

“APLICACIÓN DE MÉTODOS ESTADÍSTICOS MULTIVARIADOS EN EL ESTUDIO DE CALIDAD DE ENMIENDAS ORGÁNICAS SÓLIDAS Y LIQUIDAS PREPARADAS EN LAS PROVINCIAS DE GUAYAS, LOS RÍOS Y EL ORO”

Pamela Crow Santos¹, Omar Ruiz Barzola²

¹Ingeniero en Estadística Informática ICM-ESPOL 2006; e-mail: pcrow@espol.edu.ec

²Director de Tesis de Grado, Ingeniero en Estadística Informática, Escuela Superior Politécnica del Litoral (ESPOL), 2002. Profesor del ICM-ESPOL desde 2006, e-mail: oruiz@espol.edu.ec

Resumen: El fin del presente trabajo, fue desarrollar un estudio referencial de las técnicas estadísticas, posibles de utilizar para la valoración de la calidad de productos orgánicos sólidos o líquidos elaborados en fincas bananeras del Litoral Ecuatoriano. Los datos utilizados corresponden a índices químicos, físicos y microbiológicos, de las enmiendas orgánicas tanto sólidas como líquidas. Estos elementos naturales fueron evaluados estadísticamente para luego determinar cual o cuales de ellos están influenciando en la calidad del producto final obtenido. Para evaluar la calidad de las enmiendas orgánicas, se tomo como referencia rangos internacionales. Las muestras para este estudio fueron recolectadas por parte del Centro de Investigación Biotecnológica del Ecuador (CIBE) en diferentes haciendas ubicadas en las provincias de: Los Ríos, El Oro y Guayas. Dentro de los métodos estadísticos utilizados están: (1) el análisis de componentes principales. Esta construye nuevas variables resultado de la combinación lineal de las variables originales, reduciendo así el número de variables y analizándolas como componentes, para la explicación de un todo. El otro método multivariado es (2) el análisis de conglomerados, esta pretende encontrar patrones para clasificar datos multivariados en grupos homogéneos denominados conglomerados. Como último método se tiene (3) el análisis discriminante, este transforma las variables originales en un número pequeño de variables compuestas, denominadas funciones o variables canónicas, que maximizan la variación entre los grupos y minimizan la variación dentro de ellos. Este se lo realiza con dos fines, el descriptivo y el predictivo.

Summary: The main objective of this work was to develop a referential study of the statistical techniques that can be used to quantify organic products' quality (solid or liquid product) made in banana farms from the Ecuadorian coastal region. The data used were chemical, physical and biological indicators of solid and liquid organic fertilizers. These natural elements were evaluated statistically to later determinate which of them are affecting the quality of the final product. To evaluate the organic fertilizers' quality international ranges were taken as a reference. CIBE-ESPOL took the samples from different farms located in following provinces: Los Rios, El Oro and Guayas. The method applied was Principal Components Analysis, which is used to make new variables as a result of linear combination from original variables to reduce the numbers of variables and analyze them. Another method is Clusters Analysis; this pretends to obtain patterns in order to classify Multivariate Data into homogeneous groups named clusters. The last methods used is Discriminant Analysis, this changes the original variables into few combined variables named functions or canonical variables which maximize the variation among the groups and minimize the variation between them. This is all done with two ends: descriptive and predictive.

1. INTRODUCCIÓN

Debido al uso constante, los suelos de las haciendas con actividad agrícola van perdiendo nutrientes. Con la aplicación de las enmiendas orgánicas se trata de recuperar dichos nutrientes para mejorar los suelos.

Por otro lado existen algunas provincias del país en donde las haciendas bananeras orgánicas y convencionales están produciendo y realizando aplicaciones de compuestos naturales denominados “enmiendas orgánicas”.

Existen varios elementos utilizados para la composición de las enmiendas. Estas enmiendas orgánicas, ya sean sólidas o líquidas que son elaboradas y utilizadas en diferentes provincias fueron evaluadas de manera química, física y microbiológica, para determinar la calidad del producto final obtenido.

El uso de la estadística multivariada en los resultados químicos, físicos y microbiológicos obtenidos de las enmiendas orgánicas utilizadas en las diferentes provincias, dará al CIBE una referencia de cual de esos elementos o grupos de elementos, son los que se encuentran en las enmiendas orgánicas que adquieren o preparan los bananeros. Y si estos están aportando a la mejora de la calidad del suelo para los cultivos.

2. Análisis de Componentes Principales.

El Análisis de Componentes Principales (ACP) es una técnica estadística que sintetiza la información, es decir, construye nuevas variables resultado de la combinación lineal de las variables originales, reduciendo así el número de

variables y analizándolas como componentes, para la explicación de un todo.

