
| *Jasminum schimperi = eminii* | 0.681206363 | 0.46404211 | 0.04 | X51 | -5.3110084 | 0.14171469 | 2.48 |
| *Teclea nobilis* A | 0.668866734 | 0.44738271 | 0.04 | X50 | -3.9845708 | 0.13806026 | 2.4 |
| **Species** | **coord2** | **Cr%2** | **P 2** | **Relevés** | **coord2** | **Cr%2** | **P2** |
| *Loudetia simplex* | -0.81872329 | 0.67030783 | 0 | K | -16.323878 | 0.39597085 | 0.18 |
| *Parinari curatellifolia* J | -0.63174381 | 0.39910024 | 0.19 | F | -17.390699 | 0.30019082 | 0.74 |
| *Dalbergia nitidula* J | -0.62866935 | 0.39522515 | 0.16 | G | -12.813812 | 0.2596882 | 0.87 |
| *Parinari curatellifolia* A | -0.62866935 | 0.39522515 | 0.12 | H | -10.307321 | 0.23760869 | 1.09 |
| *Kalanchoe integra* | -0.61784043 | 0.3817268 | 0.03 | I | -8.6931168 | 0.23121891 | 0.94 |
| *Lannea fulva* A | -0.59487191 | 0.35387259 | 0.16 | E | -7.2096235 | 0.1145078 | 3.4 |
| *Rhoicissus tridentata* | -0.5892173 | 0.34717703 | 0.55 | D | -4.2463931 | 0.07066902 | 6.83 |
| *Canthium vulgare* A | -0.58619733 | 0.34362732 | 0.98 | X19 | 3.91441116 | 0.06376302 | 7.8 |
| *Ficus glumosa* J | -0.56236411 | 0.3162534 | 0.48 | X16 | 3.60828686 | 0.05775999 | 9.89 |
| *Rubiaceae* | -0.56010456 | 0.31371711 | 0.69 | X21 | 3.66426113 | 0.05341233 | 10.4 |
| **Species** | **coord3** | **Cr%3** | **P3** | **Relevés** | **coord3** | **Cr%3** | **P3** |
| *Olea europaea var. africana* A | 0.582813921 | 0.33967207 | 0.44 | F | 18.0923783 | 0.3249037 | 0.35 |
| *Pappea capensis* A | 0.582813921 | 0.33967207 | 0.6 | C | -5.4961781 | 0.17834744 | 1.61 |
| *Allophyllus africanus* J | 0.582813921 | 0.33967207 | 0.6 | A | -9.5192302 | 0.17292535 | 1.84 |
| *Indigofera dendroides* | 0.582813921 | 0.33967207 | 0.48 | B | -6.3635959 | 0.16079222 | 1.7 |
| *Hyparrhenia dissoluta* | 0.58065047 | 0.33715497 | 0.01 | X36 | 4.40467955 | 0.11517091 | 3.4 |
| *Macrotyloma maraguense* | 0.578835396 | 0.33505042 | 0.01 | K | -8.4370581 | 0.1057788 | 3.67 |
| *Ziziphus mucronata* A | 0.578215356 | 0.334333 | 0.23 | X70 | -3.0494483 | 0.10467571 | 3.96 |
| *Haplocoelum gallense* J | 0.548415449 | 0.3007595 | 0.43 | I | -5.6393765 | 0.09730485 | 4.65 |
| *Canthium vulgare* A | 0.509390495 | 0.25947868 | 0.94 | X68 | -3.4573595 | 0.08909659 | 5.09 |
| *Biophytum petersianum* | 0.478255931 | 0.22872874 | 1.19 | X29 | 3.81848752 | 0.08884279 | 4.85 |
| **Species** | **coord4** | **Cr%4** | **P4** | **Relevés** | **coord4** | **Cr%4** | **P4** |
| *Achyranthes aspera* var. *sicula* | 0.537389617 | 0.2887876 | 0.05 | F | 15.2722335 | 0.2315093 | 0.66 |
| *Pappea capensis* A | 0.518576842 | 0.26892194 | 0.84 | A | 10.3689279 | 0.20517418 | 0.8 |
| *Indigofera dendroides* | 0.518576842 | 0.26892194 | 0.65 | B | 6.25909813 | 0.15555478 | 1.78 |
| *Olea europaea* var. *africana* A | 0.518576842 | 0.26892194 | 0.54 | X33 | -5.3492606 | 0.12875056 | 2.76 |
| *Allophyllus africanus* J | 0.518576842 | 0.26892194 | 0.61 | G | -8.9258375 | 0.12600665 | 2.66 |
| *Setaria barbata* | 0.48654066 | 0.23672181 | 0.54 | X1 | 6.59384762 | 0.10740686 | 3.53 |
| *Ziziphus mucronata* A | 0.481913345 | 0.23224047 | 0.88 | X2 | 6.02991072 | 0.0753779 | 6.39 |
| *Acacia polyacantha* A | 0.458508569 | 0.21023011 | 0.75 | C | 3.55923754 | 0.0747927 | 6.53 |
| *Phyllanthus fraternus* | 0.453356598 | 0.2055322 | 0.65 | X7 | 5.33248403 | 0.06782763 | 7.79 |
| *Sporobolus stapfianus* | -0.45198988 | 0.20429485 | 0.37 | X36 | -3.0499031 | 0.05521871 | 10.5 |

Table 4. First ten species and first ten relevés arranged in descending order of relative contributions. Axes 1 to 4.

Figure 5. Principal component analysis of the Akagera file. Coordinates of relevés on axes 1 and 2.

PCA



Figure 6. Principal component analysis of the Akagera file. Coordinates of relevés on axes 1 and 2..

This analysis presents no advantage over that carried out with raw data.
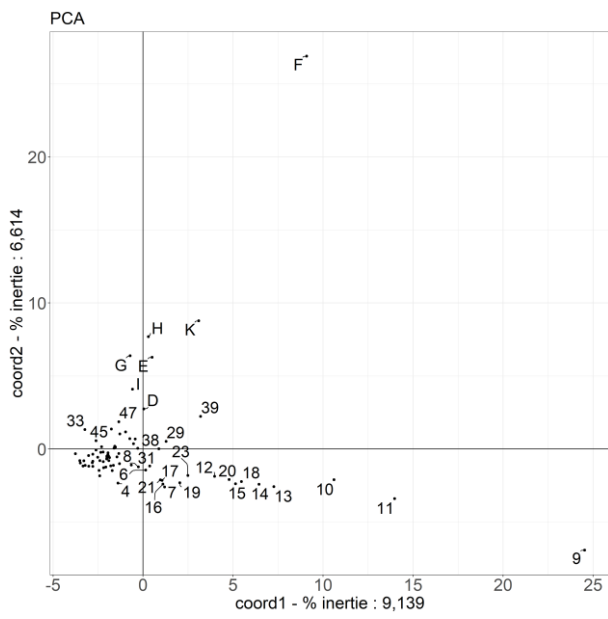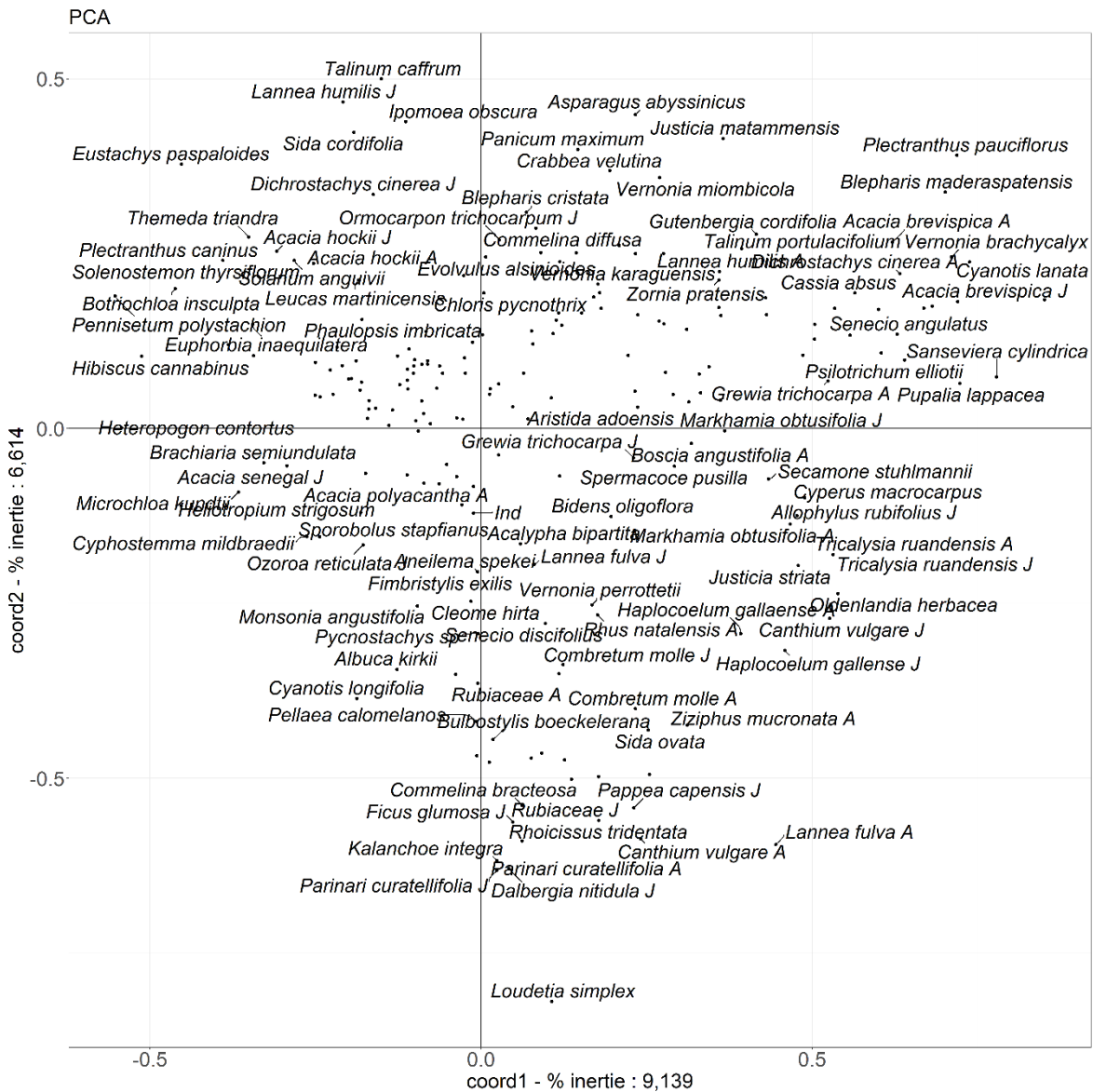
## Correspondence analysis of ordinal data

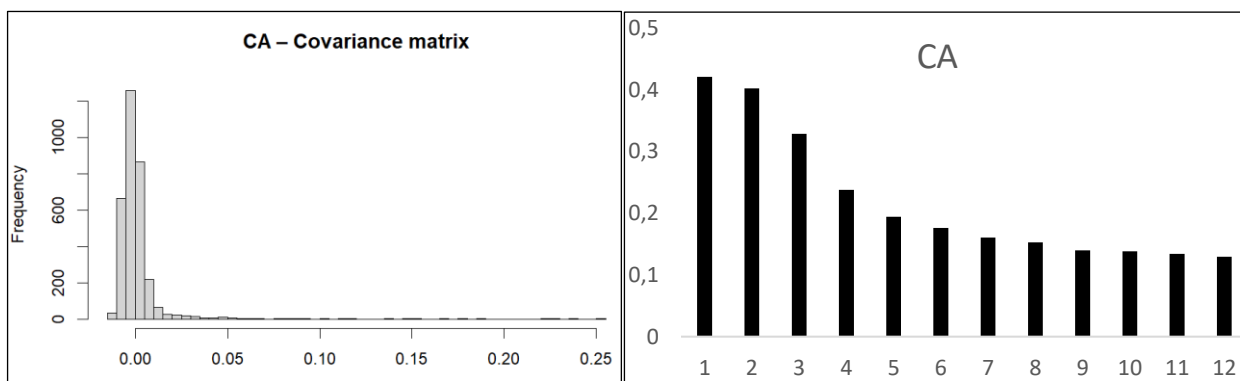The main results appear in table 5 and figures 7 to 9.

Figure 7. Distribution of covariance coefficients into 40 classes and the first twelve eigenvalues.

The distribution of covariance coefficients is highly asymmetric with a very low proportion of high values and is characteristic of the analysis of correspondences of floristic tables. The first twelve eigenvalues are significant.

| Species | coord1 | Cr%1 | P1 | Relevés | coord1 | Cr%1 | P1 |
|---|---|---|---|---|---|---|---|
| *Acacia polyacantha* A | -5.224924062 | 30.6225631 | 13.2 | A | -3.33475792 | 24.94825616 | 1.2 |
| *Acalypha bipartita* | -3.247728097 | 5.91576484 | 2.8 | C | -4.74184334 | 19.61695093 | 1.5 |
| *Setaria barbata* | -1.81964889 | 5.7775311 | 1.1 | B | -3.7523013 | 19.30310424 | 1 |
| *Loudetia simplex* | 2.007963615 | 4.27139433 | 6.4 | I | 2.932765302 | 9.647975487 | 2.8 |
| *Cissampelos mucronata* | -4.635691421 | 4.01753214 | 0.2 | K | 1.601864956 | 7.51551466 | 3.4 |
| *Dalbergia nitidula* A | 3.504596204 | 3.82696988 | 7.9 | H | 1.400051297 | 3.908831435 | 5.8 |
| *Bidens pilosa* | -3.993636823 | 2.98172412 | 1.5 | G | 0.971946969 | 3.237851276 | 5.4 |
| *Vigna macrorhyncha* | -3.664540235 | 2.51055252 | 2.9 | X2 | -0.84797273 | 2.150869574 | 6.8 |
| *Cynodon dactylon* | -2.331718164 | 2.37169708 | 25.6 | F | 0.695372396 | 2.109317718 | 6.9 |
| *Cissus petiolata* | -2.450311808 | 2.24493114 | 0.7 | E | 0.838636359 | 1.709310848 | 9.1 |
| **Species** | **coord2** | **Cr%2** | **P2** | **Relevés** | **coord2** | **Cr%2** | **P2** |
| *Acacia polyacantha* A | -3.213357061 | 12.1192487 | 30.3 | I | -4.02455289 | 19.01047653 | 1.6 |
| *Dalbergia nitidula* A | -5.064421216 | 8.36209048 | 3 | K | -2.48131084 | 18.86884964 | 1.5 |
| *Loudetia simplex* | -2.668488592 | 7.89341327 | 1 | A | -1.99086045 | 9.303992665 | 4.4 |
| *Acalypha bipartita* | -2.467076372 | 3.57184757 | 5.8 | C | -3.18815363 | 9.278798333 | 3 |
| *Dalbergia nitidula* J | -4.21176247 | 3.47003787 | 0.4 | G | -1.51054823 | 8.183085739 | 3.1 |
| *Parinari curatellifolia* A | -4.21176247 | 3.47003787 | 0.5 | B | -2.17931661 | 6.813159498 | 4.3 |
| *Parinari curatellifolia* J | -3.202087729 | 3.34288684 | 2.9 | H | -1.70766989 | 6.084754174 | 4.5 |
| *Ficus glumosa* A | -3.910986085 | 2.99212006 | 26.7 | F | -1.10292785 | 5.552358024 | 5 |
| *Kalanchoe integra* | -3.005429652 | 2.94488512 | 0.4 | E | -0.63080415 | 1.011903051 | 16.9 |
| *Cleome hirta* | -6.343409335 | 2.62379979 | 2.6 | X9 | -0.37680486 | 0.67583561 | 26.4 |
| **Species** | **coord3** | **Cr%3** | **P3** | **Relevés** | **coord3** | **Cr%3** | **P3** |
| *Acacia brevispica* J | 1.450030912 | 3.52279223 | 1.1 | X9 | 1.549086652 | 13.97614819 | 2 |
| *Grewia trichocarpa* A | 1.205310131 | 3.36131199 | 39.8 | X11 | 1.378952558 | 9.70938756 | 1.7 |
| *Acacia brevispica* A | 0.984211722 | 2.93679665 | 4.7 | F | 1.245907326 | 8.669274063 | 2.5 |
| *Dalbergia nitidula* A | -2.696796784 | 2.90121331 | 10.4 | I | -2.29733803 | 7.579412103 | 2.5 |
| *Sanseviera cylindrica* | 1.8108351 | 2.87782053 | 1.3 | X10 | 1.130579999 | 7.342573555 | 2.4 |
| *Cyanotis lanata* | 0.906602806 | 2.29517273 | 1 | X18 | 1.050440061 | 4.841929433 | 5.4 |
| *Ziziphus mucronata* A | 1.987386612 | 2.20584998 | 50.5 | X20 | 0.87484964 | 4.274428804 | 5.1 |
| *Pupalia lappacea* | 1.603733862 | 2.05200215 | 1 | X13 | 0.95149919 | 4.261692109 | 8.3 |
| *Panicum brevifolium* | 2.011106612 | 1.93613061 | 1.6 | X14 | 0.973056503 | 4.003733867 | 6.6 |
| *Vernonia brachycalyx* | 1.731597589 | 1.91380262 | 0.7 | X15 | 0.838858739 | 2.863261775 | 9.2 |
| **Species** | **coord4** | **Cr%4** | **P4** | **Relevés** | **coord4** | **Cr%4** | **P4** |
| *Ziziphus mucronata* A | -4.201910497 | 13.644314 | 9.3 | F | -2.85676322 | 63.06764315 | 0 |
| *Canthium vulgare* A | -4.231451883 | 7.90676582 | 5 | I | 2.311204726 | 10.61474121 | 1.5 |
| *Dalbergia nitidula* A | 3.195287686 | 5.63572777 | 5.6 | X9 | 0.756315756 | 4.609875827 | 7.5 |
| *Olea europaea var. africana* | -5.858909216 | 3.78960207 | 1.3 | E | -0.68977506 | 2.04851741 | 11.7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Pappea capensis* A | -5.858909216 | 3.78960207 | 0.7 | X11 | 0.537759824 | 2.04322762 | 11.5 |
| *Allophyllus africanus* J | -5.858909216 | 3.78960207 | 1.2 | H | -0.69148847 | 1.68919556 | 14.5 |
| *Indigofera dendroides* | -5.858909216 | 3.78960207 | 1.7 | K | 0.539001519 | 1.507432615 | 17.1 |
| *Haplocoelum gallense* J | -1.987249119 | 3.48782498 | 5.1 | D | -0.78600516 | 1.500491531 | 17.4 |
| *Rubiaceae* | -3.638538253 | 2.92310191 | 1.5 | X18 | 0.492686642 | 1.473887502 | 16.7 |
| *Sporobolus marginatus* | -3.636781253 | 2.92027954 | 1.6 | X13 | 0.43231407 | 1.217337002 | 21.9 |

Table 5. First ten species and first ten relevés arranged in descending order of relative contributions. Axes 1 to 4.

The analysis being dual, the species and relevés are often combined in the graphic representations. We leave them separated for easier reading.

The scores of the transformed table vary between -1 and 210.722 and the highest scores correspond to species with only one or only a few presences in very localized sites, that is to say near the lake or the rocky hill. The first four eigenvalues are respectively 8.394, 8.021, 6.556 and 4.738% of the matrix trace.

