

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE AGRONOMÍA
ESCUELA DE AGRONOMÍA
DEPARTAMENTO DE INGENIERÍA AGRÍCOLA

**APLICACIÓN DEL ANÁLISIS DISCRIMINANTE PARA REALIZAR
RECOMENDACIONES DE DOSIS DE NUTRIENTES EN MAÍZ (*Zea
mays*)**

LUIS M. PÉREZ Y.

Maracay, Mayo de 1998.

**APLICACIÓN DEL ANÁLISIS DISCRIMINANTE PARA REALIZAR
RECOMENDACIONES DE DOSIS DE NUTRIENTES EN MAÍZ (*Zea mays*)**

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE AGRONOMÍA
ESCUELA DE AGRONOMÍA
DEPARTAMENTO DE INGENIERÍA AGRÍCOLA

**APLICACIÓN DEL ANÁLISIS DISCRIMINANTE PARA REALIZAR
RECOMENDACIONES DE DOSIS DE NUTRIENTES EN MAÍZ (*Zea mays*)**

LUIS M. PÉREZ Y.

TUTOR: LUSBI. A. HERRERA. S.

Trabajo presentado como parte de los requisitos para optar al título de Ingeniero Agrónomo
Mención Ingeniería Agrícola que otorga la Universidad Central de Venezuela.

Maracay, Mayo de 1998.



Nosotros los abajo Firmantes, miembros del Jurado Examinador del Trabajo de Grado Titulado : **APLICACIÓN DEL ANÁLISIS DISCRIMINANTE PARA REALIZAR RECOMENDACIONES DE DOSIS DE NUTRIENTES EN MAÍZ (*Zea mays*)**. Cuyo autor es el Bachiller Pérez Ybarra, Luis Manuel, Cédula de Identidad N° 12.195.023. Certificamos que lo hemos leído y que en nuestra opinión reúne las Condiciones necesarias de Adecuada Presentación y es enteramente satisfactorio en alcance y Calidad como Trabajo de Grado para Optar al Título de **INGENIERO AGRÓNOMO MENCIÓN INGENIERÍA AGRÍCOLA**.

Herrera, Lusbi

CI : 3.255.060

Tutor - Coordinador

Rodriguez, Belkis

CI : 4.566.825

Jurado Principal

Figueroa, Rosana

CI : 4.271.380

Jurado Principal

*“Ut potero expliabo: nec tamen, ut Pythius Apollo,
certa ut sint et fixa, quae dixero:
sed, ut homunculus, probabilia conjectura sequens”.*

(“Lo explicaré como pueda,
pero no como cosas ciertas y rígidas de un omnisciente Apolo,
sino como cosas de un mortal,
que sigue las conjeturas más probables”).

Cicerón.

DEDICATORIA.

A mi madre,

A mi padre,

A mis hermanos: Marianela y Carlos Enrique,

A mis familiares y amigos,

Y a la memoria de mis Abuelas hoy día desaparecidas:

María, María Vidalia, Josefa y Yolanda.

AGRADECIMIENTOS.

Deseo expresar mi más sincero y profundo agradecimiento a todas las personas e instituciones que hicieron posible la realización de este trabajo.

Al Dr. Lusbi Herrera: tutor, profesor y amigo, por su apoyo, confianza y asesoramiento; así como por estar siempre pendiente de mi formación profesional.

A la Ing. Agrónomo Belkis Rodríguez, por su invaluable apoyo, confianza, comprensión y paciencia; por todo el material que me facilitó, y por todo el tiempo que le dedicó de una manera desinteresada a lograr que este trabajo fuera realizado.

Al Dr. Rafael Álvarez, la Lic. Josefa Medina y la Sra. Inelda Cisneros por su solidaridad y ayuda incondicional.

A todas aquellas personas que de una u otra forma colaboraron con la realización de este trabajo, como son muchas y las páginas no alcanzarían, a todos muchas gracias.

RESUMEN.

El propósito del presente trabajo es aplicar el análisis discriminante como una estrategia metodológica que permita extraer la mayor información de los datos experimentales para desarrollar modelos de predicción que contribuyan a mejorar el uso y manejo de los fertilizantes en los estados Apure, Guárico, Monagas, Portuguesa, Táchira y Yaracuy y pueda ser utilizado como una metodología para realizar recomendaciones de nutrientes en maíz (*Zea mays*).

Los resultados de un grupo de 68 experimentos llevados a cabo en los estados Apure, Guárico, Monagas, Portuguesa, Táchira y Yaracuy, que fueron conducidos por el Fondo Nacional de Investigaciones Agropecuarias entre los años 1981 y 1984, utilizando diseño en bloques al azar con dos repeticiones en un diseño de tratamiento compuesto central rotatable y que fueron sembrados con maíz, híbrido Arichuna, aplicando distintas dosis de nitrógeno fósforo y potasio, se utilizaron como el material para la aplicación de esta metodología. A partir de los resultados obtenidos en estos experimentos, se procedió a dividir las respuestas de rendimientos en kg/ha en tres categorías a saber: *Bajo*, *Medio* y *Alto*. Luego se encontraron las ecuaciones para discriminar entre cada uno de los grupos para cada dosis aplicada considerando como variables de estas ecuaciones a las características edáficas medidas en los experimentos; esto se realizó utilizando el programa SAS versión 6.04.

Los ecuaciones encontradas permiten predecir a que categoría de rendimiento pertenece una observación proveniente de algunas de las regiones estudiadas para cada una las dosis de fertilizantes aplicadas.

A partir de las dosis de nutrientes aplicadas se procedió a encontrar las combinaciones de fertilizantes óptimas para cada dosis utilizando la metodología de la programación lineal utilizando el programa MPS versión 2.1.

Los datos provenientes del estado Portuguesa fueron descartados por poseer muy pocas observaciones, es decir, el número de observaciones fue inferior al número de variables consideradas. Las ecuaciones discriminantes encontradas para el estado Yaracuy presentaron baja probabilidad de clasificación equivocada para la categoría de rendimientos *Alto* (menor a 0.5), lo cual las hace muy útiles para realizar recomendaciones. Para el resto de los estados estudiados las ecuaciones discriminantes presentaron altas probabilidades de mala clasificación (mayor a 0.5) por lo cual se consideró poco factible la aplicación del análisis discriminante para realizar recomendaciones de nutrientes en esas entidades.

Las dosis óptimas encontradas utilizando el programa *MPS* versión 2.1, presentaron como productos comerciales más económicos a la urea, el fosfato especial y el cloruro de potasio, y se recomienda aquella dosis en la cual además de presentar rendimientos Altos, presente además el mayor beneficio económico.

PALABRAS CLAVES.

Discriminar, análisis discriminante, ecuaciones discriminantes, programación lineal, función de producción, fertilización, rendimiento, rendimiento relativo, calibración de suelos, variables edáficas, beneficio.

TABLA DE CONTENIDO.

	<i>Página</i>
<i>Dedicatoria.</i>	<i>vi</i>
<i>Agradecimientos.</i>	<i>vii</i>
<i>Resumen.</i>	<i>viii</i>
<i>Palabras claves.</i>	<i>x</i>
<i>Tabla de contenido.</i>	<i>xi</i>
<i>Índice de cuadros.</i>	<i>xiv</i>
<i>Índice de figuras.</i>	<i>xx</i>
<i>1. Introducción.</i>	<i>1</i>
<i>1.1. Generalidades.</i>	<i>1</i>
<i>1.2. Justificación.</i>	<i>3</i>
<i>1.3. Objetivo.</i>	<i>4</i>
<i>1.4. Hipótesis.</i>	<i>4</i>
<i>2. Revisión de literatura.</i>	<i>5</i>
<i>2.1. Evolución histórica de los estudios de fertilidad de suelos en Venezuela.</i>	<i>5</i>
<i>2.2. Modelos usados para estudiar la respuesta de los cultivos a los fertilizantes.</i>	<i>8</i>
<i>2.2.1. Modelos curvilíneos continuos.</i>	<i>8</i>
<i>2.2.2. Modelos de respuesta lineal.</i>	<i>10</i>

2.2.3. Modelo discontinuo de respuesta lineal y de rendimiento máximo estable.	10
2.2.3.1 Rendimiento.	10
2.2.3.2. Rendimiento con el nutriente al mínimo.	11
2.2.3.3. Rendimiento máximo estable.	11
2.2.3.4. Rendimiento relativo.	12
2.2.3.5. Correlación del análisis de suelo con la respuesta al fertilizante.	12
2.2.3.6. El nivel crítico de Cate-Nelson.	13
2.2.3.7. Objeciones al uso del rendimiento relativo como base para elaborar recomendaciones de fertilizantes.	15
2.2.4. La función superficie de respuesta.	17
2.3. Análisis discriminante.	20
3. Conceptos básicos de análisis discriminante.	21
3.1. Introducción.	21
3.2. Discriminación cuando las poblaciones son conocidas.	22
3.2.1. Regla discriminante de máxima verosimilitud.	22
3.2.2. Regla discriminante de Bayes.	25
3.2.3. Propiedades óptimas.	26
3.3. Discriminación bajo estimación.	27
3.3.1. La regla de discriminación muestral.	27
3.3.2. ¿Es la discriminación importante?.	31
3.3.3. Función discriminante lineal de Fisher.	32
3.4. Probabilidades de mala clasificación.	34

3.4.1. Probabilidades cuando los parámetros son estimados.	34
3.4.2. Método de la resustitución.	36
3.5. Descarte de variables.	36
3.6. ¿Cuándo la correlación mejora la discriminación?.	40
4. Materiales y métodos.	41
4.1. Caracterización de la información.	41
4.2. Aplicación de la metodología.	55
5. Resultados y discusión.	66
5.1. Combinaciones de nutrientes utilizadas en el análisis discriminante.	66
5.2. Ecuaciones discriminantes.	72
5.2.1. Estado Yaracuy.	72
5.2.2. Estado Apure.	86
5.2.3. Estados Guárico y Táchira.	93
5.2.4. Estado Monagas.	98
5.3. Obtención de las dosis de nutrientes más adecuadas.	104
5.4. Recomendación de las combinaciones de productos comerciales más económicas.	122
6. Conclusiones y recomendaciones.	124
7. Referencias bibliográficas.	126
8. Anexos	132

ÍNDICE DE CUADROS.

	<i>Página</i>
<i>Cuadro 1. Esquema del análisis de varianza multivariado.</i>	<i>31</i>
<i>Cuadro 2. Coeficientes discriminantes para lo datos de iris y sus valores de F respectivos.</i>	<i>38</i>
<i>Cuadro 3. Coeficientes discriminantes para lo datos de iris y sus valores de F respectivos.</i>	<i>39</i>
<i>Cuadro 4. Ubicación de los experimentos en las distintas localidades del estado Yaracuy.</i>	<i>43</i>
<i>Cuadro 5. Ubicación de los experimentos en las distintas localidades del estado Portuguesa.</i>	<i>43</i>
<i>Cuadro 6. Ubicación de los experimentos en las distintas localidades del estado Yaracuy.</i>	<i>44</i>
<i>Cuadro 7. Ubicación de los experimentos en las distintas localidades del estado Monagas.</i>	<i>45</i>
<i>Cuadro 8. Ubicación de los experimentos en las distintas localidades del estado Apure.</i>	<i>46</i>

Cuadro 9.	<i>Ubicación de los experimentos en las distintas localidades del estado Guárico.</i>	47
Cuadro 10.	<i>Ubicación de los experimentos en las distintas localidades del estado Táchira.</i>	47
Cuadro 11.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Yaracuy.</i>	48
Cuadro 12.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Portuguesa.</i>	49
Cuadro 13.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Yaracuy.</i>	50
Cuadro 14.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Monagas.</i>	51
Cuadro 15.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Apure.</i>	52
Cuadro 16.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Guárico.</i>	53
Cuadro 17.	<i>Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Táchira.</i>	54

<i>Cuadro 18. Listado de productos comerciales utilizados en el cálculo de las dosis óptimas de productos comerciales para las combinaciones de tratamientos utilizados en el análisis discriminante.</i>	64
<i>Cuadro 19. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en el estado Yaracuy en el primer grupo de experimentos.</i>	67
<i>Cuadro 20. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites las categorías de discriminación en el estado Yaracuy en el segundo grupo de experimentos.</i>	68
<i>Cuadro 21. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en el estado Monagas.</i>	69
<i>Cuadro 22. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en el estado Apure.</i>	70

Cuadro 23.	<i>Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en los estados Guárico y Táchira.</i>	71
Cuadro 24.	<i>Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en el estado Yaracuy.</i>	73
Cuadro 25.	<i>Medias muestrales de las variables de suelo para la dosis 90-60-50 en el estado Yaracuy.</i>	74
Cuadro 26.	<i>Coeficientes de discriminación correspondientes a las ecuaciones de la dosis 90-60-50 en el estado Yaracuy.</i>	75
Cuadro 27.	<i>Medias muestrales de las variables de suelo para la dosis 180-60-50 en el estado Yaracuy.</i>	75
Cuadro 28.	<i>Coeficientes de discriminación correspondientes a las ecuaciones de la dosis 180-60-50 en el estado Yaracuy.</i>	76
Cuadro 29.	<i>Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 120-0-50 en el estado Yaracuy.</i>	77
Cuadro 30.	<i>Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 120-80-50 en el estado Yaracuy.</i>	78

Cuadro 31.	<i>Ecuaciones discriminantes correspondientes a las dosis 180-60-50, 120-80-50 y 120-80-0 en el estado Yaracuy.</i>	80
Cuadro 32.	<i>Resultados de las pruebas de medias univariadas para la dosis 90-60-50 en el estado Yaracuy.</i>	81
Cuadro 33.	<i>Resultados de las pruebas de medias univariadas para la dosis 49-128-20 en el estado Yaracuy.</i>	82
Cuadro 34.	<i>Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en el estado Apure.</i>	87
Cuadro 35.	<i>Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 140-30-20 en el estado Apure.</i>	89
Cuadro 36.	<i>Medias muestrales de las variables de suelo para las dosis 0-30-20, 57-12-8 y 140-30-20 en el estado Apure.</i>	91
Cuadro 37.	<i>Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en los estados Guárico y Táchira.</i>	93
Cuadro 38.	<i>Medias muestrales de las variables edáficas de las combinaciones de nutrientes 0-120-60 y 115-120-120 en los estados Guárico y Táchira.</i>	95

<i>Cuadro 39. Coeficientes de discriminación correspondientes a las ecuaciones de las dosis 0-120-60 y 115-120-120 en los estados Guárico y Táchira.</i>	96
<i>Cuadro 40. Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en el estado Monagas.</i>	98
<i>Cuadro 41. Medias muestrales de las variables edáficas de las combinaciones de nutrientes 0-100-70 y 144-159-112 en el estado Monagas.</i>	100
<i>Cuadro 42. Coeficientes de discriminación correspondientes a las ecuaciones de las dosis 0-100-70 y 144-159-112 en el estado Monagas.</i>	101
<i>Cuadro 43. Resultados de las pruebas de medias univariadas para la dosis 144-159-112 en el estado Monagas.</i>	102
<i>Cuadro 44. Resultados de las pruebas de medias univariadas para la dosis 36-41-112 en el estado Monagas.</i>	102
<i>Cuadro 45. Resultados del análisis de suelo correspondiente a una muestra (ficticia) traída por un productor de maíz del valle medio del estado Yaracuy.</i>	105

ÍNDICE DE FIGURAS.

	<i>Página</i>
<i>Figura 1. Formato del plástico transparente que se utiliza para obtener los valores críticos de los análisis de suelos.</i>	13
<i>Figura 2. Esquema de las declaraciones necesarias para ser corridas por el programa SAS, para realizar el análisis discriminante.</i>	60

1. INTRODUCCIÓN.

1.1. GENERALIDADES.

La eliminación del subsidio a los fertilizantes y el estancamiento de la productividad del sector agrícola en los últimos años entre otros factores, ha generado que los márgenes de ganancia de diferentes cultivos no hayan experimentado un crecimiento proporcional al incremento en los precios de los insumos agrícolas.

La productividad, el precio del producto y los costos de producción, determinan la rentabilidad de cualquier cultivo. En este sentido, el desarrollo de tecnologías que permitan aumentar la producción por unidad de superficie y al mismo tiempo promuevan el uso eficiente de los recursos, incidirán de una manera positiva en el un incremento del margen de ganancia de los productores al disminuir los costos de producción racionalizando el uso de los insumos agrícolas.

El Fondo Nacional de Investigaciones Agropecuarias (FONAIAP), como institución piloto en el diagnóstico de fertilidad de suelos y recomendaciones de fertilizantes a nivel nacional, inició a partir de la década de los 70 un proyecto nacional de calibración de suelos para los rubros de mayor importancia económica en las principales áreas agrícolas del país, lo cual permitió la generación de un gran volumen de información relativa a la respuesta de dichos rubros a la aplicación de fertilizantes en términos de rendimiento, así como de las características agroecológicas de los sitios experimentales.

El presente trabajo pretende mostrar la utilidad del análisis discriminante como alternativa metodológica para determinar requerimientos nutricionales en maíz en los Estados Apure, Guárico, Monagas, Portuguesa y Táchira, que permitan establecer un sistema de recomendaciones en diferentes condiciones agroecológicas y niveles de respuesta.

En este orden de ideas, existe un conjunto de procedimientos metodológicos asociados al área de estadística como la Función Superficie de Respuesta, cuya aplicación a resultados provenientes de experimentos relacionados con la respuesta a la aplicación de fertilizantes, permiten el desarrollo de modelos que estimen el rendimiento de los cultivos a partir de las características edáficas, conformando así relaciones funcionales entre las variables edáficas y los rendimientos esperados.

