International Journal of Research Studies in Biosciences (IJRSB)
Volume 2, Issue 11, December 2014, PP 85-94
ISSN 2349-0357 (Print) & ISSN 2349-0365 (Online)
www.arcjournals.org

Evaluation of the Statistical Methods for the Classification of *Plantain* Hybrids According to their Sensibility to the Black Leaf Streak Disease (BLSD)

Fabrice Monkam Tchamaha [1], Charles Carnot Asseng*[1]

[1]Université de Douala (Cameroun)
Faculté des Sciences
Laboratoire de Biologie et Physiologie des Organismes Végétaux, Douala.
**carnotass@yahoo.fr, carnot@mail.ru.*

Josué Ngando Essoh[2], Pascal Noupadja[2]

[2]CARBAP (Centre Africain de Recherches sur Bananiers et Plantains)
Douala-Cameroun.

Abstract: *Preliminary evaluation of the varietal sensibility allowed us to classify, in a statistical approach, various varieties of hybrids according to their sensibility in the Black Leaf Streak Disease (BLSD). The present study was realized in CARBAP (Centre Africain de Recherches sur Bananiers et Plantains). Some epidemiological parameters such as: younger leaf with streaks (YLS), younger leaf necrotic spot (YLN), number of leaf (NL) constituted the database for the classification. The statistical methods of classification used in this study were adapted to one size (univariate) and several size (multivariate). The several size method was subjected of a principal components analysis (PCA) and a hierarchical clustering (AHC). Of this study emerges that on all the compared and classified hybrids, there is obtaining three classes: a class of the sensitive hybrids, a class of the partially resistant hybrids and a class of the resistant hybrids which has to benefit particular observations.*

Keywords: *Varietal sensibility, BLSD, PCA, AHC.*

1. INTRODUCTION

The Cameroonian production of bananas and plantains is limited by attacks of phytopathogenic such as: viruses, bacteria and fungi. Among fungi we notice the constant presence of Mycosphaerella fijensis, Morelet causal agent of the BLSD which is the main parasitic constraint over the world area of banana and plantain production. This disease cause necrotic spot and dry out leaves reducing their area. Such damages can lead losses of output of the order of 50 to 100% (Moulioum Pefoura, on 1996, 2001b ; and Jones, 2000). The chemical struggle is considered often but this one is very costly and the culture in association systems return chemical struggle maladjusted in peasant environment. The resistant varieties in the BLSD create in the stemming from the hybridization process by the lined pure parents, go successively, during the following generations, undergo a series of evaluation and selection, on the basis of arbitrary statistical analyses of the data and the epidemiological parameters in the field.

The main objective of this study aims at making a contribution in the purification of the plantain banana tree, with the aim of increasing its productivity while selecting the hybrids resisting the BLSD in a statistical approach. This contribution aims specifically at:

- Classifying a statistical approach of the various varieties of hybrids according to their sensibility in the BLSD.
- Estimating the main statistical methods of classification used.

2. MATERIAL AND METHODS

The hybrids of plantains bananas used within the framework of this study are produced in the Laboratory of Genetic Improvement of the CARBAP.

CARBAP has a collection of musacée containing 700 varieties of plantains bananas belonging to diverse genomic group and native of diverse geographical regions in the world. Among these accessories, we count more than 150 varieties of plantains bananas. It under group expresses a big morphological variability in spite of its genetic homogeneity (Noupadja, 2004). All these cultivars of plantains bananas are sensitive to the BLSD. 176 listed hybrids were set up according to the two strategies of the conventional improvement of plantains bananas (Kodjo Tomekpe, et al. 2004).

The first and oldest (1920) is the strategy 3x/2x exploited to create some resistant hybrids tétraploïd to diseases and having better values agronomic in pollinated of the plantains bananas sensitive triploïd with fertile and resistant male diploid.

The second is strategy 4x/2x where hybrid triploid is create by hybridization between a parent diploid and a parent tétraploid gotten previously by doubling chromosomal to the colchicines of an ancestor diploid or a hybrid diploid. The experimental plot of land of 2500 m2 of surface, on which were planted hybrids separated by a distance of 2 m on 3 m, does not present inoculum spores of mushrooms and no fungicide treatment was made (figure 1).