Axis 1 (table 5 and figures 8 and 9) separates species such as *Acacia polyacantha* A, *Acalypha bipartita*, … and the gallery forest relevés and species such as *Loudetia simplex* and *Dalbergia nitidula* and the relevés on shallow soil. Axis 2 brings them together, which gives a V-shaped image usual for this type of analysis. Axis 3 highlights some records of xerophilic groves. Axis 4 separates surface soil surveys into two small groups.
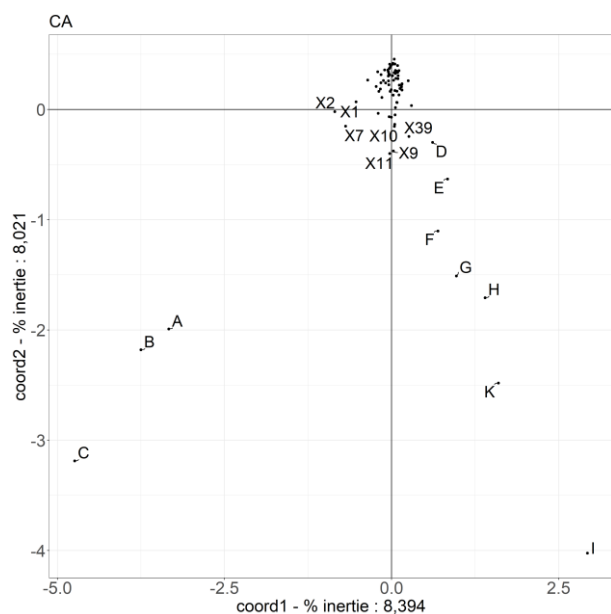


Figure 8. Correspondence analysis of file Akagera ordinal. Coordinates of relevés on axes 1 and 2.
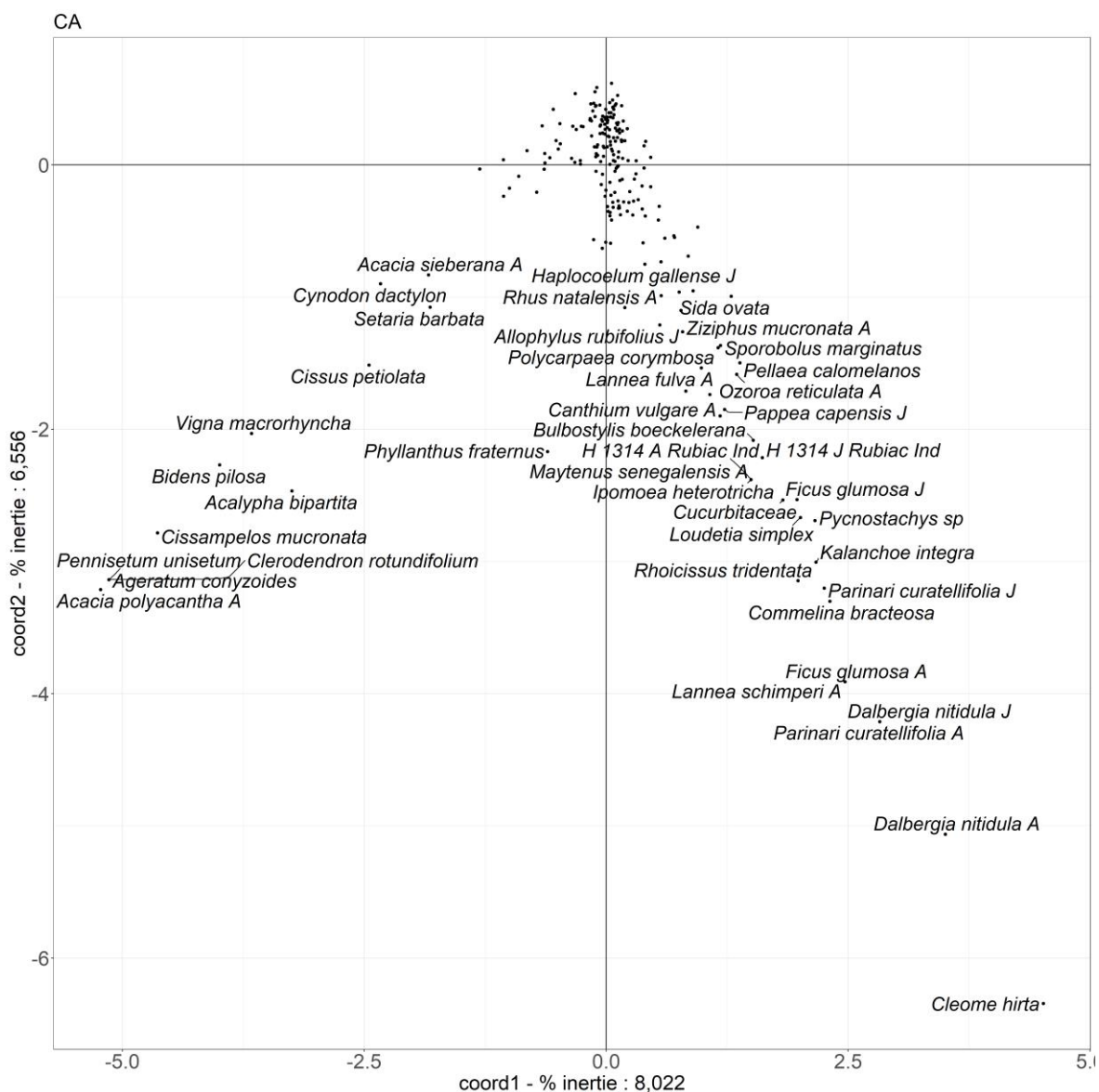
Figure 9. Correspondence analysis of file Akagera ordinal. Coordinates of relevés on axes 1 and 2.

The results of this analysis are symptomatic of correspondence analysis which favors infrequent species and localized relevés. It does not lend itself well to phytociological analysis.

## Principal coordinates analysis of ordinal data

The principal coordinates analysis starts with a distance matrix between relevés. We tested two of the most commonly used distances, namely the Euclidean distance and the chord distance. This analysis gives the eigenvalues of the distance matrix and the coordinates of the relevés. Existing software does not produce any interpretation aid such as relative contributions.

### With euclidean distance

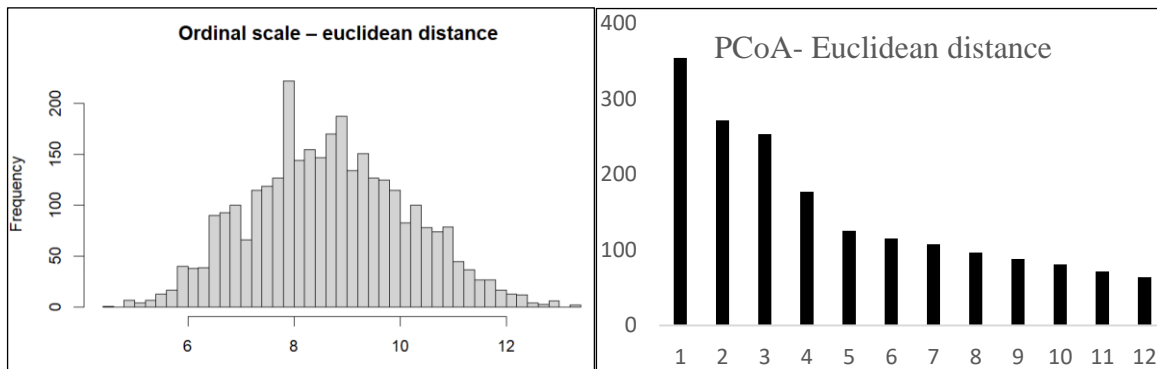The main results appear in table 6 and figures 10 and 11.



Figure 10. Distribution of euclidean distances into 40 classes and the first twelve eigenvalues.

|     | coord1 | coord2 | coord3 | coord4 | coord5 | coord6 |
|-----|--------|--------|--------|--------|--------|--------|
| Cr% | 11.57460606 | 8.866620745 | 8.2794365 | 5.784167651 | 4.089101845 | 3.761848914 |
| A | -1.78063105 | 2.918819119 | -3.811091664 | -5.093437819 | 0.432814545 | -0.933516571 |
| B | -2.062567781 | 3.523156589 | -3.610337866 | -5.033908274 | 0.630946863 | -1.452776756 |
| X1 | 0.802006664 | -1.180524779 | -0.971970099 | -1.87804676 | 0.648448127 | 0.08557696 |
| X2 | 0.549851061 | 0.521122959 | -1.15940105 | -2.824141891 | -0.031888058 | -0.259018851 |
| X3 | 1.532469015 | -1.427122073 | 0.877773059 | -1.009840159 | 0.605341079 | 0.563883549 |
| C | -1.685759575 | 3.861853614 | -2.744683887 | -3.425247529 | 0.298243959 | -2.210739172 |
| X4 | 1.306918333 | -2.535661797 | -1.201394271 | -0.095047924 | 2.096382465 | -2.647171088 |
| X5 | 1.434036756 | -2.57543379 | -1.80163057 | -0.677534676 | 2.1995954 | -1.715345853 |
| X6 | -0.089210838 | -1.825705361 | 0.316728262 | -0.674837153 | 1.931217997 | 0.775069856 |

Table 6. Partial results of the principal coordinates analysis of the Akagera file with ordinal data and the Euclidean distance, for the first nine relevés. Coord: coordinate on the axes.

The scores of the distance matrix are distributed following a Gaussian curve (figure 10). The first four eigenvalues are respectively worth 11.574, 8.867, 8.279 and 5.784% of the trace of the distance matrix with a sudden decrease for the fifth value. The calculation of distances is obviously clearly influenced by the high abundances of species and the relevés with abundant species that then are highlighted. However, the dispersion of the relevés is much greater than with a PCA or a CA and the results can lead to a richer interpretation (figure 11).
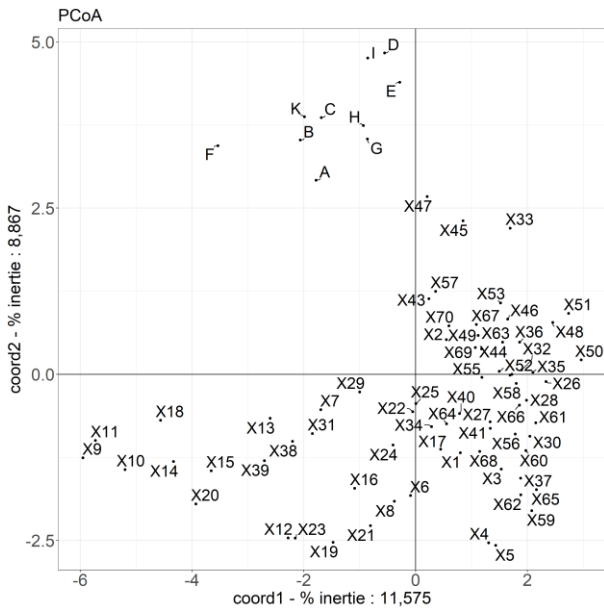
Figure 11. Principal coordinates analysis of the file Akagera with ordinal data and the euclidean distance, coordinates 1 and 2

Axis 1 contrasts relevés of thickets and surrounding areas with relevés of wooded savannah (with *Acacia hockii* or *Acacia senegal*) or herbaceous savannah.

Axis 2 isolates a mixed group of three gallery forest relevés and surface soil relevés.

Axis 3 separates the wooded savannah relevés with *Acacia polyacantha* (forest gallery) and *Acacia senegal* from other less densely wooded or grassy formations.

Axis 4 isolates a shrub savannah relevés from the three gallery forest relevés.

## *With chord distance*

PCoA is repeated with chord distance. The main results appear in table 7 and figures 12 and 13.
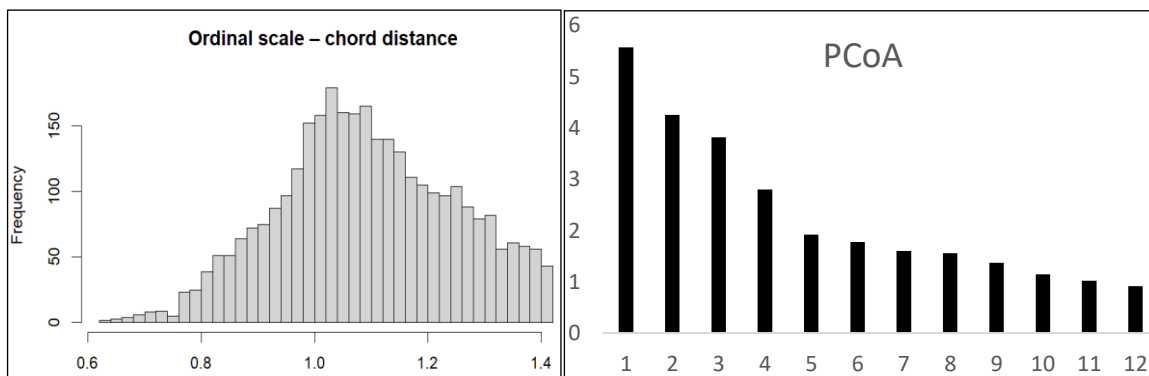


Figure 12. Distribution of chord distances into 40 classes and the first twelve eigenvalues.

|  | coord1 | coord2 | coord3 | coord4 | coord5 | coord6 |
|---|---|---|---|---|---|---|
| Cr% | 11.5881316 | 8.84664043 | 7.934641 | 5.82939817 | 3.99252643 | 3.69045722 |

| A | 0.45686316 | -0.1349234 | 0.35523253 | -0.628872 | 0.00958839 | 0.13484783 |
|---|---|---|---|---|---|---|
| B | 0.51584364 | -0.0854913 | 0.37731889 | -0.6784195 | 0.06241839 | 0.19013745 |
| X1 | -0.0807726 | -0.2130418 | -0.0654457 | -0.2372469 | 0.0019012 | -0.1299279 |
| X2 | 0.01675789 | -0.097986 | 0.0803001 | -0.3816422 | -0.0463316 | -0.0024833 |
| X3 | -0.2328349 | -0.0455137 | -0.1710282 | -0.1241955 | -0.0507581 | -0.136172 |
| C | 0.54795815 | 0.00147261 | 0.44093107 | -0.6745913 | 0.19965784 | 0.22304108 |
| X4 | -0.1230382 | -0.2634471 | -0.0706228 | -0.0503504 | 0.41513178 | -0.1403256 |
| X5 | -0.1052084 | -0.3080524 | -0.0144355 | -0.1372276 | 0.27766753 | -0.1713962 |

Table 7. Partial results of the principal coordinates analysis of the Akagera file with ordinal data and the chordbdistance, for the first eight relevés. Coord: coordinate on the axes.

The distribution of distances is a little asymmetrical and the first four eigenvalues stand out again.
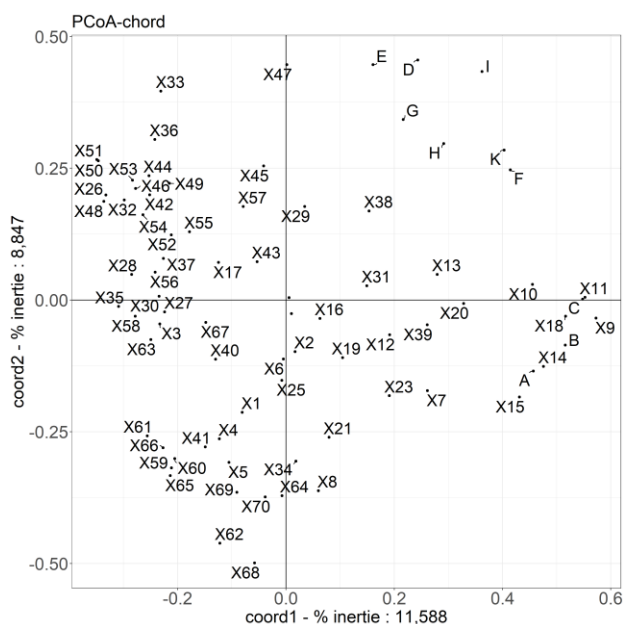


Figure 13. Principal coordinates analysis of the Akagera file with ordinal data et chord distance.

On axis 1, a group of gallery forest relevés, xerophilous thickets and relevés on surface soil is opposed to a group of wooded savannah relevés, mainly with *Acacia hockii*.

On axis 2, the relevés with *Acacia senegal* savannah are opposed to a mixture of several formations.
The other axes are quite redundant and do not provide any clear interpretation.

### Non-metric-multidimensional scaling with ordinal data

The calculation of the non-metric multidimensional scaling also starts with the same euclidean distance matrix and is done here with six axes. The software only gives the coordinates of relevés on the axes. As in PCoA, there is no help for interpretation.

The technique provides divergent results depending on whether the software starts with the coordinates of the relevés produced by another technique or by a random choice of relevés. The numerical results also depend on the number of axes retained; the coordinates of the relevés are not the same on the first or second axis, for example, if we choose to do the calculations with four or six axes. Many choices are then arbitrary.

The results are thus presented with six axes and random location of the relevés.

### With the euclidean distance

The results appear in table 8 and figure 14.

| | coord1 | coord2 | coord3 | coord4 | coord5 | coord6 |
|---|---|---|---|---|---|---|
| A | -1.0593515 | 3.12760937 | -4.7860808 | -5.4018757 | -1.0867468 | 0.10703587 |
| B | -1.9014348 | 3.79151737 | -4.2093278 | -5.3799414 | -0.3059905 | -1.1228015 |
| X1 | 1.07924878 | -1.090534 | -0.9320908 | -2.129708 | 0.93276033 | 0.24782789 |
| X3 | 1.26822522 | -1.3246027 | 0.59853233 | -0.8256296 | 0.57303183 | 0.50193654 |
| C | -1.1492332 | 3.74730939 | -2.7688435 | -3.0081023 | -0.66203 | -1.3189415 |
| X4 | 2.1618686 | -2.8867863 | -1.0103541 | -0.1777476 | 1.845727 | -2.2600538 |
| X5 | 2.49535727 | -2.7957172 | -1.6668633 | -0.794667 | 2.35721314 | -0.4174664 |

Table 8. Partial results of non-metric multidimensional scaling of the Akagera file with ordinal data and euclidean distance, for the first eight relevés. Coord: coordinates on the axes.
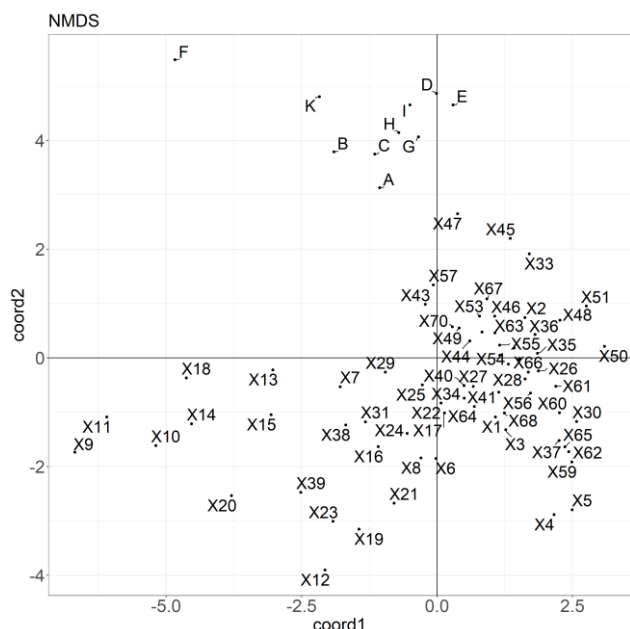


Figure 14. Non-metric multidimensional scaling of file Akagera with ordinal and euclidean distance.

Axis 1 contrasts relevés of xerophilous thickets with relevés of wooded savannah with *Acacia hockii* or *Acacia senegal*.

Axis 2 isolates the relevés on surface soil.

Axis 3 isolates a relevé of savannah with *Acacia hockii* and contrasts it with relevés of savannah with *Acacia polyacantha* and *Acacia senegal*.

Axis 4 isolates a relevé of *Acacia brevispica* savannah close to the thickets and mainly contrasts it with relevés of *Acacia polyacantha* savannah.

### With chord distance

The result appear in table 9 and in figure 15.

|    | coord1 | coord2 | coord3 | coord4 | coord5 | coord6 |
|----|--------|--------|--------|--------|--------|--------|
| A  | 0.50056657 | -0.3228975 | 0.36586493 | -0.6770017 | -0.0541512 | 0.09112633 |
| B  | 0.65689707 | -0.2601893 | 0.38063064 | -0.7125544 | -0.0790779 | 0.24022396 |
| X1 | -0.1042833 | -0.2181797 | -0.0945688 | -0.3113142 | 0.02639866 | -0.0661905 |
| X2 | -0.0835822 | -0.0958902 | 0.07457619 | -0.5844833 | -0.0216646 | 0.03071737 |
| X3 | -0.1837797 | -0.0474034 | -0.134469 | -0.1039278 | -0.0112858 | -0.0933054 |
| C  | 0.75462553 | -0.1900815 | 0.51534941 | -0.7669592 | 0.22949597 | 0.22437117 |
| X4 | -0.2255728 | -0.3004013 | -0.1779963 | -0.0169587 | 0.26948347 | -0.1308528 |
| X5 | -0.1592072 | -0.3291422 | -0.0794542 | -0.1391825 | 0.16499064 | -0.170016 |

Table 9. Partial results of non-metric multidimensional scaling of the Akagera file with ordinal data and chord distance, for the first eight relevés. Coord: coordinates on the axes.
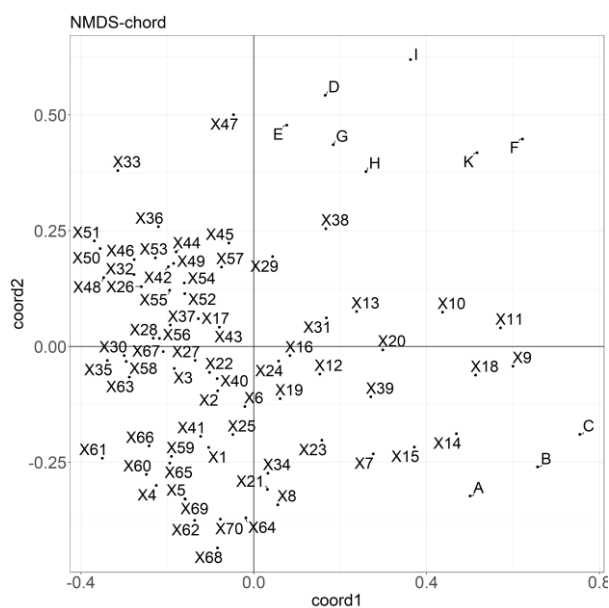


Figure 15. Non-metric multidimensional scaling of file Akagera with ordinal and chord distance. The results are once again quite comparable with those of the euclidean distance.

The results are once again quite comparable with those of the euclidean distance.

## Second data transformation

It therefore appears that the techniques adapted to analyse a vegetation relevé table with both abundances and simple presence data do not produce enough satisfactory results. A solution is to change the original table into a simple disjunctive table (several lines four each species, one for the abundances 1, the second for the abundance 2 and so on up to five) or in a complete disjunctive table (see above), which then allows the use of multivariate analyses such as CA or NSCA. Correspondence analyse with simple disjunctive tables are also named "multiple correspondence analysis". The simple disjunctive tables can also be used for PCoA. In the simple disjunctive table, 86.32 % of the cells are empty and in the complete disjunctive table, it is reduced to 56.81 %. A complete disjunctive table of vegetation data gives pride of place to absences, which are much more numerous than presences; this is not suitable for a vegetation study. Only simple disjunctive tables will be analyzed.

## Multiple correspondence analysis

CA of the simple disjunctive table has results that are comparable with those of the ordinal table.

