

Discriminación de tipos de calidad de agua de riego según atributos químicos utilizando una técnica multivariada*

Discrimination of types of irrigation water quality by chemical attributes using a multivariate technique

Javier Lorbes Medina^{1§}, Yelitza García Orellana¹, Carlos Ohep¹ y Manuel Milla Pino²

¹Universidad Centroccidental Lisandro Alvarado. Decanato de Agronomía. Cabudare, estado Lara. Tel. +582512592323. (yelitzagarcia@ucla.edu.ve; carloshep@ucla.edu.ve). ²Instituto Universitario de Tecnología de Yaracuy. San Felipe, estado Yaracuy. Tel. +584147935295. (memilla22@yahoo.com.mx). [§]Autor de correspondencia: javierlorbes@ucla.edu.ve.

Resumen

Con el objetivo de discriminar tipos de calidad de agua para riego según sus atributos químicos se consideró un experimento donde se evaluó el efecto de tres tipos de calidad de agua en la estructura de los suelos de la depresión de Quíbor, estado Lara. Los tres tipos de aguas consideradas en la investigación fueron: agua del embalse Yacambú (YAC), agua del embalse Dos Cerritos (DCE) y agua de pozo (POZO). Se tomaron 12 muestras de cada una y se midieron sobre ellas las variables: conductividad eléctrica (Ce) y los contenidos de calcio (Ca), magnesio (Mg), sodio (Na), potasio (K), bicarbonatos (HCO_3), cloruros (Cl), sulfatos (SO_4) y pH. Se utilizó el método del análisis discriminante canónico con dos técnicas gráficas multivariadas como la gráfica hipótesis-error (HE) y la gráfica de estructura discriminante canónica para evaluar los datos. Los resultados muestran diferencias significativas entre los tres tipos de agua y las variables evaluadas más influyentes en su discriminación fueron, Mg, Cl, Ce, SO_4 , Na y Ca. Las técnicas gráficas muestran que se puede interpretar las diferencias entre las aguas y relaciones entre variables y observaciones de manera fácil y sencilla ofreciendo una buena alternativa para analizar e interpretar datos.

Palabras clave: análisis discriminante canónico, gráfica HE, gráfica discriminante canónica.

Abstract

In order to discriminate types of irrigation water quality according to their chemical attributes was considered an experiment where we evaluated the effect of three types of water quality on soil structure of depression Quibor, Lara state. The three types of waters covered by the research were: reservoir water Yacambú (YAC), water reservoir Dos Cerritos (DCE) and well water (WELL). 12 samples were taken from each of them were measured variables: electrical conductivity (EC) and the calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), bicarbonate (HCO_3), chlorides (Cl), sulfate (SO_4) and pH. We used canonical discriminant analysis method with two multivariate techniques such as graphical charts and error hypothesis (HE) and canonical discriminant structure plot to evaluate the data. The results show significant differences between the three types of water and the most influential variables evaluated in discrimination were Mg, Cl, Ce, SO_4 , Na and Ca graphs show techniques that can interpret the differences between the waters and relationships between variables and observations easily and simply offering a good alternative to analyze and interpret data.

Key words: canonical discriminant analysis, graphical HE, canonical discriminant plot.

* Recibido: junio de 2013
Aceptado: noviembre de 2013

Tanto la calidad del agua de riego como el manejo adecuado es esencial para la producción, ya que la calidad afecta los rendimientos y las condiciones físicas del suelo, debido a que el tipo de agua que se utilice puede tener efectos importantes.

Los diferentes parámetros que se consideran en la determinación de la calidad del agua de riego, contemplan características físicas, químicas y biológicas que definen su adecuación y por ello se requiere de métodos de análisis que tomen en cuenta las interrelaciones entre las distintas variables de manera simultánea, lo cual se puede lograr cuando se analiza desde un punto de vista multivariante.

Muchas metodologías subyacen en el ámbito de la estadística multivariante donde el problema de reducción de la dimensionalidad de los datos es fundamental para cada aplicación experimental. La limitación fundamental de las técnicas multivariantes es la complejidad en la presentación de los resultados y la interpretación de los mismos, sobre todo cuando se tienen muchas variables, debido a las interrelaciones entre ellas y a la variabilidad de características asociadas a los modelos, lo cual ha conducido a muchos investigadores a realizar análisis univariantes para cada una de las variables por separado, con los consecuentes errores u omisiones en la interpretación de sus resultados (Amaro *et al.*, 2004).