- **Strategy 3x/2x**

Plantain AAB (sensitive) x banana diploïd AA (resistant) → hybrid tétraploïd AAAB (resistant)

- **Strategy 4x / 2x**

Development of tétraploïd

AA_{w}* x AA_{cv}**
AA_{cv} x AA_{cv} } → AA_{cv} $\xrightarrow{\text{Colchicine}}$ $AAAA_{cv}$

AA_{cv} x BB_{w} AB_{w} → $\xrightarrow{\text{colchicine}}$ $AABB_{cv}$

Development of triploïd

AA_{cv} x $AAAA_{cv}$ → AAA_{cv}
AA_{w} x $AAAA_{CV}$ → AAA_{CV}
BB_{w} X $AAAA_{CV}$ → AAB_{CV}
AA_{CV} X $AABB_{CV}$ → AAB_{CV}
AA_{w} X $AABB_{CV}$ → AAB_{CV}

** : w = wild ** : cv = cultivar*

Figure1. *Strategies of conventional improvement of plantains bananas (Kodjo Tomekpé, et al., 2004).*

For every hybrid, no rehearsal is made but 6 observations in the various dates were made on the ground in 9 months. During these observations, 3 principal parameters which will allow to determine the degree of hybrids sensibility to the BLSD are found:

- younger leaf with streaks (YLS): at the stage 3 of the BLSD;
- younger leaf necrotic spot (YLN): younger leaf presenting the necrosis tasks ;
- number of leaf (NL): leaf capable of realizing the photosynthesis.

The first part consists in making a univariate analysis concerning the qualitative variables (YLS and YLN) and the mode (modality having the biggest staff) of every hybrid is pointed. When there are 2 modes for the same modality, find the central class and take out again a hybrid-mode table.

Central class $= \frac{\text{Mode1} \ + \text{Mode 2}}{2}$.

Afterward, a deduction of the general mode of all the hybrids is made and every mode of each hybrids is compared with the general mode. Of this comparison a board presenting the frequencies of the various modalities of every parameter is drawn up. Then a sorting in increasing order on all the values of every parameter is made. From this sorting we proceed to a distribution of hybrids in 2 classes (low variable, high variable) by basing itself on the general mode. And finally, every stage is illustrated by graphs (Camembert and diagram in stick). For the quantitative variable (NL), the average of the NL is found for every hybrid and the construction of hybrid - average NL table is made.

$$\text{average} = \frac{\sum \text{Observations in the various dates}}{\text{total number of observations date}}$$

In the same way as previously, the comparison of the averages for every hybrid to the overall average of NL is realized; then a sorting in the increasing order is made on all the data of the NL. Then a distribution of hybrids in two classes stood out by basing itself on the overall average and finally the sorting is Illustrated by a graph.

In a second part, a multivariate analysis is made by making in the first place a test of correlation on the various parameters. Then, a representation of the groups, from the point of view of an analysis of the main components (AMC), is illustrated for a visualization of the correlations between the various parameters, And other one of hybrids, in a space in two or three dimensions, to identify homogeneous groups of hybrids, or unlike the a typical hybrids. Finally a hierarchical organization by the hierarchical ascending classification which ends in a tree of classification (dendrogram) by successive objects aggregations.

Analysis software: the software used for our work are mainly

- the spreadsheet Excel (Microsoft office, 2007) which allowed us to treat our data and to rerelease a descriptive analysis from it;
- the statistical R software (i386 3.0.1) with which we made the tests of correlation
- the software of analysis XLSTAT (version 2013) which allowed us to make discriminates multivariate analysis: the main constituents analysis and the hierarchical ascending classification.

3. Results

3.1. Univariate Analysis

3.1.1. Classification with Regard to the Younger Leaf with Streaks

The general mode YLS around of which is made the classes distribution is 7. The minimal value YLS is 3 and the maximal value 15 (Table 1).

Table1. *Values of central distributions and scattering YLS.*

General mode YLS	7
Minimal value	3
Maximal value	15
1st Quartile	6.37
Median	7.75
3rd quartile	9
Standard déviation	2.11

The presented frequencies distribution shows a proportion of 55 % hybrids to high YLS and 45 % hybrids to weak YLS (figure 2).