Para verificar si la técnica de componentes principales puede ser aplicada a un grupo de datos, se debe aplicar la prueba estadística de Bartlett en la cual se plantea el siguiente contraste de hipótesis:

$$H_0: \Sigma = \begin{bmatrix} \sigma_{11} & 0 & \Lambda & 0 \\ 0 & \sigma_{22} & \Lambda & 0 \\ M & M & O & M \\ 0 & 0 & \Lambda & \sigma_{pp} \end{bmatrix}$$

Vs

$$H_1: \text{No es verdad } H_0$$

$$\text{Sea } u = \frac{\det S}{S_{11} + S_{22} + \Lambda + S_{pp}} = \det R, \text{ siendo } S = \hat{\Sigma} \text{ y } R = \hat{\rho}$$

Entonces la región crítica está definida a través

$$\text{de } u' = - \left[\nu - \frac{2p+5}{6} \right] \ln u, \text{ donde}$$

ν : grados de libertad de la matriz de datos = n-1

y u' es aproximadamente $\chi^2(f)$ con

$$f = \frac{p(p+1)}{2}$$

Con (1- α) 100% de confianza se rechaza H_0 a favor de H_1 si, $u' \geq \chi^2_{\alpha, f}$.

H_0 es construido bajo supuestos de normalidad de X , donde $\sigma_{ij} = 0$ garantiza que X_i y X_j son independientes, por lo que no sería aconsejable la aplicación de Componentes Principales.

UTILIZACIÓN DEL BIPLLOT EN EL ACP

Un biplot es una representación gráfica de los datos de una matriz nxp, de donde se seleccionan dos variables a ser representadas.

Con varias variables uno puede construir un gráfico de dispersión de un arreglo matricial, pero este no es un simple gráfico de unidades muestrales. Por otra parte, un gráfico bi-dimensional de unidades muestrales puede ser obtenido al graficar las dos primeras componentes principales. El fundamento del biplot es añadir la información acerca de las variables al gráfico de las componentes principales

3. Análisis de Conglomerado

El "Análisis de Conglomerados" es una técnica multivariada que se utiliza para agrupar observaciones, variables o entidades de un

conjunto de datos en base a sus semejanzas o diferencias. Esta tiene como finalidad ubicar los objetos en grupos o clusters de forma sugerida por los datos, no definidos "a priori", tal que los objetos en un grupo dado tiendan a ser semejantes en algún aspecto y los objetos en diferentes grupos tiendan a ser distintos.

Se utilizó el método jerárquico, el cual parte de n observaciones multivariadas, donde cada una de ellas se considera un conglomerado, a fin de que utilizando distancias entre cada uno de ellos se vaya reduciendo el número de conglomerados hasta llegar a uno solo.

Luego, mediante el método aglomerativo de Ward se determinan progresivamente los conglomerados. Y la representación gráfica de estos conglomerados se denomina "Dendograma".

4. Análisis Discriminante

La técnica multivariada de Análisis Discriminante permite describir algebraicamente las relaciones entre dos o más poblaciones de manera tal que las diferencias entre ellas se maximicen o se hagan más evidente.

El análisis discriminante se realiza con dos enfoques diferentes, ya sea con fines predictivos o explicativos.

En el análisis discriminante predictivo esta relacionados a la clasificación, ya sea de nuevas observaciones u observaciones sobre las cuales no se conoce a qué grupo pertenecen. Esta observación nueva (no usada en la construcción de la regla de clasificación), se asignará al grupo en el cual tienen más probabilidad de pertenecer en base a sus características medidas. Para tal asignación es necesario definir reglas de clasificación.

A diferencia del anterior, en el análisis discriminante descriptivo estamos más interesados en las variables empleadas para diferenciar los grupos, y lo que deseamos es determinar cuáles de esas variables son las que más diferencian a los grupos, cuales son importantes y cuales no a efectos de clasificar los sujetos.

SELECCIÓN DE VARIABLES

En el cálculo de las funciones discriminantes todas las variables (X_1, \dots, X_p) son incluidas simultáneamente, sin considerar la capacidad discriminatoria de cada una de ellas. Sin embargo, si se utiliza un método de selección de variables

para definir las funciones discriminantes, las variables (X1,...,Xp) se irán incluyendo en el modelo una a una, según sea su capacidad discriminatoria. De esta forma, es muy probable que, finalmente, no todas las variables originales se utilicen para construir las funciones discriminantes: únicamente se incluirán aquellas que más contribuyan a separar los grupos, con respecto a un cierto criterio, y se desestimarán aquellas cuya contribución sea escasa, bien porque no son relevantes o porque la información que aportan ya esté recogida por alguna (o varias) de las variables ya incluidas en el modelo y resultan redundantes.

Entre los criterios para decidir qué variables entran o no en la definición de la función discriminante utilizaremos el de minimizar el estadístico Lambda de Wilks.

El estadístico lambda de Wilks también lo utilizaremos para evaluar si las funciones discriminantes canónicas contribuyeron significativamente en la separación de los grupos. Este estadístico mide el poder discriminante de un conjunto de variables y viene dada por:

$$\Lambda = \frac{|W|}{|W + B|} = \frac{1}{\prod_{i=1}^{\min(q-1,p)} (1 + \lambda_i)}$$

y toma valores entre 0 y 1 de forma que, cuanto más cerca de 0 esté, mayor es el poder discriminante de las variables consideradas y cuanto más cerca de 1, menor es dicho poder. Este estadístico tiene una distribución lambda de Wilks con p, q-1 y n-q grados de libertad si se verifica la hipótesis nula:

$$H_0: \lambda_1 = \dots = \lambda_{\min\{q-1,p\}} = 0$$

Vs
H1: No es verdad H0

Donde p es el número de variables que intervienen en la calificación y q el número de grupos.