Una técnica estadística multivariada bien desarrollada que puede ser usada como ayuda o complemento en la interpretación de los resultados es el análisis discriminante canónico (ADC), la cual permite la identificación de diferencias entre grupos (clases o tratamientos) a partir de medidas de características de los individuos pertenecientes a esos grupos y facilitar la comprensión de las relaciones entre las variables evaluadas dentro de los mismos (Afifi y Clark 1996; Vaylay y Santen, 2002; Onofre y Ciriciofolo, 2004).

Recientemente han surgido una variedad de nuevos métodos gráficos en el contexto del modelo lineal multivariado, entre los que se encuentran las gráficas HE y de estructura canónica (Friendly, 2006; 2007; Egesel *et al.*, 2011), que proveen una comparación visual directa de las matrices de covarianza de la hipótesis y error, y su uso en aproximaciones o proyecciones de bajo rango de datos multivariados para desplegar una vista en baja dimensión de los resultados basado sobre un análisis discriminante canónico.

El objetivo principal de este trabajo es discriminar tres tipos de calidad de agua de riego desde el punto de vista de nueve atributos químicos basado en el análisis discriminante

The quality of irrigation water and the proper management is essential for the production and quality affects yields and soil physical conditions, because the type of water used can have important effects.

The different parameters considered in determining the quality of irrigation water, provide physical, chemical and biological defining adequacy and therefore requires analysis methods that take into account the relationships between different variables simultaneously, which can be achieved when analyzed from a multivariate perspective.

Underlie many methodologies in the field of multivariate statistics where the problem of reducing the dimensionality of the data is crucial for each experimental application. The fundamental limitation of multivariate techniques is the complexity in the presentation of the results and the interpretation of them, especially when you have many variables, due to the linkages between them and the variability of characteristics associated with the models, which has led many researchers to perform univariate analysis for each of the variables separately, with consequent errors or omissions in the interpretation of results (Amaro *et al.*, 2004).

A well-developed multivariate statistical technique that can be used as an aid or supplement to the interpretation of the results is the canonical discriminant analysis (ADC), which allows the identification of differences between groups (classes or treatments) from measurements of characteristics individuals belonging to these groups and facilitate understanding of the relationships between the variables assessed within the same (Afifi and Clark 1996; Vaylay and Santen, 2002; Onofre and Ciriciofolo, 2004).

Recently there have been a variety of new methods in the context of graphics multivariate linear model, including graphs are canonical structure and HE (Friendly, 2006, 2007; Egesel *et al.*, 2011), which provide a direct visual comparison of the covariance matrices of the hypothesis and error and use approximations or projections under multivariate data range to display a low-dimensional view of the results based on a canonical discriminant analysis.

The main objective of this paper is to discriminate three types of irrigation water quality from the point of view of nine attributes based chemicals canonical discriminant analysis with two multivariate graphical techniques, the graph HE canonical and canonical discriminant structure.

canónico con dos técnicas gráficas multivariantes, la gráfica HE canónica y de estructura discriminante canónica.

Para la discriminación de los tipos de agua se evaluó el efecto de tres tipos de calidad de agua en la estructura de los suelos de la depresión de Quíbor, estado Lara, Venezuela. Los tres tipos de aguas consideradas en la investigación fueron: agua del embalse Yacambú (YAC), agua del embalse Dos Cerritos (DCE) y agua de pozo (POZO), en las cuales se tomaron 12 muestras de cada una y se midieron sobre ellas las siguientes variables: conductividad eléctrica (Ce), en ds m^{-1} ; y los contenidos de calcio (Ca), magnesio (Mg), sodio (Na), potasio (K), bicarbonatos (HCO_3), cloruros (Cl) y sulfatos (SO_4), todas medidas en miliequivalentes por litro (meq l^{-1}), y pH. El análisis estadístico fue hecho basado en el modelo estadístico multivariante siguiente:

$$Y_{ijh} = \mu_h + \tau_{ih} + \varepsilon_{ijh}$$

Donde: $i=1,2,3$; $j=1,2,\dots,12$; $h=1,2,\dots,9$; y_{ijh} es el vector multivariado de observaciones sobre la variable h para el tratamiento i , en la observación j ; μ_h es el vector multivariado de medias general sobre la variable h ; τ_{ih} es el vector multivariado de efectos en el tratamiento i sobre la variable h , y ε_{ijh} es el vector multivariado para errores aleatorios asociado con el vector de observaciones y_{ijh} .