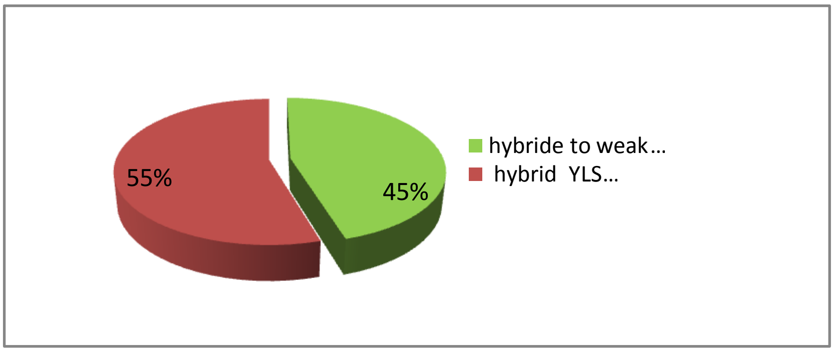

Figure 2. *Circular diagram illustrating the frequency of the hybrids to weak YLS and the one of the hybrids in YLS elevated.*

3.1.2. Classification with Regard to the Younger Leaf Necrotic Spot

The general mode YLN around of which is made the classes distribution is 9. The minimal value YLN is 5 and the maximal value 16,5 (table 2).

Table2. *Values of central distributions and scattering YLN.*

General mode YLN	9
Minimal value	5
Maximal value	16.5
1st Quartile	7.5
Median	9
3rd quartile	10.12
Standard déviation	2.21

The presented frequencies distribution shows a proportion of 43 % hybrids to high YLN and 57 % of hybrids to weak YLN (figure 3).

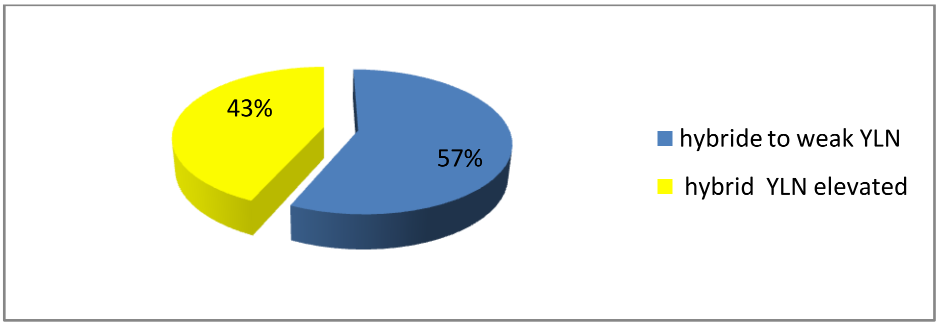

Figure3. *Circular diagram illustrating the frequency of the hybrids to weak YLN and the one of the hybrids in YLN elevated.*

3.1.3. Classification with Regard to the Number of Leaf

The overall average NL around of which is made the classes distribution is 9,58. The minimal value NL is 5,67 and the maximal value 15 (table 3).

Table3. *Values of central distributions and scattering NL.*

General means	9,58
Minimal value	5.67
Maximal value	15
1st Quartile	8.45
Median	9.33
3rd quartile	10.66
Standard déviation	1.60

The frequencies distribution presented in figure 4 show a proportion of 46 % hybrids to high NL and 54 % hybrids to weak NL.

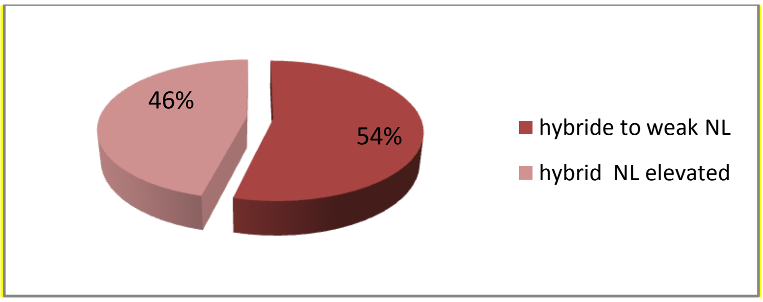

Figure4. *Circular diagram illustrating the frequency of the hybrids to weak NL and the one of the hybrids in NL elevated*

3.2. Multivariate Analyses

3.2.1. Test of Correlation: Test of Pearson

3.2.1.1. Between YLS and YLN

The result of the correlation test at the threshold of 5 % is significant (P-value< 2.2e-16), and the coefficient of correlation R2 = 0.84 gets closer to 1, thus we have a very good positive correlation between the YLS and The YLN; the right of regression inferred by this correlation is given by the following relation: Y= 0.5686X + 0.8209 (figure 5).