CLASIFICACION PARA DOS POBLACIONES

La idea de Fisher fue maximizar la distancia estadística entre las medias de los grupos, a través de una selección apropiada del vector de coeficientes de la combinación lineal. La combinación lineal del vector de observaciones ui y el vector de coeficientes Yi es conocida como función lineal discriminante de Fisher de la forma:

$$D = u_1Y_1 + u_2Y_2 + \dots + u_pY_p = u^T Y$$

Es necesario tener en cuenta que, el método de Fisher no requiere del supuesto de normalidad multivariada, también asume matrices de covarianzas homogéneas entre grupos y usa la métrica de Mahalanobis para la discriminación.

CLASIFICACION PARA MAS DE DOS POBLACIONES

El ADC transforma las variables originales en un número pequeño de variables compuestas, denominadas funciones o variables canónicas, que maximizan la variación entre los grupos y minimizan la variación dentro de ellos. La combinación lineal para una función discriminante (DC) puede ser representada por:

$$DC = \mu_1 Y_1 + \mu_2 Y_2 + \dots + \mu_i Y_i = u^T Y$$

Donde μ_i es el coeficiente canónico e Y_i son las variables independientes medidas.

5. Resultados Obtenidos.

ANALISIS DE COMPONENTES PRINCIPALES

Ubicación - Fuente de Microorganismos: Químicos y físicos

Realizada la prueba de significancia estadística de Bartlett, se tiene un valor $p=0.001$, por lo que se concluye que al menos algunas de las variables están correlacionadas y consecuentemente se puede proceder con el análisis de Componentes Principales.

Prueba de Bartlett	
Estadístico de prueba	29.998
Grados de libertad	10
Valor P	0.001

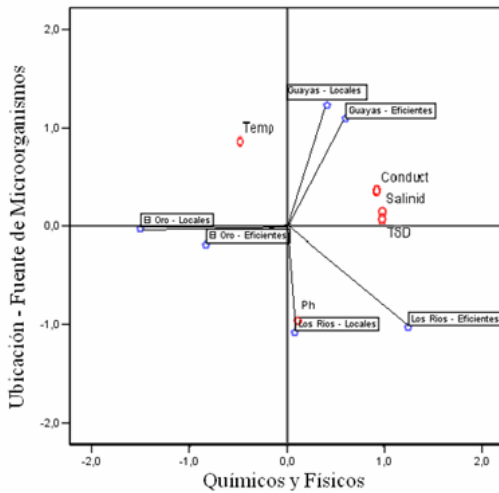
Las componentes, los valores propios, porcentajes de varianza obtenidos y el porcentaje de explicación acumulado para cada componente se muestran a continuación:

Componente	Valores Propios		
	λ	Porción de Varianza Explicada	Porción acumulada
1	3.076	61.524	61.524
2	1.781	35.629	97.153
3	0.130	2.599	99.752
4	0.012	0.233	99.985
5	0.001	0.015	100.000

Se tomó las dos primeros componente con la que explica un 97.15% de la varianza total.

La variable que explica la primera componente es “Químicos y Físicos”, para la segunda componente principal “Ubicación - Fuente de microorganismos”.

Ubicación - Fuente de Microorganismos: Químicos y Físicos
Biplot de las dos primeras componentes



Se puede constatar que tanto para la provincia del Guayas, El Oro y Los Ríos existe una asociación entre los dos tipos de microorganismos (locales y eficientes) para cada una de estas ubicaciones. Las variables: conductividad eléctrica, salinidad y total de sólidos disueltos están asociados con los microorganismos eficientes de de la provincia del Guayas y Los Ríos.

Ubicación - Meses de preparación: Nutrientes

Realizada la prueba de significancia estadística de Bartlett, se tiene un valor $p=0.$, por lo que se concluye que al menos algunas de las variables son correlacionadas y consecuentemente se procedió con el análisis de Componentes Principales

Prueba de Bartlett	
Estadístico de prueba	52.991
Grados de libertad	28
Valor P	0.003

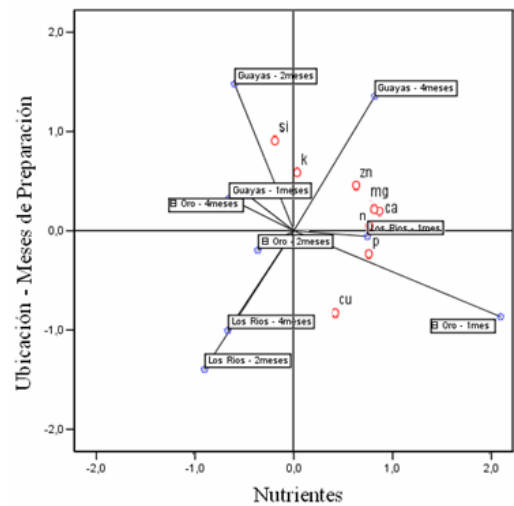
Las componentes, los valores propios, porcentajes de varianza obtenidos y el porcentaje de explicación acumulado para cada componente se muestran a continuación:

Componente	Valores Propios		
	λ	Porción de Varianza Explicada	Porción acumulada
1	3,178	39,729	39,729
2	2,202	27,522	67,251
3	1,402	17,526	84,777
4	0,630	7,873	92,650
5	0,362	4,523	97,173
6	0,166	2,074	99,247
7	0,060	0,748	99,996
8	0,000	0,004	100,000

Se tomó las dos primeros componente con la que explica un 67.25% de la varianza total.