Estas metodologías tienen la ventaja de que pueden analizar más de un nutriente a la vez, por lo cual las recomendaciones representan una opción diferente a la planteada por el proyecto de calibración de suelos, en el cual se analizaba cada nutriente por separado y genera además una gama de recomendaciones que pueden ser consideradas por el productor como viables con miras a la optimización del uso de sus recursos.

En este trabajo se plantea la aplicación del análisis discriminante como una alternativa metodológica que permite realizar recomendaciones de nutrientes a partir de las condiciones edáficas existentes en una zona en particular y la obtención de combinaciones de productos comerciales que suplan estos requerimientos al más bajo costo utilizando la técnica de la programación lineal.

1.2. JUSTIFICACIÓN.

El sector agrícola en general ha sido uno de los más perjudicados por la mala administración y políticas económicas de los gobiernos del país. Para los rubros vegetales, una de las medidas que causó un impacto negativo fue el de la eliminación del subsidio al precio de los fertilizantes, la cual se realizó en los inicios de la década de los años 90. Hasta 1984 el subsidio representaba el 50% del precio real de los fertilizantes; para 1989 cerca del 90%, año en el cual el Ministerio de Agricultura y Cría junto con el Ministerio de Fomento resuelven por decreto la eliminación progresiva desde 1989 hasta 1993.

Esto conllevó a que los costos de producción aumentaran y de esta realidad no escapó el cultivo del maíz, debido a que el costo de los fertilizantes representa alrededor del 25-30% de los costos de producción del cultivo del maíz, por lo tanto es necesario hacer un uso más eficiente de tales insumos.

En este trabajo se pretende desarrollar una herramienta que permita realizar recomendaciones sobre dosis de fertilizantes acorde con los niveles de nutrientes presentes en el suelo y, los requerimientos del cultivo y de algunas características edáficas, de tal manera que contribuya a disminuir los costos de producción, ahorro de divisas por concepto de importación de fuentes potásicas y de ácido fosfórico y de igual manera permita integrar los procedimientos estadísticos en beneficio del sector agrícola.

1.3. OBJETIVO.

- Evaluar la aplicación del análisis discriminante como una alternativa metodológica para realizar recomendaciones de fertilizantes en el cultivo de maíz (*Zea mays*) y presentar cuadros con varias alternativas de diferentes combinaciones de productos comerciales que suministren los requerimientos de nutrientes (N, P, K) al mínimo costo, a través del uso de las técnicas de la programación lineal, indicando intervalos de rendimientos esperados.

1.4. HIPÓTESIS.

“Se espera que a partir de datos experimentales pertenecientes a diferentes poblaciones de respuesta en términos de rendimientos en kg/ha de maíz, bajo diversas condiciones agroecológicas, la aplicación del análisis discriminante resulte un procedimiento eficiente para la realización de recomendaciones de fertilizantes”.

2. REVISIÓN DE LITERATURA.

2.1. EVOLUCIÓN HISTÓRICA DE LOS ESTUDIOS DE FERTILIDAD DE SUELOS EN VENEZUELA.

Si bien es cierto que ha habido estudios de suelo desde el siglo pasado, Medina (1975), citado por Bastidas *et al* (1984), es a partir de este siglo cuando en realidad se comienza a buscar alternativas para la elaboración y recomendación de dosis de fertilizantes.

Los primeros estudios de fertilidad en Venezuela comenzaron a aparecer en la década de 1940, ya en la década del 50 al 60 comienzan a reportarse algunos ensayos de campo, notándose un intento en agruparlos por cultivos y por áreas geográficas. El enfoque fundamental se hace sobre la búsqueda de la mejor dosis de nitrógeno, fósforo y potasio en términos de incremento de rendimientos con relación al testigo no abonado y la estimación de las ganancias netas debidas a la aplicación de abonos (Bastidas *et al*, 1984).

Según Ramírez (1995), estas primeras investigaciones respondían a un criterio de estudios puntuales con poco valor de extrapolación.

Avilán citado por Batidas *el al* (1984), afirma que en el período de 1976 hasta 1984, comienza el establecimiento del programa cooperativo entre el Ministerio de Agricultura y Cría (MAC) y el *International Soil Testing Project* de la Universidad de Carolina del Norte de los Estados Unidos, para calibración de análisis de suelos, el cual significó el primer intento de generar información sobre recomendaciones de fertilizantes.

Este programa de calibración de análisis de suelos fue llevado a cabo por el MAC y el Centro de Investigaciones agropecuarias (CIA) durante los primeros años de la realización de este programa (1966-1970), y los principales esfuerzos fueron concentrados en la evaluación de los diferentes métodos para determinar las concentraciones de los nutrientes aprovechables en el suelo mediante ensayos en invernadero. (Bastidas *et al*, 1984).

Ramírez (1995), afirma que el método de extracción de Olsen presenta ventajas en suelos neutros y el método del Bray en suelos ácidos. Como resultado de estos esfuerzos se han

logrado obtener guías de recomendación para diferentes cultivos en las principales zonas productoras de Venezuela.

A partir del año 1976 con la creación del FONAIAP (Fondo Nacional de Investigaciones Agropecuarias), el programa de calibración de análisis de suelos se amplía, con la finalidad de abarcar un mayor número de regiones agroecológicas y de cultivos. Este programa es llevado a cabo a través de los centros de investigaciones regionales (Bastidas *et al*, 1984).

En este período también es necesario destacar el aporte de las diferentes universidades nacionales a la investigación sobre fertilidad de suelos en el país, particularmente las escuelas o facultades de Agronomía de la Universidad Central de Venezuela, la Universidad de Oriente, la Universidad del Zulia, la Universidad Centro Occidental “Lisandro Alvarado” y la Universidad Nacional Experimental de los Llanos “Ezequiel Zamora”. En general, durante este período, la Universidad Central de Venezuela orientó sus estudios hacia una investigación de tipo básico más que hacia una investigación aplicada mientras que las otras universidades han dedicado sus esfuerzos a la solución de problemas que se encuentren dentro de su radio de influencia, es decir, más que todo problemas de índole regional (Bastidas *et al*, 1984).

Entre los cultivos estudiados se encuentran maíz, maní, sorgo y yuca para los nutrientes fósforo y potasio. (González, 1977a y 1977b; González *et al*, 1978a y 1978b; Ramírez, 1989a y 1989b; Ramírez *et al*, 1987; Ramírez *et al*, 1988a y 1988b; Ramírez *et al*, 1989a, 1989b y 1989c).

Ramírez (1995), afirma que este programa de calibración de suelos ha sido el primer programa aplicado en Venezuela con la finalidad de obtener datos que pudieran ser usados para realizar recomendaciones sobre dosis de fertilizantes y que afortunadamente ha sido mantenido en el tiempo. A pesar de ello, plantea que queda mucho por hacer, ya que la disminución en el consumo de fertilizantes experimentada en el país en los últimos años, no ha influido en los rendimientos obtenidos, por ello el consumo de fertilizantes en el país y las recomendaciones formuladas no han redundado en uso más racional del recurso que constituyen los fertilizantes comerciales. Opina además, que la información sobre fertilidad de suelos en Venezuela es insuficiente y que las recomendaciones sobre el uso de fertilizantes no

han sido las más adecuadas y que todavía falta por investigar en el manejo de los fertilizantes con los sistemas de siembra de labranza mínima, en los casos donde se presentan déficit de humedad en el perfil, como en los llanos centrales y donde se presenten condiciones de anoxia como en el caso de los llanos occidentales.

2.2. MODELOS USADOS PARA ESTUDIAR LA RESPUESTA DE LOS CULTIVOS A LOS FERTILIZANTES.

2.2.1. Modelos curvilíneos continuos.

En la literatura la mayoría de las discusiones sobre la respuesta a los fertilizantes están basadas en la interpretación de los ensayos de campo utilizándose alguna versión de la Ley de los Rendimientos Decrecientes o ley de Liebig, esto es, funciones curvilíneas continuas en las cuales incrementos adicionales iguales de fertilizantes resultan en respuestas cada vez menores. (National Academy of Sciences, 1961; Waugh *et al*, 1973).

Los modelos curvilíneos más importantes de uso actual son el cuadrático, la raíz cuadrada y las funciones logarítmicas. El modelo de Mitscherlich añade constantes a la función logarítmica con el fin de ajustar la interpretación de la respuesta al fertilizante de acuerdo con las variaciones de la fertilidad natural del suelo (Waugh *et al*, 1973).

El modelo de Mitscherlich queda expresado por la ecuación siguiente:

$$Y = A(1 - e^{-LX})$$

Donde:

- Y : Rendimiento.
- X : Cantidad de nutriente.
- A : Rendimiento máximo posible.
- L : Constante de proporcionalidad (Helyar y Godden, 1977; Hernández, 1993).

La función de Mitscherlich constituye uno de los primeros intentos de establecer una relación funcional curvilínea entre el rendimiento y los aportes de nutrientes. (National Academy of Sciences, 1961).

No existen reglas establecidas para determinar el modelo de respuesta más apropiado a ser utilizado en la interpretación de las repuestas a los fertilizantes en experimentos de campo.

Si se ha de evaluar un experimento individual, el investigador puede analizar los datos y ajustarlos a varios modelos lineales o curvilíneos y seleccionar la función que dé el mejor coeficiente de determinación (R^2) para la regresión (Waugh *et al*, 1973).

Los principales problemas que se presentan con este tipo de modelos, según Waugh *et al* (1973), son los siguientes:

Sesgo del modelo de interpretación:

La mayoría de los modelos tienen una tendencia a sesgarse bajo cierta clase de condiciones. Para el modelo cuadrático se presenta el problema de que sobrestima las respuesta a las dosis de fertilizantes, si la respuesta en el rendimiento es muy marcada con los primeros incrementos de fertilizante. Esto genera que las predicciones de dosis de nutrientes para obtener rendimientos máximos son exageradamente altas.

Variabilidad observada en el rendimiento cuando se aplican dosis elevadas de nutrientes:

La mayoría de los modelos curvilíneos de respuesta continua son muy sensitivos a fluctuaciones en el rendimiento debidas al uso de dosis altas; por ejemplo, si el fertilizante ocasiona quemado de raíces y por consiguiente una merma en el rendimiento, el modelo cuadrático se ajustará mejor a esta situación, si la respuesta a dosis altas tiende a estabilizarse los modelos de raíz cuadrada, logarítmicos y el modelo de Mitscherlich tienen mejor ajuste, sin embargo, este ajuste depende del experimento, es decir, que no necesariamente un mismo modelo presentará buen ajuste para una serie de experimentos.

Información disponible insuficiente para aplicar correctamente el modelo de respuesta a los fertilizantes:

Los modelos matemáticos que predicen el rendimiento deben asignar valores a cada una de las variables que intervienen en el modelo y hacer constantes las condiciones restantes, esto genera un gran problema cuando se pretende hacer predicciones en sitios en los cuales se tiene poca información o la información es insuficiente.

2.2.2. Modelos de respuesta lineal.

Swanson, citado por Waugh *et al* (1973), propuso en 1963 un modelo alternativo basado en la Ley del Mínimo de Liebig. Este modelo postula una respuesta lineal para el principal elemento limitante, la cual se detiene repentinamente cuando otro factor se hace limitante, pero que recupera su ascenso lineal cuando se corrige dicha limitación. Eventualmente, el rendimiento es limitado por la capacidad genética de la planta cuando todos los factores externos que la limitaban han quedado controlados.

Boyd (1970), citado por Waugh *et al* (1973), resumió una cantidad considerable de datos de la respuesta a los fertilizantes cubriendo diferentes cultivos y ha demostrado que la mayoría de los datos se ajusta muy bien a la Ley del Mínimo de Liebig. En su estudio, la mayoría de los ensayos de fertilizantes pudieron ser caracterizados por una línea recta en la porción ascendente de la curva y otra línea aproximadamente horizontal representando el rendimiento máximo.

2.2.3. Modelo discontinuo de respuesta lineal y de rendimiento máximo estable.

Este modelo tiene como base los siguientes conceptos: rendimiento con el nutriente al mínimo, rendimiento máximo estable y rendimiento relativo y es a partir de este modelo que se lleva a cabo la calibración de suelos. En Waugh *et al* (1973), se encuentran las siguientes consideraciones:

2.2.3.1 Rendimiento.

El rendimiento es una función de muchos factores que afectan el crecimiento. Genéticamente cada especie vegetal tiene cierto potencial total para transformar en materia seca los nutrientes minerales, luz, CO₂ y agua. Cuán bien una planta puede transformar estos ingredientes básicos en rendimiento depende del grado y balance en el cual se encuentran reunidos todos los requerimientos del crecimiento. Agronómicamente, el rendimiento de los cultivos puede ser relacionado con un amplio rango de factores tal como se muestra en la siguiente ecuación

$$\text{Rendimiento} = f(\text{Cultivo, suelo, fertilización, clima, } \dots, \text{ manejo})$$

Teniendo en cuenta esta ecuación, puede observarse que la fertilización por si sola no es el único factor que afecta al rendimiento, evidentemente la contribución de los otros factores deberá cuantificarse si esto fuera necesario.

Es decir que las recomendaciones de fertilización sólo se refieren a un factor constituyente de la ecuación, sin embargo, esto no impide que se puedan predecir rendimientos dentro de unos límites más manejables, ya que la ecuación considera muchos factores que no necesariamente son cuantificables y los rendimientos esperados a partir de las recomendaciones de fertilizantes constituyen una aproximación a los rendimientos teóricos esperados.

2.2.3.2. Rendimiento con el nutriente al mínimo.

Podemos asumir que cada suelo tiene cierto nivel de disponibilidad para cada nutriente y que este nivel marca el límite más bajo de la función de la respuesta, es decir, el punto de origen desde el cual se mide la respuesta. Se sugiere el término “rendimiento con el nutriente al mínimo” como un medio conveniente para describir el rendimiento limitado por un solo nutriente ya que este marca el punto de partida de donde la respuesta a ese nutriente deberá ser tomada en cuenta.

2.2.3.3. Rendimiento máximo estable.

El rendimiento máximo estable implica que el rendimiento se encuentra en el punto más elevado para ciertas condiciones particulares y permanece constante sobre cierto rango de adiciones del nutriente bajo estudio. Este es un término que tiene significado para describir el rendimiento en el cual la respuesta a cierto nutriente cesa.

Así tenemos que, para un nutriente dado, suelo y cultivo en particular, existe un límite definitivo mínimo, donde la respuesta a la fertilización se inicia (rendimiento con el nutriente al mínimo) y un límite definitivo máximo, donde la respuesta tiende a estabilizarse (rendimiento máximo estable).

2.2.3.4. Rendimiento relativo.

El rendimiento relativo es una medida de la respuesta del rendimiento a un solo nutriente cuando los otros nutrientes son suministrados en cantidades adecuadas pero no excesivas y cuando las otras variables del lugar (clima, por ejemplo) se mantienen constantes. La justificación principal para su uso es que siendo una proporción en vez de una cantidad absoluta, es posible comparar el rendimiento relativo de sitios diferentes.

El rendimiento relativo de un cultivo para unas condiciones de suelos dada viene dado por la siguiente ecuación

$$\text{Rendimiento relativo } (Y)(\%) = \frac{\text{Rendimiento con el nutriente } X \text{ al mínimo}}{\text{Rendimiento máximo estable para } X} \times 100$$

El rendimiento relativo de un cultivo en un sitio específico expresa el rendimiento con el nutriente al mínimo como el porcentaje del rendimiento máximo estable. Esto significa que el punto de partida de la respuesta está siendo comparado con el punto final de la misma. Cuando el rendimiento inicial es bajo, comparado con el rendimiento final, obviamente existirá una respuesta alta en el rendimiento debida a la adición del nutriente bajo estudio.

2.2.3.5. Correlación del análisis de suelo con la respuesta al fertilizante.

Para que un análisis de suelos tenga significado para cualquier nutriente, éste deberá estar asociado con la disponibilidad de dicho nutriente para la planta. Se han ensayado diversas técnicas para relacionar los análisis de los suelos con el crecimiento de las plantas, fluctuando desde intentos directos para tratar de extraer químicamente del suelo la misma cantidad que las raíces extraen biológicamente, hasta los de correlaciones indirectas o empíricas entre los nutrientes extraídos y el crecimiento vegetal. Los métodos indirectos de correlación han resultado ser los medios más prácticos para una evaluación, a pesar de la dificultad de controlar los factores de rendimiento.

A pesar de ser cierto que la asimilación de nutrientes, o la cantidad de nutrientes en un tejido vegetal, frecuentemente se encuentran bien correlacionados con el análisis de suelo, no existe una manera conveniente de distinguir entre el requerimiento de la planta y el consumo de lujo.