In the transformed table, the scores vary between -1 and 379.67. The distribution of the covariance score is always highly asymmetric. The first four eigenvalues equal 6.125, 5.87, 4.513 and 4.064 % of the matrix trace and the next ones decrease regularly and slowly. The first axes are again highly influenced by unfrequent and located species (i.e adult *Acacia polyacantha* and thicket species) and no advantage is found.

## Non-symmetrical multiple correspondence analysis

With NSCA of the simple disjunctive table, one enters another result category. The distribution of covariance score is slightly asymetrical and regular. Only the first five eigenvalues are significant; that clearly indicates the synthetic character of that analysis (table 10 and figure 16).
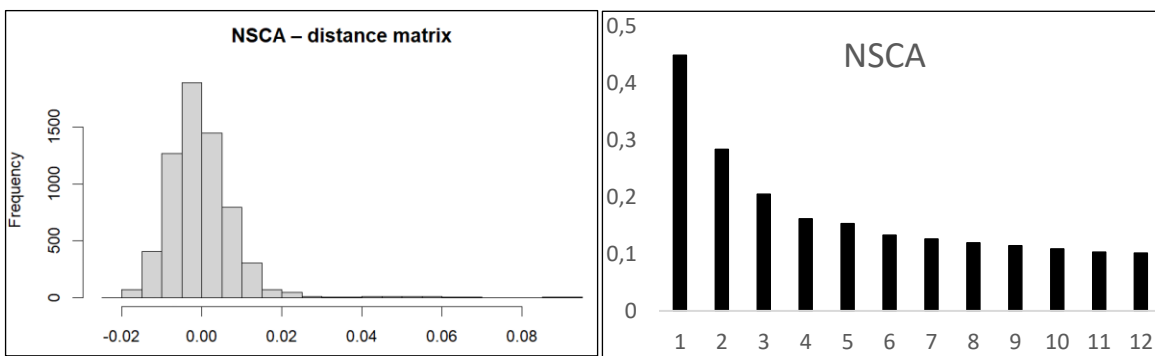


Figure 16. Distribution des covariances en 40 classes et les douze premières valeurs propres.

The results appear in table 10 and in figures 17 and 18.

| Species | coord1 | Cr%1 | P1 | Relevés | coord1 | Cr%1 | P1 |
|---|---|---|---|---|---|---|---|
| *Hibiscus cannabinus* | 0.21161587 | 4.47812784 | 7.2 | X9 | -1.7365976 | 5.81141295 | 2.8 |
| *Heteropogon contortus* | 0.21086076 | 4.44622592 | 5.4 | X15 | -2.0228431 | 5.73462547 | 3.2 |
| *Pennisetum polystachion* | 0.19760028 | 3.90458723 | 8.6 | X14 | -1.997668 | 5.35974194 | 3.8 |
| *Albuca kirkii* | 0.19352699 | 3.74526959 | 12.9 | X11 | -1.7500253 | 5.18628121 | 3.5 |
| *Botriochloa insculpta* | 0.18796352 | 3.53302859 | 9.1 | X18 | -1.7578419 | 4.60118023 | 4.5 |
| *Panicum atrosanguineum* | 0.16008675 | 2.5627766 | 14.5 | X10 | -1.3730275 | 3.57776049 | 8.5 |
| *Indigofera colutea* | 0.15630264 | 2.44305167 | 15.4 | X13 | -1.4058361 | 3.23144544 | 11.3 |
| *Microchloa kundtii* | 0.15456958 | 2.3891755 | 16.6 | X51 | 1.46390263 | 2.94077408 | 14 |
| *Plectranthus pauciflorus* | -0.1474352 | 2.17371294 | 9 | X12 | -1.2977361 | 2.70442469 | 14.4 |
| *Solenostemon thyrsiflorum* | 0.14650623 | 2.14640769 | 17.6 | X63 | 1.68436415 | 2.56787036 | 17.5 |
| **Species** | **coord2** | **Cr%2** | **P2** | **Relevés** | **coord2** | **Cr%2** | **P2** |
| *Achyranthes aspera var sicula* | -0.3050809 | 9.30743511 | 0.5 | X62 | -2.4919716 | 6.16459475 | 2.4 |
| *Commelina benghalensis* | -0.223288 | 4.98575161 | 5.7 | X70 | -2.9897854 | 5.74172177 | 3.2 |
| *Spermacoce pusilla* | 0.2129926 | 4.53658489 | 7.3 | X59 | -2.0336598 | 5.07160409 | 4.2 |
| *Hyparrhenia dissoluta* | 0.18600733 | 3.45987262 | 13.2 | E | 2.05881846 | 4.08403517 | 7.3 |
| *Justicia ruwenzoriensis* | -0.1773634 | 3.14577815 | 0.2 | X1 | -1.665733 | 3.72655941 | 8.5 |
| *Panicum maximum 2* | -0.1727271 | 2.98346677 | 7.2 | X47 | 1.62131944 | 3.6839849 | 8.7 |
| *Panicum maximum* | 0.17232925 | 2.9697369 | 14.6 | X64 | -1.7494567 | 3.66378247 | 8.4 |
| *Pennisetum polystachion* | -0.1672342 | 2.79672908 | 13.2 | X34 | -1.6443993 | 3.63171554 | 9.5 |

| Species | coord | Cr% | P | Relevés | coord | Cr% | P |
|---|---|---|---|---|---|---|---|
| *Solenostemon thyrsiflorum* | -0.1599391 | 2.55805083 | 16.4 | A | -2.1283562 | 3.30649647 | 10.9 |
| *Plectranthus caninus* | -0.1599391 | 2.55805083 | 15.9 | X5 | -1.6356383 | 3.20256394 | 10.3 |
| **Species** | **coord3** | **Cr%3** | **P3** | **Relevés** | **coord3** | **Cr%3** | **P3** |
| *Lannea humilis J* | -0.1929802 | 3.72413474 | 8.1 | K | 3.40772896 | 14.2402891 | 0.1 |
| *Cyanotis longifolia* | 0.19081255 | 3.6409431 | 0.1 | I | 5.40454751 | 11.0866785 | 0.3 |
| *Ipomoea obscura* | -0.1904212 | 3.6260222 | 6.4 | H | 3.16686959 | 8.19894199 | 1 |
| *Bidens oligoflora* | 0.18047158 | 3.25699928 | 10.1 | G | 2.34733243 | 8.04375114 | 0.9 |
| *Blepharis cristata* | -0.1772897 | 3.1431621 | 3.6 | X32 | -1.6561589 | 3.60376073 | 9.2 |
| *Asparagus abyssinicus* | -0.1724299 | 2.97320601 | 8.2 | X8 | -1.5478956 | 2.86818149 | 12.6 |
| *Vernonia miombicola* | -0.1697874 | 2.88277779 | 14.3 | X4 | -1.5307827 | 2.66827837 | 13.8 |
| *Digitaria longiflora* | -0.1572147 | 2.47164757 | 18.5 | C | 3.09475618 | 2.51671949 | 19.1 |
| *Acacia hockii J* | -0.1569337 | 2.4628185 | 11.9 | A | 1.72316586 | 2.16737268 | 18.9 |
| *Kalanchoe integra* | 0.15181364 | 2.30473822 | 0 | F | 1.11079922 | 2.05345604 | 19.5 |
| **Species** | **coord4** | **Cr%4** | **P4** | **Relevés** | **coord4** | **Cr%4** | **P4** |
| *Justicia matammensis* | -0.2236196 | 5.00057325 | 4.6 | B | -6.189709 | 13.423357 | 0.1 |
| *Crabbea velutina* | 0.21594148 | 4.66307233 | 5.3 | D | -5.3578267 | 13.4102454 | 0.1 |
| *Sporobolus stapfianus* | 0.19801573 | 3.92102291 | 7.4 | A | -3.6817695 | 9.89447212 | 0.2 |
| *Digitaria longiflora* | -0.1975248 | 3.90160423 | 9.1 | X45 | -1.9619388 | 4.94496274 | 4 |
| *Markhamia obtusifolia J* | 0.1940661 | 3.76616507 | 8.2 | C | -4.2068516 | 4.65046442 | 5.5 |
| *Brachiaria semiundulata* | -0.1821637 | 3.31836235 | 0 | X61 | 2.33401317 | 4.13541766 | 6.1 |
| *Acacia polyacantha A 5* | -0.1804481 | 3.2561518 | 5.3 | X2 | -1.7710208 | 4.12096805 | 5.7 |
| *Heliotropium strigosum* | -0.1788344 | 3.19817517 | 0 | I | 2.49679426 | 2.36618279 | 18.5 |
| *Commiphora africana J* | 0.16613102 | 2.75995148 | 15.6 | X24 | 1.23145585 | 2.12529078 | 20.6 |
| *Brachiaria brizantha* | -0.1579148 | 2.49370855 | 0.2 | X55 | 1.32859764 | 2.06151439 | 21.1 |

Table10. First ten species and first ten relevés arranged in descending order of relative contributions. Axes 1 to 4.

Axis 1 now opposes the relevés and species of closed sites to those of open savannah. Taking the relative contributions into account, the results oppose the relevés 9 to 15, 18 and 20 (most of them are significant) with *Plectranthus pauciflorus* (P = 8.3) and *Blepharis maderaspatensis* (P = 1.2) to relevés 46, 48, 50, 51 53, 58, 61and 63 with open savannah species (*Albuka kirkii, Botriochloa insculpta, Heteropogon contortus, Hibiscus cannabinus, Pennisetum polystachion*), but without any significant relative contribution.

On the second axis, two groups of species and relevés are clearly opposed : the first with several relevés and species (*Achyranthes aspera* var. *sicula, Commelina benghalensis*, *Justicia ruwenzoriensis* and *Commicarpus grandiflorus* with the significant relative contributions) located at the foot of the hill (wooded savannah) and at the edge or near the lake (forest gallery) and the second with frequent species in the other sites.

The third axis is dominated by four relevés of the summit of the transect, on the rocky hill, with a significant contribution; several species are highly characteristic: *Bulbostylis boeckelerana, Cleome hirta*, adult *Combretum molle*, *Cyanotis longifolia,* adult *Dalbergia nitidula, Kalanchoe integra, Loudetia simplex (*abundance 3*)* and adult and young *Parinari curatellifolia*.

With axis 4, on notes *Crabbea velutina* a very common species, and several other species, absent from the three relevés of the lake border or from one rocky site.

 With axis 5 *Solenostemon thyrsiflorum* et *Plectranthus caninus*, adult *Acacia hockii* (the first two species with a significant relative contribution) are absent from the three relevés with large *Acacia polyacantha*, the relevés in the thicket zone and on the rocky hill; they are opposed to *Cassia falcinella, Commelina erecta, Indigofera colutea, Justicia matammensis, Talinum caffrum, Vernonia miombicola*, young *Acacia senegal*, some of them with a significant relative contribution. These species occupy sites or a portion of the three sites listed for the first group.
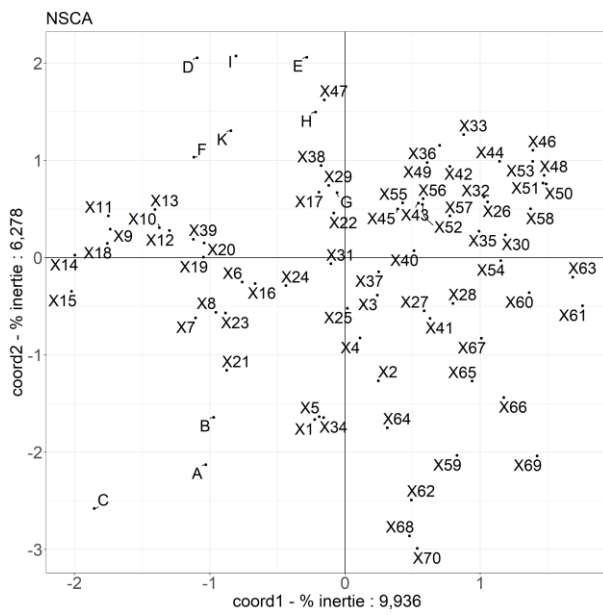
Figure 17. Non-symmetrical correspondence analysis of the Akagera file with simple disjunctive table. Coordinates of relevés on axes 1 and 2.
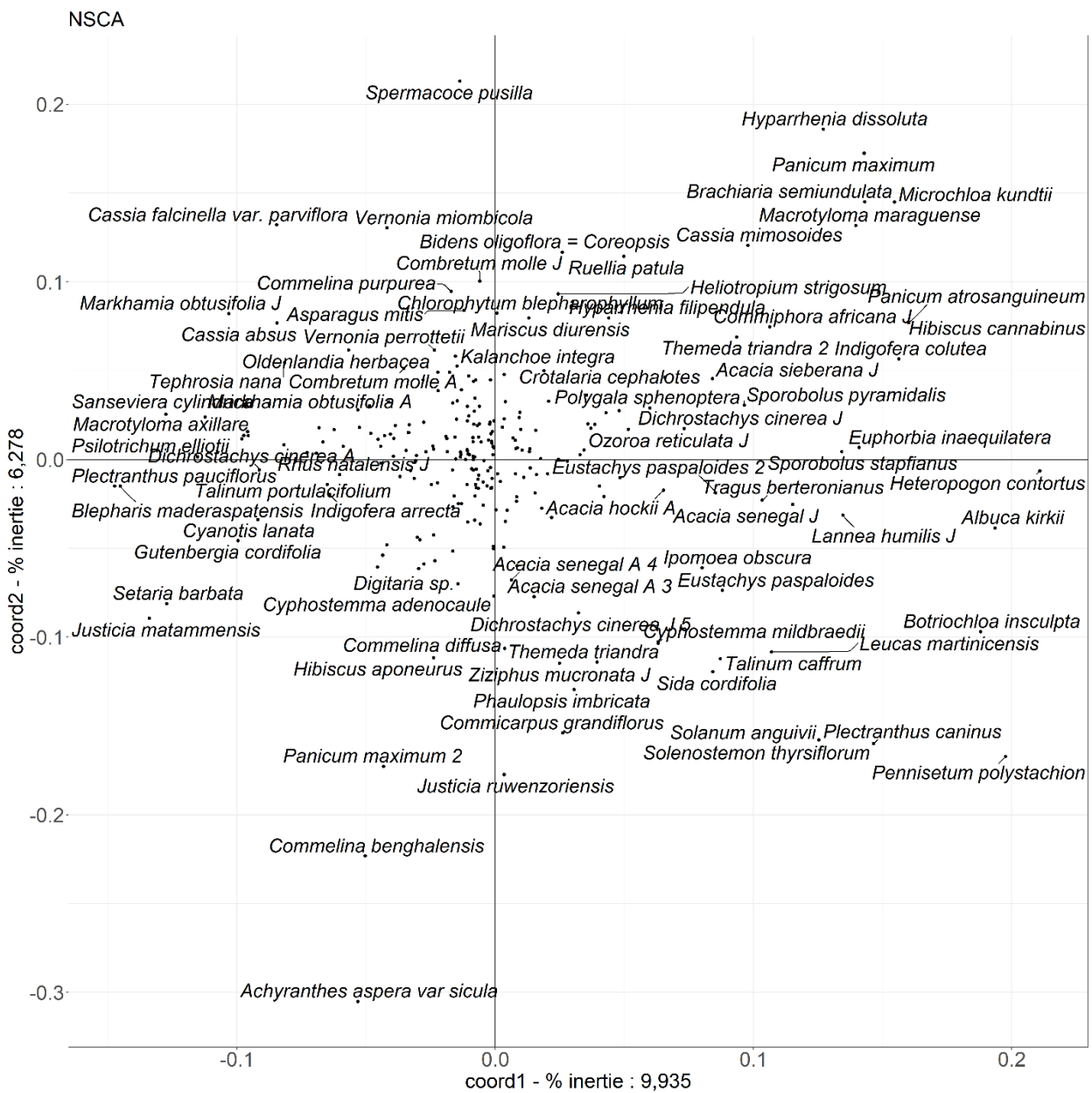
NSCA



Figure 18. Non-symmetrical correspondence analysis of the Akagera file with simple disjunctive table. Coordinates of species on axes 1 and 2.

The first two axes provide maximum information on the main plant groupings and the following axes highlight the relevés of less important groupings in terms of number of relevés.

## Principal coordinates analysis of the simple disjunctive table

The principal coordinates analysis is also tested with the simple disjunctive table.
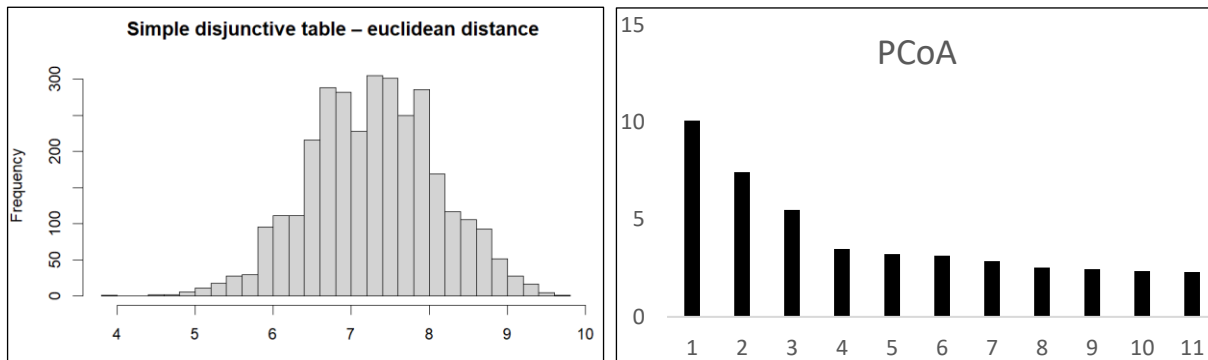
### With euclidean distance



Figure 19. Distribution of the euclidean distances into 40 classes and the first twelve eigenvalues.

PCoA is calculated with a matrix of euclidean distances and the distribution of distances has a shape close to a Gaussian distribution but with a weak asymmetry (figure 19). The first four eigenvalues are respectively 10.080, 7.444, 5.500 and 3.487% of the matrix trace and the decay continues in a slow and regular manner. The main results appear in figure 20.

Axis 1 contrasts the surveys of thickets and surrounding areas with those of *Acacia hockii* and *Themeda triandra* savannah.

Axis 2 contrasts relevés of more or less wooded savannah with relevés of gallery forests and those of herbaceous or woody vegetation on surface soil.

Axis 3 mainly highlights relevés on surface soil.

Axis 4 does not generate any clear interpretation.

Only the first three axes provide a clear interpretation of the vegetation.
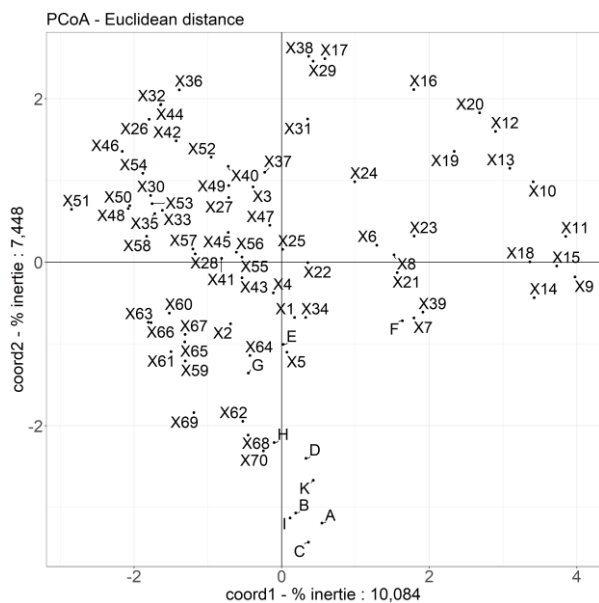


Figure 20. Principal coordinates of the file Akagera with the simple disjunctive table and euclidean distance. Coordinates of relevés on axes 1 and 2.

### With chord distance

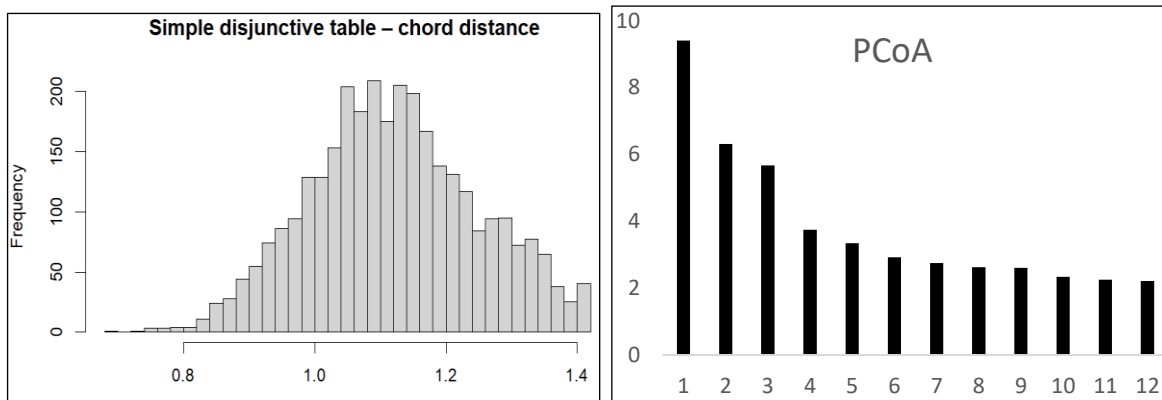The analysis is repeated with the chord distance (Figures 21 and 22).