La variable que explica la primera componente es “Nutrientes”, para la segunda componente principal “Ubicación - Meses de preparación”.

Ubicación - Meses de Preparación: Nutrientes
Biplot de las dos primeras componentes



El Oro y Los Ríos en un mes de preparación y Guayas en cuatro meses de preparación, están asociados con los nutrientes: potasio, zinc, magnesio, calcio, nitrógeno, fósforo y cobre. Nótese que en la provincia de Los Ríos en un mes de preparación, tiene una fuerte asociación con cinco de los ocho nutrientes a diferencia del tratamiento de dos y cuatro meses de preparación en la misma provincia.

Ubicación - Meses de preparación: Químicos y físicos

Realizada la prueba de significancia estadística de Bartlett, se tiene un valor $p=0.000$, por lo que se concluye que al menos algunas de las variables son correlacionadas y consecuentemente se puede proceder con el análisis de Componentes

Principales

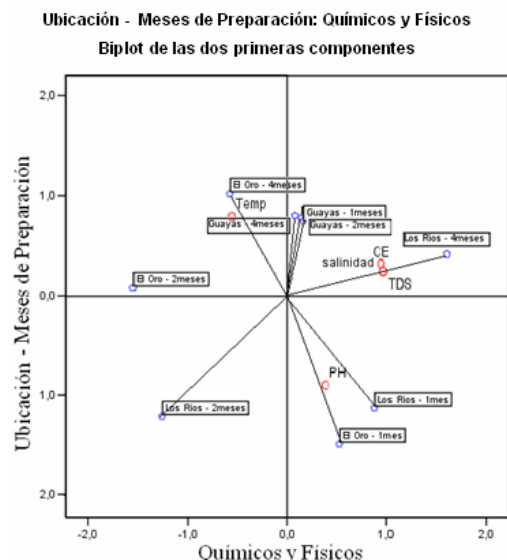
Prueba de Bartlett	
Estadístico de prueba	74.131
Grados de libertad	10
Valor P	0.000

Las componentes, los valores propios, porcentajes de varianza obtenidos y el porcentaje de explicación acumulado para cada componente se muestran a continuación:

Varianza Explicada por las Componentes Principales			
Componente	Valores Propios		
	λ	Porción de Varianza Explicada	Porción acumulada
1	3.219	64.371	64.371
2	1.668	33.359	97.729
3	0.107	2.141	99.870
4	0.006	0.122	99.992
5	0.000	0.008	100.000

Se tomó las dos primeros componente con la que explica un 97.72% de la varianza total.

La variable que explica la primera componente es “Químicos y Físicos”, para la segunda componente principal “Ubicación - Meses de preparación”.

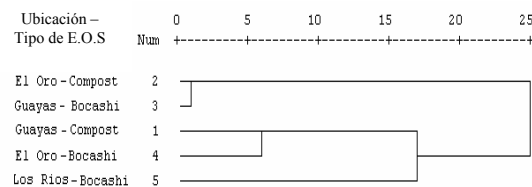


Se puede constatar que para las ubicación: en los tres diferentes meses de preparación en Guayas, en Los Ríos y El Oro en cuatro meses de preparación están asociados con los parámetros: conductividad eléctrica, total de sólidos disueltos, salinidad y temperatura.

ANÁLISIS DE CONGLOMERADOS

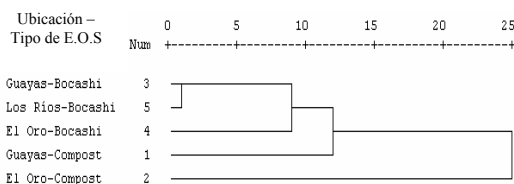
Enmiendas Orgánicas Sólida

Ubicación - Tipo de Enmienda Orgánica Sólida Nutrientes:



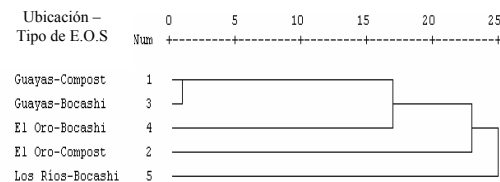
Con un 90% de similaridad en cada caso, se observa dos grupos que están conformados: el primero por la provincia de El Oro con el tipo de enmienda orgánica sólida “compost” y la provincia del Guayas con el tipo de enmienda orgánica sólida “bocashi”. El segundo grupo esta formado por la provincia del Guayas con el tipo de enmienda orgánica sólida “compost” y la provincia de El Oro con el tipo de enmienda orgánica sólida “bocashi”.

Químicos:



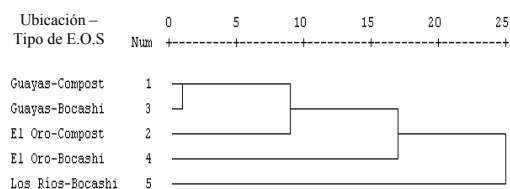
Con un 90% de similaridad en cada caso, se observa un sólo conglomerado que está conformado por: la provincia del Guayas y Los Ríos, ambos con el tipo de enmienda orgánica sólida “bocashi” y la provincia de El Oro con el tipo de enmienda orgánica sólida “bocashi” dentro del mismo grupo.