2.2.3.6. El nivel crítico de Cate-Nelson.

Estos investigadores observaron que un diagrama de dispersión entre el rendimiento relativo y el análisis de suelos podía ser dividido gráficamente en dos poblaciones pudiéndose identificar así un nivel crítico del análisis de suelo para un nutriente dado. En esta técnica se utiliza un plástico transparente sobrepuesto, el cual tiene dibujados sobre sí dos líneas perpendiculares que forman cuatro cuadrantes que tienen aproximadamente el mismo tamaño relativo que los que se muestran en la Figura 1. El plástico se coloca sobre el gráfico de manera que el máximo número de observaciones de rendimientos relativos para unos contenidos dados de nutrientes en el suelo resulten ubicados en los cuadrantes positivos y el mínimo en los cuadrantes negativos; luego, aquel valor de contenido en el suelo del nutriente estudiado que permita realizar esta división será considerado como el nivel crítico. En la práctica, el diagrama de dispersión de Cate-Nelson provee una técnica simple, pero altamente efectiva, para evaluar los métodos analíticos de suelos. Un buen extractante del suelo para cualquier nutriente deberá proveer una división perfecta entre las diferentes categorías de suelo-cultivo (Waugh *et al*, 1973; Cate y Nelson, 1971).

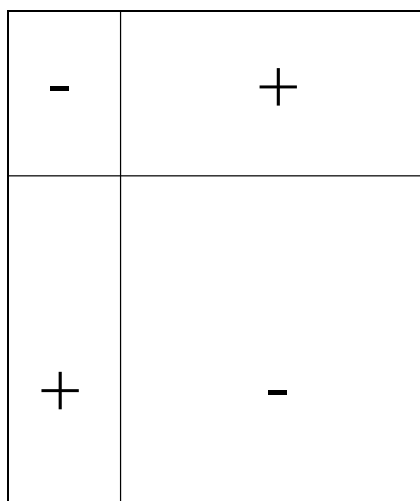


Figura 1. Formato del plástico transparente que se utiliza para obtener los valores críticos de los análisis de suelos.

Fuente: Cate and Nelson. 1971.

Los investigadores Cate y Nelson (1971), propusieron un procedimiento estadístico para obtener el nivel crítico del análisis de suelo para un nutriente dado a partir del diagrama de dispersión entre el rendimiento relativo y el análisis de suelos. Este método consiste en dividir las observaciones en dos poblaciones, y obtener el límite entre ambas cuando los coeficientes de ajuste R^2 se hagan máximos para cada población. Este modelo presenta los coeficientes de ajuste más altos que cualquier otro modelo para las observaciones menores que el nivel crítico, y fue a partir de este procedimiento que plantearon la aproximación mencionada anteriormente.

Es a partir de estos datos que se procede a realizar la calibración de suelos, la cual consiste en encontrar curvas en las que se puedan estimar las mejores dosis de fertilizantes para una cantidad dada de nutriente en el suelo y de los rendimientos relativos obtenidos a partir de cada una de las dosis de fertilizante añadidas (Rouse, 1968).

El modelo discontinuo de respuesta lineal y de rendimiento máximo estable, fue el utilizado para llevar a cabo el programa de calibración de suelos en Venezuela. El CIA realizó innumerables ensayos para determinar cuales eran los métodos más adecuados para el análisis de fósforo y potasio en el suelo (González et al, 1978). Luego, el FONAIAP continuó con el programa de calibración de análisis de suelos; entre los cultivos estudiados se encuentran maíz, maní, sorgo y yuca para los nutrientes fósforo y potasio. (González, 1977a y 1977b; González et al, 1978a y 1978b; Ramírez, 1989a y 1989b; Ramírez et al, 1987; Ramírez et al, 1988a y 1988b; Ramírez et al, 1989a, 1989b y 1989c).

Las curvas de calibración obtenidas a partir de la metodología de los modelos discontinuos presentan una serie de ventajas que les permiten ser una herramienta muy útil a la hora de realizar recomendaciones de dosis de nutrientes, entre ellas podemos citar:

- No requiere de mucha información disponible, es decir, no es necesario realizar un gran número de ensayos para obtener la cantidad de información necesaria para poder aplicar esta metodología. Por esta razón el MAC se adscribió al *International Soil Testing Project* de la Universidad de Carolina del Norte de los Estados Unidos para calibración de análisis de suelos en el país, porque debido a la poca información existente para esa

época en los países latinoamericanos, la Universidad de Carolina del Norte desarrolló este proyecto a partir de esta metodología.

- No se requiere de un equipo sofisticado para realizar los cálculos pertinentes a esta metodología, por ello se presentaba como una alternativa muy viable de aplicar en Latinoamérica en ese entonces, debido al poco desarrollo computacional de la época.

2.2.3.7. Objeciones al uso del rendimiento relativo como base para elaborar recomendaciones de fertilizantes.

A pesar de las ventajas operativas que ofrece el método de calibración de suelos para estimar el requerimiento de fertilizantes y lo extendido de este método, Colwell *et al* (1988), presentaron las siguientes objeciones:

Objeción 1:

Los valores del rendimiento relativo no proveen una base para estimar dosis económicas de fertilizantes. Esta objeción proviene del hecho de que los productores derivan sus ingresos de los rendimientos reales de sus cultivos y no de los rendimientos relativos, y se debe justificar el costo de las aplicaciones de fertilizantes por los resultados en incrementos reales más bien que por el hecho de lograr un porcentaje del rendimiento máximo posible.

Objeción 2:

El rendimiento máximo estable usado en el cálculo del rendimiento relativo, muchas veces es definido en forma deficiente por los datos experimentales y debe ser estimado por un procedimiento altamente subjetivo. Frecuentemente las observaciones experimentales indican que los más altos rendimientos pueden ser logrados aplicando dosis mayores de fertilizantes y entonces el valor máximo se estima por extrapolación o simplemente es estimado por el investigador a partir de su experiencia en la zona de estudio. Los valores de rendimiento relativo generados de esta manera pueden ser poco confiables y pueden atentar contra la precisión y la capacidad predictiva del modelo.

Objeción 3:

La relación funcional entre el rendimiento y los análisis de suelo no sólo presentan variaciones proporcionales entre las curvas de respuesta para un mismo lugar sino que también difieren en la forma de la curva de respuesta y pueden comportarse de manera muy distinta para las diferentes localidades y no simplemente presentarán variaciones proporcionales al rendimiento máximo. Experimentos llevados a cabo por Colwell en años sucesivos en el mismo sitio con condiciones de desarrollo del cultivo diferentes mostraron diferencias entre los rendimientos obtenidos a pesar de poseer las mismas condiciones edáficas y de fertilización, además obtuvo respuestas que no corresponden a variaciones proporcionales del rendimiento máximo, lo cual revela diferentes relaciones funcionales entre el rendimiento del cultivo y los niveles de fertilizantes aplicados.

Objeción 4:

El cálculo del rendimiento relativo puede producir sesgos estadísticos. Esta objeción estadística se origina porque al dividir los rendimientos con el nutriente al mínimo entre el rendimiento máximo estable, aumenta la desviación del error del rendimiento máximo lo cual puede afectar la estimación de la regresión para estimar la relación funcional entre el rendimiento relativo y el contenido de nutriente en el suelo.

Objeción 5:

Las pruebas de significación para la regresión estimadas a partir de la combinación de datos provenientes de diferentes experimentos son inválidas. Esto ocurre porque las observaciones presentan diferentes tipos de error, un error propio de cada experimento y otro error generado por las condiciones de la localidad donde los experimentos se llevaron a cabo.

Estas objeciones muestran que a pesar de que el modelo discontinuo de respuesta lineal y de rendimiento máximo estable, es una metodología de fácil aplicación, se deben buscar otras opciones que permitan realizar recomendaciones de fertilizantes, estimar valores esperados de la respuesta de los cultivos, realizar otros análisis posteriores, como lo son los análisis económicos y que arrojen resultados que no presenten errores estadísticos sin controlar.

2.2.4. La función superficie de respuesta.

Myers (1971), define la función superficie de respuesta (FSR) como una aproximación de la función de producción a un polinomio de bajo orden.

Chacín (1983), plantea que la aproximación a un polinomio de bajo orden permite el establecimiento de relaciones funcionales de fácil manejo. La respuesta que se logra observar, está determinada no sólo por la relación funcional que presentan las variables, sino también por un componente aleatorio que se refiere a todos aquellos factores que no pueden ser controlados por el experimentador tales como:

- Variaciones en las condiciones y manejo del experimento.
- Errores en la medición en la variable respuesta.
- Variaciones intrínsecas del material experimental.
- Variaciones en las condiciones ambientales.

Agrega además que el establecimiento de la función de producción aproximada y su posterior análisis permite principalmente:

- Determinar variables que están incidiendo en la variable respuesta.
- Encontrar niveles de las variables que produzcan la máxima o mínima respuesta.
- Realizar un análisis económico que puede incluir isocuantas, isoclinas, etc.

Villasmil *et al* (1977), plantean que si se establece como uno de los objetivos fundamentales estimar la curva de respuesta, se debe pensar en un diseño que sin sacrificar información, permita obtener dicha estimación con el menor número de combinaciones. Los modelos de la FSR generalmente utilizan diseños de tratamientos para estimar respuestas de segundo orden.

Box y Hunter (1957), citados por Chacín (1980), hacen énfasis en la necesidad de escoger diseños de tratamientos que cumplan una serie de requisitos, estos serían:

- Aquellos que permitan la estimación de los parámetros de una FSR polinomial con una precisión satisfactoria dentro de una región de interés.
- Que el número de tratamientos no sea grande.
- Que el diseño posibilite una verificación de la exactitud, el diseño es eficiente siempre y cuando el grado del polinomio sea el adecuado.
- Que posibilite el uso del diseño experimental en bloques.

Villasmil *et al* (1977), plantean que es deseable que los diseños de tratamientos para estimar la FSR contemplen la utilización de un testigo y señala que la desventaja de los diseños compuestos utilizados en la industria es que no presentan esta característica.

Son muchos los diseños utilizados para estimar la FSR, entre los cuales se encuentran los experimentos factoriales, diseños ortogonales y rotables; el diseño San Cristóbal no ortogonalizado y el diseño San Cristóbal ortogonalizado, diseños hipercubos múltiples y los diseños doble estrella ortogonales y rotables, presentando cada uno de ellos sus ventajas y desventajas para las estimaciones realizadas a partir de ellos. (Chacín, 1993).

El diseño de tratamiento central rotatable y los diseños de tratamiento doble estrella ortogonales y rotables son los que presentan en general mejor comportamiento, ya que minimizan la varianza generalizada de los parámetros a estimar y la máxima varianza de cualquier error estimado. (Chacín, 1983; Machado, 1992; Ascanio, 1995).

Machado (1992), señala que si se desea obtener la mejor precisión en la estimación de los efectos lineales y cuadráticos de un polinomio de segundo orden el diseño más adecuado es el diseño central rotatable doble estrella.

Bandes *et al* (1991), aplicando el diseño central rotatable para estimar la FSR en ajo porro, no encontraron diferencias significativas entre los efectos estimados; por ello, señalan la importancia de realizar análisis de suelos previos al experimento para evitar cometer este tipo de errores.

Hernández (1993), aplicó esta metodología a varias localidades del Estado Yaracuy, logrando establecer funciones de producción y obteniendo las correspondientes dosis óptimas, encontrando además que la respuesta obtenida en las localidades estudiadas respondían bien al nitrógeno, pero no mostraron respuesta a los diferentes niveles de fósforo y potasio. En este caso, como se presentaron resultados absurdos para los nutrientes fósforo y potasio, se planteó la importancia de escoger como mucho cuidado el tratamiento central, para de esta manera obtener observaciones dentro de las cuales o muy cercana a ellas se encuentre el óptimo de los rendimientos y no solamente puntos de silla, ya que resulta muy costoso la elaboración de estos ensayos.

Aunque Hernández (1993), sólo logró establecer dosis óptimas para el nitrógeno a partir de las funciones de producción, nada impide obtener funciones de producción múltiples, como en el caso de Machado (1992), lo cual es una ventaja muy importante sobre los otros métodos de estimación vistos anteriormente, ya que la FSR puede considerar más de un nutriente al mismo tiempo.

La metodología de la FSR presenta la ventaja con respecto a las anteriores de poder estudiar más de un factor a la vez y de presentar dosis óptimas completas, sin embargo sólo genera una opción catalogada como óptima.

Es necesario recordar que este tipo de análisis no era aplicado anteriormente por lo complicado y engorroso de los cálculos, pero gracias a el avance de la informática se cuenta hoy en día con programas adecuados y con los equipos necesarios para poder evaluar otras alternativas diferentes a las planteadas anteriormente.

2.3. ANÁLISIS DISCRIMINANTE.

Los primeros intentos sistemáticos de intentar clasificar entre poblaciones se remontan a 1920 con los trabajos de Pearson, y luego a 1926 con los trabajos de Mahalanobis. Sin embargo el primer trabajo de lo que hoy conocemos como análisis discriminante se debe a Tildesley en 1921. (Montiel, 1986).

Montiel (1986), aplicó tres métodos: análisis de componentes principales, análisis de agrupamiento y análisis discriminante, para la identificación y caracterización de ambientes en Mantecal, Venezuela, además de evaluar el impacto de los módulos de apure en los paisajes y de la distribución de la flora y la fauna de los llanos venezolanos, siendo este un ejemplo de la aplicación de estos métodos multivariados en ecología.

Mata (1989), clasificó especímenes de cangrejos de las especies *Mithrax caribbaeus*, *Mithrax verrucosus* y *Mithrax forceps* utilizando el análisis por conglomerados y el análisis discriminante, hallando dos funciones discriminantes de cuatro variables para identificar a las especies estudiadas. Al utilizar análisis por conglomerados encontró que la capacidad de clasificación y de creación de grupos fue inferior a la presentada por el análisis discriminante.

Entre las áreas en las cuales se ha utilizado el análisis discriminante se pueden contar, taxonomía zoológica y botánica, economía, medicina, psicología, etc.

Montiel (1896) y Mata (1989), citan en sus trabajos un número considerable de investigaciones realizadas en análisis discriminante en distintas áreas ligadas a la biología.

En el área de la recomendación de fertilizantes de suelos no se encontraron trabajos en los cuales se aplique la metodología del análisis discriminante, por lo cual a pesar de que la existencia de esta metodología data desde hace bastante tiempo, el presente trabajo se muestra novedoso en el sentido de ser una aplicación diferente del análisis discriminante.

3. CONCEPTOS BÁSICOS DE ANÁLISIS DISCRIMINANTE.

Luego de revisar varios libros que contienen lo relativo al análisis discriminante, se optó por utilizar el Multivariate analysis de Mardia, Kent y Bibby (1979), del cual se extrajo lo siguiente:

3.1. Introducción.

Considérense g poblaciones o grupos $\Pi_1, \dots, \Pi_g, g \geq 2$. Supóngase que asociada con cada una de las poblaciones Π_j hay una función de densidad de probabilidad (*f.d.p*) $f_j(\mathbf{x})$ en R^p (espacio real p -dimensional), de tal manera que si una observación o individuo proviene de una población Π_j , ésta tiene la función de densidad de probabilidad $f_j(\mathbf{x})$. El objetivo del análisis discriminante es ubicar un individuo en alguno de esos g grupos a partir de la mediciones de las variables de \mathbf{x} .

Una regla discriminante d corresponde a una división de R^p en regiones disjuntas R_1, \dots, R_g ($\cup R_j = R^p$). La regla d es definida por:

Ubicar \mathbf{x} en Π_j , si $\mathbf{x} \in R_j$, para $j = 1, \dots, g$

La discriminación será más precisa si Π_j tiene más de su probabilidad concentrada en R_j para cada j .

Usualmente, no se posee información *a priori* acerca de cual población es probable que provenga una observación. Sin embargo, si tal información está disponible, puede ser incorporada en una estimación Bayesiana.

La situación donde las funciones de densidad de probabilidad $f_j(\mathbf{x})$ son conocidas es la más simple de analizar teóricamente, aunque en la práctica es la más difícil de hallar.

Una variante de esa situación ocurre cuando la forma de la distribución de probabilidad es conocida, pero hay parámetros que deben ser estimados. La estimación está basada en una matriz de datos $\mathbf{X}_{(n \times p)}$ cuyas filas están particionadas en g grupos,

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_g \end{bmatrix}$$

La matriz $\mathbf{X}_{j(n_j \times p)}$ representa una muestra de n_j individuos provenientes de la población Π_j .

Finalmente, existe una aproximación empírica al análisis discriminante donde no se asume ninguna forma matemática particular para las poblaciones Π_j , sino por una regla que nos permita discriminar entre ellas. Esta regla está basada en la función lineal discriminante de Fisher.

3.2. Discriminación cuando las poblaciones son conocidas.

3.2.1. Regla discriminante de máxima verosimilitud.

Considérese la situación donde las distribuciones de probabilidad de las poblaciones Π_1, \dots, Π_g son conocidas. El punto de partida del análisis es la intuitivamente plausible regla de la máxima verosimilitud. Se escribirá la *f.d.p.* de la j -ésima población como $f_j(\mathbf{x}) = L_j(\mathbf{x})$ para enfatizar que se está pensando en la verosimilitud de un punto \mathbf{x} como una función de j .

Definición 3.2.1.:

La regla discriminante de máxima verosimilitud (RDMV) para asignar una observación \mathbf{x} en alguna de las poblaciones Π_1, \dots, Π_g , consiste en ubicar a \mathbf{x} en aquella población que posea la mayor verosimilitud para \mathbf{x} .

Esto es, x se ubicará en Π_j , si:

$$L_j(\mathbf{x}) = \max_i L_i(\mathbf{x})$$

Si algunas funciones de máxima verosimilitud toman el mismo máximo valor, cualquiera de ellas puede ser escogida.