En la gráfica HE, la elipse de la hipótesis (H) es obtenida de las sumas de cuadrados y productos para la hipótesis (tratamientos), mientras que la elipse del error (E) es obtenida de las sumas de cuadrados y productos para el error. Por otro lado, la gráfica de estructura discriminante canónica provee un resumen visual compacto de las características resaltantes de los datos, ya que muestra todas las observaciones, las medias de los grupos sobre las dos variables canónicas más grandes, círculos de confianza para estas medias, y vectores variables mostrando las correlaciones de las variables respuestas con las funciones canónicas (Egesel *et al.*, 2011).

Para generar las gráficas HE y de estructura canónica con el ADC se utilizaron los procedimientos PROC GLM y PROC CANDISC del programa estadístico SAS (SAS, 1999), conjuntamente con los programas macros CANPLOT, HEPLLOT y HECAN para las representaciones gráficas (Friendly, 2006; 2007).

For the discrimination of the types of water, the effect of three types of water quality on soil structure of depression Quíbor, Lara state, Venezuela. The three types of waters covered by the research were: reservoir water Yacambú (YAC), water reservoir Dos Cerritos (DCE) and well water (WELL), in which 12 samples were taken from each of them were measured the following variables: electrical conductivity (EC), ds m^{-1} , and the calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), bicarbonate (HCO_3), chlorides (Cl) and sulfate (SO_4), all measured in milliequivalents per liter (meq l^{-1}), and pH. Statistical analysis was performed based on the multivariate statistical model follows:

$$Y_{ijh} = \mu_h + \tau_{ih} + \varepsilon_{ijh}$$

Where: $i = 1, 2, 3$, $j = 1, 2, \dots, 12$, $h = 1, 2, \dots, 9$; Y_{ijh} is the multivariate vector of observations on the variable h for treatment i at observation j ; μ_h is the mean vector general multivariate variable h ; τ is the vector ih multivariate effects in the treatment i on the variable h , and ε is the vector ijh multivariate random errors associated with the vector of observations y_{ijh} .

In the graph HE, the ellipse of the hypothesis (H) is obtained from the sums of squares and products for the hypothesis (treatments), while the error ellipse (E) is obtained from the sums of squares and products for error. On the other hand, the graph of canonical discriminant structure provides a compact visual summary of the salient features of the data, as it shows all observations, the means of the groups on the two canonical variables larger circles of trust for these measures, variables and vectors showing the correlations of the variables answers canonical functions (Egesel *et al.*, 2011).

To generate graphs and canonical structure HE the ADC were used PROC GLM and PROC procedures CANDISC SAS statistical program (SAS, 1999), together with macro programs CANPLOT, HEPLLOT and HECAN for graphs (Friendly, 2006; 2007).

In both graphs the order and separation of treatment means on each canonical variable, indicating how the linear combination of the responses discriminate between types of water and vector variables on the graph indicate the correlation of each of the observed variables canonical dimensions. With axes equal, the relative length of each vector is proportional to its contribution to the

En ambas gráficas el orden y separación de las medias de los tratamientos sobre cada variable canónica, indicará cómo esa combinación lineal de las respuestas discrimina entre los tipos de agua y los vectores variables sobre la gráfica indicarán la correlación de cada una de las variables observadas con las dimensiones canónicas. Con ejes igualados, la longitud relativa de cada vector será proporcional a su contribución a la discriminación entre los tratamientos, como también, los ángulos entre los vectores indicarán aproximadamente las correlaciones entre las diferencias medias de los grupos, basado sobre la matriz H estandarizada proyectada en el espacio de las dimensiones canónicas.