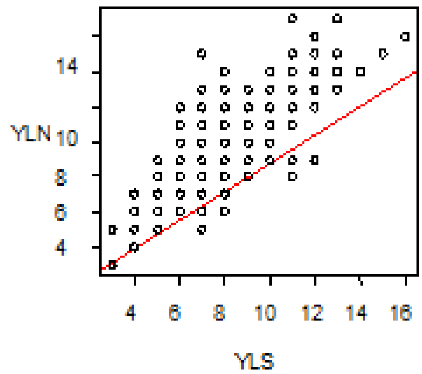

Figure5. *Illustration of the correlation and the regression between the YLS and the YLN*

3.2.1.2. Between YLS and NL

The test of Correlation gives P-value< 2.2e-16 and coefficient of correlation R2= 0.63. There is thus a positive correlation between the NL and the YLS but less good than the previous one. And the right of regression inferred by this correlation is given by the relation Y = 0.6489X + 1.9216 (figure 6).

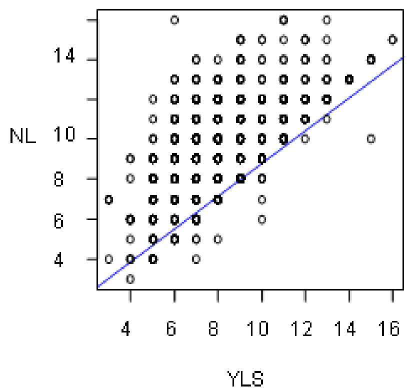

Figure6. *Illustration of the Correlation and the regression between the NL and the YLS*

3.2.1.3. Between NL and YLN

The results of the test show the existence of a positive correlation between the YLN and the NL (p - value < 2.2e-16, R2 = 0.77); the right of regression inferred by this correlation is given by the relation Y = 0.8178 X + 1.3864 (figure 7).

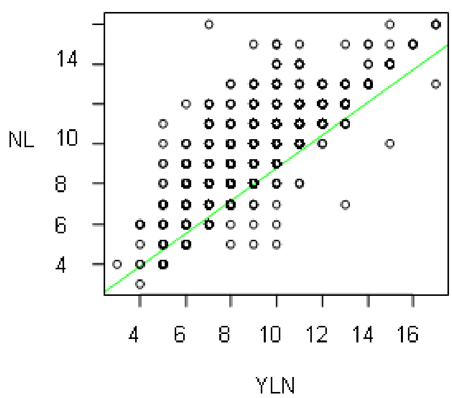

Figure7. *Illustration of the Correlation and the regression between the NL and the YLN*

In seen to the various correlation tests made, it emerges from it that there is a stronger relation between the YLN and the NL. It would thus be obvious that during a preliminary observation, only one of these two parameters is taken into account.

3.2.2. Analysis of the Main Components (AMC)

This analysis presents results according to 2 axes (Figure 8):

- axis F1 classifies hybrids according to them coordinated; those who have negative coordinates (K8, K102...) and those who have positive coordinates (K21, K105...);
- axis F2 classifies hybrids according to the various variables; those who gets closer more to the YLS and the YLN (K20, J695...) and those who gets closer more to the NL (K104, K220...).

Besides the visualization of the correlations appears very clearly and we can from this analysis, confirm the strong correlation which exists between YLN and YLS. Nevertheless the AMC does not show us a net separation between classes and does not present either a hierarchical organization between hybrids. This analysis will thus be completed by the HAC (Hierarchical Ascending Classification).