Ejemplo 3.2.1.1.:

Supóngase que $\Pi_1 \sim N(\mu_1, \sigma_1^2)$ y $\Pi_2 \sim N(\mu_2, \sigma_2^2)$. La función de máxima verosimilitud $L_i (i = 1, 2)$ es:

$$L_i(x) = (2\pi\sigma_i^2)^{-1/2} e^{\left\{ -\frac{1}{2} \left(\frac{x-\mu_i}{\sigma_i} \right)^2 \right\}}$$

$L_1(x)$ supera a $L_2(x)$ si:

$$\frac{\sigma_2}{\sigma_1} e^{\left\{ -\frac{1}{2} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - \left(\frac{x-\mu_2}{\sigma_2} \right)^2 \right] \right\}} > 1$$

Aplicando logaritmos y agrupando la desigualdad se obtiene:

$$x^2 \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) - 2x \left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} \right) + \left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} \right) < 2 \log \left(\frac{\sigma_2}{\sigma_1} \right)$$

Ejemplo 3.2.1.2.:

Un caso especial muy importante del ejemplo previo ocurre cuando $\sigma_1 = \sigma_2$, para el cual $L_1(x)$ supera a $L_2(x)$ cuando:

$$|x - \mu_2| > |x - \mu_1|$$

En otras palabras, si $\mu_2 > \mu_1$ entonces la regla discriminante de máxima verosimilitud asignará \mathbf{x} a Π_2 si $x > \frac{1}{2}(\mu_1 + \mu_2)$, y asignará \mathbf{x} a Π_1 en caso contrario. Si $\mu_1 > \mu_2$ la regla se invierte.

La generalización multivariada de este resultado es dada a continuación:

Teorema 3.2.1.:

(a) Si Π_i es una población $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, $i = 1, 2, \dots, g$, y $\boldsymbol{\Sigma} > \mathbf{0}$, entonces la regla discriminante de máxima verosimilitud asignará \mathbf{x} a Π_j , donde $j \in \{1, \dots, g\}$ si ese valor de j minimiza la distancia de Mahalanobis

$$(\mathbf{x} - \boldsymbol{\mu}_i)' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_i)$$

(b) Cuando $g = 2$, la regla asigna \mathbf{x} a Π_1 si:

$$\boldsymbol{\alpha}'(\mathbf{x} - \boldsymbol{\mu}) > 0$$

Donde $\boldsymbol{\alpha} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$ y $\boldsymbol{\mu} = \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$, y lo asignará a Π_2 en otro caso.

Demostración:

La i -ésima función de verosimilitud es

$$L_i(\mathbf{x}) = |2\pi\boldsymbol{\Sigma}|^{-1/2} e^{\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right\}}$$

Ésta es maximizada cuando el exponente es minimizado, lo cual provee la parte (a) del teorema.

Por otra parte, nótese que $L_1(\mathbf{x}) > L_2(\mathbf{x})$ si y sólo si

$$(\mathbf{x} - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_1) < (\mathbf{x} - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_2)$$

Cancelando y simplificando se obtiene la condición formulada en el teorema.

3.2.2. Regla discriminante de Bayes.

En ciertas situaciones es factible suponer que las poblaciones tienen probabilidades *a priori*. Por ejemplo, en diagnóstico médica se puede considerar a un estado gripal como intrínsecamente más probable que la polio. Esta información puede ser incorporada en el análisis usando la regla de la discriminación bayesiana (RDB). Para simplificar se supondrá que todas las probabilidades *a priori* π_j son estrictamente positivas, $j = 1, \dots, g$.

Definición 3.2.2.:

Si las poblaciones Π_1, \dots, Π_g tienen probabilidades *a priori* $(\pi_1 \dots \pi_g) = \boldsymbol{\pi}'$, entonces la regla discriminante de Bayes (con respecto a $\boldsymbol{\pi}$) asignará una observación \mathbf{x} a la población para la cual

$$\pi_j L_j(\mathbf{x}) \quad (3.2.2)$$

es maximizada.

La función en (3.2.2) puede ser considerada como proporcional a la probabilidad *a posteriori* de Π_j dados los datos \mathbf{x} . Nótese que la RDMV es un caso especial de la regla de Bayes, cuando todas las probabilidades *a priori* son iguales.

En el caso de la discriminación entre $g = 2$ poblaciones, el efecto de introducir probabilidades *a priori* es simplemente el de mover el valor crítico por una cantidad $\text{Log}(\pi_2/\pi_1)$. La regla queda:

$$\text{Asignar } \mathbf{x} \text{ a } \Pi_1 \text{ si } h(\mathbf{x}) > \text{Log}(\pi_2/\pi_1)$$

y a Π_2 en caso contrario.

3.2.3. Propiedades óptimas.

La RDB descrita anteriormente (incluyendo la RDMV), tiene ciertas propiedades óptimas. La primera, es que si $\mathbf{x}_1 = \mathbf{x}_2$, entonces \mathbf{x}_1 y \mathbf{x}_2 serán asignadas a la misma población.

Definición 3.2.3.1.:

Una regla aleatoria discriminante d implica asignar una observación \mathbf{x} a una población j con probabilidad $\Phi_j(\mathbf{x})$, donde Φ_1, \dots, Φ_g , son funciones no negativas definidas sobre R^p , que satisfacen $\sum \Phi_j(\mathbf{x}) = 1$ para toda \mathbf{x} .

Está claro que una regla de asignación determinística es un caso especial de la regla de asignación aleatoria, obtenida de hacer $\Phi_j(\mathbf{x}) = 1$ para $\mathbf{x} \in R_j$ y $\Phi_j(\mathbf{x}) = 0$ en otro caso. Por ejemplo, la regla de Bayes con respecto a las probabilidades *a priori* π_1, \dots, π_g está definida por:

$$\Phi_j(\mathbf{x}) = \begin{cases} 1 & \text{si } \pi_j L_j(\mathbf{x}) = \max_j \pi_j L_j(\mathbf{x}) \\ 0 & \text{en otro caso.} \end{cases}$$

Excepto para aquellas \mathbf{x} donde el máximo es obtenido por más de una población.

La probabilidad de colocar un individuo e la población Π_j , cuando en realidad pertenece a Π_j , está dada por:

$$p_{ij} = \int \Phi_i(\mathbf{x}) L_j(\mathbf{x}) d\mathbf{x} \quad (3.2.3)$$

Particularmente, si una observación proviene de Π_i , la probabilidad de realizar una asignación correcta es p_{ii} y la probabilidad de realizar una mala clasificación es $1 - p_{ii}$. El desempeño de la regla discriminante puede ser resumido por los números p_{11}, \dots, p_{gg} . La siguiente definición da un orden parcial en el establecimiento de reglas discriminantes.

Definición 3.2.3.2.:

Una regla discriminante d con probabilidades de correcta asignación $\{p_{ii}\}$ es tan buena como otra regla d' con probabilidades de correcta asignación $\{p'_{ii}\}$ si:

$$p_{ii} \geq p'_{ii}, \text{ para toda } i = 1, \dots, g$$

Se dice que d es mejor que d' si al menos una de las inecuaciones es exacta. Si d es una regla para la cual no hay otra regla mejor, se dice que esta regla d es admisible.

Nótese que no siempre se pueden comparar 2 reglas de asignaciones usando este criterio, por ejemplo, si $p_{11} > p'_{11}$ pero $p_{22} < p'_{22}$. No obstante, se puede probar la siguiente propiedad óptima de la RDB.

Teorema 3.2.3.1.:

Toda RDB (incluyendo RDMV) es admisible.

Teorema 3.2.3.2.:

Si las poblaciones Π_1, \dots, Π_g tienen probabilidades a priori π_1, \dots, π_g , entonces ninguna regla de discriminación tiene una probabilidad a posteriori de asignación correcta mayor que la de la regla de discriminación de Bayes con respecto a esa probabilidad a priori.

3.3. Discriminación bajo estimación.

3.3.1. La regla de discriminación muestral.

La regla de discriminación de máxima verosimilitud muestral (RDMVM) es útil cuando la forma matemática de la distribución de Π_1, \dots, Π_g es conocida, pero sus parámetros deben ser estimados de una matriz de datos $\mathbf{X}_{(n \times p)}$. Supóngase que las filas de \mathbf{X} están particionadas en g grupos, $\mathbf{X}' = (\mathbf{X}'_1 \dots \mathbf{X}'_g)$ y que \mathbf{X}_i contiene n_i observaciones provenientes de Π_i .

Por ejemplo, supóngase que los grupos son asumidos como provenientes de distribuciones multinormales con diferentes medias y la misma matriz de varianza-covarianza. Sean $\bar{\mathbf{x}}_i$ y \mathbf{S}_i la media y la matriz varianza-covarianza muestrales del i -ésimo grupo. Entonces los estimadores insesgados de $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_g$ y $\boldsymbol{\Sigma}$ son $\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_g$ y $\mathbf{S}_u = \sum n_i \mathbf{S}_i / (n - g)$ respectivamente. La RDMVM es entonces obtenida al sustituir estos estimadores en el teorema 3.2.1, particularmente, cuando $g = 2$, la regla asigna \mathbf{x} a Π_1 si y sólo si:

$$\mathbf{a}' \left\{ \mathbf{x} - \frac{1}{2}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2) \right\} > 0 \quad (3.3.1)$$

Donde $\mathbf{a} = \mathbf{S}_u^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$.

Ejemplo 3.3.1.1.:

Considérense las $n_1 = n_2 = 50$ observaciones en 2 especies de iris. *Iris setosa* e *I. Versicolor* (Mardia *et al*, 1979, p. 6.). Se discriminará entre ellas basados en 2 variables: longitud y ancho de los sépalos. Las medias muestrales y las varianzas para cada grupo están dadas por:

$$\bar{\mathbf{x}}_1 = (5.006 \quad 3.428)', \quad \bar{\mathbf{x}}_2 = (5.936 \quad 2.770)',$$

$$\mathbf{S}_1 = \begin{pmatrix} 0.1218 & 0.0972 \\ & 0.1408 \end{pmatrix}, \quad \mathbf{S}_2 = \begin{pmatrix} 0.2611 & 0.0835 \\ & 0.0965 \end{pmatrix}$$

$$\begin{aligned} \mathbf{a} &= \left[(50\mathbf{S}_1 + 50\mathbf{S}_2) / 98 \right]^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \\ &= \begin{pmatrix} 0.1953 & 0.0922 \\ 0.0922 & 0.1211 \end{pmatrix}^{-1} \begin{pmatrix} -0.930 \\ 0.658 \end{pmatrix} = \begin{pmatrix} -11.436 \\ 14.143 \end{pmatrix} \end{aligned}$$

y la regla discriminante está dada para asignar \mathbf{x} a Π_1 si:

$$h(\mathbf{x}) = (-11.436 \quad 14.143) \begin{pmatrix} x_1 - \frac{1}{2}(5.006 + 5.936) \\ x_2 - \frac{1}{2}(3.428 + 2.770) \end{pmatrix}$$

$$h(\mathbf{x}) = -11.436x_1 + 14.143x_2 + 18.739 > 0$$

y a Π_2 en caso contrario.

Considere que nos traen una flor de iris cuyas medias son $x_1 = 5.296$ y $x_2 = 3.213$. Debido a que $h(\mathbf{x}) = 3.615 > 0$, la clasificamos como *Iris setosa*.

Ejemplo 3.3.1.2.:

Extendiendo el último ejemplo para incluir las $n_3 = 50$ observaciones de la tercera especie de iris, *I. virginica*, la cual tiene la siguiente media y varianza muestral:

$$\bar{\mathbf{x}}_3 = \begin{pmatrix} 6.588 \\ 2.974 \end{pmatrix}, \mathbf{S}_3 = \begin{pmatrix} 0.3963 & 0.0919 \\ & 0.1019 \end{pmatrix}$$

En este caso Σ es estimada por:

$$\mathbf{S}_u = (50\mathbf{S}_1 + 50\mathbf{S}_2 + 50\mathbf{S}_3)/147 = \begin{pmatrix} 0.2650 & 0.0927 \\ 0.0927 & 0.1154 \end{pmatrix}$$

y la discriminación está basada en 3 funciones:

$$\begin{aligned} h_{12}(\mathbf{x}) &= (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{S}_u^{-1} \left\{ \mathbf{x} - \frac{1}{2}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2) \right\} \\ &= -7.657x_1 + 11.856x_2 + 5.153 \\ h_{13}(\mathbf{x}) &= -10.220x_1 + 12.147x_2 + 20.362 \\ h_{23}(\mathbf{x}) &= -2.562x_1 + 0.291x_2 + 15.208 \end{aligned}$$

Nótese que $h_{12}(\mathbf{x})$ no es idéntica a la función discriminante $h(\mathbf{x})$ del ejemplo 3.3.1.1, porque se ha utilizado una estimación de Σ diferente.

Entonces las regiones de asignación están definidas por:

$$\text{Ubicar } \mathbf{x} \text{ a } \begin{cases} \Pi_1 & \text{si } h_{12}(\mathbf{x}) > 0 \text{ y } h_{13}(\mathbf{x}) > 0 \\ \Pi_2 & \text{si } h_{12}(\mathbf{x}) < 0 \text{ y } h_{23}(\mathbf{x}) > 0 \\ \Pi_3 & \text{si } h_{13}(\mathbf{x}) < 0 \text{ y } h_{23}(\mathbf{x}) < 0 \end{cases}$$

La regla de decisión se obtiene al sustituir los valores de los vectores de medias en las ecuaciones para discriminar entre dos poblaciones y se observa el signo que toma la ecuación.

Al sustituir los valores de los vectores de medias de *Iris setosa* en las ecuaciones de $h_{12}(\mathbf{x})$ y $h_{13}(\mathbf{x})$, se observa que para $h_{12}(\mathbf{x}) = 7.464 > 0$ y para $h_{13}(\mathbf{x}) = 10.406 > 0$, como las ecuaciones toman estos signos para esta población, de allí se originan las reglas de decisión. De manera similar es el procedimiento para establecer las reglas de decisión para *I. virginica* e *I. versicolor*.

Se escribirá:

$$h_{ij}(\mathbf{x}) = (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)' \mathbf{S}_u^{-1} \mathbf{x} - \frac{1}{2} \bar{\mathbf{x}}_i' \mathbf{S}_u^{-1} \bar{\mathbf{x}}_i + \frac{1}{2} \bar{\mathbf{x}}_j' \mathbf{S}_u^{-1} \bar{\mathbf{x}}_j$$

Para $i \neq j$, es fácil de verificar que las funciones discriminantes están linealmente relacionadas por:

$$h_{12}(\mathbf{x}) + h_{23}(\mathbf{x}) = h_{13}(\mathbf{x})$$

Así la frontera consiste en tres rectas coincidentes en el punto donde:

$$h_{12}(\mathbf{x}) = h_{23}(\mathbf{x}) = h_{13}(\mathbf{x}) = 0$$

3.3.2. ¿Es la discriminación importante?

Considérese la situación donde se quiere discriminar entre g poblaciones multinormales con la misma matriz de varianza-covarianza, y los parámetros son estimados a partir de la matriz de datos $X = \left(X_1', \dots, X_g' \right)$. Si todas las medias poblacionales son iguales, μ_1, \dots, μ_g , entonces no tiene sentido tratar de discriminar entre estos grupos sobre la base de esas variables.

Sin embargo, algunas veces, aunque las medias poblacionales son iguales, las medias muestrales $\bar{x}_1, \dots, \bar{x}_g$ son diferentes unas de otras, por lo cual un aparentemente plausible análisis discriminante puede ser llevado a cabo. Así, para probar si el análisis discriminante es o no importante, es interesante evaluar las hipótesis $\mu_1 = \dots = \mu_g$ dado que $\Sigma_1 = \dots = \Sigma_g$. Esta prueba de hipótesis es exactamente el análisis de varianza multivariado para el diseño completamente aleatorizado.

Recordando que 2 posibles pruebas de esta hipótesis son obtenidas por la partición de matriz de suma de cuadrados y productos totales $T = X'HX$, como $T = W + B$, donde W y B son las matrices de suma de cuadrados y productos dentro de grupos y entre grupos, respectivamente. Un esquema de esta partición se presenta en el cuadro 1.

Cuadro 1. Esquema del análisis de varianza multivariado.

<i>Fuente de variación</i>	<i>Grados de libertad</i>	<i>Matriz</i>
<i>Categorías</i>	$t - 1$	W
<i>Error</i>	$n - t$	B
<i>Total</i>	$n - 1$	T

Fuente: Mardia *et al.* 1979.

Entonces la prueba Λ de Wilk y la prueba de la mayor raíz característica están dadas como funciones de los autovalores de $\mathbf{W}^{-1}\mathbf{B}$. Particularmente, si $g = 2$, entonces $\mathbf{W}^{-1}\mathbf{B}$ tiene un sólo autovalor diferente de cero y las pruebas son iguales y equivalentes a la prueba T^2 de Hotelling para dos muestras. Luego, bajo la hipótesis nula,

$$\left\{ \frac{n_1 n_2 (n-2)}{n} \right\} \mathbf{d}' \mathbf{W}^{-1} \mathbf{d} \sim T^2_{(p, n-2)} = \left\{ \frac{(n-2)p}{n-p-1} \right\} F_{(p, n-p-1)}$$

donde $\mathbf{d} = \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2$.

La hipótesis nula es rechazada para valores altos de este estadístico.