Así, la longitud relativa de cada vector será un indicativo de la proporción de varianza explicada por esa variable y el ángulo entre cada vector con cada función canónica será proporcional a su correlación; por tanto los coeficientes de estructura canónica total medirán la correlación lineal simple bivariada entre una variable respuesta original y la variable canónica, por lo que el coeficiente canónico reflejará la varianza que la variable observada comparte con la variable canónica, y puede ser interpretada en la valoración de la contribución relativa de cada variable a cada función canónica.

Si existe una correlación positiva entre dos variables, los vectores tendrán similar dirección, y mostrarán diversas direcciones en el caso de una correlación negativa. La longitud del vector relacionado indica qué tan bien la variación de la variable puede ser explicada en el espacio bidimensional. En otras palabras, cuanto más largo es un vector, mejor interpretación es posible sobre la variable que representa. Además, las formas de las elipses de la hipótesis (H) y del error (E) en las gráficas HE brinda información acerca de la distribución de los datos y las dimensiones canónicas. Si el círculo de la hipótesis es anular significa que la variación total es explicada por ambas dimensiones canónicas igualmente, mientras que si es un elipsoide plano entonces significa que la primera dimensión explica una proporción mayor que la segunda dimensión. El diámetro del círculo del error indica la tasa de error de la hipótesis, y los niveles de los factores ubicados dentro de este círculo tienen una alta tasa de error.

Los valores obtenidos para cada una de las nueve respuestas estudiadas fueron sometidos a análisis exploratorios de datos. Los resultados demostraron que ninguna variable presentó desvíos significativos con relación a las suposiciones de normalidad y homogeneidad de varianzas, indicando ausencia de limitaciones para la realización del análisis de la

discrimination between treatments, as well, the angles between the vectors indicate correlations approximately differences group means, and based on the matrix H standardized projected the canonical space dimensions.

Thus, the relative length of each vector is indicative of the proportion of variance accounted for by the variable and the angle between each vector with each canonical function is proportional to the correlation, and therefore the overall structure coefficients measured canonical correlation simple linear bivariate between an original response variable canonical variable, so the canonical coefficient variance reflect the observed variable shares with the canonical variable, and can be interpreted in assessing the relative contribution of each variable to each canonical function.

If there is a positive correlation between two variables, the vectors will have similar vein, and show various directions in the case of a negative correlation. Related vector length indicates how well the variation of the variable can be explained in two-dimensional space. In other words, the longer a vector, the best performance possible on the variable representing. Furthermore, the shape of the ellipses of hypothesis (H) and the error (E) in the graphs HE provides information about the distribution of data and the canonical dimensions. If the circle of the hypothesis is void means that the total variation is explained by both dimensions equally canonical, while if it is a flat ellipsoid then it means that the first dimension explains a higher proportion than the second dimension. The diameter of the circle of error indicates the error rate of the hypothesis, and the levels of factors located within this circle have a high error rate.

The values obtained for each of the nine responses studied underwent exploratory data analysis. The results showed that significant deviations presented no variable in relation to the assumptions of normality and homogeneity of variance, indicating no limitations to the performance of the analysis of variance of the data. It was found that the error vector is multinormal distribution, verified by Mardia tests and Henze-Multnorm Zirkler using the SAS macro version 8 (SAS, 1999), which also reports the Shapiro-Wilk test for normality and univariate multivariate tests of Mardia Skewness and Kurtosis and statistical T-Zirkler Henze.

ADC with canonical functions were generated which allowed the separation of the three types of water which can be verified through the Mahalanobis distance D^2 between the mean (centroid) of water types. They are shown in Table 1.

varianza de los datos. Se comprobó que el vector de errores tiene distribución multinormal, verificado por las pruebas de Mardia y Henze-Zirkler usando el macro Multnorm del SAS versión 8 (SAS, 1999), el cual reporta además las pruebas de Shapiro-Wilk, para normalidad univariada y las pruebas multivariadas de Asimetría y Curtosis de Mardia y el estadístico T de Henze-Zirkler.

Con el ADC se generaron funciones canónicas que posibilitaron la separación de los tres tipos de aguas, lo cual puede ser verificado a través de la distancia D^2 de Mahalanobis entre las medias (centroides) de los tipos de aguas. Las mismas se muestran en el Cuadro 1.