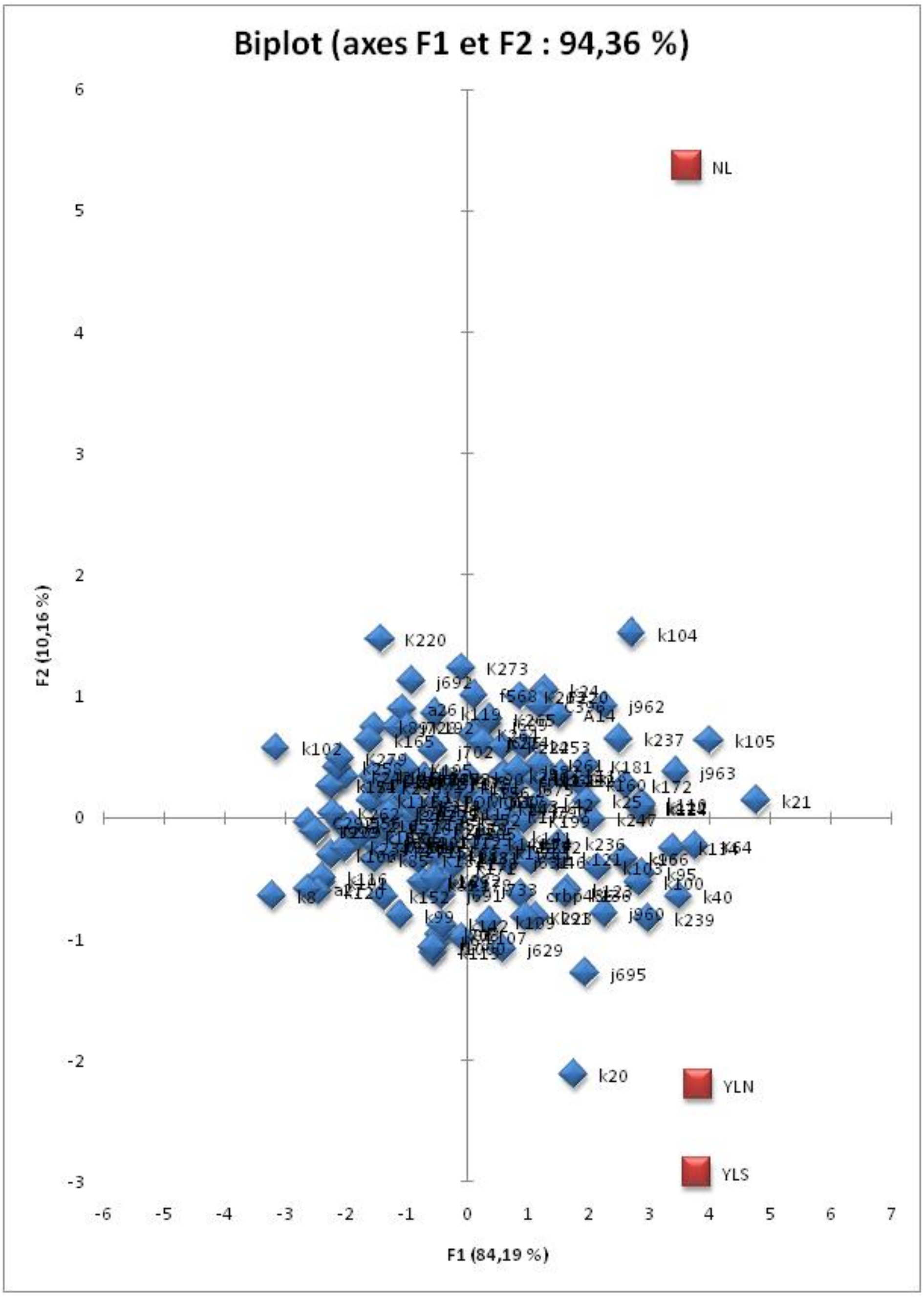

Figure8. *Graph of the Principal component analysis*

3.2.3. Hierarchical Ascending Classification (HAC)

Descriptive results obtained after the HAC application present a partition of hybrids in three classes according to their scores obtained from three analyzed parameters. It emerges from it that the class one contains 62 hybrids, class 2 contains 88 hybrids and finally the class 3 contains 26 hybrids.