Figure 21. Distribution of chord distances into 40 classes and the first twelve eigenvalues.
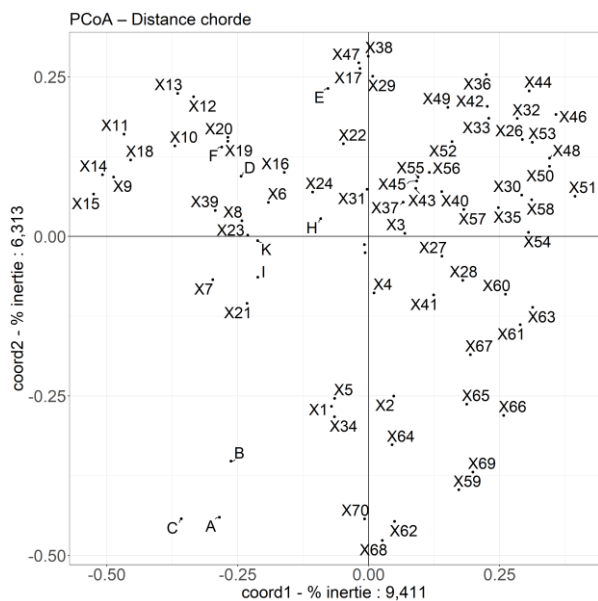


Figure 22. Principal coordinates analysis of the file Akagera with the simple disjunctive file and chord distance. Coordinates of relevés on axes1 and 2.

Axis 1 contrasts relevés of more or less wooded savannah with a set of relevés on surface soil or gallery forest.

Axis 2 isolates the *Acacia senegal* wooded savannah relevés and those of the *Acacia polyacantha* forest gallery.

With axis 3, the relevés on surface soil that are isolated.

Axis 4 isolates the relevés on surface soil with also those of gallery forests.

But as already seen previously, there is no help for interpretation in the available software, such as the cos² on the axes. Additionally, PCoA is calculated only for relevés.

# A third practical way with the transformation into a disjunctive table 113 and into a logarithm of abundances

The transformation into a disjunctive table with a possible line for each degree of abundance seems best suited and most rational to the data in the phytosociological tables. However, it makes the interpretation more cumbersome and is likely to discourage more than one user who is not accustomed to this type of analysis.

Phytosociologists favor presence data but abundance, an expression of biomass, sometimes makes it possible to discriminate certain sites. We therefore tested the following solution: for each row of a table, there are three possible rows:

- A first in which all the abundances (+, i, 1, 2, 3, 4, 5) become equal to 1, the specific name appears alone:

- A second in which abundances greater than 1 become equal to 1, the specific name is followed by >1;

- A third in which abundances greater than 3 become equal to 1, the specific name is followed by >3.

These files end with "113".

Another even simpler solution is proposed, taking into account the particularities of NSCA. In a phytosociological table with abundances transformed into scores 0.1 or 0.5 for r or + abundances, and into scores 1, 2, 3, 4, 5 for other abundances, NSCA is dominated by abundant species. In order to reduce this preponderance, it is enough to transform the abundances into logarithms. There is then only one line per species.

## Transformation 113

The transformation 113 is now applied with non-symmetrical correspondence analysis



Figure 23. Distribution of covariances into 40 classes and the first twelve eigenvalues.

The first eight eigenvalues are significant.
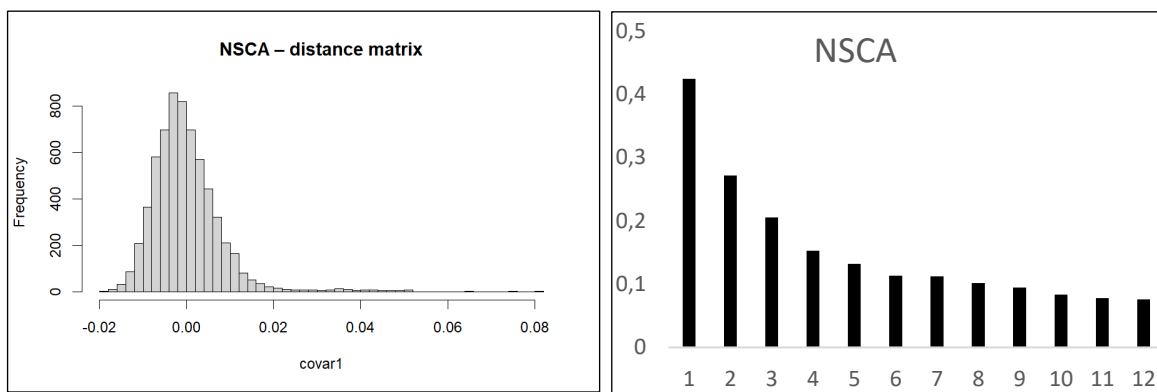
| Espèces | coord1 | Cr%1 | P | Relevés | coord1 | Cr%1 | P1 |
|---|---|---|---|---|---|---|---|
| *Pennisetum polystachion* | -0.205654 | 4.229356757 | 7.4 | X9 | 1.931746786 | 7.279287087 | 0.6 |
| *Hibiscus cannabinus* | -0.19457641 | 3.785997997 | 10 | X11 | 1.896731568 | 6.043094547 | 2.7 |

| Espèces | coord | Cr% | P | Relevés | coord | Cr% | P |
|---|---|---|---|---|---|---|---|
| *Heteropogon contortus* | -0.19391402 | 3.760264825 | 8.9 | X18 | 1.90386914 | 5.303027796 | 3.3 |
| *Botriochloa insculpta* | -0.17567716 | 3.086246343 | 13 | X14 | 1.948410712 | 5.245503101 | 3.1 |
| *Eustachys paspaloides* | -0.17247503 | 2.974763604 | 0.8 | X15 | 1.917336658 | 5.07952243 | 4.5 |
| *Cyanotis lanata* | 0.151068063 | 2.282155958 | 15 | X10 | 1.500473981 | 4.330831045 | 6.4 |
| *Panicum atrosanguineum* | -0.14585373 | 2.127330928 | 19 | F | 1.30081292 | 2.934037174 | 14 |
| *Themeda triandra* | -0.14417659 | 2.078688907 | 24 | X13 | 1.314346486 | 2.714586501 | 14 |
| *Microchloa kundtii* | -0.14338215 | 2.055844003 | 20 | X61 | -1.76454138 | 2.446344688 | 17 |
| *Setaria barbata* | 0.142631628 | 2.034378125 | 12 | X20 | 1.136825992 | 2.380964153 | 19 |
| **Espèces** | **coord2** | **Cr%2** | **P 2** | **Relevés** | **coord2** | **Cr%2** | **P2** |
| *Achyranthes aspera* var *sicula* | -0.29649337 | 8.790831752 | 0.7 | X68 | -2.45938489 | 5.899449087 | 2.5 |
| *Panicum maximum* >1 | -0.23178046 | 5.372217916 | 4.9 | X47 | 1.900134963 | 4.890968408 | 4.9 |
| *Commelina benghalensis* | -0.2085065 | 4.347496109 | 9.6 | E | 2.140714012 | 4.593830663 | 3.9 |
| *Spermacoce pusilla* | 0.200329963 | 4.013209412 | 8 | X62 | -1.95519421 | 4.142817031 | 6 |
| *Hyparrhenia dissoluta* | 0.171534636 | 2.94241312 | 16 | I | 3.020812315 | 3.955700694 | 8.4 |
| *Microchloa kundtii* | 0.16503794 | 2.723752157 | 14 | X70 | -2.42736541 | 3.831223722 | 6.7 |
| *Acacia senegal* A | -0.15652509 | 2.450010399 | 4.9 | A | -2.05688627 | 3.553351805 | 10 |
| *Justicia ruwenzoriensis* | -0.14658204 | 2.148629557 | 1.6 | D | 2.62206297 | 3.53912409 | 9.9 |
| *Acacia senegal* A >1 | -0.14350156 | 2.059269714 | 1.1 | X33 | 1.664920475 | 3.379523393 | 10 |
| *Brachiaria semiundulata* | 0.139663016 | 1.950575798 | 15 | X64 | -1.57330265 | 2.950755184 | 14 |
| **Espèces** | **coord3** | **Cr%3** | **P3** | **Relevés** | **coord3** | **Cr%3** | **P3** |
| *Albuca kirkii* | -0.20082014 | 4.032872862 | 6.3 | I | -4.91201218 | 10.45911184 | 0.1 |
| *Acacia hockii* J | 0.192182098 | 3.693395863 | 10 | K | -2.7815723 | 9.223363737 | 0.4 |
| *Ormocarpon trichocarpum* J | 0.18186171 | 3.307368163 | 13 | G | -2.13411921 | 6.663264697 | 1.2 |
| *Asparagus abyssinicus* | 0.176820801 | 3.126559559 | 8 | H | -2.85677371 | 6.633288565 | 1.8 |
| *Lannea humilis* J | 0.171565833 | 2.94348351 | 12 | C | -3.48272441 | 3.614821521 | 9.1 |
| *Loudetia simplex* | -0.17082224 | 2.918023813 | 0 | A | -2.03928024 | 3.492781921 | 8.6 |
| *Blepharis cristata* | 0.15622873 | 2.440741594 | 4.9 | X69 | -2.00399413 | 3.046539937 | 14 |
| *Loudetia simplex* >1 | -0.15509669 | 2.405498437 | 0 | X32 | 1.468564246 | 2.687816945 | 15 |
| *Ipomoea obscura* | 0.153544795 | 2.357600413 | 13 | B | -2.46888443 | 2.642271618 | 14 |
| *Vernonia miombicola* | 0.151890874 | 2.307083759 | 20 | F | -1.18295331 | 2.42644883 | 16 |
| **Espèces** | **coord4** | **Cr%4** | **P4** | **Relevés** | **coord4** | **Cr%4** | **P 4** |
| *Acacia polyacantha* A | 0.239734653 | 5.747270377 | 0 | B | 6.092445228 | 16.09011709 | 0 |
| *Setaria barbata* | 0.222355623 | 4.944202309 | 0.8 | C | 6.542512286 | 12.75668212 | 0.1 |
| *Acacia polyacantha* A >3 | 0.191574746 | 3.670088313 | 0 | A | 3.518284528 | 10.39631825 | 0 |
| *Brachiaria brizantha* | 0.18452798 | 3.405057557 | 0 | X2 | 2.311738273 | 6.805047157 | 1.9 |
| *Crabbea velutina* | -0.18290449 | 3.34540529 | 10 | X70 | -2.15794219 | 3.027936819 | 11 |
| *Cissampelos mucronata* | 0.166762532 | 2.780974212 | 0 | X61 | -1.87372444 | 2.758451765 | 15 |
| *Cissus petiolata* | 0.164189639 | 2.695823747 | 0 | X1 | 1.426162794 | 2.755270001 | 15 |
| *Sporobolus stapfianus* | -0.15317559 | 2.346276275 | 17 | X69 | -1.69170935 | 2.171028299 | 19 |
| *Albuca kirkii* | -0.15247468 | 2.324852832 | 20 | X68 | -1.48125323 | 2.140016273 | 19 |
| *Acacia hockii* J | 0.151291024 | 2.288897403 | 19 | X51 | 1.141877958 | 1.730977466 | 27 |

Table 11. First ten species and first ten relevés arranged in decreasing order of relative contributions. Axes 1 to 4.

The first axis (table 11 and figure 24) clearly contrasts species and relevés of grassy savannah (positive coordinates) with species and relevés of xerophilous groves; we therefore have, on the one hand, the most open herbaceous formations opposed to the most densely woody thicket formations.

Axis 2 contrasts species and relevés of surface soils with species and relevés of savannah with *Acacia senegal* and *Acacia polyacantha*, *i.e.* the densest wooded savannah vegetations.

The following axes relate to various groupings represented by a smaller number of relevés.
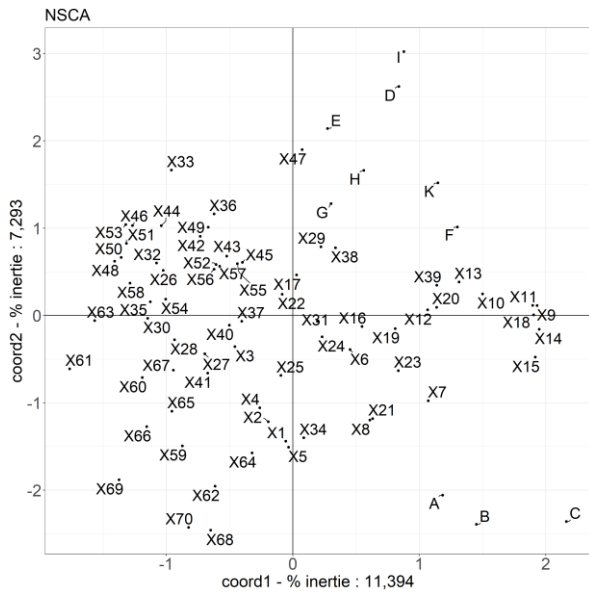
Figure 24. Non-symmetrical correspondence analysis of the disjunctive 113 Akagera file. Coordinates of relevés on axes 1 and 2.
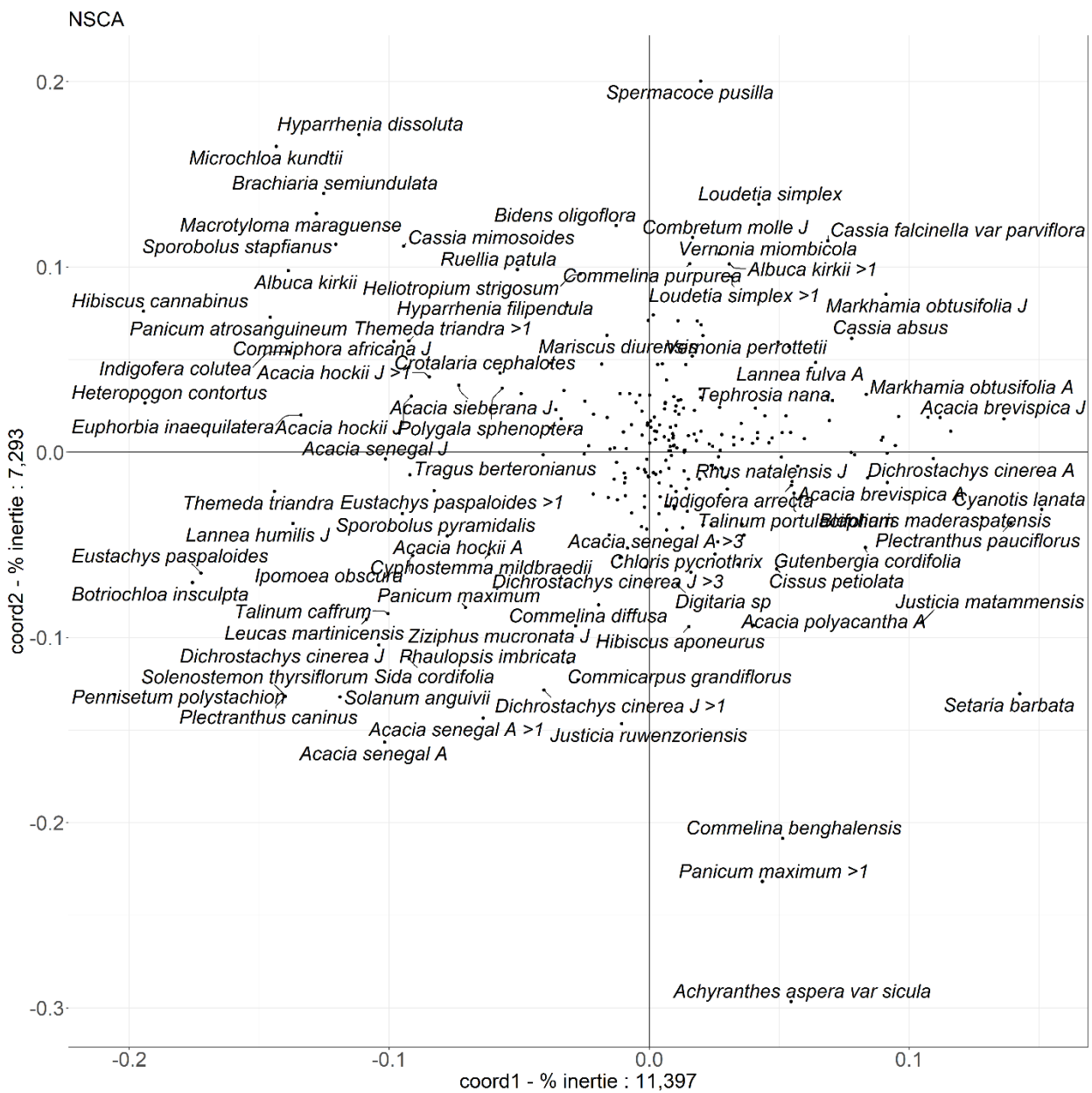
Figure 25. Non-symmetrical correspondence analysis of the disjunctive 113 file Akagera. Coordinates of species on axes 1 and 2.

This technique provides a good basis leading to an in-depth interpretation of the vegetation.

## Logarithmic transformation

The non-symmetrical analysis with abundance data of type 12345 is not applicable because it is too strongly influenced by the abundant species which overwrite all other data. The logarithmic transformation is tested as intermediate

form between simple presence data and abundance data (Figures 26 and 27, table 11).



Figure 26. Distribution of covariances into 40 classes and the first twelve eigenvalues.

The first eigth eigenvalues are significant.

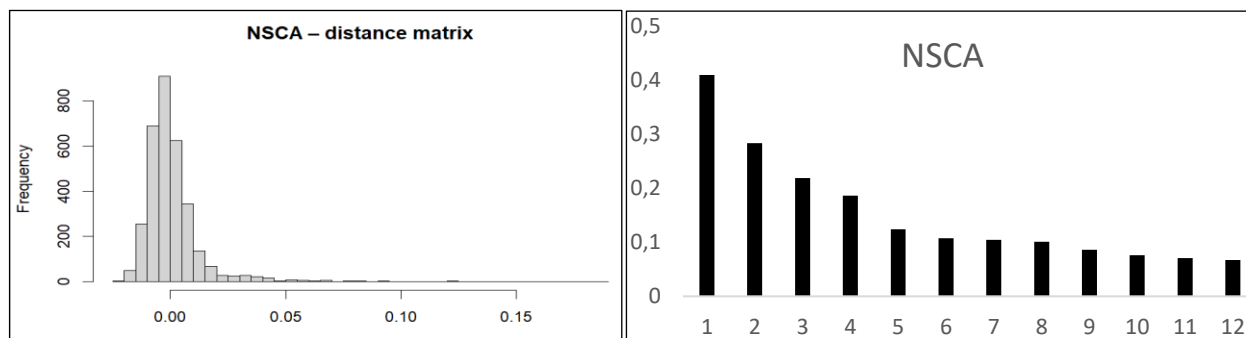| Species | coord1 | Cr%1 | P1 | Relevés | coord1 | Cr%1 | P1 |
|---|---|---|---|---|---|---|---|
| *Eustachys paspaloides* | -0.21339249 | 4.553635498 | 0.5 | X9 | 1.952533953 | 7.367856611 | 1.8 |
| *Pennisetum polystachion* | -0.20511014 | 4.207016862 | 1 | X11 | 1.909471698 | 6.122380793 | 2.2 |
| *Dichrostachys cinerea* J | -0.19309178 | 3.728443555 | 64.6 | X18 | 1.8603817 | 5.086075828 | 4.6 |
| *Themeda triandra* | -0.18710563 | 3.500851597 | 25.1 | X14 | 1.886364105 | 4.893273176 | 4.4 |
| *Cyanotis lanata* | 0.177081351 | 3.1357805 | 13 | X10 | 1.531669499 | 4.452027512 | 6.3 |
| *Setaria barbata* | 0.176206754 | 3.104882028 | 6.3 | X15 | 1.758069429 | 4.250308548 | 6.1 |
| *Heteropogon contortus* | -0.17399798 | 3.027529808 | 2.9 | F | 1.434897127 | 3.604385453 | 9.7 |
| *Hibiscus cannabinus* | -0.17141107 | 2.938175565 | 5.2 | X61 | -1.96419565 | 3.003778085 | 12.6 |
| *Acacia hockii* J | -0.16620387 | 2.762372679 | 85.2 | C | 2.948027705 | 2.683033182 | 12.8 |
| *Acacia senegal* A | -0.16547459 | 2.738183838 | 70.9 | X20 | 1.182870945 | 2.593939431 | 17.2 |
| **Species** | **coord2** | **Cr%2** | **P2** | **Relevés** | **coord2** | **Cr%2** | **P2** |
| *Achyranthes aspera* var *sicula* | 0.315712561 | 9.967442107 | 0.1 | B | 4.291176443 | 8.22949935 | 1.1 |
| *Acacia senegal* A | 0.251774941 | 6.339062093 | 43.8 | A | 2.661511406 | 5.939949577 | 2.9 |
| *Acacia polyacantha* A | 0.25118667 | 6.309474308 | 0.2 | I | -3.49130735 | 5.2504266 | 3.7 |
| *Panicum maximum* | 0.220742022 | 4.872704045 | 7.2 | D | -3.09781429 | 4.929298421 | 3.9 |
| *Setaria barbata* | 0.217179173 | 4.716679337 | 3.6 | X68 | 2.239873514 | 4.746260437 | 4.6 |
| *Dichrostachys cinerea* J | 0.201490279 | 4.059833273 | 51.9 | E | -2.15561563 | 4.591882107 | 5.1 |
| *Commelina benghalensis* | 0.19509359 | 3.806150877 | 2.9 | C | 3.777524666 | 4.405320806 | 6 |
| *Loudetia simplex* | -0.19334108 | 3.738077281 | 1.3 | X47 | -1.76151289 | 4.216808347 | 6.8 |
| *Spermacoce pusilla* | -0.17369181 | 3.016884354 | 4 | X62 | 1.828634006 | 3.519061607 | 8.3 |
| *Sporobolus stapfianus* | -0.16584379 | 2.750416248 | 11.2 | X33 | -1.57779386 | 3.039064015 | 11.3 |
| **Species** | **coord3** | **Cr%3** | **P3** | **Relevés** | **coord3** | **Cr%3** | **P 3** |
| *Albuca kirkii* | 0.320772136 | 10.28947635 | 7.6 | I | 4.918342153 | 10.41971651 | 0.5 |
| *Acacia polyacantha* A | 0.268158828 | 7.190915702 | 0.1 | C | 4.988695638 | 7.683108024 | 1.1 |
| *Loudetia simplex* | 0.258640933 | 6.689513211 | 0 | B | 4.016534795 | 7.209808647 | 1.8 |
| *Acacia hockii* J | -0.19817944 | 3.927509129 | 41.2 | D | 3.336491382 | 5.718134541 | 2.6 |
| *Ormocarpon trichocarpum* J | -0.19423327 | 3.772656268 | 21.6 | A | 2.523450027 | 5.339683458 | 3.5 |
| *Sporobolus stapfianus* | 0.176529656 | 3.116271936 | 11.1 | K | 1.964411181 | 4.755168009 | 4.8 |
| *Acacia senegal* A | 0.169496043 | 2.87289087 | 52.6 | H | 2.355974997 | 4.602303709 | 6.7 |
| *Asparagus abyssinicus* | -0.14677704 | 2.154349816 | 9.7 | G | 1.75548669 | 4.470948348 | 5.8 |
| *Plectranthus pauciflorus* | -0.13858614 | 1.920611896 | 8.8 | E | 1.854664298 | 3.399215554 | 9.6 |
| *Acalypha bipartita* | 0.138441191 | 1.916596336 | 0.6 | X12 | -1.22065085 | 2.395488469 | 14.9 |
| **Species** | **coord4** | **Cr%4** | **P4** | **Relevés** | **coord4** | **Cr%4** | **P4** |
| *Acacia polyacantha* A | -0.41020576 | 16.82687666 | 0 | B | -5.77213975 | 14.88998795 | 0 |
| *Acacia senegal* A | 0.317230669 | 10.06352975 | 13.2 | C | -6.48694008 | 12.99100336 | 0.1 |
| *Acacia hockii* J | -0.28321225 | 8.020917793 | 15 | A | -2.84251455 | 6.775346131 | 1.7 |
| *Setaria barbata* | -0.25613971 | 6.560755264 | 1.3 | X68 | 2.307287883 | 5.036260053 | 5.6 |
| *Dichrostachys cinerea* J | 0.220299849 | 4.853202356 | 29.1 | X70 | 2.60541073 | 4.424867295 | 6.4 |
| *Panicum maximum* | 0.190616031 | 3.633447115 | 16.5 | X69 | 2.265682377 | 3.972544313 | 8.9 |
| *Albuca kirkii* | 0.180305763 | 3.251016817 | 36 | X50 | -1.83469447 | 3.791576906 | 7.5 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Brachiaria brizantha* | -0.14319256 | 2.050410909 | 0.3 | X62 | 1.714753917 | 3.094403142 | 11.4 |
| *Acalypha bipartita* | -0.14120115 | 1.993776489 | 0.5 | X64 | 1.58127545 | 2.993887655 | 11.2 |
| *Cissampelos mucronata* | -0.12907242 | 1.665968992 | 0 | X2 | -1.43397922 | 2.632600802 | 15.7 |