3.3.3. Función discriminante lineal de Fisher.

Otra aproximación al problema de la discriminación basado en una matriz de datos \mathbf{X} puede ser hecho sin asumir ninguna forma paramétrica particular para las distribuciones poblacionales Π_1, \dots, Π_g , pero sí por la búsqueda de una regla “sensible” para discriminar entre ellas. La sugerencia de Fisher era el buscar para la función lineal $\mathbf{a}'\mathbf{x}$ la cual haga máxima la razón entre la suma de cuadrados “entre grupos” y “dentro de grupos”; esto es,

$$\text{sea } \mathbf{y} = \mathbf{X}\mathbf{a} = \begin{bmatrix} \mathbf{X}_1 \mathbf{a} \\ \vdots \\ \mathbf{X}_g \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_g \end{bmatrix}$$

una combinación lineal de las columnas de \mathbf{X} . Entonces \mathbf{y} tiene una suma de cuadrados total:

$$\mathbf{y}'\mathbf{H}\mathbf{y} = \mathbf{a}'\mathbf{X}'\mathbf{H}\mathbf{X}\mathbf{a} = \mathbf{a}'\mathbf{T}\mathbf{a}$$

la cual puede ser particionada como una suma de cuadrados dentro de grupos,

$$\sum \mathbf{y}_i' \mathbf{H}_i \mathbf{y}_i = \sum \mathbf{a}' \mathbf{X}_i' \mathbf{H}_i \mathbf{X}_i \mathbf{a} = \mathbf{a}' \mathbf{W} \mathbf{a}$$

más la suma de cuadrados entre grupos,

$$\sum n_i (\bar{y}_i - \bar{y})^2 = \sum n_i \{ \mathbf{a}' (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}) \}^2 = \mathbf{a}' \mathbf{B} \mathbf{a}$$

donde, \bar{y}_i es la media del i -ésimo sub-vector y_i de \mathbf{y} , y \mathbf{H}_i es la $(n_i \times n_i)$ matriz central.

El criterio de Fisher es intuitivamente atractivo porque es un facilitador para separar los grupos si la suma de cuadrados para \mathbf{y} es alta en relación con la suma de cuadrados dentro de los grupos. La razón está dada por:

$$\mathbf{a}' \mathbf{B} \mathbf{a} / \mathbf{a}' \mathbf{W} \mathbf{a} \quad (3.3.3.1)$$

Si \mathbf{a} es el vector que maximiza (3.3.3.1), la función lineal $\mathbf{a}' \mathbf{x}$ se denomina función lineal discriminante de Fisher (FLDF).

Teorema 3.3.3.1.:

El vector \mathbf{a} en la FLDF es el autovector de $\mathbf{W}^{-1} \mathbf{B}$ correspondiente al mayor autovalor.

Una vez que la FLDF ha sido calculada, una observación \mathbf{x} puede ser asignada a alguna de las g poblaciones sobre la base de su “valor discriminante” $\mathbf{a}' \mathbf{x}$. Las medias muestrales $\bar{\mathbf{x}}_i$ tienen valores $\mathbf{a}' \bar{\mathbf{x}}_i = \bar{y}_i$. Entonces \mathbf{x} es asignado a aquella población cuya media está más cercana a $\mathbf{a}' \mathbf{x}$; esto es, asignar \mathbf{x} a Π_j si:

$$| \mathbf{a}' \mathbf{x} - \mathbf{a}' \bar{\mathbf{x}}_j | < | \mathbf{a}' \mathbf{x} - \mathbf{a}' \bar{\mathbf{x}}_i |, \text{ Para todo } i \neq j$$

La FLDF es importante en el caso especial de $g = 2$ grupos. Entonces \mathbf{B} tiene rango 1 y puede ser escrito como:

$$\mathbf{B} = \left(\frac{n_1 n_2}{n} \right) \mathbf{d} \mathbf{d}',$$

Donde $\mathbf{d} = \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2$. Así, $\mathbf{W}^{-1} \mathbf{B}$ tiene solamente un autovalor diferente de cero que puede ser calculado explícitamente. Este autovalor es igual:

$$\text{tr} \mathbf{W}^{-1} \mathbf{B} = \left(\frac{n_1 n_2}{n} \right) \mathbf{d}' \mathbf{W}^{-1} \mathbf{d},$$

y el correspondiente autovector es:

$$\mathbf{a} = \mathbf{W}^{-1}\mathbf{d}$$

Entonces la regla discriminante queda:

$$\text{asignar } \mathbf{x} \text{ a } \Pi_1 \text{ si } \mathbf{d}'\mathbf{W}^{-1}\left\{\mathbf{x} - \frac{1}{2}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)\right\} > 0 \quad (3.3.3.2)$$

y a Π_2 , en caso contrario.

Nótese que la regla de asignación dada por (3.3.3.2) es exactamente igual a la RDMVM (3.3.1) para 2 grupos provenientes de una distribución multinormal con la misma matriz de varianza-covarianza. Sin embargo, las justificaciones para esta regla son completamente diferentes en los dos casos. En (3.3.1) hay una explícita condición de normalidad, mientras que en (3.3.3.2), tenemos sólo una regla “sensible” basada en una función lineal de \mathbf{x} . Así, se puede esperar que esa regla sea apropiada para poblaciones donde la hipótesis de normalidad no está satisfecha.

Para $g \geq 3$ grupos, la regla de asignación basada en la FLDF y RDMVM para poblaciones multinormales con la misma matriz de varianza-covarianza no son las mismas, salvo que las medias muestrales sean colineales (aunque las dos reglas serán similares si las medias son cercanamente colineales).

3.4. Probabilidades de mala clasificación.

3.4.1. Probabilidades cuando los parámetros son estimados.

Formalmente, las probabilidades de mala clasificación p_{ij} están dadas en (3.2.3). Si los parámetros de las distribuciones bajo estudio son estimados a partir de la matriz de datos, entonces se pueden estimar las probabilidades \hat{p}_{ij} .

Por ejemplo, considérese el caso de dos poblaciones normales $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ y $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$.

Si $\boldsymbol{\mu} = \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$, entonces cuando \mathbf{x} proviene de Π_1 ,

$\boldsymbol{\alpha}'(\mathbf{x} - \boldsymbol{\mu}) \sim N_p\left[\frac{1}{2}\boldsymbol{\alpha}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2), \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\right]$. Dado que la función discriminante está dada por

$h(\mathbf{x}) = \boldsymbol{\alpha}'(\mathbf{x} - \boldsymbol{\mu})$ con $\boldsymbol{\alpha} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$, se observa que si \mathbf{x} proviene de Π_1 ,

$h(\mathbf{x}) \sim N\left(\frac{1}{2}\Delta^2, \Delta^2\right)$, donde:

$$\Delta^2 = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$$

Es el cuadrado de la distancia de Mahalanobis entre las poblaciones. De forma similar, si \mathbf{x} proviene de Π_2 , $h(\mathbf{x}) \sim N\left(-\frac{1}{2}\Delta^2, \Delta^2\right)$.

Así, la probabilidad de mala clasificación está dada por:

$$\begin{aligned} p_{12} &= P(h(\mathbf{x}) > 0 \mid \Pi_2) \\ &= \Phi\left(-E(h)/\sqrt{V(h)}\right) = \Phi\left(-\frac{1}{2}\Delta\right) \end{aligned}$$

Donde Φ es la función de la distribución normal estándar, de forma similar, también

$$p_{21} = \Phi\left(-\frac{1}{2}\Delta\right).$$

Si los parámetros son estimados de la matriz de datos, entonces un estimador de Δ^2 es:

$$D^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathcal{S}_u^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2),$$

y la probabilidad de mala clasificación estimada es $\hat{p}_{12} = \hat{p}_{21} = \Phi\left(-\frac{1}{2}D\right)$.

Desafortunadamente, este método tiende a ser optimista, es decir, subestima las verdaderas probabilidades de mala clasificación o clasificación equivocada.

Ejemplo 3.4.1.:

En la discriminación entre *I. setosa* e *I. versicolor* en el ejemplo (3.3.1.1), tenemos que $D^2 = 19.94$. Por lo tanto la probabilidad de mala clasificación estimada es:

$$\Phi\left(-\frac{1}{2}D\right) = 0.013 \text{ ó } 1.3\%$$

3.4.2. Método de la resustitución.

Supóngase que la discriminación está basada en la matriz de datos X de la cual n_j individuos provienen de la población j . Si la regla discriminante está definida por las regiones de asignación R_i , sea n_{ij} el número de individuos de Π_j que pertenecen a R_i , (así $\sum_i n_{ij} = n_j$). Entonces $\hat{p}_{ij} = n_{ij}/n_j$ es un estimador de p_{ij} . Desafortunadamente este método también tiende a ser optimista para las probabilidades de mala clasificación.

Ejemplo 3.4.2.:

Considérese la matriz de datos de *I. setosa* (Π_1) e *I. versicolor* (Π_2) nuevamente. Si se conoce que una observación de Π_2 fue asignada a Π_1 , quedará:

$$\hat{p}_{12} = 0, \hat{p}_{21} = 1/50$$

Y sí se asume que $p_{12} = p_{21}$, entonces se tiene que el único estimador es $\hat{p}_{12} = \hat{p}_{21} = 1/100$, el cual está en general de acuerdo con el ejemplo (3.4.1). Sin embargo, nótese que para tal pequeña probabilidad de mala clasificación, esas mediciones muestrales son pequeñas para dar estimaciones exactas.

3.5. Descarte de variables.

Considérese el problema de la discriminación entre 2 poblaciones normales multivariadas con medias μ_1, μ_2 y una matriz común de varianza-covarianza Σ . Los coeficientes de la función discriminante de máxima verosimilitud poblacional $\alpha'x$ están dados por:

$$\alpha = \Sigma^{-1}\delta, \text{ donde } \delta = \mu_1 - \mu_2$$

En la práctica estos parámetros son estimados por $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2$ y $\mathcal{S}_u = m^{-1}(n_1\mathcal{S}_1 + n_2\mathcal{S}_2) = m^{-1}\mathbf{W}$, donde $m = n_1 + n_2 - 2$. Haciendo $\mathbf{d} = \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2$, los coeficientes de la función discriminante de máxima verosimilitud muestral, quedan:

$$\mathbf{a} = m\mathbf{W}^{-1}\mathbf{d}$$

Particionando $\boldsymbol{\alpha}' = (\boldsymbol{\alpha}'_1, \boldsymbol{\alpha}'_2)$ y $\boldsymbol{\delta}' = (\boldsymbol{\delta}'_1, \boldsymbol{\delta}'_2)$, donde $\boldsymbol{\alpha}_1$ y $\boldsymbol{\delta}_1$ tienen k componentes, y suponiendo que $\boldsymbol{\alpha}_2 = 0$; es decir, suponer que las variables $\mathbf{x}_{k+1}, \dots, \mathbf{x}_p$ no tienen poder de discriminación una vez que las variables han sido tomadas en cuenta, y por lo tanto pueden ser descartadas. Nótese que la hipótesis $\boldsymbol{\alpha}_2 = 0$ es equivalente a $\boldsymbol{\delta}_{2,1} = 0$, donde $\boldsymbol{\delta}_{2,1} = \boldsymbol{\delta}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\delta}_1$. Es equivalente a $\Delta_p^2 = \Delta_k^2$, donde:

$$\Delta_p^2 = \boldsymbol{\delta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}, \quad \Delta_k^2 = \boldsymbol{\delta}'_1\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\delta}_1$$

Es decir, la distancia de Mahalanobis entre las poblaciones es la misma, esté basada en los primeros k componentes como en todos los p componentes.

Una prueba de la hipótesis $H_0: \boldsymbol{\alpha}_2 = 0$ usando la distancia de Mahalanobis muestral

$$D_p^2 = m\mathbf{d}'\mathbf{W}^{-1}\mathbf{d} \quad \text{y} \quad D_k^2 = m\mathbf{d}'_1\mathbf{W}_{11}^{-1}\mathbf{d}_1,$$

ha sido propuesto por Rao (1973). Esta prueba usa el estadístico:

$$\left\{ \frac{(m-p+1)}{(p-k)} \right\} c^2 \left(D_p^2 - D_k^2 \right) / \left(m + c^2 D_k^2 \right) \quad (3.5.1)$$

donde $c^2 = n_1 n_2 / n$. Bajo la hipótesis nula, (3.5.1) sigue la distribución de F con $(p-k, m-p+1)$ grados de libertad, y rechazaremos H_0 para valores grandes del estadístico propuesto.

La aplicación más importante de esta prueba ocurre con $k = p - 1$, donde se prueba la importancia de una variable en particular con respecto a las demás variables tomadas en cuenta. En este caso el estadístico (3.5.1) puede ser simplificado usando la inversa de la matriz de la suma de cuadrados y productos total $\mathbf{T}^{-1} = (t^{ij})$. Considerando la hipótesis $H_0: \alpha_i = 0$, donde i es fijo. Entonces (3.5.1), con D_k ahora representando la distancia de Mahalanobis basada en todas las variables excepto la i -ésima, nos queda:

$$(m - p + 1)c^2 a_i^2 / \left\{ m t^{ij} (m + c^2 D_p^2) \right\} \quad (3.5.2)$$

Este estadístico tiene distribución de F con $(1, m - p + 1)$ grados de libertad, cuando $\alpha_i = 0$.

Desde luego, el estadístico (3.5.2) es estrictamente válido solamente si la variable i es seleccionada antes de realizar el análisis. Sin embargo, es conveniente muchas veces observar el valor de ese estadístico para ver cuales variables son más importantes.

Ejemplo 3.5.1.:

Considérense las dos especies de iris del ejemplo (3.3.1.1), usando las dos primeras variables, longitud y ancho de los sépalos. Aquí $n_1 = n_2 = 50$, $m = 100 - 2 = 98$, y la función discriminante tiene coeficientes $\mathbf{a}' = (-11.436 \quad 14.143)$. Es fácil ver que:

$$\mathbf{T}^{-1} = \begin{pmatrix} 0.02562 & 0.00707 \\ & 0.04602 \end{pmatrix}$$

Y que $D_p^2 = 19.94$. Así las dos pruebas de F para la importancia de cada una de las variables dada la otra de (3.5.2), nos quedan, tal y como se muestra en el Cuadro 2:

Cuadro 2. Coeficientes discriminantes para los datos de iris y sus valores de F respectivos.

<i>Variable</i>	<i>a_{ij}</i>	<i>F</i>
a_1	-11.436	211.8
a_2	14.143	108.3

Fuente: Mardia *et al.* 1979.

Dado que $F_{(1, 97, 0.01)} = 7.0$, se puede concluir que ambas variables son altamente útiles en la discriminación.

Sin embargo, una conclusión diferente se sigue si se discrimina basados en las cuatro variables: longitud y ancho de sépalos, longitud y ancho de pétalos. Entonces los coeficientes de discriminación y sus F correspondiente están dados en el Cuadro 3.

Cuadro 3. Coeficientes discriminantes para los datos de iris y sus valores de F respectivos.

<i>Variable</i>	<i>a_{ij}</i>	<i>F</i>
a_1	3.053	0.716
a_2	18.023	25.7
a_3	-21.766	24.6
a_4	-30.844	10.7

Fuente: Mardia *et al.* 1979.

Como $F_{(1, 97, 0.01)} = 7.0$, el estadístico F para longitud de sépalo no es significativo, así se concluye que la longitud del sépalo es redundante si todas las demás variables están presentes. Nótese que los estadísticos F del Cuadro 3, son todos menores que los valores de F expuestos en Cuadro 2. Este comportamiento refleja el hecho empírico de que como el número de variables aumenta, la información aportada por cualquier variable no aportada por otras variables, tiende a decrecer.

3.6. ¿Cuándo la correlación mejora la discriminación?

Podría pensarse que una combinación lineal de dos variables podría proporcionar una mejor discriminación si ellas estuvieran correlacionadas, que si no lo estuvieran. Sin embargo, esto no es necesariamente así, como se mostrará en el siguiente ejemplo:

Ejemplo 3.6.:

Sean Π_1 y Π_2 dos poblaciones normales bivariadas (Cochran, 1962). Supóngase que $\Pi_1 \sim N_2(0, \Sigma)$ y $\Pi_2 \sim N_2(\boldsymbol{\mu}, \Sigma)$, donde $\boldsymbol{\mu}' = (\mu_1 \quad \mu_2)$ y:

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Ahora la distancia de Mahalanobis entre Π_1 y Π_2 nos queda:

$$\Delta^2 = \boldsymbol{\mu}' \Sigma^{-1} \boldsymbol{\mu} = (\mu_1^2 + \mu_2^2 - 2\rho\mu_1\mu_2)/(1 - \rho^2)$$

Si las variables no están correlacionadas nos queda:

$$\Delta_0^2 = \mu_1^2 + \mu_2^2$$

Ahora, la correlación mejorará la discriminación, si y sólo si $\Delta^2 > \Delta_0^2$. Esto ocurre si y sólo si:

$$\rho\{(1 + f^2)\rho - 2f\} > 0$$

Donde $f = \mu_2/\mu_1$.

En otras palabras, la discriminación se verá mejorada, salvo si ρ toma valores entre cero y $2f/(1 + f^2)$, pero un valor bajo de ρ puede ser perjudicial para la discriminación. Note que si $\mu_1 = \mu_2$, cualquier valor positivo del coeficiente de correlación reduce el poder de discriminación.

4. MATERIALES Y MÉTODOS.

4.1. CARACTERIZACIÓN DE LA INFORMACIÓN.

La aplicación del análisis discriminante para estimar rendimientos en maíz (*Zea mays*) con la aplicación de los nutrientes nitrógeno (N), fósforo (P) y potasio (K) se realizó a partir de la base de datos generada por el proyecto de calibración de suelos efectuado por el FONAIAP y coordinado por el Dr. Ricardo Ramírez, acompañado por un grupo importante de investigadores. (Ramírez *et al*, 1988a y 1988b).