The distribution of listed hybrids in every class is presented in tables 4, 5, 6 following ones:

Table4. *Class 1 of the hybrids to weak score exit of the HAC classified by increasing order*

CLASS 1																
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
K78	k84	k85	k89	k92	k101	k102	k106	k111	k116	k151	j692	k294	j702	k234	j716	k238
18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
k120	k157	j726	j728	k242	j985	a27	k245	k135	k165	k167	K52	k8	j556	a26	k145	k170
35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
k174	k258	k259	K262	K268	K271	J577	GALT	K176	K183	J999	C292	J671	K275	K276	K278	K279
52	53	54	55	56	57	58	59	60	61	62						
K192	K195	K198	K201	D248	K209	K213	K214	K216	K220	k207						

Table5. *Class 2 of the hybrids to score middle exit of the HAC classified by increasing order.*

CLASS 2																	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
k79	k80	k83	k86	k90	k91	k98	k99	k107	k108	k109	k112	k113	k117	k148	k152	j685	j687
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
j689	j691	j721	j703	k233	k236	k118	k119	k155	k156	k158	k160	k161	j733	j790	k243	j983	k244
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
j962	k246	k24	k131	k136	crbp613	crbp436	k252	j631	k253	k46	k42	k138	k140	k141	k142	k143	k169
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	77	
k171	j629	J261	K263	K264	K265	K272	K273	F220	K274	K177	K178	K179	K181	K182	J1000	f568	
72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	
C396	J675	A14	Y3	K282	A9	K199	K202	K205	K206	K74	J682	J795	TOMOLO	K211	K223	K224	

Table6. *Class 3 of the hybrids high-scoring exit of the HAC classified by increasing order.*

CLASS 3								
1	2	3	4	5	6	7	8	9
k87	k95	k96	k100	k103	k104	k105	k110	k114
10	11	12	13	14	15	16	17	18
j695	k237	k239	k121	k122	k123	j963	j960	k247
19	20	21	22	23	24	25	26	
k25	k134	k166	k21	k20	k40	k172	K64	

The dendrogram presents three classes differentiated between them by different colors (figure 9). By being interested in the general results of this classification, the report shows that hybrids which have a weak score belong to the class 1. Then, the class 2 possesses hybrids which have an average score. Finally class 3 is the one which possesses hybrids which have a big score. This classification (HAC) thus brings a modification to those made previously according to individual parameters. It's the one which brings the major information in our classification.