Cuadro 1. Distancias D^2 de Mahalanobis, significancia para los contrastes entre las tres aguas y coordenadas centroides para agua.

Table 1. Mahalanobis D^2 distances, significance for the contrasts between the three waters and water centroid coordinates.

Agua	Distancias ^s D^2 y significancia ^t			Centroides	
	DCE	POZO	YAC	Can 1	Can 2
DCE		292.38	23.64	-4.28	-1.36
POZO	<0.0001		450.79	12.75	0.27
YAC	<0.0001	<0.0001		-8.47	1.09

^sValores de distancia cuadrada de Mahalanobis se muestran en la diagonal superior. ^tNivel de significancia P, se muestra en la diagonal inferior.

Los valores de D^2 indican que la mayor distancia ocurrió entre las aguas de YAC y POZO y la menor entre YAC y DCE, y aunque la distancia entre las aguas de YAC y DCE fue la más pequeña ($D^2=23.64$), las medias entre las tres aguas difieren significativamente entre sí ($p<0.01$), lo que indica que el grado de separación entre las mismas es relativamente grande.

El análisis multivariante reveló un efecto significativo para los tratamientos. El contraste global basado en el estadístico Lambda de Wilks tiene un valor de 0.005, el cual se distribuye como una F con 18 y 50 grados de libertad y un valor de $p<0.01$; indicando que el contraste multivariante para diferenciar entre los tipos de agua es altamente significativo considerando las nueve respuestas simultáneamente. En el Cuadro 2, se resumen los resultados del análisis discriminante canónico.

Se puede ver que la proporción explicada por el primer eje canónico es de 0.988; indicando esto que la primera variable canónica (Can 1) representa 98.8% de la varianza total, lo cual coincide con la premisa básica del ADC que establece que la primera función canónica debe expresar la máxima variación entre los tratamientos. Se presentan también los valores de correlación canónica de cada variable canónica con los tipos de agua, resultando ambas significantes ($p<0.01$)

D^2 values indicate that the greatest distance occurred between YAC and WELL waters and the smaller YAC and DCE, and although the distance between the waters of YAC and DCE was the smallest ($D^2=23.64$), stockings between three waters differ significantly ($p<0.01$), indicating that the degree of separation between them is relatively large.

Multivariate analysis revealed a significant effect for treatments. The overall contrast based on the Wilks' Lambda statistic has a value of 0.005, which is distributed as an F with 18 and 50 degrees of freedom and a p -value <0.01 , indicating that the multivariate contrast to differentiate between types

of water is highly significant considering the nine responses simultaneously. Table 2 summarizes the results of the canonical discriminant analysis.

Cuadro 2. Resumen del análisis discriminante canónico.

Table 2. Summary of canonical discriminant analysis.

Variable canónica	Correlación canónica	Valor propio	Proporción	Nivel de significancia
Can 1	0.99	91.81	0.988	<0.0001
Can 2	0.73	1.14	0.012	0.0052

We can see that the proportion explained by the first canonical axis is 0.988, indicating that the first canonical variable (Can 1) represents 98.8% of the total variance, which is consistent with the basic premise of the ADC which provides that the first function canonical should express the maximum variation between treatments. Also presented canonical correlation values of each variable with the types of canonical water, resulting both significant ($p<0.01$) indicating that the two canonical variables may explain the differentiation of three waters, being Can 1 (0.99) canonical function having the highest discriminating power.

indicando que las dos variables canónicas pueden explicar la diferenciación de las tres aguas, siendo Can 1 (0.99) la función canónica que posee el mayor poder discriminante.

En el Cuadro 3, se reportan los coeficientes de la estructura canónica total de las 9 variables evaluadas sobre las dos funciones canónicas para los tipos de agua. La matriz de estructura canónica total contiene las correlaciones entre las variables y las funciones discriminantes canónicas, donde se examinan las correlaciones simples entre las variables originales y las funciones canónicas. Los pesos canónicos reflejan la varianza que la variable observada comparte con la función canónica, y puede ser interpretada en la evaluación de la contribución relativa de cada variable a cada función canónica.

Cuadro 3. Coeficientes de la estructura canónica total de las dos funciones canónicas y variación explicada (R²) por cada vector variable.