Físicos:



Con un 90% de similaridad en cada caso, se observa un sólo conglomerado que está conformado por: la provincia del Guayas con los dos tipos de enmienda orgánica sólida “bocashi” y “Compost”.

Microbiológicos:

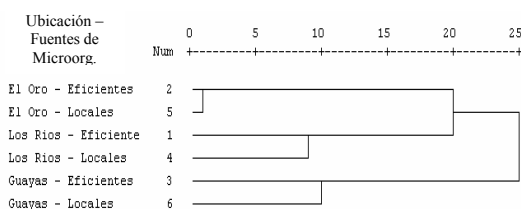


Con un 90% de similaridad en cada caso, se observa un solo conglomerado que está conformado por: la provincia del Guayas con los dos tipos de enmienda orgánica sólida “compost” y “bocashi” y la provincia de El Oro con el tipo de enmienda orgánica sólida “compost” dentro del mismo grupo.

Enmiendas Orgánicas Líquidas

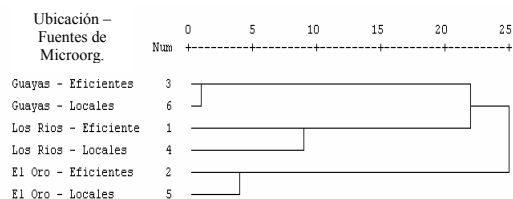
Ubicación - Fuente de Microorganismos

Nutrientes:



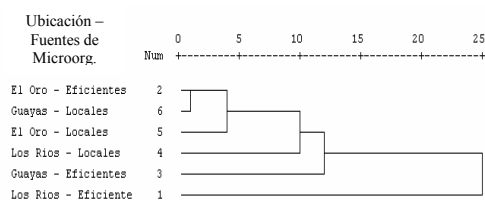
Con un 90% de similaridad en cada caso, observamos tres conglomerados que están conformados por las Ubicación de El Oro, Los Ríos y Guayas con los dos diferentes microorganismos en cada caso que son: eficientes y locales.

Químicos y Físicos:



Con un 90% de similaridad en cada caso, se observa tres grupos diferentes que están conformados por las provincias de Guayas, Los Ríos y El Oro en los dos tipos de microorganismos (locales y eficientes) para cada una de las provincias mencionadas.

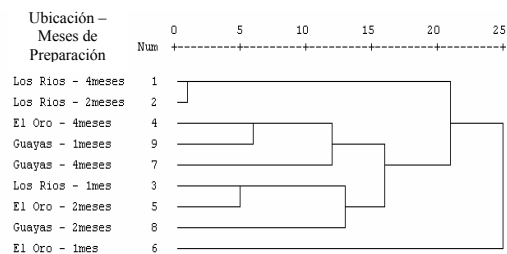
Microbiológicos:



Con un 90% de similaridad en cada caso, se observa un solo conglomerado que esta conformado por: la provincia de El Oro en los dos tipos de microorganismos (locales y eficientes), la provincia del Guayas dentro del mismo grupo y todos estos con Los Ríos en microorganismos locales.

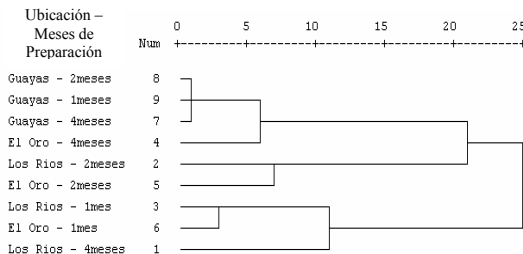
Ubicación - Meses de Preparación

Nutrientes:



Con un 90% de similaridad en cada caso, se observa tres grupos que están conformados: el primero por la provincia de Los Ríos en dos y cuatro meses de preparación. El segundo por: El Oro en cuatro meses y Guayas en un mes de preparación. Y finalmente el tercer grupo por: Los Ríos en un mes y El Oro en dos meses de preparación.

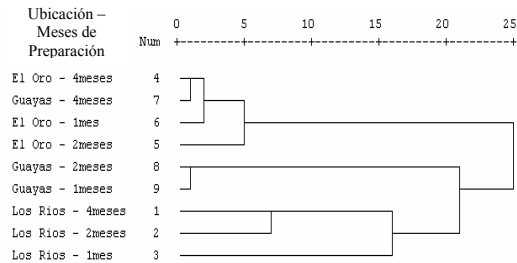
Químicos y Físicos:



Con un 90% de similaridad en cada caso, observamos tres grupos o conglomerados que están conformados: el primero por la provincia del Guayas en los tres diferentes meses de preparación y la provincia de El Oro cuatro meses de preparación dentro del mismo grupo. El

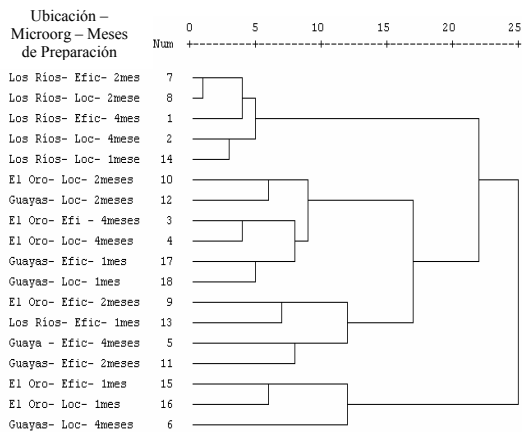
segundo grupo esta determinado por: la provincia de Los Ríos y El Oro, ambas en dos meses de preparación. Y finalmente el tercer grupo por: Los Ríos y El Oro, ambos en un mes de preparación.