Tableau 11. First ten species and first ten relevés arranged in order of decreasing eigenvalues. Axes 1 to 4.

On axis 1, we find on one side species of Poaceae and other grassy savannahs species and on the other, species and relevés of dense xerophilous thickets.

Axis 2 opposes species and relevés of densely wooded savannah with *Acacia polyacantha* and *Acacia senegal* and on the other, species of surface soils.

Axis 3 contrasts these elements of axis 2 with the species and relevéss of shrub savannah.

Axis 4 contrasts the relevéss and species of gallery forest and those of *Acacia senegal* savannah.
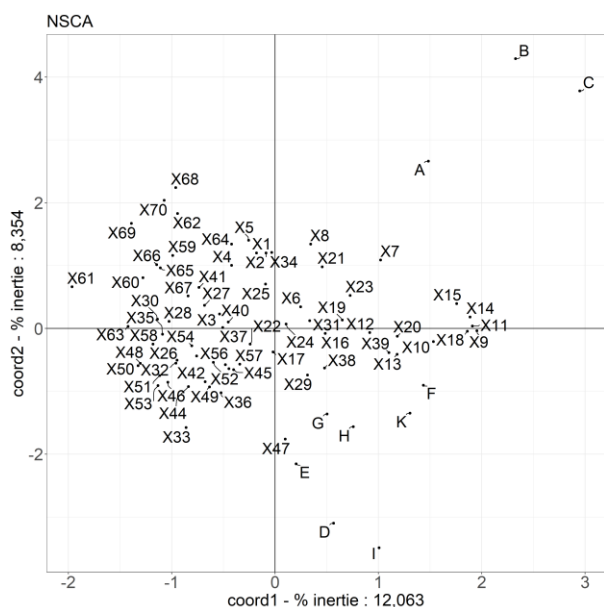


Figure 27. Non-symmetrical correspondence analysis of the file Akagera, logarithmic transformation. Coordinates of relevés on axes 1 and 2.
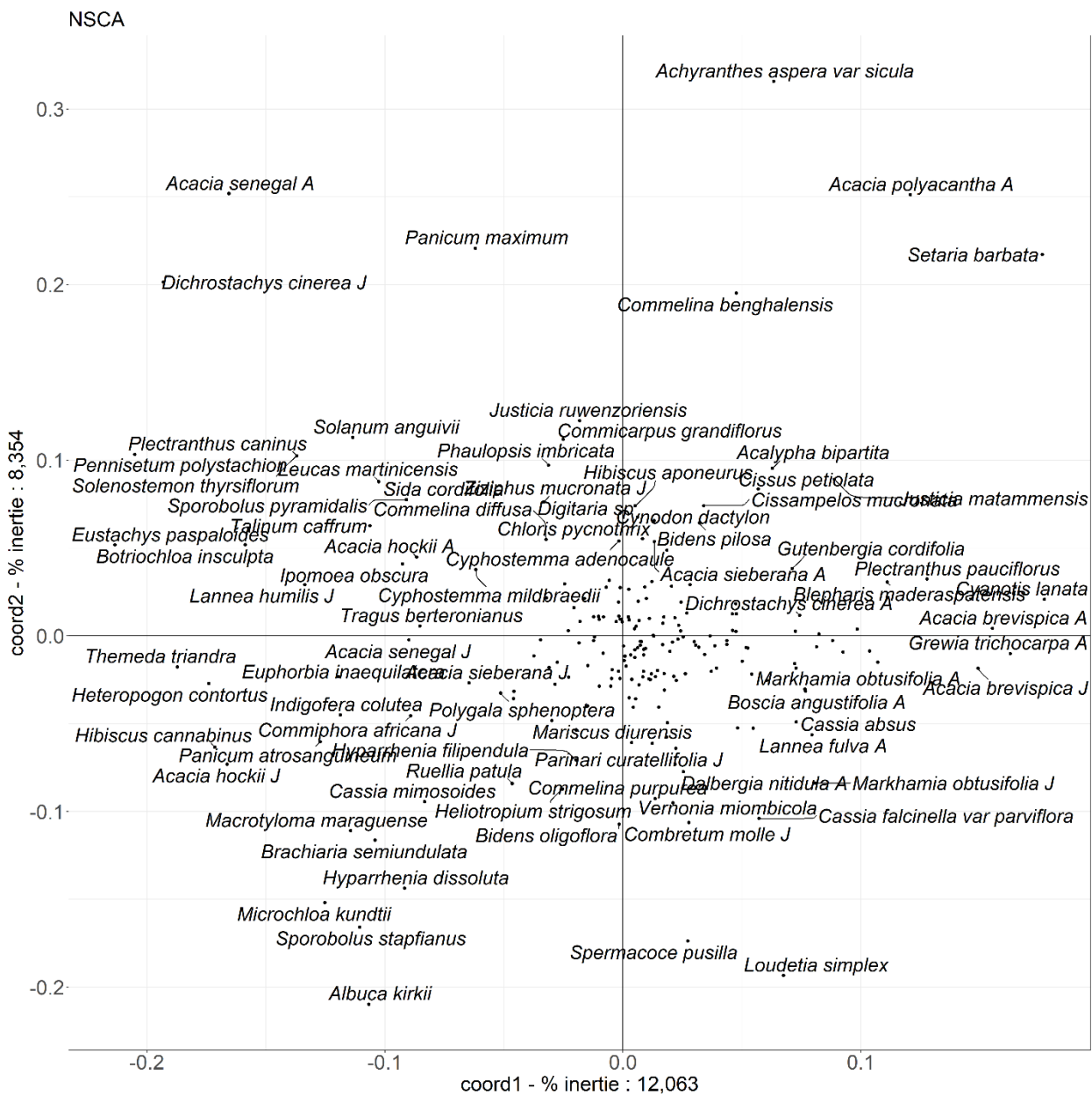
Figure 28. Non-symmetrical analysis of the file Akagera with logarithmic transformation. Coordinates of species on the axes 1 and 2.

This lorarithmic transformation of abundances provides results very close to those of transformation 113.

## Multiple factor analyses

Another problem is the presence, in a same table, of small herbaceous and very large ligneous species with very different dispersion patterns. In a savannah vegetation, for instance, the woody plants, of *Poaceae* and *Cyperaceae* forming dense clumps and of other herbaceous shows a high range of patterns, varying from small aggregates, clumps, density variations, gradients and complex ones. If every category of plants is separately submitted to a multivariate analysis, it appears that every subset has its own structure. This difference in structure can be expressed, for example, by the first eigenvalue of a CA and a NSCA. If these first values are very different, the choice of a multiple factor analysis is necessary to rebalance the subsets.

## Simple disjunctive file

The analysis is carried out with three subtables : one with 116 lines for the woody species, one with 57 lines for the *Poaceae* and *Cyperaceae* and a third one for the other herbaceous species. The densities of the subtables are 0.07424569, 0.17982456, and 0.17230216, respectively; this means that woody plants are favored over herbaceous ones. The main results are presented in figures 29 to 31 and in table 12.
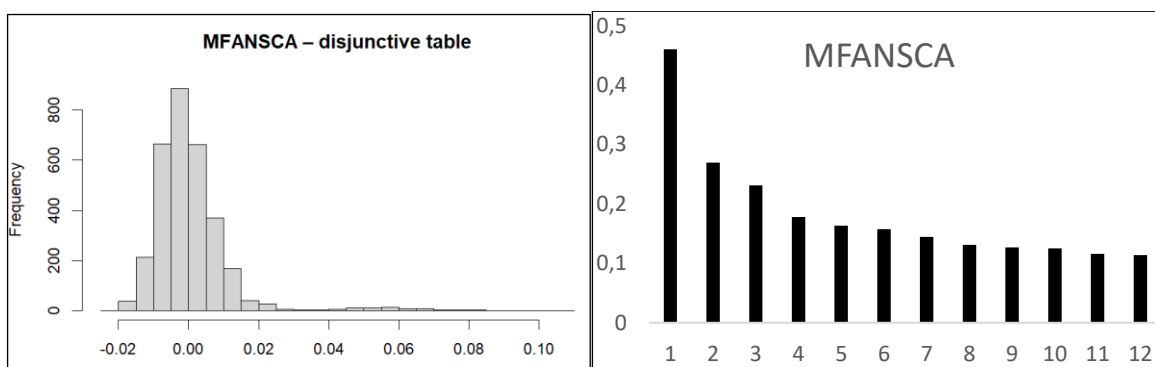


Figure 29. Distribution of covariances into 40 classes and the first twelve eignevalues.

| Species | coord 1 | CR 1 | P 1 | Relevés | coord1 | CR1 | P1 |
|---|---|---|---|---|---|---|---|
| *Lannea humilis* J | 0.228837702 | 5.236669379 | 0 | X9 | -1.69735155 | 5.911751981 | 2.4 |
| *Hibiscus cannabinus* | 0.189627734 | 3.595867733 | 0.2 | X15 | -1.94795826 | 5.349843569 | 3.5 |
| *Pennisetum polystachion* | 0.181837929 | 3.306503247 | 0.7 | X18 | -1.82717711 | 5.056154637 | 4.3 |
| *Heteropogon contortus* | 0.180926039 | 3.27342316 | 0.6 | X14 | -1.93909954 | 5.0495039 | 4 |
| *Albuca kirkii* | 0.172709029 | 2.982840887 | 1.3 | X11 | -1.59936695 | 4.531080513 | 5.3 |
| *Botriochloa insculpta* | 0.167252781 | 2.797349259 | 0.9 | X10 | -1.25767857 | 3.137842808 | 12 |
| *Commiphora africana* J | 0.15996519 | 2.558886205 | 1.8 | F | -1.30470719 | 2.991867615 | 13 |
| *Acacia senegal* J | 0.156040086 | 2.434850835 | 0.1 | X13 | -1.25829634 | 2.544641739 | 13 |
| *Acacia brevispica* J | -0.15347631 | 2.355497799 | 0 | X61 | 1.80302273 | 2.495695742 | 17 |
| *Markhamia obtusifolia* J | -0.15158521 | 2.297807691 | 1.9 | X39 | -1.28710761 | 2.488907276 | 16 |
| Species | coord2 | CR2 | P2 | Relevés | coord2 | CR2 | P2 |
| *Achyranthes aspera* | 0.26539892 | 7.043658699 | 0 | K | -2.20166576 | 6.934855355 | 1.5 |
| *Commelina benghalensis* | 0.202324818 | 4.093533212 | 0.1 | X68 | 2.31181371 | 4.815076888 | 5 |
| *Spermacoce pusilla* | -0.20106872 | 4.042862821 | 0.2 | I | -3.31819606 | 4.550712087 | 4.7 |
| *Combretum molle* J | -0.18251188 | 3.33105851 | 0 | H | -2.3482996 | 4.538522492 | 5.9 |
| *Panicum maximum* | -0.16498708 | 2.722073685 | 1.2 | X1 | 1.82599952 | 4.277440785 | 6.9 |
| *Hyparrhenia dissoluta* | -0.15980916 | 2.55389661 | 2.1 | X70 | 2.56061586 | 4.261242898 | 4.3 |
| *Ziziphus mucronata* J | 0.150224349 | 2.256735512 | 0 | X62 | 2.05288651 | 4.020057363 | 7 |
| *Justicia ruwenzoriensis* | 0.150012464 | 2.250373927 | 0 | X5 | 1.80130522 | 3.926828505 | 8.7 |
| *Panicum maximum* 2 | 0.149699483 | 2.240993511 | 1 | X59 | 1.73550218 | 3.632982171 | 8.1 |
| *Markhamia obtusifolia* J | -0.14775075 | 2.183028419 | 2.7 | E | -1.96365876 | 3.565049738 | 10 |
| Species | coord3 | CR3 | P3 | Relevés | coord3 | CR3 | P3 |
| *Lannea humilis* J | -0.24837846 | 6.169185773 | 0 | K | 2.42066137 | 8.383062603 | 0.5 |
| *Ormocarpon trichocarpum* J | -0.22122611 | 4.894099095 | 0 | G | 1.9903612 | 6.112347677 | 1.8 |
| *Acacia hockii* J | -0.20206673 | 4.08309643 | 0 | I | 3.40633722 | 4.795683921 | 4.5 |
| *Acacia senegal* J | 0.187756997 | 3.525268984 | 0 | X32 | -1.77281425 | 4.035795282 | 6.3 |
| *Vernonia miombicola* | -0.16109209 | 2.595066165 | 1.2 | H | 2.19377105 | 3.96086539 | 6.6 |
| *Blepharis cristata* | -0.15563974 | 2.422372935 | 0 | X68 | 2.09050257 | 3.9373042 | 7.9 |
| *Boscia angustifolia* J | -0.15402882 | 2.372487741 | 0 | C | 3.7106775 | 3.854565547 | 9.3 |

| Species | coord4 | CR4 | P4 | Relevés | coord4 | CR4 | P4 |
|---|---|---|---|---|---|---|---|
| *Commiphora africana* J | -0.14934833 | 2.230492445 | 3.3 | X70 | 2.35094842 | 3.59197769 | 9.1 |
| *Ruellia patula* | -0.14373713 | 2.066036268 | 3.5 | A | 2.25915251 | 3.507742517 | 9.3 |
| *Asparagus abyssinicus* | -0.13917043 | 1.936840831 | 2.2 | X62 | 1.91269199 | 3.489734982 | 10 |
| **Species** | **coord4** | **CR4** | **P4** | **Relevés** | **coord4** | **CR4** | **P4** |
| *Markhamia obtusifolia* J | -0.35678743 | 12.72972681 | 0 | K | -2.19215992 | 6.875101195 | 1.9 |
| *Commiphora africana* J | -0.27523249 | 7.575292375 | 0 | I | -4.01389826 | 6.658984326 | 2.4 |
| *Dichrostachys cinerea* J | 0.267925698 | 7.178417991 | 0 | X45 | 2.21807026 | 5.976539014 | 2 |
| *Acacia hockii* A | -0.25245989 | 6.373599492 | 0 | D | 3.54650167 | 5.291053076 | 4.2 |
| *Dichrostachys cinerea* J 2 | -0.1970746 | 3.883839617 | 0 | X55 | -2.06475049 | 5.214376457 | 4 |
| *Justicia matammensis* | 0.164505674 | 2.706211675 | 1 | X36 | 1.72288125 | 4.398930065 | 5.9 |
| *Microchloa kundtii* | 0.163733475 | 2.680865097 | 1.1 | G | -1.55685362 | 3.739726277 | 8.8 |
| *Markhamia obtusifolia* A | -0.14796055 | 2.189232467 | 0.1 | X31 | 1.36029284 | 3.043730388 | 12 |
| *Indigofera colutea* | 0.145747741 | 2.124240413 | 1.4 | X11 | 1.2918838 | 2.956326859 | 13 |
| *Dalbergia nitidula* J | -0.13144247 | 1.727712249 | 0 | B | 2.95900816 | 2.895118005 | 13 |

Table 12. The first ten species and the first ten relevés arranged in decreasing order of eigenvalues. Axes 1 to 4.

With axis 1, we notice the opposition between grassy savannah species and xerophilous thickets but this time several woody species are highlighted such as *Lannea humilis* J, *Commiphora africana* J in open sites and *Acacia brevispica* J, *Markhamia obtusifolia* and *Acacia senegal* in closed sites.

In axis 2 which opposes the vegetation of surface soils and wooded savannah with *Acacia senegal*, two woody species are also highlighted, namely *Combretum molle* J and *Markhamia obtusifolia*. We can also notice *Panicum maximum* with abundance 2 which develops well below *Acacia senegal*.

Axis 3 is strongly influenced by young woody species (J) which accompany several characteristic herbaceous species in open environments while young *Acacia senegal* prefer wooded savannahs.