Los resultados de rendimiento en grano, expresado en kg/ha, correspondieron a sesenta y ocho (68) grupos de experimentos en maíz llevados a cabo entre los años 1981 a 1984, utilizando híbrido “Arichuna” distribuidos en diversas localidades de los estados Apure, Guárico, Monagas, Portuguesa, Táchira y Yaracuy. (Cuadros 4, 5, 6, 7, 8, 9 y 10).

Las matrices de las combinaciones de N, P y K, que constituyeron los tratamientos fueron generadas bajo un diseño de superficie de respuesta con arreglo de tratamiento central rotatable en un diseño en bloques al azar con dos repeticiones, las cuales se mantuvieron fijas para cada estado, independientemente de la ubicación de los experimentos. Con la excepción de Yaracuy en donde fueron probados 30 dosis, en cada estado se utilizaron 15 dosis o tratamientos, es decir, 15 combinaciones de N, P y K. (Cuadros 11, 11, 12, 13, 14, 15, 16 y 17).

La aplicación del análisis discriminante se realizará para cada combinación de nutrientes de cada estado, debido a que las dosis fueron específicas para cada uno de estos y solo coincidieron las de los estados Táchira y Guárico. En estas últimas dos entidades, el análisis discriminante se realizará en conjunto para las dosis de ambos estados, ya que las combinaciones de nutrientes aplicadas en cada uno fueron las mismas. (Cuadros 16 y 17).

Las razones que condujeron a seleccionar esta base de datos fueron las siguientes:

- Las observaciones disponibles representan respuestas a la aplicación de fertilizantes en diversas condiciones agroecológicas, donde se produce comercialmente el cultivo del maíz en Venezuela.

- Disponibilidad de una información de difícil obtención, la cual se puede utilizar para generar alternativas para la recomendación de fertilizantes.
- Las observaciones fueron obtenidas a partir de un diseño de tratamiento central rotatable, lo cual garantiza un número bastante considerable de combinaciones de fertilizantes para cada uno de los estados estudiados, es decir se presentan combinaciones que varían desde dosis con bajos niveles de nutrientes, hasta dosis con elevados niveles.
- Disponibilidad de información relativa a la caracterización edáfica de los sitios experimentales.

Cuadro 4. Ubicación de los experimentos en las distintas localidades del estado Yaracuy.

<i>Distrito</i>	<i>Localidad</i>	<i>Responsable</i>
<i>Sucre</i>	<i>Tibana II Marrón</i>	<i>Elsa Meléndez</i>
<i>Sucre</i>	<i>Iboa</i>	<i>Elsa Meléndez</i>
<i>Yaritagua</i>	<i>Radio Faro</i>	<i>Omar Colmenares</i>
<i>Yaritagua</i>	<i>La Tapa</i>	<i>Omar Colmenares</i>
<i>Bruzual</i>	<i>Chara Vieja</i>	<i>Omar Colmenares</i>
<i>Bruzual</i>	<i>La Palma</i>	<i>Omar Colmenares</i>
<i>Sucre</i>	<i>Tibana I (Negro)</i>	<i>Omar Colmenares</i>
<i>Urachiche</i>	<i>Belisa</i>	<i>Omar Colmenares</i>
<i>Bruzual</i>	<i>El Ceibal</i>	<i>Omar Colmenares</i>
<i>Bruzual</i>	<i>La Bartola</i>	<i>Dunja Beg, Pedro Marcal</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 5. Ubicación de los experimentos en las distintas localidades del estado Portuguesa.

<i>Distrito</i>	<i>Localidad</i>	<i>Responsable</i>
<i>Ospino</i>	<i>La Aparición</i>	<i>Samuel Cabrera</i>
<i>Turén</i>	<i>Villa Bruzual</i>	<i>Samuel Cabrera</i>
<i>Turén</i>	<i>Villa Bruzual</i>	<i>Samuel Cabrera</i>
<i>Turén</i>	<i>Santa Rosalía</i>	<i>Samuel Cabrera</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 6. Ubicación de los experimentos en las distintas localidades del estado Yaracuy.

<i>Distrito</i>	<i>Localidad</i>	<i>Responsable</i>
	<i>Jobo Liso</i>	<i>Dunja Beg</i>
<i>Urachiche</i>	<i>El Matico</i>	<i>Dunja Beg</i>
	<i>San Ramón</i>	<i>Dunja Beg</i>
	<i>Tartazal</i>	<i>Dunja Beg</i>
<i>Sucre</i>	<i>Iboa</i>	<i>Dunja Beg</i>
	<i>Yamaro</i>	<i>Dunja Beg</i>
<i>Yaritagua</i>	<i>El Palmar</i>	<i>Omar Colmenares</i>
<i>Yaritagua</i>	<i>Las Velas</i>	<i>Omar Colmenares</i>
<i>Yaritagua</i>	<i>La Tapa</i>	<i>Omar Colmenares</i>
<i>Urachiche</i>	<i>Sabana de Parra</i>	<i>Dunja Beg</i>
	<i>Camunare</i>	<i>Dunja Beg</i>
	<i>La Virgen</i>	<i>Dunja Beg</i>
	<i>Sabana Larga</i>	<i>Dunja Beg</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 7. Ubicación de los experimentos en las distintas localidades del estado Monagas.

<i>Distrito</i>	<i>Localidad</i>	<i>Responsable</i>
	<i>Tejero</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Sta. Bárbara</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Merecure B.</i>	<i>Jesús Tenías</i>
	<i>Viento Fresco</i>	<i>Jesús Tenías</i>
	<i>Medina</i>	<i>Jesús Tenías</i>
	<i>Potrerito</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>San Ramón</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Sta. Bárbara I.</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Sta Bárbara II</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Casupal</i>	<i>Jesús Tenías</i>
	<i>Viento Fresco</i>	<i>Jesús Tenías</i>
	<i>Punta Gorda</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Sta. Bárbara Gris</i>	<i>Jesús Tenías</i>
<i>Maturín</i>	<i>Sta. Bárbara Rojo</i>	<i>Jesús Tenías</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 8. Ubicación de los experimentos en las distintas localidades del estado Apure.

<i>Distrito</i>	<i>Municipio</i>	<i>Responsable</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>
<i>Achaguas</i>	<i>Queseras del Medio</i>	<i>César Hernández</i>
<i>Achaguas</i>	<i>Queseras del Medio</i>	<i>César Hernández</i>
<i>Achaguas</i>	<i>Queseras del Medio</i>	<i>César Hernández</i>
<i>Achaguas</i>	<i>Queseras del Medio</i>	<i>César Hernández</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>
<i>Achaguas</i>	<i>Queseras del Medio</i>	<i>César Hernández</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>
<i>Achaguas</i>	<i>Queseras del Medio</i>	<i>César Hernández</i>
<i>San Fernando</i>	<i>Biruaca</i>	<i>César Hernández</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 9. Ubicación de los experimentos en las distintas localidades del estado Guárico.

<i>Distrito</i>	<i>Localidad</i>	<i>Responsable</i>
<i>Zaraza</i>	<i>Zaraza</i>	<i>Alberto Riera</i>
<i>Ribas</i>	<i>Tucupido</i>	<i>Alberto Riera</i>
<i>Infante</i>	<i>Chaguaramas</i>	<i>Alberto Riera</i>
<i>Infante</i>	<i>Valle de la Pascua</i>	<i>Alberto Riera</i>
<i>Zaraza</i>	<i>Zaraza</i>	<i>Edgar Colmenares</i>
<i>Ribas</i>	<i>Tucupido</i>	<i>Edgar Colmenares</i>
<i>Infante</i>	<i>Chaguaramas</i>	<i>Edgar Colmenares</i>
<i>Infante</i>	<i>Valle de la Pascua</i>	<i>Edgar Colmenares</i>
<i>Ribas</i>		<i>Edgar Colmenares</i>
<i>Infante</i>	<i>Las Mercedes del Llano</i>	<i>Edgar Colmenares</i>
<i>Infante</i>	<i>Chaguaramas</i>	<i>Edgar Colmenares</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 10. Ubicación de los experimentos en las distintas localidades del estado Táchira.

<i>Distrito</i>	<i>Localidad</i>	<i>Responsable</i>
<i>Libertador</i>	<i>El Tropezón. Isla de Vetancourt</i>	<i>Pedro Yáñez</i>
<i>Libertador</i>	<i>Las Cruces. Isla de Vetancourt</i>	<i>Pedro Yáñez</i>
<i>Libertador</i>	<i>Guafitas</i>	<i>Pedro Yáñez</i>

Fuente: Base de Datos del FONAIAP. Maracay.

Cuadro 11. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Yaracuy.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
37	24	20
37	24	80
37	96	20
143	24	20
143	96	20
143	24	80
37	96	80
143	96	80
0	60	50
180	60	50
90	0	50
90	120	50
90	60	0
96	60	100
90	60	50¹

Fuente: Base de Datos del FONAIAP. Maracay.

¹ La combinación resaltada representa el Tratamiento central.

Cuadro 12. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Portuguesa.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
30	20	16
30	20	64
30	80	16
120	20	16
120	80	16
120	20	64
30	80	64
120	80	64
0	50	40
150	50	40
75	0	40
75	100	40
75	50	0
75	50	80
75	50	40 ²

Fuente: Base de Datos del FONAIAP. Maracay.

² La combinación resaltada representa el Tratamiento central.

Cuadro 13. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Yaracuy.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
49	32	20
49	32	80
49	128	20
191	32	20
191	128	20
191	32	80
49	128	80
191	128	80
0	80	50
240	80	50
120	0	50
120	160	50
120	80	0
120	80	100
120	80	50³

Fuente: Base de Datos del FONAIAP. Maracay.

³ La combinación resaltada representa el Tratamiento central.

Cuadro 14. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Monagas.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
36	41	28
36	41	112
36	159	28
144	41	28
144	159	28
144	41	112
36	159	112
144	159	112
0	100	70
180	100	70
90	0	70
90	200	70
90	100	0
90	100	140
90	100	70⁴

Fuente: Base de Datos del FONAIAP. Maracay.

⁴ La combinación resaltada representa el Tratamiento central.

Cuadro 15. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Apure.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
57	12	8
57	12	32
57	48	8
223	12	8
223	48	8
223	12	32
57	48	32
223	48	32
0	30	20
280	30	20
140	0	20
140	60	20
140	30	0
140	30	40
140	30	20⁵

Fuente: Base de Datos del FONAIAP. Maracay.

⁵ La combinación resaltada representa el Tratamiento central.

Cuadro 16. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Guárico.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
47	49	24
47	49	96
47	191	24
183	49	24
183	191	24
183	49	96
47	191	96
183	191	96
0	120	60
230	120	60
115	0	60
115	240	60
115	120	0
115	120	120
115	120	60⁶

Fuente: Base de Datos del FONAIAP. Maracay.

⁶ La combinación resaltada representa el Tratamiento central.

Cuadro 17. Combinaciones de nutrientes (tratamientos) utilizados en las distintas localidades del estado Táchira.

<i>Dosis (kg/ha)</i>		
<i>N</i>	<i>P</i>	<i>K</i>
47	49	24
47	49	96
47	191	24
183	49	24
183	191	24
183	49	96
47	191	96
183	191	96
0	120	60
230	120	60
115	0	60
115	240	60
115	120	0
115	120	120
115	120	60⁷

Fuente: Base de Datos del FONAIAP. Maracay.

⁷ La combinación resaltada representa el Tratamiento central.

4.2. APLICACIÓN DE LA METODOLOGÍA.

Con la finalidad de obtener los grupos de discriminación, a partir de las combinaciones de nutrientes (tratamientos del diseño central rotatable) se partitionaron las observaciones en tres categorías de respuesta:

1. *Rendimiento alto*, si: $\text{Rendimiento} > \bar{x}_i + s_i$.
2. *Rendimiento medio*, si: $\bar{x}_i - s_i < \text{Rendimiento} < \bar{x}_i + s_i$.
3. *Rendimiento bajo*, si: $\text{Rendimiento} < \bar{x}_i - s_i$.

Donde \bar{x}_i es la media de producción y s_i es la desviación estándar de los rendimientos para la i -ésima combinación de dosis de nutrientes aplicada como tratamiento.

Esta partición proviene del hecho de que para aplicar el análisis discriminante es necesario particionar un espacio completo p -dimensional en dos (2) o más grupos o categorías de discriminación y luego, a partir de las características muestrales ubicar una observación x en alguno de esos grupos o categorías de discriminación. (Giri, 1977; Linares, 1986; Mardia *et al*, 1979).

El criterio de discriminación utilizado es la regla de discriminación paramétrica basada en la distancia de Mahalanobis calculada a partir de la matriz de varianza-covarianza y con una aproximación bayesiana estimada a partir de las probabilidades *a priori* de pertenecer a alguna de las tres (3) categorías de discriminación establecidas.

Las probabilidades *a priori* de pertenecer a alguna de las tres (3) categorías de discriminación se estimaron a partir de la proporción muestral de las observaciones provenientes de cada una de las categorías.

Las ecuaciones generadas por este análisis son ecuaciones discriminantes lineales, donde una observación x es clasificada dentro alguna de las tres categorías de discriminación, si esta observación x posee la menor distancia de Mahalanobis para alguna de las tres categorías,

esto ocurre si dicha observación obtiene el mayor valor en la ecuación discriminante para esa categoría.

La probabilidad de mala clasificación se estimó a partir de las proporciones de las observaciones clasificadas en una categoría incorrecta, es decir, esto constituye una estimación de la probabilidad de clasificar las observaciones en categorías incorrectas y coincide con el método de la resustitución descrito en la revisión de literatura.

Se descartaron aquellas combinaciones de nutrientes para las cuales se obtuvieron pocas observaciones, entendiendo por pocas observaciones las que presentaron menos de $p+1$ observaciones, donde p es el número de variables edáficas utilizadas en el análisis discriminante; también se descartaron las combinaciones en las cuales no se pudieron establecer las tres categorías de discriminación, es decir, en aquellos tratamientos en los cuales sólo se pudieron crear dos categorías.

Además, entre todas las variables edáficas medidas en los diferentes ensayos se escogieron aquellas que se consideran en un análisis rutinario de suelo, esto con la finalidad de facilitar el desarrollo de la metodología propuesta, para estimar el requerimiento de fertilizantes a partir de muestras de suelo con fines de fertilidad que el productor lleva al laboratorio.

Las variables que se consideraron para la realización del análisis discriminante son las siguientes:

- *Fósforo* medido en ppm (P_{olsen}), extraído por el método de Olsen.
- *Potasio* medido en ppm (K_{olsen}), extraído por el método de Olsen.
- *pH* (pH), medido en solución 1:2.5.
- *Carbono orgánico* (C.O.), medido en porcentaje (%).
- *Conductividad eléctrica* (C.E.), en dMhos/cm.
- *Capacidad de Intercambio Catiónico* (C.I.C.), medida en meq./100 gr. de suelo.

- *Calcio* (Ca), intercambiable medido en meq./100 gr. de suelo en el estado Yaracuy; y disponible medido en ppm en los estados Apure, Guárico, Monagas y Táchira.
- *Magnesio* (Mg), intercambiable medido en meq./100 gr. de suelo en el estado Yaracuy; y disponible medido en ppm en los estados Apure, Guárico, Monagas y Táchira.
- *Arena* (a), medida en porcentaje (%).
- *Arcilla* (A), medida en porcentaje (%).

Con las observaciones correspondientes a cada una de las dosis consideradas, se procedió a la realización del análisis discriminante utilizando el programa *SAS Versión 6.04*, programa con el cual fueron descartadas aquellas variables que no generaron información importante para la discriminación en cada uno de los tratamientos, dejando sólo las que en realidad aportaban información en las ecuaciones discriminantes.

Para ello se utilizó el procedimiento DISCRIM del programa SAS, el cual es específico para realizar análisis discriminante, además de utilizar el procedimiento FORMAT para etiquetar adecuadamente a cada una de las clases de discriminación con nombres fácilmente identificables al ser impresos los resultados.

Dentro de las declaraciones para el desarrollo de la sintaxis correcta en el análisis de la matriz de datos, se utilizaron las siguientes: INPUT, FORMAT y LABEL.

- En la declaración INPUT se introdujeron los nombres de cada una de las variables en el mismo orden en el cual aparecen en la matriz de datos.
- En la declaración FORMAT dentro de la caracterización de la base de datos utilizando la declaración DATA se especificó que la variable CATEG sea etiquetada con los valores especificados en el procedimiento FORMAT.
- En la declaración LABEL se especifica cuales variables edáficas identifica cada variable, de manera que al aparecer impresos los resultados en la salida del programa, aparezca también la significación de cada una de las variables analizadas, esto con la finalidad de identificar fácilmente a las variables utilizadas para realizar la discriminación.