Table 3. Canonical structure coefficients entire two canonical functions and variance explained (R²) for each vector variable.

Variable	Can 1	Can 2	R ²
Ce	0.87	-0.19	0.77
Ca	0.49	0.38	0.31
Mg	0.99	0.03	0.97
Na	0.84	-0.33	0.76
K	0.11	-0.13	0.02
HCO ₃	0.33	0.06	0.11
Cl	-0.88	0.20	0.80
SO ₄	0.86	0.16	0.74
pH	-0.24	-0.01	0.06

Se observa que la primera variable canónica está dominada por un valor grande para Mg (0.99), seguido por Cl (-0.88), Ce (0.87), SO₄ (0.86), Na (0.84) y Ca (0.49). La segunda variable canónica está dominada por Ca (0.38) y Na (-0.33), evidenciando que la calidad de las tres aguas difiere principalmente en Mg, Cl, Ce, SO₄, Na y Ca, y son responsables de una porción importante de la discriminación entre los tres tipos de agua.

Los coeficientes R² que indican la proporción de variabilidad explicada de cada vector variable, siendo Mg con un valor de 0.97 el vector variable con mayor proporción de explicación, seguido de Cl, Ce, Na y SO₄. Esta explicación se ve reflejada en las gráficas HE canónica y de estructura discriminante canónica ilustrada en las Figuras 1 y 2,

In Table 3, we report the coefficients of the total canonical structure 9 variables evaluated on the two canonical functions for types of water. The Total canonical structure matrix contains the correlations between variables and canonical discriminant functions, which examines the simple correlations between the original variables and canonical functions. Canonical weights reflect the variance observed variable shares the canonical function, and can be interpreted in the evaluation of the relative contribution of each variable to each canonical function.

It is observed that the first canonical variable is dominated by a large value for Mg (0.99), followed by Cl (-0.88), Ce (0.87), SO₄ (0.86), Na (0.84) and Ca (0.49). The second canonical variety is dominated by Ca (0.38) and Na (-0.33), showing that the three water quality differs mainly Mg, Cl, Ce, SO₄, Na and Ca, and are responsible for a major portion of the discrimination between the three types of water.

R² coefficients that indicate the proportion of variability explained by each vector variable, being Mg 0.97 with a value of the vector variable with the greatest proportion of explanation, followed by Cl, Ce, Na, and SO₄. This explanation is reflected in the graphs HE canonical discriminant canonical structure illustrated in Figures 1 and 2, in which the canonical discriminant projection takes into account the largest variation between the averages of the three types of water relative to the error.

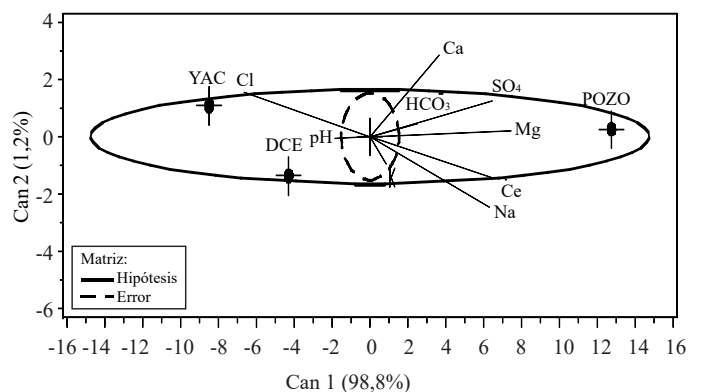


Figura 1. Gráfica HE en el espacio discriminante canónico bidimensional (2D) para los tres tipos de agua.

Figure 1. HE graphical canonical discriminant space (2D) for the three types of water.

en las que la proyección discriminante canónica toma en cuenta la más grande variación entre las medias de los tres tipos de agua relativa al error.

Es aparente que la estructura de las medias es esencialmente unidimensional con 98.8% de la variación, explicada por grandes diferencias entre las tres aguas, principalmente sobre las variables Mg, Ce, SO₄ y Na, y en menor medida, Ca, las cuales tienen correlación positiva con la primera variable canónica y Cl, que tiene correlación negativa. Es decir, 98.8% de la variación en las medias de las tres aguas es explicada por una sola dimensión canónica, que corresponde a valores grandes para POZO, y pequeñas para YAC y DCE sobre Mg, Cl, Ce, SO₄ y Na y moderado para Ca.