Axis 4 is strongly influenced by *Acacia hockii* wooded savannah species, significantly less densely covered than *Acacia senegal* or *Acacia polyacantha* savannahs.
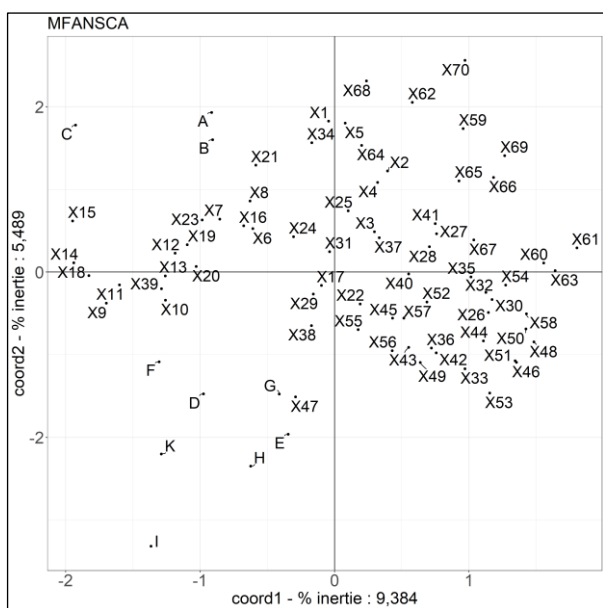


Figure 30. Multiple factor analysis based on non-symmetrical correspondence analysis of the simple disjunctive Akagera file. Coordinates of relevés on axes 1 and 2.
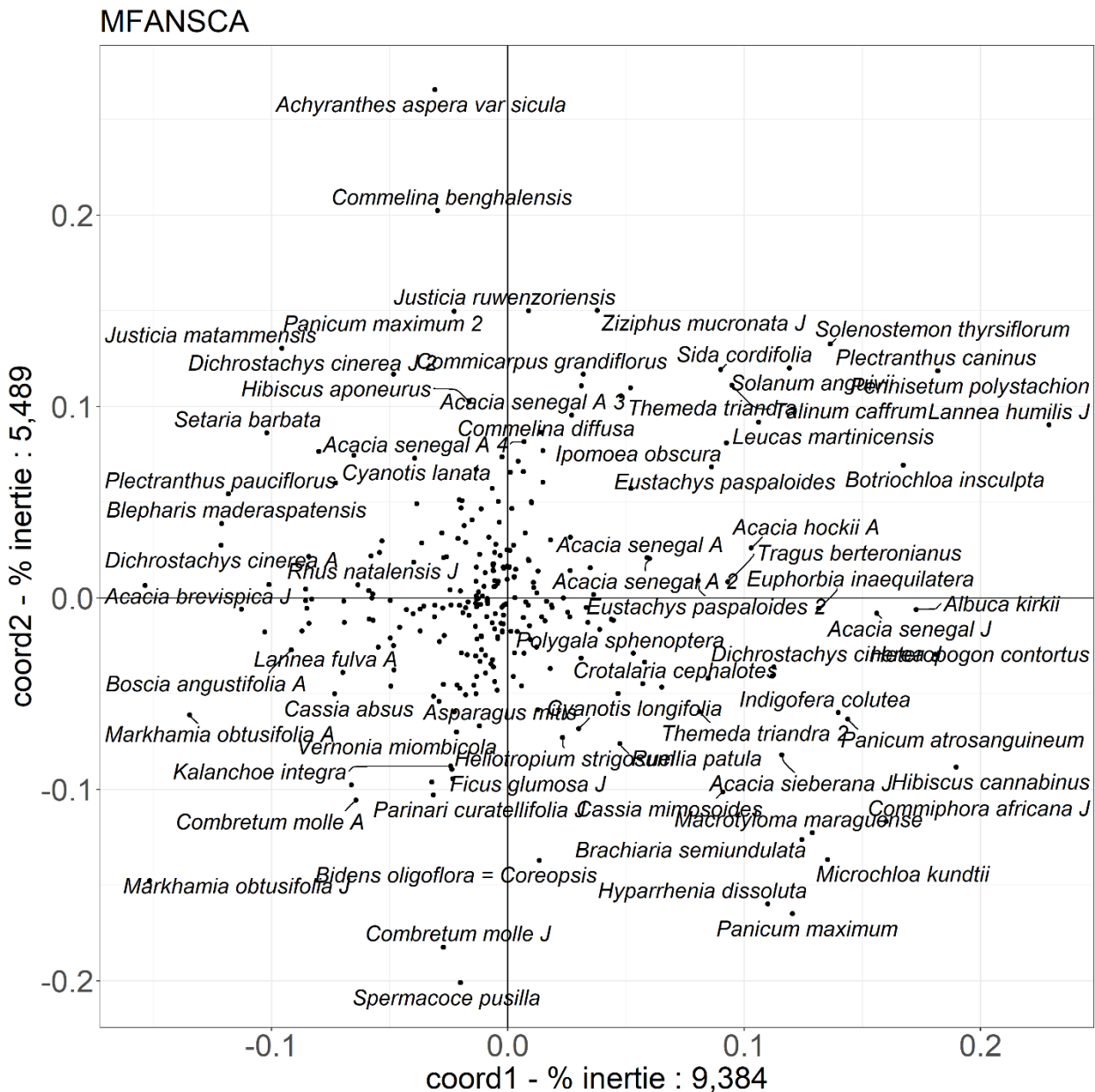
Figure 31. Multiple factor analysis based on non-symmetrical correspondence analysis of the simple disjunctive Akagera file. Coordinates of species on axes 1 and 2.

## File 113

The analysis is carried out with three subtables: one of 105 lines for woody plants, one of 52 lines for *Poaceae* and *Cyperaceae* and one of 136 lines for other herbaceous plants. The densities of the subtables are 0.09952381, 0.21995192, and 0.17830882, respectively; The woody plants file is clearly favored over herbaceous species, and among these herbaceous plants other than *Poaceae* and *Cyperaceae* are also somewhat favored. The main results are presented in figures 32 to 34 and in table 13.

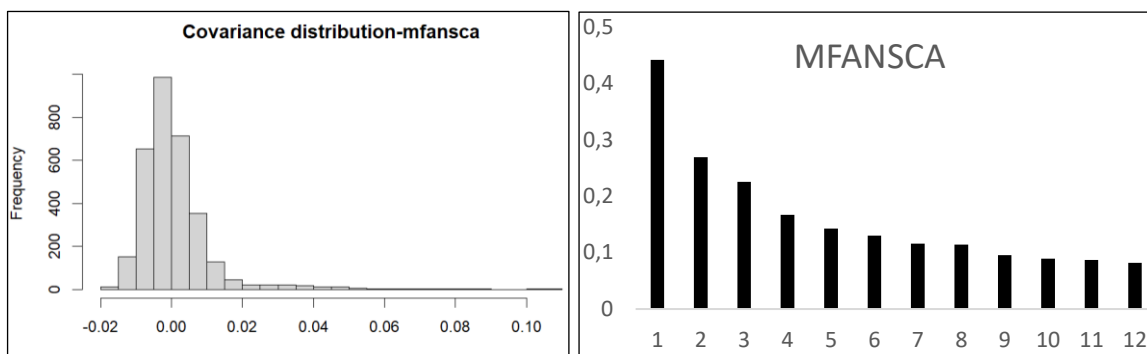The first eight eigenvalues are significant.

Figure 32. Distribution of covariances into 40 classes and the first twelve eigenvalues.

| Species | Coord1 | CR1 | P1 | Relevés | Coord1 | CR1 | P1 |
|---|---|---|---|---|---|---|---|
| *Lannea humilis* J | -0.19580723 | 3.83404713 | 0.3 | X9 | 1.922770785 | 7.566415081 | 1.2 |
| *Pennisetum polystachion* | -0.18126604 | 3.285737582 | 0.1 | X11 | 1.821345519 | 5.815800319 | 1.6 |
| *Hibiscus cannabinus* | -0.17822665 | 3.176473905 | 0.5 | X18 | 1.914920945 | 5.421226903 | 1.9 |
| *Acacia brevispica* J | 0.17037193 | 2.902659457 | 0 | X10 | 1.507844311 | 4.549129674 | 5.3 |
| *Heteropogon contortus* | -0.15735716 | 2.476127559 | 1.9 | X14 | 1.796394474 | 4.526544571 | 4.6 |
| *Dichrostachys cinerea* J | -0.15492261 | 2.400101361 | 0 | X15 | 1.732678895 | 4.184965157 | 6.6 |
| *Eustachys paspaloides* | -0.15437563 | 2.383183386 | 0 | F | 1.454645365 | 3.857749558 | 8.7 |
| *Botriochloa insculpta* | -0.14976169 | 2.24285623 | 2.2 | K | 1.498154028 | 2.972721932 | 13 |
| *Grewia trichocarpa* A | 0.145613794 | 2.120337708 | 0 | X39 | 1.280153875 | 2.733856701 | 17 |
| *Acacia senegal* A | -0.14246327 | 2.029578362 | 0.3 | X20 | 1.157426961 | 2.587225225 | 16 |
| **Species** | **Coord2** | **CR2** | **P2** | **Relevés** | **Coord2** | **CR2** | **P2** |
| *Achyranthes aspera* var. *sicula* | 0.286930682 | 8.2329216 | 0 | X68 | 2.410010117 | 5.676863802 | 2.5 |
| *Commelina benghalensis* | 0.208527626 | 4.348377064 | 0.2 | X70 | 2.619458378 | 4.476846065 | 5.8 |
| *Acacia senegal* A | 0.205037708 | 4.204046187 | 0 | X62 | 2.034383213 | 4.4566812 | 5.6 |
| *Panicum maximum* >1 | 0.198629824 | 3.945380715 | 0 | X47 | -1.74845967 | 4.013447238 | 7.3 |
| *Acacia senegal* A >1 | 0.191705225 | 3.675089327 | 0 | I | -2.84318778 | 3.691682954 | 7.9 |
| *Spermacoce pusilla* | -0.18626242 | 3.469368907 | 0.6 | E | -1.94021047 | 3.583206121 | 9.2 |
| *Dichrostachys cinerea* J >1 | 0.164647188 | 2.710869641 | 1.1 | A | 2.099624985 | 3.560241665 | 9.9 |
| *Hyparrhenia dissoluta* | -0.15825073 | 2.504329402 | 2.7 | K | -1.56482723 | 3.243203086 | 11 |
| *Combretum molle* J | -0.15157021 | 2.297352711 | 0 | X33 | -1.66258394 | 3.217648424 | 12 |
| *Microchloa kundtii* | -0.1452494 | 2.109738753 | 3.6 | X64 | 1.63837519 | 3.200054841 | 13 |
| **Species** | **Coord3** | **CR3** | **P3** | **Relevés** | **Coord3** | **CR3** | **P3** |
| *Acacia hockii* J | -0.23925453 | 5.724273189 | 0 | K | 2.527092123 | 8.458308139 | 0.5 |
| *Ormocarpon trichocarpum* J | -0.23801289 | 5.665013683 | 0 | I | 4.269082524 | 8.323048164 | 0.5 |
| *Lannea humilis* J | -0.20047461 | 4.019007086 | 0 | G | 2.128827922 | 6.913838487 | 1.5 |
| *Albuca kirkii* | 0.199639776 | 3.985604011 | 0.1 | H | 2.62448696 | 5.618466675 | 3.8 |
| *Acacia senegal* A >1 | 0.160798042 | 2.585601017 | 0 | X69 | 2.208788398 | 3.575643925 | 11 |
| *Acacia senegal* A | 0.159359458 | 2.53954368 | 0 | X68 | 1.771017296 | 3.065606078 | 13 |
| *Asparagus abyssinicus* | -0.15225308 | 2.318100076 | 1.5 | X62 | 1.594906788 | 2.739155292 | 14 |
| *Acacia hockii* A | -0.14717191 | 2.165957161 | 0.8 | X70 | 2.011577156 | 2.640113194 | 14 |
| *Blepharis cristata* | -0.14156728 | 2.004129545 | 0.3 | X32 | -1.4487688 | 2.539299422 | 14 |
| *Plectranthus pauciflorus* | -0.1344135 | 1.806698981 | 1.2 | C | 2.818127445 | 2.506484055 | 17 |
| **Species** | **Coord4** | **CR4** | **P4** | **Relevés** | **Coord4** | **CR4** | **P 4** |
| *Acacia polyacantha* A | -0.31237972 | 9.75810925 | 0 | B | -6.03624455 | 15.18214296 | 0 |
| *Acacia polyacantha* A >3 | -0.24032899 | 5.775802392 | 0 | C | -6.86901302 | 14.89129155 | 0 |
| *Acacia hockii* A | -0.1998283 | 3.993135007 | 0 | A | -3.74618424 | 11.33376696 | 0 |
| *Acacia hockii* J | -0.19585114 | 3.835767094 | 0.6 | X2 | -2.20770499 | 5.857079683 | 3.1 |
| *Setaria barbata* | -0.17852863 | 3.187247085 | 0 | X50 | -1.61767367 | 2.94866308 | 12 |
| *Acacia senegal* A | 0.170728809 | 2.914832634 | 0 | X1 | -1.4957051 | 2.931221497 | 14 |
| *Cissampelos mucronata* | -0.15528265 | 2.411270007 | 0 | X7 | -1.42820648 | 2.827697034 | 15 |
| *Crabbea velutina* | 0.152539492 | 2.32682965 | 2.5 | X61 | 1.869004586 | 2.754013813 | 14 |
| *Albuca kirkii* | 0.151659897 | 2.300072442 | 2.5 | X70 | 1.929537251 | 2.42915648 | 17 |
| *Cissus petiolata* | -0.15133498 | 2.290227474 | 0 | X5 | -1.35619787 | 2.338373745 | 19 |

Table 13. First ten species and first ten relevés narreanged in order of decreasing eigenvalues. Axes 1 to 4.

On axis 1, we find the opposition between species and relevés of xerophilous thickets (*Grewia trichcarpa* A and *Acacia brevispica* J) and savannah species with young individuals of *Lannea humilis*, *Dichrostachis cinerea* or adult *Acacia senegal*.

On axis 2, there are on one side the relevés of dense wooded savannah with *Acacia senegal* present and abundant and *Acacia polyacantha* present and abundant with *Achyranthes aspera* var. sicula present in both formations and on the other side of the surface soil relevés. Several woody species are still evident.

On axis 3, young individuals of *Acacia hockii*, *Ormocarpon trochocarpon* and *Lannea humilis* in association with adult *Acacia hockii* oppose adult *Acacia senegal*.

On axis 4, it is mainly the savannah with *Acacia senegal* which is highlighted, in opposition to the gallery forest with *Acacia polyacantha*.
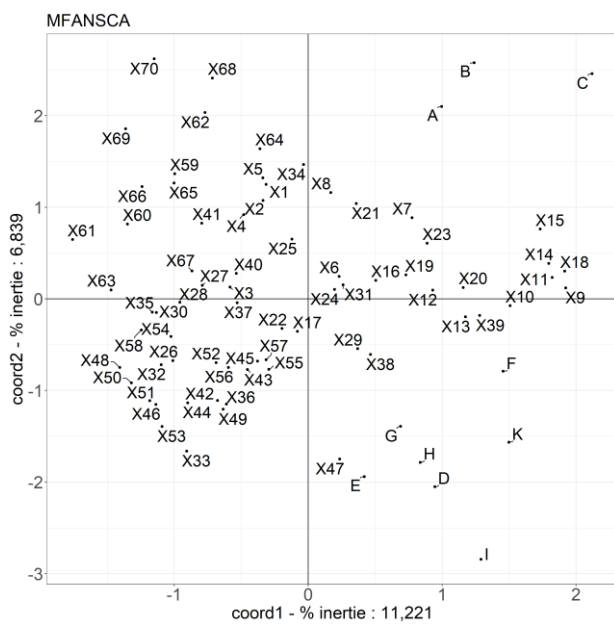


Figure 33. Multiple factor analysis of the disjunctive 113 Akagera file. Coordinates of relevés on axes 1 and 2.
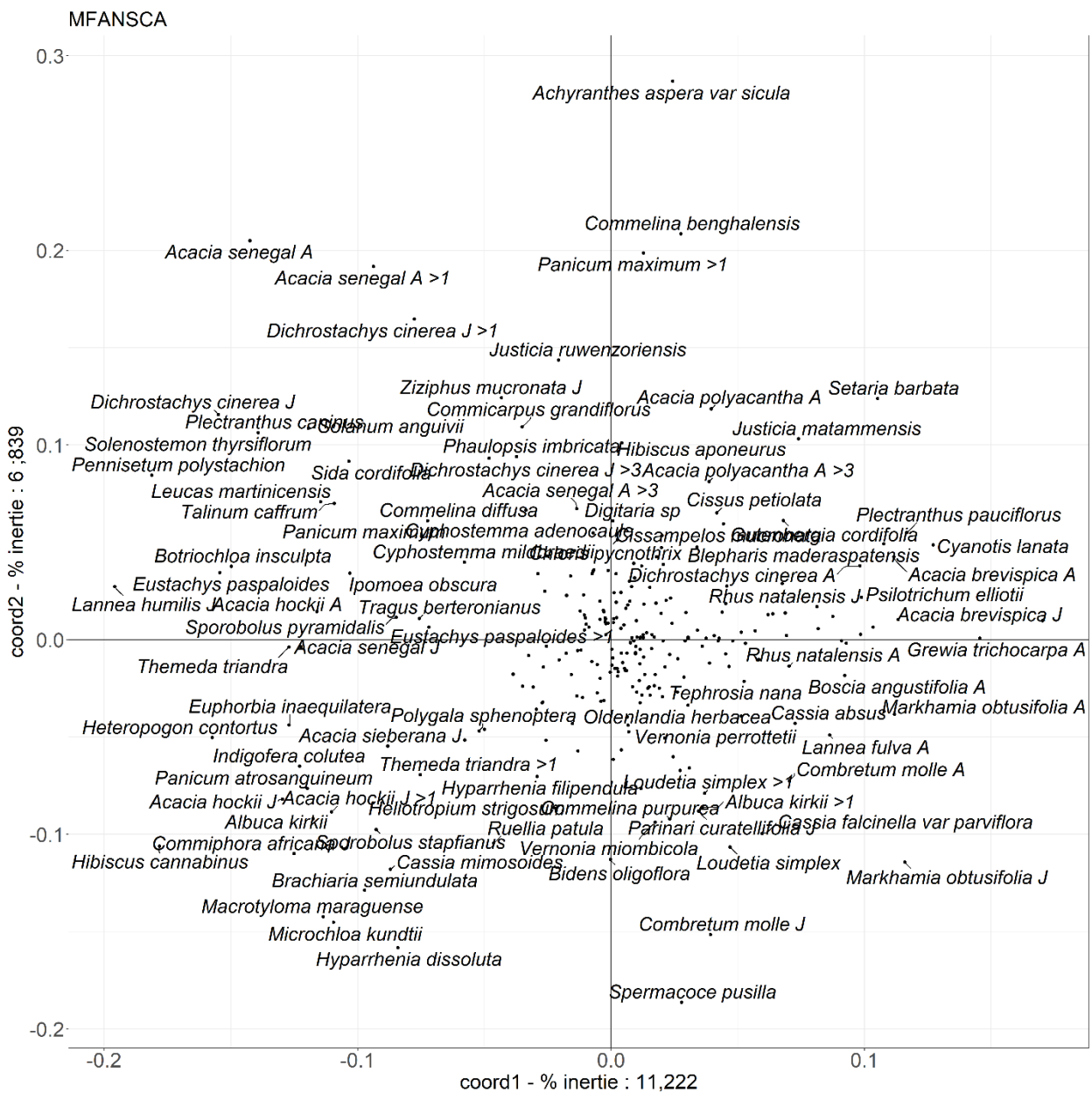
Figure 34. Multiple factor analysis of the disjunctive 113 Akagera file. Coordinates of species on axes 1 and 2.

## Logarithmic file

The densities of the subarrays are respectively 0.09807021 for woody plants, 0.201106 for *Poaceae* and *Cyperaceae* and 0.13081321 for other herbaceous plants.

The first eight eigenvalues are significant. The main results appear in figures 35 to 37 and table 14.
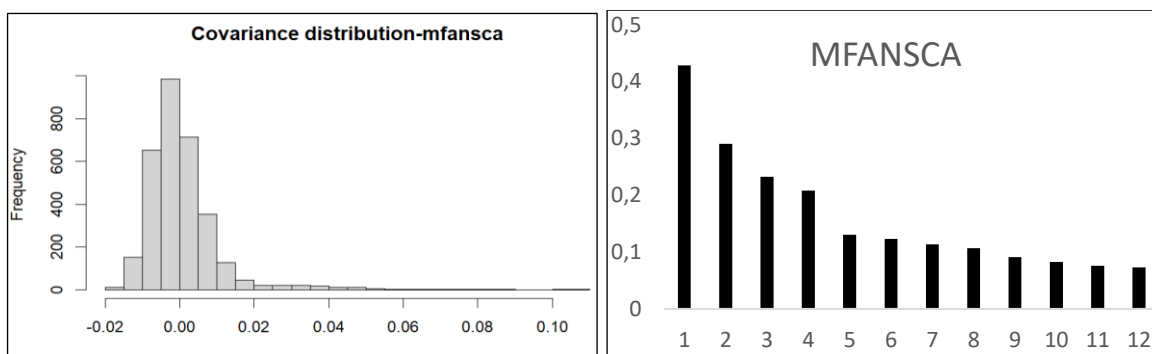
Figure 35. Distribution of covariances into 40 classes and then first twelve eigenvectores