Microbiológicos:



Con un 90% de similaridad en cada caso, observamos tres grupos o conglomerados que están conformados: el primero por la provincia de El Oro y Guayas, ambos en cuatro meses de preparación, El Oro en un mes de preparación dentro del mismo grupo y todos estos con la provincia de El Oro en dos meses de preparación como un solo conglomerado. El segundo grupo esta determinado por: la provincia del Guayas, en dos y un mes de preparación. Y finalmente el tercer grupo por: Los Ríos en dos y un mes de preparación.

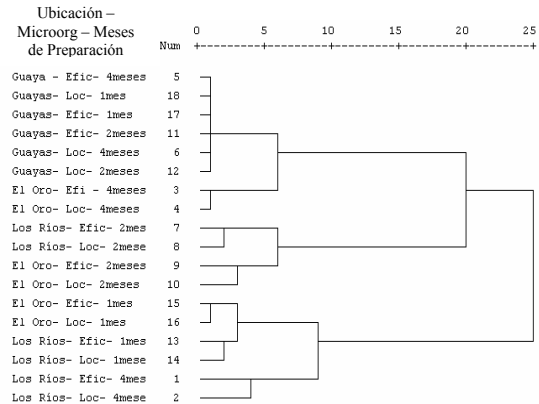
Ubicación - Microorganismos - Meses de Preparación Nutrientes:



Con un 90% de similaridad en cada caso, observamos cinco grupos o conglomerados que están conformados: el primero por los casos 7, 8, 1, 2 y 14 que representan el primer conglomerado. El segundo grupo esta determinado por los casos: 10, 12, 3, 4, 17 y 18 que representan el segundo conglomerado. El tercer grupo esta determinado por los casos: 9 y 13 que representan el tercer

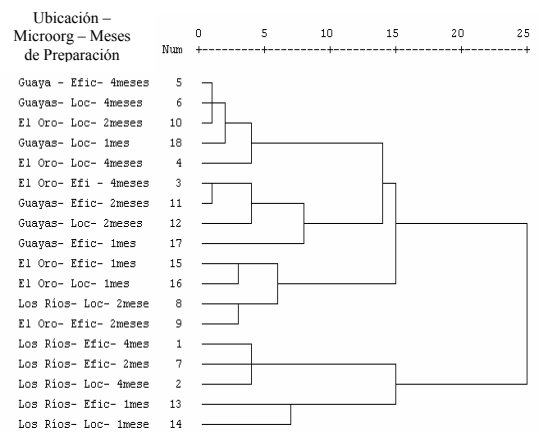
conglomerado. El cuarto grupo esta determinado por los casos: 5 y 11 que representan el cuarto conglomerado. Y finalmente el quinto grupo representado por los casos: 15 y 16 que representan el quinto y último conglomerado.

Químicos y Físicos:



Con un 90% de similaridad en cada caso, observamos tres grupos o conglomerados que están conformados: el primero por los casos 5, 18, 17, 11, 6, 12, 3 y 4 que representan el primer conglomerado. El segundo grupo esta determinado por los casos: 7, 8, 9 y 10 que representan el segundo conglomerado. Y finalmente el tercer grupo esta determinado por los casos: 15, 16, 13, 14, 1 y 2 que representan el tercer y último conglomerado.

Microbiológicos:



Con un 90% de similaridad en cada caso, observamos cinco grupos o conglomerados que están conformados: el primero por los casos 5, 6, 10, 18 y 4 que representan el primer conglomerado. El segundo grupo esta

determinado por los casos: 3, 11, 12 y 17 que representan el segundo conglomerado. El tercer grupo esta determinado por los casos: 15, 16, 8 y 9 que representan el tercer conglomerado. El cuarto grupo esta determinado por los casos: 1, 7 y 2 que representan el cuarto conglomerado. Y finalmente el quinto grupo representado por los casos: 13 y 14 que representan el quinto y último conglomerado.

ANÁLISIS DISCRIMINANTE

Meses de preparación (uno / dos / cuatro meses)

Para contrastar la hipótesis de igualdad de media entre los grupos (meses de preparación), se calculó el estadístico lambda de Wilks y el estadístico F. Cuando el valor de F para una determinada variable es grande y el valor de significancia (valor p) es menor que 0.10 se rechaza la hipótesis nula de igualdad de medias entre los grupos, por tanto hay diferencias significativas entre los meses de preparación para esa variable, es decir, que las variables que tengan valores de significancia cercanos a cero tendrán un potencial de discriminación mayor que los otros.