El orden y separación de las medias de los tratamientos sobre las dos variables canónicas indica que Can 1 discrimina los tres tipos de agua, mientras que Can 2, sólo discrimina entre las aguas de YAC y DCE. Se observa que la mayor parte de la variación entre las aguas (98.8%) puede ser explicada por las diferencias entre las aguas de YAC y POZO y entre DCE y POZO, donde esta última es alta en los contenidos de Mg, Na, Ce y SO₄.

En la Figura 2, se muestra la gráfica de la estructura discriminante canónica. En la misma están representados conjuntamente los individuos, los tratamientos y las variables, así como los círculos de confianza alrededor de la media de cada tratamiento que permite comparar gráficamente las tres aguas.

Ésta gráfica muestra que la primera dimensión canónica está reflejando las diferencias entre los tratamientos, con 98.8% de la variación. Se observa que los círculos de confianza no se interceptan, lo que significa que hay diferencias entre los tres tipos de agua; principalmente sobre el primer eje canónico. Las aguas de YAC y POZO son diferentes en Mg, Cl, SO₄, Ce y Na, porque al proyectar sus círculos de confianza sobre el vector que representa a esas variables, las proyecciones son disjuntas. Lo mismo se puede decir de las aguas de DCE y POZO sobre las variables Mg, Cl, SO₄, Ce y Ca. La variable que más influye en la separación de las aguas de YAC y POZO es Mg, porque el vector que representa esta variable es casi paralelo a la recta que pasa por los centroides de ambos. Para la separación de DCE y POZO, las variables que más influyen son SO₄ y Mg, y para la separación de YAC y DCE Cl y Na. Las variables con mayor correlación con el primer eje canónico son Mg, Cl, Ce, SO₄ y Na y con el segundo eje Ca y Na. Por tanto, las

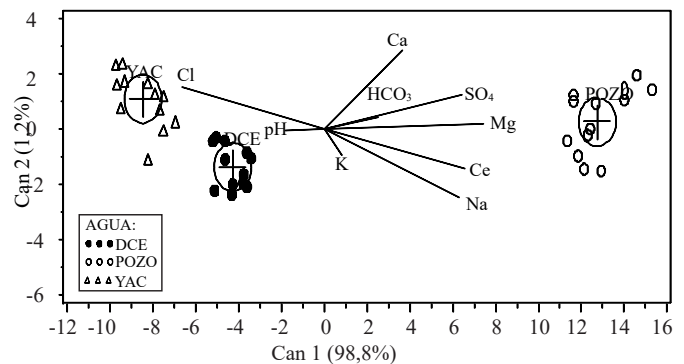


Figura 2. Gráfica de la estructura canónica de los tres tipos de agua.

Figure 2. Graph of the canonical structure of the three types of water.

It is apparent that the structure is essentially one-dimensional averages with 98.8% of the variance explained by significant differences between the three waters, mainly on variables Mg, Ce, Na, and SO₄, and to a lesser extent, Ca, which have correlation positive with the first canonical variant and Cl, which has negative correlation. I.e. 98.8% of the variation in the mean of the three water is explained by a single canonical dimension, corresponding to large values for WELL, and small for YAC and DCE on Mg, Cl, Ce, Na, and SO₄ and moderate Ca.

The order and mean separation of treatments on the two canonical variables indicates that Can one discriminate the three types of water while Can 2, only discriminates between water and DCE, YAC. It is observed that most of the variation between the waters (98.8%) can be explained by the differences between the YAC and water and between DCE and WELL WELL, where the latter is high in content of Mg, Na, Ce and SO₄.

The Figure 2 shows the graph of the canonical discriminant structure. In the same individuals are represented together, treatments and variables and trusted circles around the mean of each treatment can graphically compare the three waters.

This graph shows that the first canonical dimension is reflecting the differences between treatments with 98.8% of the variation. It is observed that no confidence circles intersect, which means that there are differences between the three types of water on the first shaft mainly canonical. The YAC and WELL waters are different in Mg, Cl, SO₄, Ce and Na, because the project their circles of trust on the vector representing these variables, the projections are

variables Mg, Cl, Ce, SO₄, Na y Ca, son las que presentan la mayor proporción de varianza explicada (R²) con los dos ejes canónicos.