| Species | coord1 | CR1 | P1 | Relevés | coord1 | CR1 | P1 |
|---|---|---|---|---|---|---|---|
| *Dichrostachys cinerea* J | -0.259425669 | 6.730167794 | 9.8 | X9 | 1.907155037 | 7.312626847 | 1.6 |
| *Acacia senegal* A | -0.217828126 | 4.744909248 | 22.8 | X11 | 1.845086894 | 5.944282292 | 3 |
| *Acacia hockii* J | -0.197421614 | 3.897529355 | 65.3 | X18 | 1.845976745 | 5.079864247 | 3.7 |
| *Grewia trichocarpa* A | 0.190712865 | 3.63713967 | 1.5 | X10 | 1.532706249 | 4.605218095 | 5.1 |
| *Eustachys paspaloides* | -0.176273823 | 3.107246078 | 0 | X14 | 1.722529192 | 4.186893032 | 6.9 |
| *Acacia brevispica* A | 0.175048854 | 3.064210133 | 6.9 | F | 1.497387597 | 4.0847892 | 7.3 |
| *Pennisetum polystachion* | -0.170839478 | 2.918612731 | 0 | X15 | 1.575643301 | 3.496978094 | 8 |
| *Acacia brevispica* J | 0.169707648 | 2.880068563 | 0 | X61 | -2.02020901 | 3.118859787 | 11 |
| *Hibiscus cannabinus* | -0.165536144 | 2.740221499 | 0 | K | 1.508590277 | 3.006253768 | 12 |
| *Lannea humilis* J | -0.164808645 | 2.716188931 | 0.1 | X20 | 1.195296654 | 2.757832704 | 11 |
| Species | coord2 | CR2 | 2 | Relevés | coord2 | CR2 | P2 |
| *Acacia polyacantha* A | -0.343403329 | 11.79258462 | 0 | B | -5.09063644 | 11.37806905 | 0.2 |
| *Achyranthes aspera* var *sicula* | -0.333722427 | 11.1370658 | 0 | A | -3.12140265 | 8.007325565 | 1 |
| *Acacia senegal* A | -0.313871549 | 9.851534909 | 1.7 | C | -4.90324743 | 7.865918784 | 1.2 |
| *Commelina benghalensis* | -0.204469025 | 4.180758229 | 0 | X68 | -2.31742992 | 5.030181317 | 4.3 |
| *Setaria barbata* | -0.19715431 | 3.886982178 | 0 | X62 | -2.05580598 | 4.444753951 | 5.7 |
| *Dichrostachys cinerea* J | -0.196007732 | 3.841903118 | 25.6 | X70 | -2.45950794 | 3.93575739 | 7.8 |
| *Acacia hockii* J | 0.170017445 | 2.890593154 | 39.5 | X47 | 1.467290986 | 2.870537676 | 16 |
| *Spermacoce pusilla* | 0.167671184 | 2.811362604 | 0.1 | E | 1.685163394 | 2.689228215 | 15 |
| *Panicum maximum* | -0.157556095 | 2.482392297 | 2.6 | I | 2.448476484 | 2.674275654 | 13 |
| *Albuca kirkii* | 0.152955054 | 2.339524843 | 17.4 | X69 | -1.85657778 | 2.593679573 | 15 |
| Species | coord3 | CR3 | P3 | Relevés | Coord3 | CR3 | P3 |
| *Albuca kirkii* | 0.356480508 | 12.70783526 | 0 | I | 5.03395195 | 11.30400726 | 0.1 |
| *Ormocarpon trichocarpum* J | -0.230098095 | 5.29451334 | 0.1 | G | 2.034947753 | 6.147117434 | 3 |
| *Loudetia simplex* | 0.22278379 | 4.963261716 | 0 | K | 2.133156416 | 6.010746783 | 2.5 |
| *Acacia polyacantha* A | 0.214859697 | 4.616468954 | 0 | D | 3.488016295 | 5.802830858 | 2.2 |
| *Acacia hockii* J | -0.204818602 | 4.195065973 | 15.3 | H | 2.53774337 | 5.360175937 | 4 |
| *Sporobolus stapfianus* | 0.164418129 | 2.703332109 | 0.5 | E | 2.275332401 | 4.902677529 | 5.8 |
| *Acacia senegal* A | 0.161418753 | 2.605601391 | 31.9 | C | 3.839022577 | 4.821955208 | 5.2 |
| *Lannea humilis* J | -0.155250311 | 2.410265904 | 0.4 | B | 2.800794373 | 3.444188019 | 9.1 |
| *Acacia brevispica* A | -0.148847552 | 2.215559386 | 17.8 | A | 1.884626482 | 2.91902945 | 14 |
| *Dalbergia nitidula* A | 0.146879675 | 2.157363896 | 0 | X21 | -1.40803445 | 2.817180358 | 15 |
| Species | coord4 | CR4 | P4 | Relevés | coord4 | CR4 | P4 |
| *Acacia polyacantha* A | -0.478723905 | 22.91765776 | 0 | C | -6.7479814 | 14.8980557 | 0 |
| *Acacia senegal* A | 0.3478979 | 12.10329488 | 0 | B | -5.81119363 | 14.82706169 | 0.1 |
| *Acacia hockii* J | -0.332564836 | 11.05993703 | 0.4 | A | -3.23460269 | 8.598640353 | 1.2 |
| *Dichrostachys cinerea* J | 0.237440143 | 5.637782157 | 5.8 | X68 | 2.202905527 | 4.545295941 | 5.6 |
| *Setaria barbata* | -0.200190687 | 4.007631099 | 0.1 | X70 | 2.435470277 | 3.85920216 | 7.1 |
| *Albuca kirkii* | 0.166775857 | 2.781418651 | 13.1 | X50 | -1.79434528 | 3.56810159 | 8.6 |
| *Acacia hockii* A | -0.157089785 | 2.467720052 | 1.2 | X69 | 2.050446365 | 3.163638515 | 9.5 |
| *Acalypha bipartita* | -0.143101459 | 2.047802765 | 0 | X62 | 1.705678588 | 3.05969312 | 12 |
| *Acacia brevispica* A | 0.131518453 | 1.729710349 | 25.6 | X64 | 1.47589798 | 2.614536218 | 16 |
| *Cissampelos mucronata* | -0.126383686 | 1.597283615 | 0 | X2 | -1.38304666 | 2.344620374 | 16 |

Table 14. First ten species and first ten relevés arranged in order of decreasing relative contributions. Axes 1 à 4.

On axis 1, the place taken by woody species is even more important than in the previous multivariate analyzes and we find the opposition between thickets and savannah, mainly wooded since this analysis favors woody species.

On axis 2, these are the most densely wooded savannahs, with *Acacia polyacantha* or *Acacia senegal* which are opposed to the most open formations on surface soil.

On axis 3, open sites with a large number of small woody individuals of *Acacia hockii*, *Lannea humilis* or *Ormocarpon trichocarpum* which are opposed to wooded savannahs with large individuals.

Axis 4 separates the *Acacia polyacantha* wooded savannah from the Acacia *senegal savannah*.
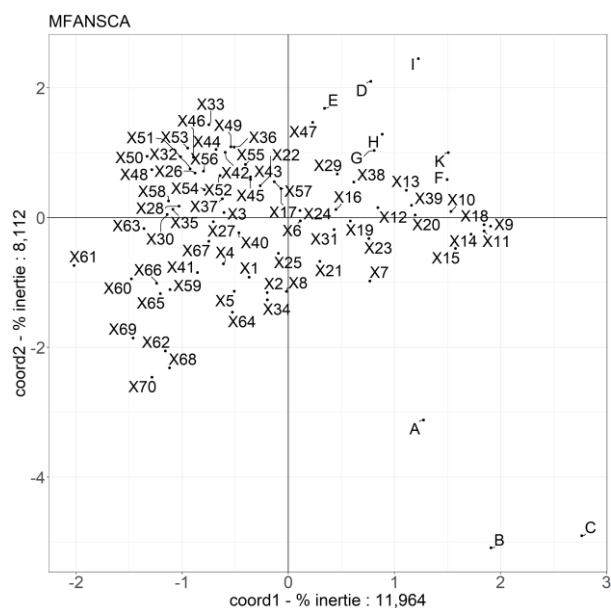


Figure 36. Multiple factor analysis of the logarithmic Akagera file. Coordinates of relevés on axes 1 and 2.
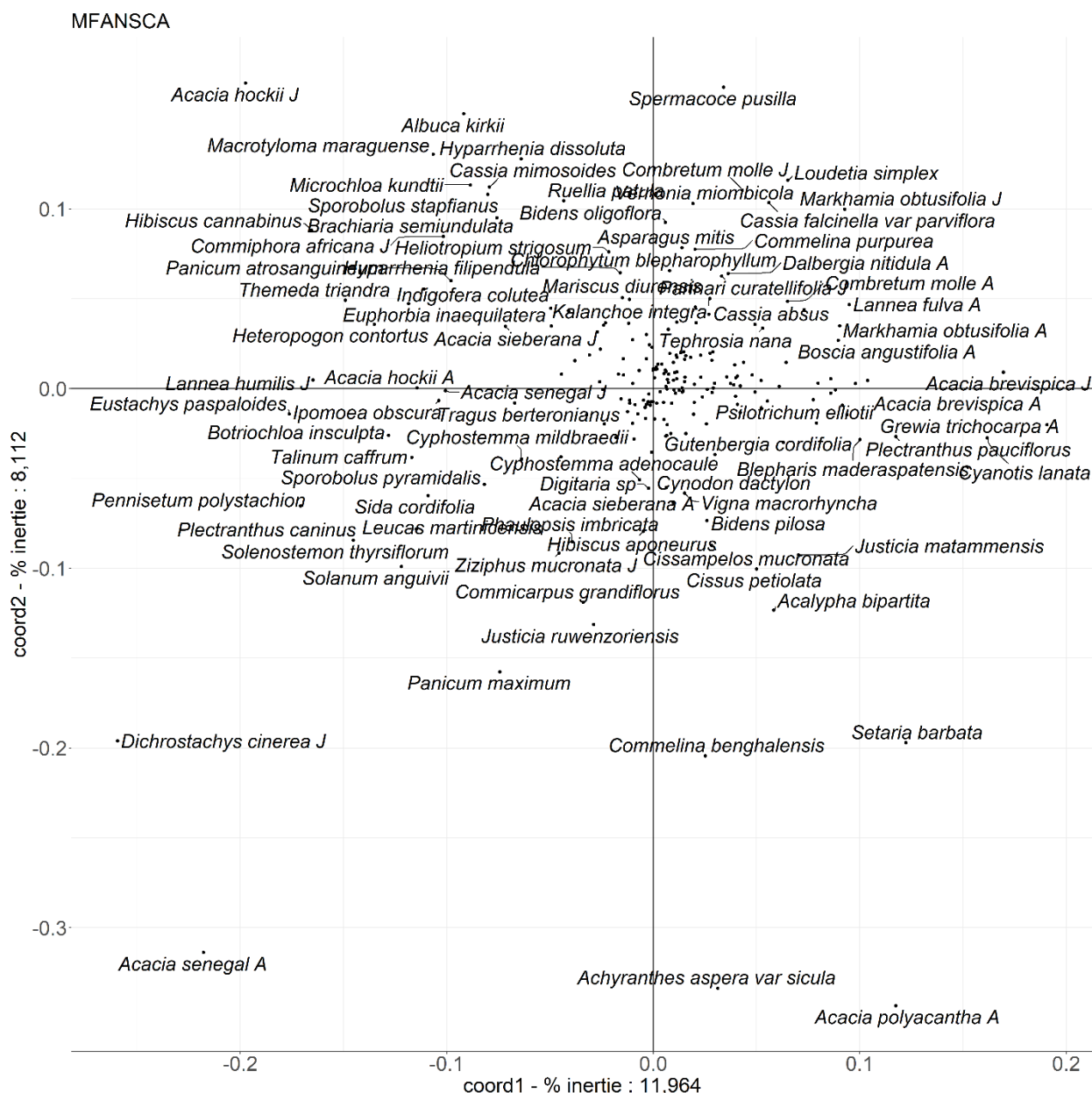
MFANSCA



Figure 37. Multiple factor analysis of the logarithmic Akagera file. Coordinates of species on axes 1 and 2.

As might be expected, the woody species dominate the analysis, compared to herbaceous species, and in particular open savannah grasses.

## Floristic and environmental files combined

As an example, the simple disjunctive floristic file and the simple disjunctive environmental file are grouped. The multiple factorial analysis is carried out based on three sub-tables, one of 116 lines for woody plants (density = 0.07424569), one of 196 lines for herbaceous plants (density = 0.1744898) and one of 118 lines for mesological parameters (density = 0.21186441). The woody plant file which is yet favored. The first twelve eigenvalues are significant.

The main results are presented in figures 38 and 39 and in table 15.
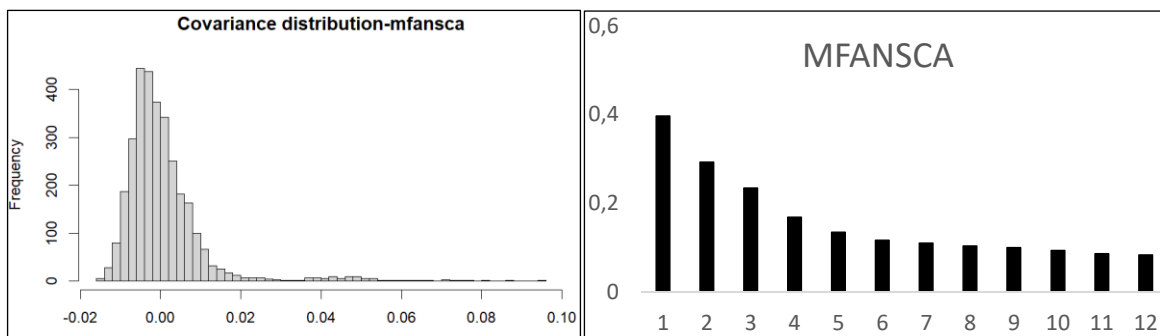
Figure 38. Distribution of covariances into 40 classes and the first twelve eigenvalues.

The first twelve eigenvalues are significant.

| Species | coord1 | CR1 | P1 | Relevés | coord1 | CR1 | P1 |
|---|---|---|---|---|---|---|---|
| *Lannea humilis* J | -0.17906878 | 3.206562962 | 0 | X9 | 1.698714859 | 5.167545557 | 3.6 |
| *Pennisetum polystachion* | -0.16009549 | 2.563056722 | 0 | F | 1.751331526 | 4.88370607 | 4 |
| S p-2 | -0.15637185 | 2.445215506 | 0.3 | K | 1.842804049 | 4.660150649 | 5.4 |
| *Markhamia obtusifolia* J | 0.155077742 | 2.404910612 | 0.1 | X18 | 1.78217821 | 4.536243008 | 5.4 |
| pH profondeur-3 | -0.14833353 | 2.200283759 | 0.5 | X14 | 1.774658643 | 4.127397941 | 7.3 |
| *Hibiscus cannabinus* | -0.14267804 | 2.035702365 | 1.3 | X15 | 1.73485618 | 4.080294259 | 7.5 |
| *Botriochloa insculpta* | -0.14167784 | 2.007261084 | 0.5 | X11 | 1.555175008 | 3.871122335 | 7.2 |
| *Heteropogon contortus* | -0.13887731 | 1.928690677 | 0.9 | X10 | 1.404684393 | 3.442007183 | 9.6 |
| pH surface-4 | -0.13313565 | 1.772510043 | 1.2 | X60 | -1.54625205 | 2.468208721 | 16 |
| *Albuca kirkii* | -0.13040711 | 1.700601561 | 2.3 | X63 | -1.56003265 | 2.445862359 | 16 |
| *Acacia brevispica* J | 0.130155418 | 1.694043281 | 0 | X61 | -1.60558242 | 2.385411394 | 17 |
| couvert2-0 | -0.12851143 | 1.65151882 | 0.1 | I | 1.866554949 | 2.383026773 | 19 |
| Vp-3 | -0.12693319 | 1.611203585 | 3.1 | X13 | 1.263814056 | 2.37805911 | 17 |
| *Acacia senegal* J | -0.12170946 | 1.481319335 | 0.1 | X39 | 1.288448102 | 2.357154852 | 17 |
| *Markhamia obtusifolia* A | 0.120982569 | 1.463678189 | 0.1 | X58 | -1.28381001 | 1.938552003 | 21 |
| **Species** | **coord2** | **CR2** | **P2** | **Relevés** | **coord2** | **CR2** | **P2** |
| Altitude 1 | -0.16540133 | 2.735760147 | 0.4 | H | 3.203618935 | 9.862919612 | 0 |
| Milieu ouvert | 0.165005852 | 2.722693121 | 0 | I | 3.663146831 | 9.178168283 | 0.3 |
| Exposition0 | -0.14282081 | 2.039778387 | 0.9 | K | 2.536547982 | 8.829331127 | 0.3 |
| pas de textre profondeur | 0.140862307 | 1.984218945 | 0 | X45 | 1.475977188 | 2.673752653 | 14 |
| Capacité rét eau P-0 | 0.140862307 | 1.984218945 | 0 | X43 | 1.561650588 | 2.601725683 | 14 |
| pH profondeur-0 | 0.140862307 | 1.984218945 | 0 | X21 | -1.29262065 | 2.296967115 | 20 |
| Sp-0 | 0.140862307 | 1.984218945 | 0 | X8 | -1.32194001 | 2.134356859 | 22 |
| Vp-0 | 0.140862307 | 1.984218945 | 0 | X15 | -1.24143831 | 2.08936742 | 18 |
| *Ormocarpon trichocarpum* J | -0.1393557 | 1.942001063 | 0.4 | X16 | -1.01045096 | 1.639108548 | 26 |
| *Combretum molle* J | 0.139198628 | 1.937625807 | 0 | X19 | -1.02120373 | 1.583645234 | 27 |
| *Microchloa kundtii* | 0.125032464 | 1.563311704 | 1.9 | X5 | -1.09592723 | 1.466921858 | 30 |
| *Plectranthus pauciflorus* | -0.12226252 | 1.494812374 | 0.4 | X13 | -0.98119768 | 1.433406301 | 33 |
| pas de structure profondeur | 0.122029888 | 1.489129351 | 0 | X14 | -1.01710669 | 1.355750301 | 35 |
| Milieu moyennement fermé | -0.11934732 | 1.424378292 | 0.4 | X24 | -0.96709703 | 1.303520004 | 33 |
| Horizon humifère-0 | 0.116701674 | 1.361928072 | 0 | X20 | -0.85549595 | 1.263651691 | 32 |
| **Species** | **coord3** | **CR3** | **P3** | **Relevés** | **coord3** | **CR3** | **P3** |
| *Achyranthes aspera* var *sicula* | -0.17387087 | 3.023107778 | 0.1 | X68 | -2.73792746 | 7.604962096 | 0.7 |
| S s-3 | -0.17018899 | 2.896429177 | 0 | X62 | -2.50616259 | 6.588565605 | 1.4 |
| S s-2 | 0.168934853 | 2.853898451 | 0.1 | X64 | -2.06833358 | 5.224980898 | 3.5 |
| Vs-4 | -0.16671148 | 2.779271605 | 0 | X69 | -2.34367106 | 4.926021544 | 4.2 |
| *Commiphora africana* J | 0.166037422 | 2.756842546 | 0.6 | C | -2.77943703 | 4.586019463 | 5.7 |
| Vp-3 | -0.16121623 | 2.599067281 | 0 | X70 | -2.31001958 | 4.506367323 | 5.3 |
| *Hyparrhenia dissoluta* | 0.156070707 | 2.435806548 | 0.3 | A | -2.20453154 | 4.225615132 | 5.6 |

| Species | coord4 | CR4 | P4 | Relevés | coord4 | CR4 | P4 |
|---|---|---|---|---|---|---|---|
| *Vernonia miombicola* | 0.14778861 | 2.184147321 | 0.5 | X65 | -1.62449467 | 3.028676302 | 11 |
| *Commelina benghalensis* | -0.13992313 | 1.957848286 | 1.5 | B | -2.16552521 | 2.945608664 | 12 |
| *Cassia mimosoides* | 0.134170115 | 1.800161964 | 1.5 | X44 | 1.509249506 | 2.860616683 | 13 |
| Vp-2 | 0.132809187 | 1.763828012 | 0.6 | X66 | -1.53380739 | 2.828601283 | 14 |
| *Lannea humilis* J | 0.130763971 | 1.709921613 | 2.2 | X46 | 1.424084148 | 2.51071491 | 17 |
| *Spermacoce pusilla* | 0.127416 | 1.62348371 | 2.7 | X51 | 1.35244285 | 2.447098411 | 16 |
| Structure surface5 | 0.125320033 | 1.570511076 | 1.8 | X38 | 1.173442186 | 2.174527384 | 20 |
| *Asparagus abyssinicus* | 0.1251553 | 1.566384922 | 1.3 | X59 | -1.32116056 | 2.129784595 | 21 |
| *Ruellia patula* | 0.122381813 | 1.497730823 | 2.8 | X42 | 1.237446892 | 2.090509917 | 20 |
| **Species** | **coord4** | **CR4** | **P4** | **Relevés** | **coord4** | **CR4** | **P4** |
| Texture profondeur5 | -0.1678977 | 2.818963631 | 0 | B | -4.15917838 | 10.86584878 | 0.4 |
| pH surface-3 | -0.16699611 | 2.788770145 | 0 | C | -3.90041273 | 9.031156027 | 0.4 |
| Texture surface5 | -0.16635183 | 2.767293005 | 0 | A | -2.60607131 | 5.905135917 | 2.7 |
| Exposition0 | -0.16339717 | 2.669863571 | 0 | X5 | -2.18215649 | 5.815878546 | 2.9 |
| Horizon humifère-1 | 0.157383284 | 2.476949803 | 0.2 | X2 | -2.14351567 | 5.721111775 | 3.4 |
| Matière organique-1 | 0.155874421 | 2.429683522 | 0 | X1 | -1.9128186 | 4.655938262 | 5.5 |
| Structure surface6 | -0.15535249 | 2.413439501 | 0 | X4 | -1.90747218 | 4.417840847 | 4.3 |
| Altitude 1 | -0.14753106 | 2.17654126 | 0.9 | X3 | -1.55951203 | 3.271190874 | 11 |
| *Acacia polyacantha* A 5 | -0.14712738 | 2.164646647 | 0 | X11 | 1.335879318 | 2.856359139 | 15 |
| *Sporobolus stapfianus* | 0.143867908 | 2.069797489 | 0.3 | D | -1.87062469 | 2.414327491 | 16 |
| Exposition2 | 0.143556393 | 2.06084381 | 0.3 | X63 | 1.526883006 | 2.343020883 | 20 |
| *Acacia hockii* A | -0.14081645 | 1.982927159 | 0 | X64 | 1.282760866 | 2.00971719 | 23 |
| *Acacia senegal* J | 0.138178151 | 1.909320131 | 0 | X10 | 1.071095003 | 2.001291009 | 20 |
| Vs-3 | -0.13557389 | 1.838027992 | 1.6 | X7 | -1.18026072 | 1.901756917 | 22 |
| *Albuca kirkii* | 0.131627 | 1.732566709 | 2.2 | X70 | 1.475432577 | 1.838374567 | 25 |

Table 15. First fifteen species and first fifteen relevés arranged in order of decreasing relative contributions. Axes 1 to 4.

On axis 1, several species of *Poaceae* such as *Pennisetum plystachion*, *Botriochloa insculpta*, *Heteropogon contortus*, with *Hibiscus cannabinus*, *Albuca kirkii* and young individuals of *Lannea humilis* and *Acacia senegal*, are found together, all associated with a content of exchangeable cations of level 2 at depth, a pH of level 3 at depth and level 4 on the surface, a base saturation rate of level 3 and a dense cover (thicket type) absent. The whole is opposed to young and adult *Markhamia obtusifolia*.