Prueba de Igualdad de la Media de los Grupos					
Variabes	Lambda de Wilks	F	gl1	gl2	Valor p
K	0,879	1,028	2	15	0,381
P	0,362	13,216	2	15	0,000*
Ca	0,738	2,668	2	15	0,102
Mg	0,755	2,428	2	15	0,122
N	0,632	4,373	2	15	0,032*
Zn	0,914	0,704	2	15	0,510
Cu	0,849	1,331	2	15	0,294
Si	0,995	0,037	2	15	0,963
pH	0,641	4,204	2	15	0,036*
Temperatura	0,697	3,262	2	15	0,067
Conductividad Eléctrica	0,590	5,209	2	15	0,019*
Total de Sólidos Disueltos	0,577	5,502	2	15	0,016*
Salinidad	0,571	5,625	2	15	0,015*
Coliforme	0,778	2,140	2	15	0,152
E. coli	0,955	0,355	2	15	0,707

Las variables: fósforo, nitrógeno, pH, conductividad eléctrica, total de sólidos disueltos y salinidad contribuyeron de manera significativa a la clasificación de los meses de preparación.

Se muestran los coeficientes de correlación canónica y el porcentaje de varianza explicada de cada una de las funciones canónicas, la cual nos permite analizar la importancia de cada función discriminante.

Valores Propios			
Función	% Relativo de la Varianza	% Acumulado de la Varianza	Correlación Canónica
1	78,750	78,750	0,990
2	21,250	100	0,965

La capacidad discriminante de la primera función es del 78.75% frente al 21.50% de explicación de la segunda función. Estos valores confirman que la importancia relativa de la primera función es superior a la segunda función.

Los valores del coeficiente de correlación canónica indican que ambas funciones cumplen el objetivo de separar los grupos, ya que para estas funciones se obtuvo un valor de correlación canónica 0.990 y 0.965 respectivamente, es decir, ambas miden fuertemente la asociación entre las funciones discriminantes y los grupos (meses de preparación).

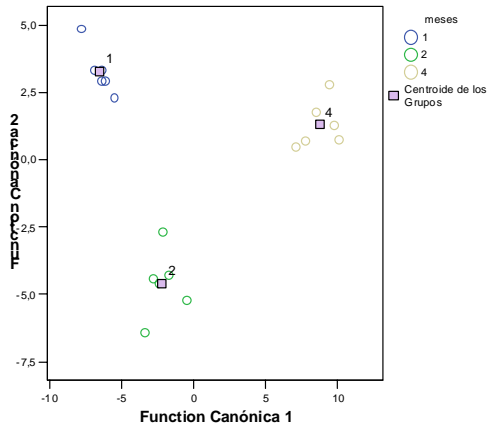
A continuación se muestran los valores del contraste de significancia secuencial basado en el estadístico lambda de Wilks.

Lambda de Wilks				
Tras la función	Lambda de Wilks	Chi-cuadrado	Grados de Libertad	Valor p
1	0,001	52,792	30	0,006
2	0,069	21,364	14	0,093

Se muestran los valores del estadístico lambda de Wilks y el valor de la prueba Chi-cuadrado a un nivel de confianza del 90%, para contrastar la significancia de cada función discriminante. Aquí se observa que es conveniente usar ambas funciones para interpretar las diferencias entre los meses de preparación.

Se presenta mediante una nube de puntos, la posición de las 18 observaciones y los tres centroides sobre los dos ejes discriminantes simultáneamente. El eje horizontal recoge las puntuaciones de la primera función discriminante y el eje vertical las de la segunda función discriminante. Los puntos se representan con un color distinto para cada grupo. Además se representa el centroide para cada mes de preparación. Las coordenadas de estos puntos son las siguientes: (-6.561, 3.285) para un mes de preparación, (-2.197, -4.594) para dos meses de preparación y (8.758, 1.309) para cuatro meses de preparación.

Gráfico de Dispersión de las Funciones Canónicas



La primera función discrimina mejor a los grupos que la segunda: la distancia entre los centroides sobre el eje de la función canónica 1 es mayor que las observadas en el eje de la función canónica 2.

Los coeficientes de la función discriminante estimada, se muestran junto con el término constante.

Coeficiente de la Función Discriminante Canónica		
Variables	Función	
	1	2
<i>K</i>	0,000	0,000
<i>P</i> *	-0,177	0,062
<i>Ca</i>	0,034	-0,014
<i>Mg</i>	0,008	0,111
<i>N</i> *	0,001	0,000
<i>Zn</i>	0,456	-0,884
<i>Cu</i>	0,498	-0,699
<i>Si</i>	0,000	0,000
<i>pH</i> *	8,464	-6,680
<i>Temperatura</i>	-0,617	0,201
<i>Conductividad Eléctrica</i> *	-2,870	-9,129
<i>Total de Sólidos Disueltos</i> *	20,378	15,506
<i>Salinidad</i> *	-12,635	5,641
<i>Coliforme</i>	-0,007	0,003
<i>E. coli</i>	-0,843	-1,638
<i>Constante</i>	-19,870	-0,829

El modelo discriminante es confiable para predecir mediante procedimientos sistemáticos la clasificación de nuevas observaciones desconocidas en alguno de los grupos considerados, y no solo, identificar que existe diferencia en una serie de grupos en las que se divide una población.