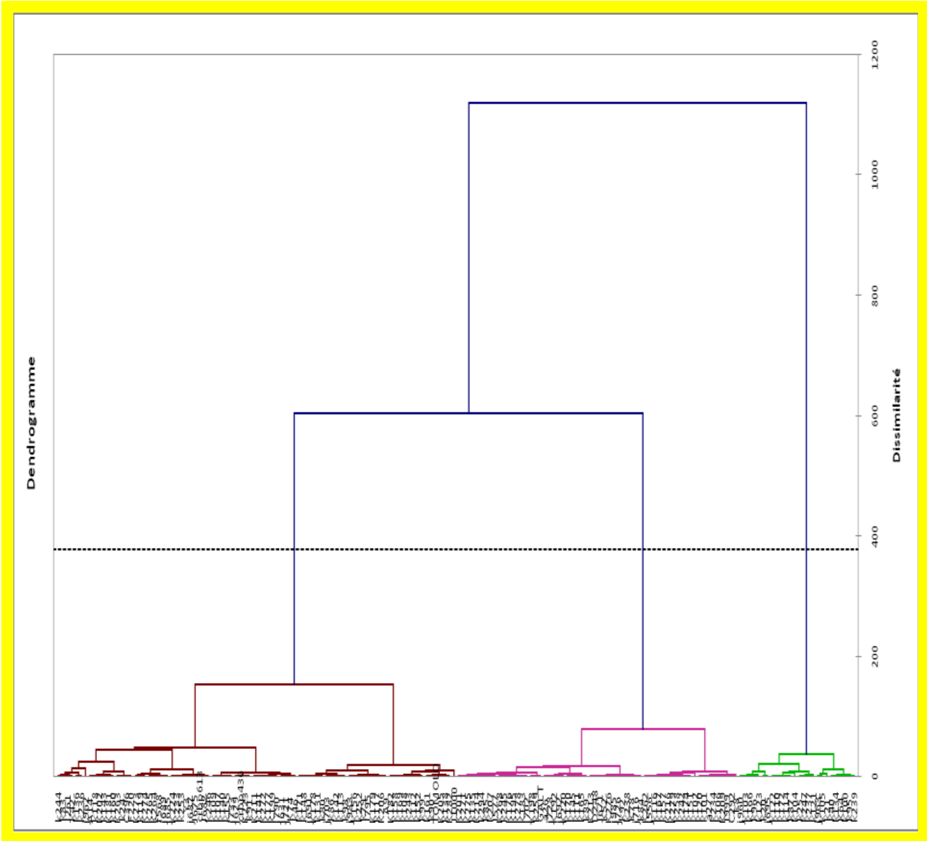

Figure9. *Dendrogram generated by the ascending hierarchical Clustering.*

4. Discussions

4.1. Determination of Sensitive and Resistant Hybrids Classes

The works of Fouré (1982) show that the reaction of diverse accessories of the kind Musa allows to distinguish two types of reactions:

The incompatible interaction which presents the very resistant phenotypes;

The compatible interaction which presents 2 types of phenotype:

- Phenotype with partial resistance (NL > 6 and YLN > 8)
- Sensitive Phenotype (very weak NL).

The results of our study join the conclusions of these works on two bases:

The univariate analysis shows on a first basis a distribution of two classes:

- The class of sensitive hybrids which corresponds to the class of hybrids to weak YLS, YLN, and NL.
- The class of resistant hybrids which corresponds to the class of hybrids to high YLS, YLN and NL.

On a second basis, from the multivariate analyses, three classify that takes out again the HAC appear as follows:

- The class of the sensitive hybrids corresponding to the class 1 of the hybrids with weak score;
- The class of the partially resistant hybrids corresponding to the class 2 of the hybrids with average score;
- The class of the very resistant hybrids corresponding to the class 3 of the hybrids with high score.

4.2. Essential Parameters of Classification

In sight of the various tests made of Correlation, it emerges from it that there is a stronger relation between the YLS and the YLN.

An accurate observation was made on classifications according to the YLS, YLN and that made by the HAC shows a similarity of about 70 % between these 3 classifications. These two parameters (YLS, YLN) can be thus considered as the ones which determine the sanitary state of the plot of land. What joins the works made by Mouliom Pefoura (1984) at the only difference which the tests of correlation allow to assert that during a preliminary observation, only one of these two parameters is taken into account.

4.3. New Resistant Hybrids to be Spread in a Variation Center of the Groups of Plantains Bananas

The hierarchical ascending classification takes out again three classes of the hybrids of which that we shall pay a particular attention (class of resistance fighters). On all the 176 hybrids of the departure 26 hybrids were listed as resistance fighters. These are going to be the object of a next more rigorous observation still called clonale phase. Then will follow successively the agronomic phase, the multi-local phase and finally the participative phase which will allow the distribution of new hybrids. Some of new varieties of bananas plantains will make surface in the world and the CARBAP centers situated in Central Africa more particularly in Cameroon will participate one more time in the variation of these groups of banana trees; characteristic of the regions of western and east Central Africa which are considered as secondary variation centers of the groups of banana trees (Ngueze, on 1994).

4.4. Methodology of Classification

The methodology used for the classification of banana tree plantain hybrids in this study constituted to make an analysis univariate followed by an analysis multivariate which was mainly the object of an AMC and a HAC. This methodology presents a difference compared with that of I. Bermúdez et al used in 2002 to select the vitroplants of ' Manzano ' (AAB) and ' Gros Michel' (AAA) resistant in the fusariose. Their statistical processing consisted of a variance analysis of simple classification called 'One way' to determine the homogeneous and/or significantly different, in 5 %, They used the multiple method of Duncan's ranks with preliminary check of one supposed homogeneity of variance and normality of the data. Should the opposite occur, they applied the test of Dunnett: uncomparison parametric of the averages.

Besides T. Moens (2002) et al in their works on the evaluation of the descent of a crossing between' Pisang Berlin ' and M. acuminata spp. burmannicoides ' Calcutta 4 ' in search of ségrégeantes populations in what concerned the resistance in the BLSD and the nématodes, compares the various lineages F1 by using analysis of variance having transformed the numbers of nématodes by log10 (x+1) and estimated the averages by the test of Waller-Duncan.