On axis 2, the environmental parameters of superficial open environments (with no soil horizon other than superficial) poor in organic matter and exchangeable cations with *Microchloa kundtii* and young *Combretum molle* are brouht together. In contrast, we find moderately closed environments in the lower part of the transect at Lake Ihéma with *Plectranthus pauciflorus* and young *Ormocarpon trichocarpon*.

On axis 3, we find *Achyranthes aspera* var. *sicula* and *Commelina benghalensis* associated with base saturation rates of level 4 on the surface and a concentration of exchangeable cations of level 3 on the surface, all opposed to concentrations of exchangeable cations of level 2 on the surface with young *Commiphora africana, Hyparrhenia dissoluta, Vernonia miombicola, Cassia mimosoides, Lannea humilis* young, *Spermacoce pusilla, Asparagus abyssinicus* and *Ruellia patula*.

On axis 4, it appears that the *Acacia polyacantha* forest gallery and some nearby relevés are found in the lower part of the transect, on soils rich in organic and mineral matter, at the edge of the lake. The whole is opposed to various more open environments.
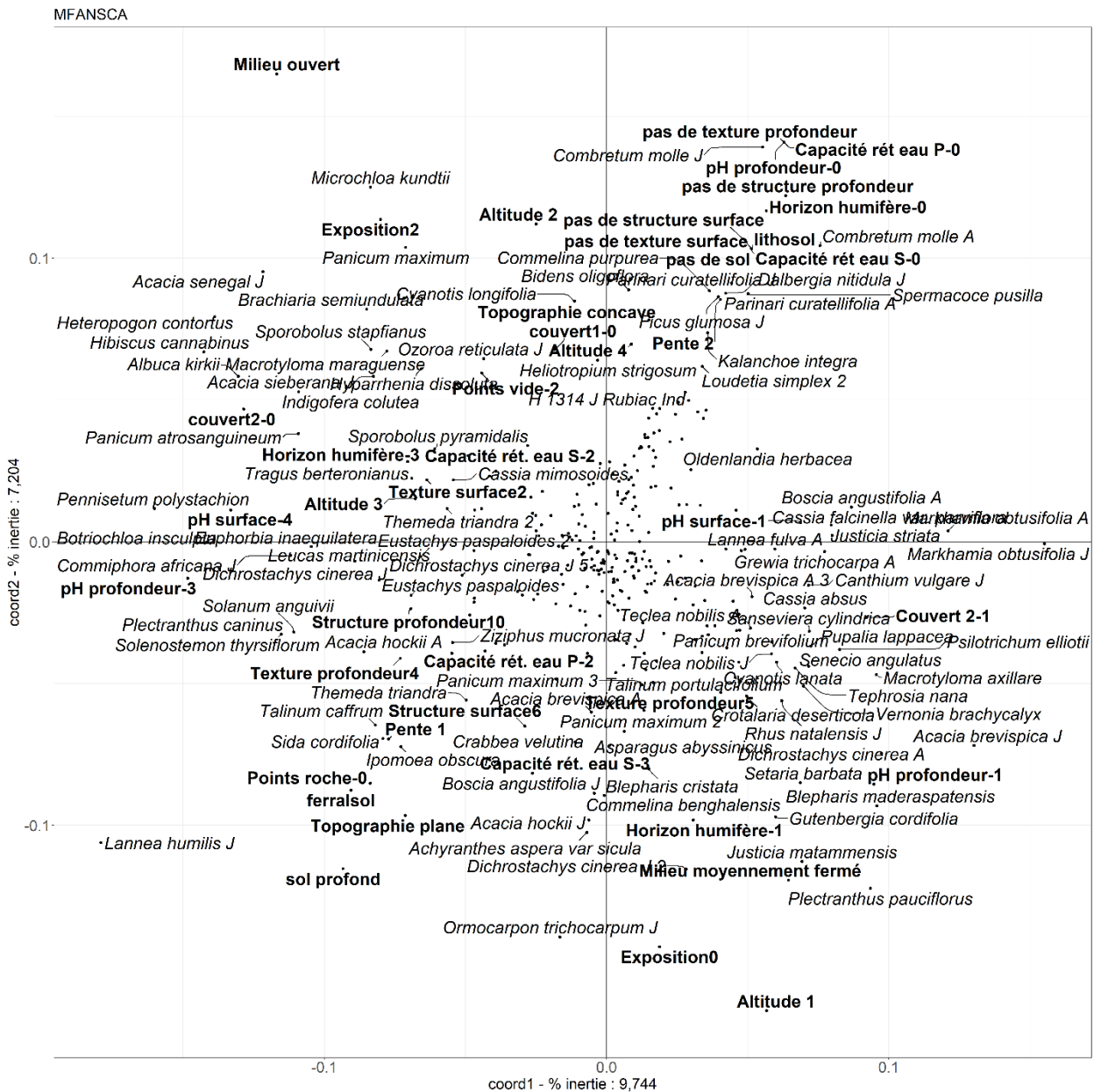
Figure 39. Multiple factor analysis of the united floristic and environmental Akagera files. Coordinates of species and mesological parameters on axes 1 and 2.

This example of analysis combining floristic and environmental variables shows the benefit of first transforming the raw tables into disjunctive tables.

# Discussion and conclusions

This example develops mainly an approach, explaining how one can gradually get interpretable results by adapting the data and in striving to find the best possible match between the data and the multivariate techniques. It clearly appears that it is never superfluous to examine all the various steps of calculation of the multivariate analyses to optimize their use: transformed table for PCA, CA or NSCA, correlation, covariance, or distance matrix for all techniques, eigenvalues and their significance, coordinates of lines or columns on the axes. From the presented data and many others, I may

conclude that the form of the distribution of the covariance, distances or dissimilarities coefficients is an essential step for deciding if the multivariate analysis is well suited to data. A very important asymmetry, with a very low proportion of extreme coefficients is a cause of non-use as a very small part of the data table is highlighted. A perfect symmetry with a Gaussian distribution of the coefficients is an ideal situation, especially when the proportion of the central values of the distribution is not too high (less than 20 %). With vegetation data, PCA and CA suffer from this situation, partly because the proportion of empty cells in the table of data is important, which is a very common situation. With NSCA of simple or complete disjunctive table, the asymmetry is less pronounced than with CA. In the tables of distance or similarity, the distribution of coefficients is often symmetric, but sometimes with low proportions of extreme values. In addition, PCA and more specifically CA are very sensitive to local species patterns; the presence of several rare species in one or a small number of relevés makes both techniques practically unusable. The performances of PCA or CA are also often associated with linear or unimodal relations between the variables (ESPINAR *et al.*, 2002), which is not very common with vegetation data, especially because of sampling techniques that consist in relevés separated by relatively large space.

The use of permutation tests is a strong appreciable complement. I recommend the use of such aid, which allows to get out the species in order of decreasing relative contribution, *i.e.* the species that play an important role in the calculations linked to the axes. What's more, when a test by permutation is associated with the relative contributions, one adds an inferential dimension to the interpretation.

Here are now the more precise conclusions on each of the developed analyzes:

A principal component analysis of raw data only provides fairly poor results, dominated by localized species. A transformation into ordinal data does not improve the results.

Correspondence analysis of ordinal data performs even worse than principal component analysis by exacerbating local variations relative to larger plant formations that are significantly overwritten in the analysis.

The principal coordinates analysis of ordinal data provides better dispersion of the relevés, but the extreme data are highlighted and interpretation is not easy. Only relevé coordinates are provided. In addition, there is no help for interpretation apart from the eigenvalues. The type of distance has little impact on interpretation.

Multidimensional non-metric analyzes of ordinal data also produce a good dispersion of relevés but again only produce the relevé coordinates and no aid to interpretation. In addition, the choice, initially of the number of axes to extract, has an impact on the results, which makes the interpretation of the results quite delicate.

With the analysis of a simple disjunctive table, correspondence analysis turns out to be completely unsuitable and it should no longer be used. The same goes for variants created like CCA or DCA, ... which also suffer from the same basic defect. They are strongly discouraged.

The non-symmetrical correspondence analysis of the simple disjunctive table produces a nice view of the entire transect and then more local variations. Interpretation aids are useful. The synthetic nature of the analysis is certain with only the first five significant eigenvalues.

The principal coordinates analysis produces a good dispersion of the relevés that is fairly easy to interpret on the first axes but does not provide coordinates for the species or help for interpretation.

A non-symmetrical correspondence analysis of a disjunctive table of type 113 also produces good results that are quite easy to interpret like those of the simple disjunctive table.

The non-symmetrical correspondence analysis of abundance data in logarithmic version always favors abundant species, but in an attenuated manner compared to an untransformed version. There is therefore a choice to make if we want to favor abundance without giving them too much weight. In such an analysis, there is only one row per species.

The multiple factorial analysis of a disjunctive table, with three subtables, including one for woody species, clearly favors the latter. There is therefore a choice to make if we wish to highlight a stratum or a category of species. The same analysis with file 113 also provides results that are easy to interpret and it is again a choice to make between the two types of disjunctive table, the last giving a little more importance to simple presences compared to abundances. With the logarithmic file, abundant species are again favored, but less clearly than with abundances 12345.

If we wish to bring together the floristic and mesological tables, a multiple factor analysis based on the non-symmetrical correspondence analysis of disjunctive tables appears to be a very powerful technique for revealing the parallelism between species and mesological factors, without prejudging the type of relationship between these parameters.

The problem with multivariate analyzes is finding adequate weighting for the subarrays. With disjunctive tables, the proportion of presences of each subtable was chosen. In this case, having one or more empty columns of a sub-table does not prevent the analysis, as is the case with the choice of the first eigenvalues as weighting parameters.

We strongly recommend non-symmetrical correspondence analysis. Its main quality is its low sensitivity to rare species; thus a complete table, with all recorded species, can be analyzed. With any ordinal scale or a Braun-Blanquet scale in particular, I always recommend transforming the data in order to avoid the weight of the table being concentrated in a small number of abundant species. The advocacy for the use of disjunctive tables is neither new nor original. BOUROCHE and SAPORTA (1980) underlined this in these terms: "Putting it in disjunctive form is much more than a mathematical convenience and this for several reasons. Since a numerical character can be transformed into a qualitative character by dividing its values, it is possible to study tables containing a mixture of numerical and qualitative characters: it is enough to make everything qualitative and to carry out a multiple correspondence analysis. The advantage is that by proceeding in this way we can take into account possible non-linear connections between characters".

Our analysis of the Akagera file does not take into account any a priori paradigm. The gradient paradigm could have been favored as the transect studied goes from the edge of a lake to a dry and rocky hill but it does not seem possible to reduce this vegetation to one or more simple gradients.

# References

AUSTIN, M.P. 1977. Use of ordination and other multivariate descriptive methods to study succession. *Vegetatio* 35: 165-175.

AUSTIN, M.P. 2012. *Vegetation and environment: discontinuities and continuities*. In Vegetation Ecology, second edition. Edited by Eddy van der Maarel and Janet Franklin. Wiley-Blackwell.

AUSTIN, M.P. 2013. Inconsistencies between theory and methodology: a recurrent problem in ordination studies. *Journal of Vegetation Science* 24: 251-268.

AUSTIN, M.P., NICHOLLS, A.O., DOHERTY, M.D. & MEYERS, J.A. 1994. Determining species response functions to an environmental gradient by means of a β-function. *Journal of Vegetation Science* 5: 215-228.

BONIN, G. & ROUX, M. 1978. Utilisation de l'analyse factorielle des correspondances dans l'étude phyto-écologique de quelques pelouses de l'Apennin lucano-calabrais. *Oecologia Plantarum* 13: 121-138.

BONIN, G., AUBERT, G., BARBERO, M., GAMISANS, J., GRUBER, M., LOISEL, R., QUEZEL, P., SANDOZ, H., THINON, M. & VEDRENNE, G. 1983. Mise en évidence de la dynamique de quelques écosystèmes forestiers et préforestiers provençaux aux étages méditerranéens S.L. à l'aide des taxons indicateurs. *Vegetatio* 54: 79-96.

BOUROUCHE, J.-M. & SAPORTA, G. 1980. *L'analyse des données*. P.U.F. Collection Que sais-je ?

BOUXIN, G. 1975a. Ordination and classification in the savanna vegetation of the Akagera park (Rwanda, Central Africa). *Vegetatio* 29: 155-167.

BOUXIN, G. 1975b. Ordination of quantitative and qualitative data in a savanna vegetation (Rwanda, Central Africa). *Vegetatio* 30:197-200.

BOUXIN, G. 1983. Multi-scaled pattern analysis: an example with savannah vegetation and a proposal for a sampling design. *Vegetatio* 52: 161-169.

BOUXIN, G. 1986. Le traitement statistique des tableaux de végétation. I. Les petits tableaux. *Biom. Praximétrie* 26: 49-72.

BOUXIN, G. 1987a. Le traitement statistique des tableaux de végétation. II. Les ensembles de tableaux et les grands tableaux. *Biom. Praximétrie* 27: 65-97.

BOUXIN, G. 1987b. La végétation herbacée des ruisselets d'un massif boisé sur roches éodévoniennes, en Belgique. *Coll. Phytosociol*. XV: 93-106.

BOUXIN, G. 1991. La végétation aquatique et du bord de l'eau dans le bassin versant du Bocq (Condroz, Belgique). *Revue des Sciences de l'Eau* 4: 185-210.

BOUXIN, G. 1995. Démarche conduisant à la mise au point d'une technique de description de la végétation. *Acta Botanica Gallica* 142: 533-540.

BOUXIN, G. 1999. Description de la végétation aquatique et du bord de l'eau dans le bassin hydrographique de la Molignée (Condroz, Belgique) par l'analyse multiscalaire des motifs. *Ecologie,* t. 30: 139-163.

BOUXIN, G. 2011. Évolution de la végétation macrophytique et trophie dans deux ruisseaux du bassin hydrographique de la Molignée (Condroz, Belgique). *Revue des Sciences de l'Eau* 24: 253-266.

BOUXIN, G. 2013**.** Végétation macrophytique, environnement et qualité d'eau dans le bassin versant du Bocq (Belgique, Wallonie). *Revue des Sciences de l'Eau* 26: 1-19.

DEBUSSCHE, M., ESCARRÉ, J., LEPART, J., HOUSSARD, C. & LAVOREL, S. 1996. Changes in Mediterranean plant succession: old-fields revisited. *Journal of Vegetation Science* 7: 519-526.

DEVINEAU, J.-L. 2001. Les espèces ligneuses indicatrices des sols dans des savanes et jachères de l'Ouest du Burkina Faso. *Phytocoenologia* 31: 325-351.

DRAY, S., DUFOUR, A.B. & CHESSEL, D. 2007. The ade4 package - II: Two-table and K-table methods. *R News* 7(2): 47-52.

ESCOFIER, B. & PAGES, J. 1984. *L'analyse factorielle multiple : une méthode de comparaison de groupes de variables.* In: Data Analysis and Informatics III. Diday, E. & Coll. (Eds.) Elsevier, North-Holland. 41-55.

ESCOFIER, B. & PAGES, J. 1986*. Le traitement des variables qualitatives et des tableaux mixtes par analyse factorielle multiple*. In: Data Analysis and Informatics IV. 179-191. Diday, E. & Coll. (Eds.) Elsevier, North-Holland.

ESCOFIER, B. & PAGES, J. 1989. *Multiple factor analysis: results of a three-year utilization*. In: *Multiway data analysis*. 277-285. Coppi, R. & Bolasco, S. (Eds.) Elsevier Science Publishers B.V., North-Holland.

ESPINAR, J.L., GARCÍA, L.V., GARCÍA MURILLO, P. & TOJA, J. 2002. Submerged macrophyte zonation in a Mediterranean salt marsh: a facilitation effect from established helophytes? *Journal of Vegetation Science* 13: 831-840.

FEHMI, J. 2010. Confusion among three common plant cover definitions may result in data unsuited for comparison. *Journal of Vegetation Science* 21: 273-279.

FOUCART, T. 1982. *Analyse factorielle. Programmation sur micro-ordinateurs*. Masson, Paris.

GAUCH, H.G., Jr. 1982. *Multivariate analysis in community ecology*. Cambridge University Press.

GAUTHIER, S., BOUCHER, D., MORISSETTE, J. & DE GRANDPRÉ, L. 2010. Fifty-seven years of composition change in the eastern boreal forest of Canada. *Journal of Vegetation Science* 21: 772-785.

GIMARET-CARPENTIER, C, CHESSEL, D. & PASCAL, J.-P. 1998. Non-symmetrical correspondence analysis: an alternative for species occurrences data. *Plant Ecology* 138: 97-112.

GOODALL, D.W. 1954. Objective methods for the classification of vegetation. III. An essay in the use of factor analysis. *Austral. J. Bot*. 2: 304-324.

GOWER, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53 : 325-338.

GREENACRE, M. 2010. Correspondence analysis of raw data. *Ecology* 91: 958-963.

GREENACRE, M. 2013. Fuzzy coding in constrained ordinations. *Ecology* 94: 280-286.

GREIG-SMITH, P. 1964. *Quantitative plant ecology (second edition)*. Butterworths, London.

HAURY, J. & BAGLINIERE, J.-L. 1990. Relations entre la population de truites communes (*Salmo trutta*), les macrophytes et les paramètres du milieu sur un ruisseau. *Bull. Fr. Pêche Piscic*. 318: 118-131.

HUSSON, F., LÊ, S. ET PAGÈS, J. 2009. *Analyse des données avec R*. Presses Universitaires de Rennes.

KNOX, R.G. & PEET, R.K. 1989. Bootstrapped ordination: a method for estimating sampling effects in indirect gradient analysis. *Vegetatio* 80: 153-165.

KRUSKAL, J.B. 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29: 1-27.

KRUSKAL, J.B. 1964b. Nonmetric multidimensional scaling: a numerical method. *Psychometrika* 29 : 115-129.

MANLY, B.F.J. 1997. *Randomization, bootstrap and Monte Carlo methods in biology*. Second edition. Chapman & Hall.

MCCUNE, B & MEFFORD, M.J. 1999. *PC-Ord. Multivariate analysis of ecological data, version 4*. MjM Software Design, Gleneden Beach, OR, US.

MINCHIN, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio* 69: 89-107.

ORLÓCI, L. & KENKEL, N.C. 1985. Introduction to data analysis, with examples from population and community ecology. International Co-operative Publishing House. *Statistical Ecology Monographs*, 1: 1-339.

PALM, R. 1993. Les méthodes d'analyse factorielle : principes et applications. *Notes Stat. Inform*. (Gembloux) 93/1.

PIELOU, E.C. 1977. *Mathematical Ecology*. Wiley-Interscience.

PILLAR, V.D.P. 1999. The bootstrapped ordination re-examined. *Journal of Vegetation Science* 10: 895-902.

PODANI, J. 2006. Braun-Blanquet's legacy and data analysis in vegetation science. *Journal of Vegetation Science* 17: 113-117.

PRENTICE, I.C. 1977. Non-metric ordination methods in ecology. *Journal of Ecology* 65: 85-94.

PRENTICE, I.C. 1980. Vegetation analysis and order invariant gradient models. *Vegetatio* 42: 27-34.

REED, R.A., PEET, R.K., PALMER, M.W. & WHITE, P.S. 1993. Scale dependence of vegetation-environment correlations: A case study of a North Carolina piedmont woodland. *Journal of Vegetation Science* 4: 329-340.

ROOT, H.T. & NELSON, P.R. 2011. Does philogenetic distance aid in detecting environmental gradients related to species composition? *Journal of Vegetation Science* 21: 1143-1148.

STEVENS, C., DUPRÉ, C., GAUDNIK, C., DORLAND, E., DISE, N., GOWING, D., BLEEKER, A., ALARD, D., BOBBINK, R., FOWLER, D., VANDVIK, V., CORCKET, E., MOUNTFORD, J.O., AARRESTAD, P.A., MULLER, S. & DIEKMANN, M. 2011. Changes in species composition of European acid grasslands observed along a gradient of nitrogen deposition. *Journal of Vegetation Science* 22: 207-215.

TALBOT, J., RICHARD, P.J.H., ROULET, N.T. & BOOTH, R.K. 2010. Assessing long-term hydrological and ecological responses to drainage in a raised bog using paleoecology and a hydrosequence. *Journal of Vegetation Science* 21: 143-156.

TROUPIN, G. 1982. *Flore des plantes ligneuses du Rwanda*. Institut national de la Recherche scientifique, Butare, République Rwandaise. Publication n° 21.

TROUPIN, G. 1983. *Flore du Rwanda. Spermatophytes, Volume II*. Musée royal de l'Afrique centrale – Tervuren, Belgique. Annales-Série IN-8°-Sciences économiques – n°13.

TROUPIN, G. 1985. *Flore du Rwanda. Spermatophytes, volume III.* Institut national de la Recherche scientifique, Butare. Publication n° 30.

VON WEHRDEN, H., HANSPACH, J., BRUELHEIDE, H. & WESCHE, K. 2009. Pluralism and diversity: trends in the use and application of ordination methods 1990-2007. *Journal of Vegetation Science* 20: 695-705.

WHITTAKER, R.H. 1967. Gradient analysis of vegetation. *Biol. Rev*. 42: 207-264.

WHITTAKER, R.H. 1973. *Ordination and classification of communities.* Handbook of Vegetation Science. Part V. Dr. W. Junk b.v. - Publishers - The Hague.

WILDI, O. 2010. *Data analysis in vegetation ecology*. Wiley-Blackwell. 211 pp.

WILDI, O. 2013. *Data analysis in vegetation ecology* (second edition). Wiley-Blackwell. 301 pp.

WILSON, J.B. 2011. Cover plus: ways of measuring plant canopies and the terms used for them. *Journal of Vegetation Science* 22: 197-206.