Conclusiones

- De acuerdo a los análisis estadísticos realizados, el método multivariado de análisis de conglomerados fue el que proporcionó mayor información respecto a los diferentes factores en estudio para las enmiendas orgánicas sólidas y líquidas, sin embargo, el análisis de componentes principales proporcionó información mas relevante acerca de estos mismos factores para las enmiendas orgánicas líquidas.
- Para las enmiendas orgánicas sólidas no fue factible la construcción del análisis de componentes principales, debido a que tanto para el Compost como para el Bocashi, ambos productos presentaron altas correlaciones entre sus indicadores.
- En el estudio de las enmiendas orgánicas líquidas, el análisis de componentes principales permitió explicar, a través de la reducción de variables, la asociación entre las variables “químicos y físicos” versus las tres provincias (Los Ríos, El Oro y Guayas) para los dos tipos de microorganismos (Locales y Eficientes) y los diferentes meses de preparación (uno, dos y cuatro meses).
- La asociación entre los “nutrientes” con las tres provincias en los diferentes meses de preparación para las enmiendas orgánicas líquidas, pudo ser explicada también con el análisis de componentes principales, no así con el tipo de microorganismos.
- El análisis de componentes principales mostró que la mayor cantidad de nutrientes en los bioproductos, están asociados a la provincia de los Ríos en un mes de preparación, demostrando así que, mejor es el biol de un mes que el de dos o cuatro meses de preparación en dicha provincia.
- A través del análisis de conglomerado se reveló un 90% de similaridad, entre los tipos de enmiendas orgánicas sólidas (compost y bocashi) versus los elementos químicos, físicos y microbiológicos para las tres provincias.
- Para los nutrientes en las enmiendas orgánicas sólidas, el análisis de conglomerado reveló un mínimo del 10% de diferencias entre el compost y el bocashi. Es decir ambas

- enmiendas orgánicas sólidas muestran una alta presencia de nutrientes.
8. Con un 90% de similaridad, el análisis de conglomerado de las enmiendas orgánicas líquidas, reveló la asociación entre los tratamientos: fuentes de microorganismos y los meses de preparación con las variables químicas, físicas y microbiológicas para las tres provincias.
 9. Para los nutrientes y otras variables físicas y químicas en las enmiendas orgánicas líquidas, el análisis de conglomerado reveló un mínimo del 10% de diferencias entre los microorganismos: eficientes y locales, para cada una de las tres provincias en estudio. Es decir que, para cada provincia es indiferente la fuente de microorganismos que utilizan para la elaboración del bioproducto, con respecto a dichas variables.
 10. La construcción de un modelo discriminante canónico que sirva como clasificador y predictor de nuevas observaciones solo fue posible realizarla para las enmiendas orgánicas líquidas.
 11. En las enmiendas orgánicas líquidas, solo para meses de preparación se encontró diferencias entre los grupos, con respecto al conjunto de variables de estudio. Dicho modelo proporciona procedimientos sistemáticos de clasificación para nuevas observaciones.
3. Merrill, R., and McKeon, J., 1998. "Organic Teas from Compost and Manure". Organic Farming Research Foundation Project Report #97-40. Santa Cruz, California.
 4. Compiladores: Martínez C. Claudia y Ramirez F. Leonel, 2000. "Lombricultura y Agricultura Sustentable". México.
 5. Wolf Benjamin and Snyder H. George, 2003. "Sustainable Soils: the Place of Organic Matter in sustaining Soils and their Productivity". Editorial Food Products Press. Binghamton, New York, Estados Unidos.
 6. Colaboradores: Pardo R. Nelson, Martínez D. Héctor, Durán N. Luis Felipe, Rincón S. Javier y Rosa R. Antonio, 2004. "Volvamos al Campo: Manual de cultivos Orgánicos y Alelopatía". Grupo Latino LTDA. Colombia.
 7. "Manual para extensionistas, promotores y productores del campo", 2001, (http://ppathw3.cals.cornell.edu/iipmweb/C_hapter7.pdf), (visitada Julio 5, 2006)
 9. N. M. Downie y R. W. Heath, 1986 "Métodos Estadísticos Aplicados", Editorial Harper & Row Publishers, Quinta Edición. New York, Estados Unidos. Capítulos: 2, 4 y 11.

REFERENCIAS BIBLIOGRAFICAS

1. Suquilanda B. Manuel, 1996. "Serie Agricultura Orgánica". Editorial Fundación para el Desarrollo Agropecuario. Quito, Ecuador, Capítulo 1.
2. Duicela G. Luis Alberto, Corral C. Rubèn, Cedeño G. Lincoln, Chòez T. Fredy, Romero R. Francisco, Palma P. Richard, Fernández A. Fabián, Macías N. Auro, Farfán T. Diana, Ramirez Joffre, Zambrano A. Luis, Reyes P. Javier, Farfán T. Daniel y Aveiga A. Tito, 2003. "Tecnologías para la producción de café arábigo orgánico", Concejo Cafetalero Nacional. Ecuador.
10. Richard A. Jonson, Dean W. Wichern, 1998. "Applied Multivariate Statistical Analysis". Editorial PRENTICE HALL. Cuarta Edición, New Jersey, Estados Unidos.
11. César Pérez López, 2001. "Técnicas Estadísticas con SPSS". Editorial PRENTICE HALL. Quinta Edición pp.447 – 488.