Dentro de las opciones del procedimiento DISCRIM, se utilizaron las siguientes:

- SIMPLE, que es la opción que arroja los estadísticos univariados para el total de los datos y para cada grupo de clasificación.
- STDMEAN, que presenta todas las medias y desviaciones estándares para cada grupo de clasificación y para el total de los datos.
- DISTANCE, presenta la distancia de Mahalanobis entre cada uno de los grupos de clasificación.
- ANOVA, presenta los análisis de varianza univariados para cada variable de cada grupo de clasificación, probando la hipótesis nula $H_0: \mu_1 = \dots = \mu_g$, contra la alternativa $H_1: \mu_i \neq \mu_j, i \neq j$.
- MANOVA, presenta los análisis de varianza multivariados para los vectores de medias de cada grupo de clasificación, probando la hipótesis nula $H_0: \mu_1 = \dots = \mu_g$, contra la alternativa $H_1: \mu_i \neq \mu_j, i \neq j$.
- METHOD = NORMAL, esta opción especifica que el método a utilizar es el análisis discriminante paramétrico y no el análisis discriminante no paramétrico o el análisis discriminante canónico.
- POOL = YES, esta opción utiliza la matriz de covarianza general para el cálculo de la distancia de Mahalanobis.
- LISTERR, presenta una lista de las observaciones mal clasificadas, ordenándolas por grupo de clasificación y además, realiza una estimación de la probabilidad de mala clasificación para cada grupo y para el modelo en general.
- OUTSTAT = CORNSTAT, en esta opción se crea una base de datos con cada uno de los estadísticos utilizados para realizar el análisis discriminante y con cada uno de los resultados del mismo, esta matriz de datos de salida es una base de datos creada por el mismo programa SAS y que sólo se puede ver utilizando el procedimiento PRINT.

El procedimiento DISCRIM, presenta otras declaraciones, de las cuales se utilizaron: CLASS, VAR, PRIORS, además de agregar tres títulos a los resultados, el primero para identificar la dosis utilizada, el segundo para identificar el estado del cual provienen los datos y el tercero para especificar que se utiliza la función discriminante lineal.

- La declaración CLASS especifica a que variables se les realizará el análisis discriminante, en nuestro caso se escribió CATEG, el cual es el nombre que identificó a las categorías de clasificación dentro de cada matriz de datos.
- La declaración VAR especifica que variables deben tomarse en cuenta para la realización del análisis discriminante, en nuestro caso éstas variables dependieron de la dosis a analizar.
- La declaración PRIORS especifica la probabilidad *a priori* de clasificación dentro de cada clase, en nuestro caso se utilizó la opción PROPORTIONAL, la cual indica que las probabilidades *a priori* serán iguales a la proporción o al “peso” de las observaciones presentes en cada grupo de clasificación con respecto al total de los datos.

Es importante señalar que las ecuaciones discriminantes sólo serán presentadas por el programa SAS, si el método es paramétrico (METHOD = NORMAL) y si además se utiliza la matriz de covarianza (POOL = YES), de lo contrario el programa sólo presentará la distancia de Mahalanobis y las observaciones mal clasificadas. Un ejemplo de como queda estructurada una matriz de datos para ser corrida por el programa SAS es presentada en la Figura 2. (SAS INSTITUTE INC., 1989).

Como consecuencia de la partición en las tres categorías mencionadas anteriormente, los límites de cada categoría son distintos y van a depender de los valores que tomen los estadísticos media aritmética (\bar{x}_i) y desviación estándar (s_i) dentro de la muestras correspondientes a cada una de las combinaciones de dosis de fertilizantes.

Luego de descartar las combinaciones de nutrientes con pocas observaciones, entre los cuales se encontraron todos los experimentos llevados a cabo en el estado Portuguesa, y las

combinaciones en las cuales no fue posible establecer las tres categorías de clasificación, se les aplicó el análisis discriminante a los tratamientos restantes.

```
PROC FORMAT;
  VALUE RENDTO
    1='ALTO'
    2='MEDIO'
    3='BAJO';
  VALUE RENDTOS
    1='A'
    2='M'
    3='B';
RUN;
DATA CORN;
  TITLE 'ANALISIS DISCRIMINANTE 0-100-70 EN MAIZ';
  INPUT N CATEG REND POLSEN KOLSEN PH11 CO CE CA;
  FORMAT CATEG RENDTO.;
  LABEL
    REND='RENDIMIENTO EN KG/HA'
    POLSEN='P EN PPM'
    KOLSEN='K EN PPM'
    PH11='PH EN SOLUCION 1:1'
    CO='CARBONO ORGANICO EN %'
    CE='COND. ELECT. EN dMHOS/CM'
    CA='CA EN PPM (SOLUCION)';
CARDS;
Observaciones
;
PROC DISCRIM DATA=CORN OUTSTAT=CORNSTAT
  SIMPLE STDMEAN DISTANCE ANOVA MANOVA
  METHOD=NORMAL POOL=YES LISTERR;
CLASS CATEG;
VAR PH11 CO CE CA;
TITLE2 'ESTADO MONAGAS';
TITLE3 'USING LINEAR DISCRIMINANT FUNCTION';
PRIORS PROPORTIONAL;
RUN;
```

Figura 2. Esquema de las declaraciones necesarias para ser corridas por el programa SAS, para realizar el análisis discriminante.

Los resultados (*Output* o salida del programa) del análisis discriminante corrido con el programa *SAS Versión 6.04* se estructuran en cada uno los siguientes apartes:

- *Class level information:* en este aparte se presenta un resumen de la información ingresada al programa; esta información incluye número de observaciones y grados de libertad, además de la frecuencia y proporción de las observaciones por cada categoría de discriminación.
- *Simple statistics:* se presentan los estadísticos univariados para las observaciones y para cada una de las clases, éstos incluyen el total de las observaciones, media aritmética, sumatoria de las observaciones, varianza y desviación estándar.
- *Total sample standardized class means:* presenta las medias de clases estandarizadas y las matrices de covarianza entre grupos.
- *Pooled covariance matrix information:* presenta el rango de la matriz de covarianza y el logaritmo natural del determinante de la matriz de covarianza.
- *Pairwise squared distance between groups:* presenta la distancia de Mahalanobis sin tomar en cuenta las probabilidades *a priori* de cada una de las categorías de clasificación.
- *Pairwise generalized squared distance between groups:* presenta la distancia de Mahalanobis tomando en cuenta las probabilidades *a priori* de cada una de las categorías de clasificación.
- *Univariate test statistics:* presenta los resultados de las pruebas de significación univariadas, probando la hipótesis nula $H_0: \mu_1 = \dots = \mu_g$, contra la alternativa $H_1: \mu_i \neq \mu_j, i \neq j$.
- *Multivariate statistics and F approximations:* presenta los resultados de las pruebas de significación multivariadas, probando la hipótesis nula $H_0: \mu_1 = \dots = \mu_g$, contra la alternativa $H_1: \mu_i \neq \mu_j, i \neq j$.

- *Linear discriminant function:* presenta las ecuaciones de discriminación lineales para cada una de las categorías de clasificación.
- *Resubstitution results for calibration data:* presenta las observaciones mal clasificadas y sus probabilidades *a posteriori* de pertenecer a cada una de las categorías de discriminación.
- *Classification summary for calibration data:* presenta un resumen de las observaciones mal clasificadas y una estimación de la probabilidad de mala clasificación total (\hat{p}_{ij}) y para cada uno de las categorías de discriminación $(\hat{p}_{ij(g)})$.

Un ejemplo de como quedan los resultados del análisis discriminante al ser corridos con el programa *SAS versión 6.04* se presentan en el anexo 9.

Luego de obtenidas las correspondientes ecuaciones de discriminación se procedió a la optimización de las dosis requeridas para cada una de las situaciones, es decir a seleccionar aquellas combinaciones de productos comerciales que suplan los requerimientos de nutrientes de las dosis al más bajo costo, esto se llevó a cabo utilizando el programa *MPS versión 2.1*, el cual optimiza utilizando la metodología de la programación lineal. (Pfeiffer, 1983).

Este programa a diferencia del programa *SAS*, no requiere de la construcción de una matriz de datos para ser introducida al programa, sino que ésta debe ser creada al mismo tiempo de ingresar los datos. El esquema de ingreso de datos sería el siguiente:

- Declarar el nombre de los productos comerciales.
- Ingresar el contenido de nutrientes de los productos comerciales.
- Ingresar el precio por cada producto comercial.
- Especificar los requerimientos de nutrientes para cada una de las dosis a optimizar.
- Se debe declarar que se desea minimizar costos.

Desgraciadamente este programa posee una interface muy deficiente, por lo cual la salida es directamente a la impresora y no se puede observar la matriz de datos al ser ingresada sino después de haberlo hecho, es decir, es menester imprimirla una vez introducida. Sin embargo en la salida (*Solution Output*) del programa se puede observar:

- *LP problem file name*: en donde se presenta el nombre del archivo a optimizar.
- *Problem type*: en donde aparece especificado la naturaleza del problema, en nuestro caso aparece *min*, ya que se quiere minimizar costos, además aparece en cuantas iteraciones fue encontrada la solución óptima.
- *Date*: aparece la fecha en la cual fue corrido el programa.
- *Algorithm start time*: aparece la hora en la cual se empezó a correr el programa.
- *Algorithm end time*: aparece la hora en la cual se terminó de correr el programa.
- *Objective function*: se presenta el costo de la combinación de productos comerciales que resultó más económica.
- *Section 1 - rows*: se presentan los requerimientos de nutrientes a optimizar, caracterizados por el límite inferior y superior de éstos.
- *Section 2 - columns*: presentan las características de los productos comerciales considerados en la optimización, se presenta además la cantidad de producto comercial para la combinación óptima y en cuanto se incrementaría el costo de la dosis en caso de utilizar un producto comercial diferente al recomendado por el programa. (Pfeiffer, 1983).

Un ejemplo de como quedan los resultados del cálculo de las combinaciones de productos comerciales más económicas al ser corridos con el programa *MPS versión 2.1* se presenta en el anexo 10.

Las dosis óptimas de producto comercial encontradas para cada una de las combinaciones de tratamientos usando el programa *MPS ver 2.1*, fueron seleccionadas a partir de la lista de los productos comerciales presentados en el Cuadro 18.

Cuadro 18. Listado de productos comerciales utilizados en el cálculo de las dosis óptimas de productos comerciales para las combinaciones de tratamientos utilizados en el análisis discriminante.

<i>Producto Comercial</i>	<i>N</i>	<i>P₂O₅</i>	<i>K₂O</i>	<i>Precio (Bs./100 kg)</i>
<i>Urea</i>	46	0	0	8880
<i>Sulfato de amonio</i>	21	0	0	5750
<i>Fosfato diamónico</i>	18	40	0	13950
<i>Fosfato especial</i>	16	42	0	12500
<i>Cloruro de potasio</i>	0	0	60	8910
<i>Sulfato de potasio</i>	0	0	50	15506
<i>Sulfato doble de potasio</i>	0	0	22	10588
<i>12-24-12 CP</i>	12	24	12	12500
<i>12-24-12/3 CP</i>	12	24	12	13656
<i>14-14-14/11 CP</i>	14	14	14	11428
<i>10-26-26 CP</i>	10	26	26	14652
<i>12-12-17/2 CP</i>	12	12	17	11808
<i>12-12-17/2 SP</i>	12	12	17	14296

Fuente: Lista de precios de fertilizantes. PEQUIVEN. 1998.

Se escogió la lista de precios de PEQUIVEN, debido a que ellos son los principales distribuidores y vendedores de fertilizantes en nuestro país.

Se tomó como unidad de producto comercial 100 kg de fertilizante, porque es equivalente a dos (2) sacos de producto comercial, y además como los requerimientos de nutrientes para

cada dosis vienen expresados en porcentaje, se decidió expresarlo en base a 100 kg para simplificar los cálculos, debido a que los resultados vendrán expresados en unidades de 100 kg de fertilizante.

El programa presenta además la cantidad de cada producto comercial necesario para suplir cada nutriente y el costo por hectárea de cada combinación.

5. RESULTADOS Y DISCUSIÓN.

5.1. COMBINACIONES DE NUTRIENTES UTILIZADAS EN EL ANÁLISIS DISCRIMINANTE.

Los resultados de la partición en tres categorías de discriminación arrojaron límites distintos para cada una de las categorías de discriminación en cada una de las dosis de nutrientes, tal y como se mencionó en el capítulo de Materiales y Métodos, ya que estos límites dependieron de los valores que tomaron los estadísticos media aritmética (\bar{x}_i) y desviación estándar (s_i) dentro de cada una de las muestras correspondientes a cada una de las combinaciones de las dosis de nutrientes estudiadas.

Las combinaciones de nutrientes descartadas con pocas observaciones ($n < p + 1$), entre las cuales se encontraron todos los experimentos llevados a cabo en el estado Portuguesa, y las combinaciones en las cuales no fue posible establecer las tres categorías de clasificación, se presentan a continuación. (Cuadros 19, 20, 21, 22 y 23).

Cuadro 19. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en el estado Yaracuy en el primer grupo de experimentos.

<i>Tratamiento</i>	<i>Nº de Obs.</i>	\bar{x}_i	s_i	<i>Límites (kg/ha)</i>	
				<i>Bajo</i> <i>(Lím. Superior)</i>	<i>Alto</i> <i>(Lím. Inferior)</i>
37-24-20	18	3329.94	1207.96	2121.98	4537.90
37-24-80	18	3411.00	1384.07	2026.92	4795.07
37-96-20	18	2924.81	1303.56	1224.28	4274.64
143-24-20	18	4061.30	1528.52	2532.79	5589.82
143-96-20	18	3788.12	1292.41	2495.70	5080.53
143-24-80	18	3903.63	1078.14	2825.49	4981.77
37-96-80	17	3271.37	1247.05	2024.31	4518.42
0-60-50	18	2749.46	1525.18	1224.28	4274.64
180-60-50	18	4359.41	1284.22	3075.19	5643.63
90-120-50	18	3571.57	1241.83	2329.74	4813.39
90-60-0	18	3874.77	1625.13	2249.64	5499.90
90-60-100	18	3732.25	1166.38	2565.88	4898.63
90-60-50	108	3637.02	1199.56	2437.47	4836.58

Cuadro 20. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites las categorías de discriminación en el estado Yaracuy en el segundo grupo de experimentos.

<i>Tratamiento</i>	<i>Nº de Obs.</i>	\bar{x}_i	s_i	<i>Límites (kg/ha)</i>	
				<i>Bajo</i> <i>(Lím. Superior)</i>	<i>Alto</i> <i>(Lím. Inferior)</i>
49-32-20	20	2568.47	1004.49	1563.98	3572.97
49-32-80	20	3235.38	1406.03	1829.35	4641.42
191-32-20	23	3918.51	1385.70	2532.81	5304.21
191-128-20	23	3707.37	1574.99	2132.38	5282.36
191-32-80	23	4110.48	1791.35	2319.12	5901.83
49-128-20	23	2755.95	1688.02	1067.93	4443.96
191-128-80	23	4069.17	1818.57	2250.60	5887.75
0-80-50	17	2702.04	1179.85	1522.19	3881.90
240-80-50	23	4122.50	1657.86	2464.64	5780.36
120-0-50	23	3501.45	1760.66	1740.79	5262.11
120-160-50	23	3561.38	1659.76	1901.62	5221.15
120-80-0	23	3338.71	1424.37	1914.33	4763.08
120-80-100	23	3558.45	1615.38	1943.07	5173.84
120-80-50	138	3701.57	1551.95	2149.62	5253.53

Cuadro 21. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en el estado Monagas.

<i>Tratamiento</i>	<i>Nº de Obs.</i>	\bar{x}_i	s_i	<i>Límites (kg/ha)</i>	
				<i>Bajo</i> <i>(Lím. Superior)</i>	<i>Alto</i> <i>(Lím. Inferior)</i>
36-41-28	28	2598.81	1781.08	817.73	4379.89
36-41-112	27	2493.42	1915.73	577.68	4409.15
36-159-28	28	2451.35	1822.62	628.73	4273.97
144-41-28	28	2798.61	1846.75	951.86	4645.36
144-159-28	28	2814.96	1631.07	1183.89	4446.03
144-41-112	28	2772.02	1765.29	1006.73	4537.31
36-159-112	28	2477.50	1668.23	809.27	4145.73
144-159-112	28	3138.25	1768.03	1370.23	4906.28
0-100-70	24	2017.97	1490.18	527.79	3508.15
180-100-70	28	3382.38	1774.05	1608.33	5156.43
90-0-70	26	3016.50	1878.29	1138.21	4894.78
90-200-70	25	3167.38	1910.33	1257.05	5077.71
90-100-0	27	3005.35	1679.77	1325.58	4685.13
90-100-140	28	2725.44	1614.67	1110.76	4340.11
90-100-70	165	2944.05	1713.17	1230.87	4657.22

Cuadro 22. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en el estado Apure.

<i>Tratamiento</i>	<i>Nº de Obs.</i>	\bar{x}_i	s_i	<i>Límites (kg/ha)</i>	
				<i>Bajo</i> (<i>Lím. Superior</i>)	<i>Alto</i> (<i>Lím. Inferior</i>)
57-12-8	21	3581.20	993.39	2587.81	4574.59
57-48-8	18	3901.16	1379.45	2521.71	5280.61
223-12-8	20	4010.49	1138.92	2871.57	5149.40
223-48-8	21	4150.34	1188.26	2962.07	5338.60
223-12-32	19	4184.72	1247.15	2937.57	5431.87
57-48-32	18	4344.33	1019.78	3324.55	5364.11
223-48-32	21	4469.98	1389.77	3080.21	5859.76
0-30-20	19	3643.18	1505.87	2137.32	5149.05
280-30-20	20	4296.15	1072.83	3223.32	5368.97
140-0-20	20	4163.28	811.78	3351.50	4975.06
140-60-20	21	4099.60	932.74	3166.86	5032.34
140-30-0	21	4203.96	1053.69	3150.26	5257.65
140-30-40	21	4173.86	1119.64	3054.22	5293.50
140-30-20	124	4152.73	1238.81	2913.91	5391.54

Cuadro 23. Combinaciones de nutrientes utilizados para la realización del análisis discriminante y límites de las categorías de discriminación en los estados Guárico y Táchira.