5. Conclusion

The study aimed at classifying on the basis of certain epidemiological parameters during a preliminary evaluation a group of hybrids according to their sensibility to the BLSD. Objective was reached by the use of certain tools of statistical analysis suited. The results obtained after analyses univariate allowed us to appreciate 2 main classes in the whole of hybrids: the class of the resistant hybrids and the class of the sensitive hybrids. Every epidemiological parameter presented a frequency different from the distribution of the classes. What entrain afterward the research for the parameters presenting similarity. It emerges from it that classifications made by the YLS and the YLN are similar about 80 %. This percentage was justified by the big positive correlation which exists between these two parameters. On the basis of the multivariate analyses, the results obtained by means of the AHC allowed us to appreciate three classes: the class of the sensitive hybrids, the class of hybrids with partially resistant and the class of the resistant hybrids. We afterward deducted that there is a similarity of about 70 % between the AHC and the classifications made according to the YLS and to the YLN. From this deduction, we thus determined the YLS and the YLN as being the essential parameters of the classification

References

Bikoï A. 1998. Les productions bananières au Cameroun : étude de cas. *In* : bananas and Food Security / Les productions bananières : Un enjeu économique majeur pour la sécurité alimentaire, proct . Int. Symp. Douala, Cameroun, 10-14. November1998 (picq, C. Fouré, E. et Frison, A. E. Eds.) CTA , Montpellier, France. Pp.89 – 101.

FAOSTAT, 2012. http:// apps.fao/default.htm. 03 mai 2013 ; 11h- 44 min

Fouré E., **198**2. Etude de la sensibilité variétale des bananiers et plantains à *Mycosphaerella fijiensis* Morelet au Gabon. I- incubation et évolution de la maladie. *Fruits*, 37 :749-771.

I. **Bermúdez, L. Herrera I.*, P. Orellana, N. Veitía, C. Romero, J. Clavelo, L. García, M. Acosta, L. Garcia R. et Y. Padrón, 2002**. Étude de variants expérimentalement induits de 'Manzano' (AAB) et 'Gros Michel' (AAA) dans le but de sélectionner ceux qui sont résistants à la fusariose. *InfoMusa*-Vol. 11-N°2, décembre 2002 pages 7-8.

Jacobsen K. 2010. The importance of *Pratylenchus goodeyi*on bananas and plantains in cameroun Highlands and development of cultural control methods. Thèse de bioscience Enginnering, FBE, Katholeke Universiteit Leuveun, Allemagne.

Jones D.R., 2000. Diseases of Banana, Abacà and Enset. CABI Publishing.

Kodjo Tomekpe, Christophe Jenny et Jean-Vincent Escalant, 2004, Revue des stratégies d'amélioration conventionnelle de Musa. La Revue Internationale sur Bananiers et Plantains *InfoMusa*-Vol. 13-N02, décembre 2004 pages 2-6.

Mouliom Pefoura A. et Lassoudière A., 1984. Mise au point d'un système d'avertissement biologique pour la lutte contre *Mycosphaerella fijiensis* Morelet, agent de la cercosporiose noire dans les bananeraies camerounaises. *Séminaire WARCOP, 3-12/12/1984*. Libreville, Gabon. 24p.

Mouliom Pefoura A., Lassoudière A., Foko J., Fontem D.A., 1996. Comparison of development of *Mycosphaerella fijiensis* and *Mycosphaerella musicola* on banana and plantain in the various ecological zones in Cameroon. *Plant Disease*, 80 : 950-954.

Noupadja P. **2004**. Evaluation clonale d'hybrides tétraploïdes de bananiers plantains (*Musa spp*.) résistant à la maladie des raies noires. Mémoire d'études approfondies, Faculté des sciences, Université de Yaoundé 1, Cameroun.41 P.

Ngueze, P.B. 1994. Banans and their management. CTA, 145p.

T. Moens, J.A. Sandoval, J.V. Escalant et D. De Waele, 2002. Evaluation de la descendance d'un croisement entre 'Pisang Berlin' et *M. acuminata* spp. *burmannicoides* 'Calcutta 4' à la recherche de populations ségrégeantes en ce qui a trait à la résistance à la maladie des raies noires et aux nématodes *InfoMusa*-Vol. 11-N°2, décembre 2002 pages 20-22.