Conclusiones

Debido a la importancia que tiene conocer cómo es el agua que se utiliza para regar, especialmente si proviene de diferentes fuentes, en el presente artículo se discriminan claramente los tres tipos de agua, según sus atributos químicos. Los resultados muestran que el análisis discriminante canónico con las gráficas HE y de estructura canónica usadas conjuntamente, permitieron identificar relaciones importantes entre las variables respuestas y los tratamientos evaluados de una manera resumida y compacta, brindando una buena alternativa para analizar datos de este tipo de estudios. La técnica permitió diferenciar claramente los tratamientos e identificar las variables más importantes en la discriminación de los tres tipos de agua. Las variables evaluadas que más influyen en la separación de las tres aguas, fueron Mg, Cl, Ce, SO₄, Na y Ca, y estas diferencias están básicamente sobre el primer eje canónico. El agua de Pozo se diferencia grandemente de las aguas de Yacambú y de la de Dos Cerritos, y está relacionada principalmente con las variables Mg, Ce, Na y SO₄.

Literatura citada

- Afifi, A. A. and Clark, V. 1996. Computer-aided multivariate analysis. 3th. Chapman and Hall, London.
- Amaro, R.; Vicente Villardón, J. L. y Galindo Villardón, M. P. 2004. MANOVA-Biplot para arreglos de tratamientos con dos factores basado en modelos lineales generales multivariantes. *Interciencia*, 29(1):26-32.
- Egesel, C. O.; Kahnman, F. and Güi, M. K. 2011. Discrimination of maize inbreds for kernel quality traits and fatty acid composition by a multivariate technique. *Acta Sci. Agron. Maringá*, 33(4):613-620.
- Friendly, M. 2006. Data ellipses, HE plots and reduced-rank displays for multivariate linear models: SAS Software and Examples. *J. Statis. Software*. 17(6):1-43.
- Friendly, M. 2007. HE plots for multivariate general linear models. *J. Comp. Graph. Statistics*. 16:1-23.
- Onofri A. e E, Ciricifolo. 2004. Characterisation of yield quality in durum wheat by canonical variate analysis. Proceedings VIII ESA Congress "European Agriculture in a global context", Copenhagen, 11-15 July 2004. 541-542 pp.

disjoint. The same can be said of the waters of DCE and WELL on variables Mg, Cl, SO₄, Ce and Ca. The most influential variable in the separation of water is WELL and YAC, because the vector that represents this variable is almost parallel to the line through the centroids of both. For the separation of DCE and WELL, the most influential variables are SO₄ and Mg, and for the separation of YAC and DCE of Cl and Na. The variables with highest correlation with the first canonical axis are Mg, Cl, Ce, SO₄ and Na and the second axis Ca and Na. Therefore, the variables Mg, Cl, Ce, SO₄, Na and Ca, are those with the highest proportion of variance explained (R²) with the two canonical axes.

Conclusions

Due to the importance of knowing what water is used for irrigation, especially if it comes from different sources, in this article clearly discriminate the three types of water, according to their chemical attributes. The results show that the canonical discriminant analysis with graphs and canonical structure HE used together, helped identify important relationships between the response variables and the treatments evaluated in a condensed and compact, providing a good alternative to analyze data from these studies. The technique allowed us to clearly differentiate the treatments and identify the most important variables in discriminating the three types of water. The evaluated variables that influence the separation of the three waters were Mg, Cl, Ce, SO₄, Na and Ca, and these differences are basically on the first canonical axis. Well water greatly differs Yacambu water and Dos Cerritos, and is mainly related to variables Mg, Ce, Na, and SO₄.

End of the English version



- Statistical Analysis System (SAS) Institute. 1999. SAS user's guide. Statistics. Version 8. SAS Inst., Cary, NC. USA. Quality, and elemental removal. *J. Environ. Qual.* 19:749-756.
- Vaylay, R. and van Santen, E. 2002. application of canonical discriminant analysis for the assessment of genetic variation in tall fescue. *Crop Sci.* 42:534-539.