<i>Tratamiento</i>	<i>Nº de Obs.</i>	\bar{x}_i	s_i	<i>Límites (kg/ha)</i>	
				<i>Bajo</i> <i>(Lím. Superior)</i>	<i>Alto</i> <i>(Lím. Inferior)</i>
<i>47-49-24</i>	<i>17</i>	2427.29	1294.05	1133.24	3721.34
<i>47-191-24</i>	<i>17</i>	2412.18	1263.88	1148.31	3676.06
<i>183-49-24</i>	<i>17</i>	3658.66	1460.53	2198.13	5119.18
<i>183-191-24</i>	<i>17</i>	3361.85	1252.86	2108.99	4614.70
<i>183-49-96</i>	<i>17</i>	3540.88	1224.57	2316.31	4765.45
<i>47-191-96</i>	<i>17</i>	2479.12	1316.54	1162.57	3795.66
<i>183-191-96</i>	<i>17</i>	3388.68	1778.83	1609.85	5167.52
<i>0-120-60</i>	<i>17</i>	1964.61	1313.41	651.20	3278.02
<i>230-120-60</i>	<i>17</i>	3898.58	1263.48	2635.10	5162.06
<i>115-0-60</i>	<i>17</i>	2935.79	1618.51	1317.29	4554.30
<i>115-120-0</i>	<i>17</i>	3147.67	1310.49	1837.18	4458.16
<i>115-120-120</i>	<i>17</i>	3270.20	1425.50	1844.70	4695.70
<i>115-120-60</i>	<i>102</i>	3049.63	1518.41	1531.22	4568.04

5.2. ECUACIONES DISCRIMINANTES.

5.2.1. Estado Yaracuy.

En el estado Yaracuy fueron factibles de analizar un total de 27 combinaciones de nutrientes. Las dosis de nutrientes estudiadas en los experimentos del estado Yaracuy mostraron los siguientes rangos, N: desde 0 Kg./Ha. hasta 240 Kg./Ha., P: desde 0 Kg./Ha. hasta 160 Kg./Ha. y K: desde 0 Kg./Ha. hasta 100 Kg./Ha. (Cuadros 19 y 20).

La amplitud de los niveles estudiados y la variabilidad de las condiciones edáficas existentes, aunado al número de variables tomadas en cuenta en las ecuaciones discriminantes en las localidades estudiadas permitió obtener observaciones de rendimientos ubicados en un amplio espectro de respuesta. Como consecuencia de este hecho, se puede suponer la existencia de una asociación entre las distintas condiciones edáficas y las distintas categorías de rendimiento, es decir, existirán condiciones de suelos asociadas a rendimientos bajos y también a rendimientos altos, esta variabilidad presente en los valores tomados por las variables de suelo en las localidades estudiadas en el estado Yaracuy aunado a la información presentada por las variables consideradas, fue lo que permitió encontrar ecuaciones de discriminación con una baja probabilidad de mala clasificación.

Como puede observarse en el Cuadro 24, las condiciones de suelo varían ampliamente, pasando de suelos muy pobres en nutrientes minerales, con un contenido de sales moderado y un pH ácido, hasta suelos ricos en nutrientes minerales, con presencia de salinidad y pH alcalino, variando también la textura, desde suelos que presentan gran cantidad de arcilla hasta suelos con un alto contenido de arena.

Cuadro 24. Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en el estado Yaracuy.

VARIABLES DE SUELO MEDIDAS EN EL ESTADO YARACUY										
Valores	P_{olsen}	K_{olsen}	pH	C.E.	C.I.C.	Ca	Mg	C.O.	a	A
	ppm		Sol 1:2.5	dMhos/cm	meq/100 gr. de suelo				%	
Mínimo	2	43	4.2	0	4.5	0.3	0.05	0.7	8	13.7
Máximo	52	302	8.3	2.68	22	41.4	3.8	2.35	65	46.5

Las ecuaciones discriminantes encontradas para las dosis consideradas en el estado Yaracuy muestran una buena discriminación, esto refleja el hecho de que las probabilidades de clasificación errada obtenidas son en la mayoría de las veces menores a 0.5, hallándose incluso numerosas dosis (13) en las cuales dichas probabilidades toman el valor de cero para las tres categorías. (Anexo 1).

Sin embargo, el hecho de tener probabilidades iguales a cero de clasificación errada para las diferentes categorías de rendimiento en numerosas dosis, no implica necesariamente que a nivel poblacional se mantenga esta tendencia, ya que según Mardia *et al* (1979), el método de la resustitución tiende a subestimar la probabilidad de clasificación errada, no se descarta pues, que en muestras más grandes, como en el caso de los tratamientos centrales, en los cuales se presenta el mayor número de repeticiones, las probabilidades de mala clasificación son diferentes y mayores que la tendencia general mostrada por el resto de las combinaciones de nutrientes, de hecho, en el tratamiento central 90-60-50, la probabilidad de clasificación errada del modelo de discriminación en general fue de 0.2056, de 0.0952 para la clasificación en la categoría *Alto* y de 0.8125 para la clasificación en la categoría *Bajo*. (Anexo 1).

En algunas dosis se presentaron errores acentuados para clasificar dentro de alguna categoría, esto ocurre debido a que las medias muestrales de las variables de suelos para dos categorías se encuentran muy cercanas entre si, como en el caso del tratamiento 90-60-50 (Cuadro 25), que muestra una probabilidad altísima para clasificar erradamente en la categoría *Bajo*. Esto ocurre porque las medias de las variables edáficas para las categorías de rendimiento *Medio* y *Bajo* se encuentran muy cercanas entre si, como lo muestran las medias

muestrales de las variables *pH*, *C.E.*, *Ca*, *Mg*, *a* y *A*, las cuales presentan medias similares entre las dos categorías pero bastante disímiles de las correspondientes a la categoría de rendimiento *Alto*. Es de esperarse que los valores de las medias muestrales de los valores de rendimiento se encuentren también cercanos entre sí.

Cuadro 25. Medias muestrales de las variables de suelo para la dosis 90-60-50 en el estado Yaracuy.

MEDIAS MUESTRALES DE LAS VARIABLES DE SUELO								
Categorías	<i>P</i>_{olsen}	<i>K</i>_{olsen}	<i>pH</i>	<i>C.E.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>
	<i>ppm</i>		<i>Sol 1:2.5</i>	<i>dMhos/cm</i>	<i>meq/100 gr. de suelo</i>		%	
<i>Alto</i>	13.95	108.57	7.63	0.60	25.92	0.56	25.04	27.18
<i>Bajo</i>	4.25	63.31	6.99	0.42	14.54	0.94	43.71	19.44
<i>Medio</i>	7.06	90.93	6.84	0.48	15.01	1.01	40.68	22.19
Total muestra	7.99	90.26	7.02	0.49	17.08	0.91	38.06	22.76

La consecuencia directa de esta situación se encuentra reflejada en los valores de los coeficientes de la ecuación de discriminación, estos coeficientes resultan similares para aquellas variables edáficas cuyas medias muestrales son parecidas, por lo cual el peso en la clasificación para dicha variable en esas categorías disminuye. Los coeficientes de las ecuaciones en las variables mencionadas anteriormente son muy similares, sobre todo en el caso de la variable *C.E.*, en la cual se observan valores muy parecidos entre los coeficientes de las ecuaciones de discriminación para las categorías de rendimiento *Medio* y *Bajo*, y una diferencia bastante notoria entre el coeficiente correspondiente a la categoría de rendimiento *Alto*. (Cuadro 26).

Cuadro 26. Coeficientes de discriminación correspondientes a las ecuaciones de la dosis 90-60-50 en el estado Yaracuy.

COEFICIENTES DE LAS ECUACIONES DE DISCRIMINACIÓN									
Categorías	Constante	P_{olsen}	K_{olsen}	pH	C.E.	Ca	Mg	a	A
<i>Alto</i>	-497.30866	-2.30263	-0.4819	60.00629	-134.14306	-0.73997	-73.3686	7.51376	21.00973
<i>Bajo</i>	-493.72696	-2.85577	-0.40003	62.78296	-147.92469	-1.03721	-64.92809	7.52997	20.154
<i>Medio</i>	-484.5423	-2.66194	-0.38491	59.45386	-148.271	-0.68133	-64.54354	7.66467	20.24524

Algo muy distinto ocurre cuando las medias muestrales de las variables edáficas están alejadas entre si (Cuadro 27 y 28), en este caso los coeficientes de discriminación tienden a ser diferentes para cada grupo de clasificación y las probabilidades de mala clasificación son bajas en cada una de las categorías. Esto se explica porque las características de suelos asociadas a una categoría de rendimiento son lo bastante diferentes como para garantizar una buena clasificación. Un buen ejemplo de este caso lo constituye la dosis 180-60-50, la cual presenta una probabilidad de mala clasificación de cero y coeficientes de discriminación bastante diferentes para cada una de las tres categorías. Como se explicó anteriormente, las probabilidades de mala clasificación pueden estar subestimadas, se espera que al aumentar las observaciones los estimadores tomen valores más parecidos a los valores poblacionales.

Cuadro 27. Medias muestrales de las variables de suelo para la dosis 180-60-50 en el estado Yaracuy.

MEDIAS MUESTRALES DE LAS VARIABLES DE SUELO				
Categorías	pH	C.E.	a	A
	Sol 1:2.5	dMhos/cm	%	
<i>Alto</i>	7.76	0.65	21.13	27.33
<i>Bajo</i>	7.13	0.45	50.5	18.1
<i>Medio</i>	6.8	0.47	38.95	22.83
Total muestra	7.02	0.49	37.91	22.86

Cuadro 28. Coeficientes de discriminación correspondientes a las ecuaciones de la dosis 180-60-50 en el estado Yaracuy.

COEFICIENTES DE LAS ECUACIONES DE DISCRIMINACIÓN								
Categorías	Constante	pH	C.E.	a	A	$\hat{P}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
Alto	-677.58477	83.61798	-415.88668	15.69668	23.0542	0.0	0.0	0.1667
Bajo	-969.79584	99.30719	-516.31414	19.15584	27.74135	0.0		0.1667
Medio	-851.35348	93.23267	-482.79062	17.90463	26.09555	0.0		0.6667

La inclusión de la estimación de las probabilidades *a priori* para cada una de las categorías se realizó con la finalidad de minimizar el error de clasificación. Es decir, cuando una observación proveniente de la categoría de rendimiento *Bajo* es clasificada de manera errónea, es más probable que el modelo la clasifique en la categoría de rendimiento *Medio* y no en la categoría de rendimiento *Alto*, esto por supuesto, hace más confiable a las ecuaciones de discriminación ya que permite esperar que las clasificaciones erradas sean ubicadas en la categoría siguiente y no en la categoría más alejada. En el caso de la combinación de nutrientes correspondientes a la dosis 120-0-50, se observa que las malas clasificaciones correspondientes a la categorías de rendimientos *Alto* son clasificadas dentro de la categoría de rendimiento *Medio*. (Cuadro 29 y Anexo 1).

En el caso de la dosis tomada como ejemplo (Cuadro 29), se observa que las probabilidades de clasificar una observación proveniente de la categoría de rendimiento *Alto* en la categoría de rendimiento *Bajo* presentan valores sumamente pequeños; esto favorece mucho a las ecuaciones discriminantes con miras a realizar recomendaciones, ya que el productor podrá confiar que en el peor de los casos, el rendimiento esperado se encontrará en la categoría siguiente a la estimada y no en la categoría más alejada.

Cuadro 29. Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 120-0-50 en el estado Yaracuy.

<i>Categoría de origen</i>	<i>Categoría de clasificación</i>	<i>Probabilidades a posteriori de clasificación</i>		
		<i>Alto</i>	<i>Bajo</i>	<i>Medio</i>
ALTO	MEDIO *	0.0719	0.2946	0.6335
MEDIO	ALTO *	0.5024	0.0000	0.4976
MEDIO	BAJO *	0.2051	0.7274	0.0675
MEDIO	ALTO *	0.9359	0.0027	0.0613
ALTO	MEDIO *	0.0094	0.0000	0.9906
MEDIO	BAJO *	0.0935	0.5211	0.3854
ALTO	MEDIO *	0.4171	0.0005	0.5825
ALTO	MEDIO *	0.0568	0.0019	0.9413

(*) Indica que es una observación clasificada de forma errónea.

En el caso de la dosis 120-80-50 (Cuadro 30), la cual presentó una probabilidad de 0.52, de mala clasificación ver Anexo 1, para la categoría de rendimiento *Alto*, se puede observar que a pesar de que esta dosis presenta tan alta probabilidad de clasificar erradamente; las observaciones provenientes de esta categoría de rendimiento presentan probabilidades *a posteriori* de clasificación que serán siempre mayores para la categoría de rendimiento *Medio*, por lo cual estas observaciones presentan mayor probabilidad de ser clasificadas erradamente dentro de la categoría de rendimiento *Medio* y no clasificadas en la categoría de rendimiento *Bajo*.

Cuadro 30. Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 120-80-50 en el estado Yaracuy.

<i>Categoría de origen</i>	<i>Categoría de clasificación</i>	<i>Probabilidades a posteriori de clasificación</i>		
		<i>Alto</i>	<i>Bajo</i>	<i>Medio</i>
ALTO	MEDIO *	0.0656	0.0388	0.8955
ALTO	MEDIO *	0.0656	0.0388	0.8955
ALTO	MEDIO *	0.0341	0.2537	0.7122
ALTO	MEDIO *	0.0341	0.2537	0.7122
ALTO	MEDIO *	0.0341	0.2537	0.7122
ALTO	MEDIO *	0.0817	0.0012	0.9172
BAJO	MEDIO *	0.0590	0.0430	0.8980
BAJO	MEDIO *	0.0590	0.0430	0.8980
BAJO	MEDIO *	0.0590	0.0430	0.8980
MEDIO	BAJO *	0.0585	0.6479	0.2936
MEDIO	BAJO *	0.0585	0.6479	0.2936
MEDIO	BAJO *	0.0585	0.6479	0.2936
MEDIO	BAJO *	0.0585	0.6479	0.2936
ALTO	MEDIO *	0.2559	0.2046	0.5396
ALTO	MEDIO *	0.2559	0.2046	0.5396
ALTO	MEDIO *	0.2559	0.2046	0.5396
ALTO	MEDIO *	0.0470	0.0098	0.9432
ALTO	MEDIO *	0.0470	0.0098	0.9432
ALTO	MEDIO *	0.0470	0.0098	0.9432
ALTO	MEDIO *	0.0470	0.0098	0.9432

(*) Indica que es una observación clasificada de forma errónea.

En lo referente a las variables consideradas en las ecuaciones de discriminación, éstas variaron dependiendo de la dosis estudiada, en teoría, todas las variables deberían tener alguna influencia en el rendimiento y por consiguiente, en la discriminación, sin embargo es de esperarse que la respuesta en rendimiento para cada una de las dosis haga superflua o redundante la información aportada por una variable en particular, y por lo tanto se pueda omitir dicha variable para realizar la discriminación. Como se observa en el caso de las dosis 180-60-50, 120-80-50 y 120-80-0 (Cuadro 31 y Anexo 1), las variables que conforman las ecuaciones discriminantes para cada una de las dosis, variaron en cada ecuación, si bien se nota una tendencia a presentar como variables discriminantes a las relacionadas con la retención de humedad (a y A), reacción del suelo (pH), y la capacidad de intercambio catiónico ($C.I.C.$); el resto de las variables que no se consideraron en la ecuación de discriminación, fueron descartadas debido a que no aportaron información importante a la ecuación, o la información aportada por estas variables también estaba presente en otra variable presente en la ecuación, así se presenta el caso de que al considerarse alguna variable textural, % de arcilla (A), por ejemplo, otra variable textural, % de arena (a), muchas veces no está presente en la ecuación de discriminación, esto ocurre sobre todo cuando las variables se encuentran correlacionadas, ya que según Cochran, 1969, citado por Mardia *et al*, 1979, la discriminación se verá afectada si existe correlación entre las variables y la capacidad de discriminar disminuirá si la correlación es positiva.

Cuadro 31. Ecuaciones discriminantes correspondientes a las dosis 180-60-50, 120-80-50 y 120-80-0 en el estado Yaracuy.

COEFICIENTES DE LA ECUACIÓN DE DISCRIMINACIÓN												
Dosis	Nivel	Constante	P_{olsen}	K_{olsen}	pH	C.O.	C.E.	C.I.C.	Ca	Mg	a	A
180-60-50	Alto	-677.58477			83.61798		-415.88668				15.69668	23.0542
	Bajo	-969.79584			99.30719		-516.31414				19.15584	27.74135
	Medio	-851.35348			93.23267		-482.79062				17.90463	26.09555
120-80-50	Alto	-38.05080				59.4683		-2.5496		-14.00635		1.27737
	Bajo	-20.83680				39.23211		-0.9877		-13.12037		0.86176
	Medio	-33.47475				64.73997		-3.04508		-9.94882		0.92848
120-80-0	Alto	-187.06121			51.10306	57.69898		-10.63795	-4.01416	-7.90208		4.24302
	Bajo	-124.64403			43.3062	36.99299		-7.12691	-3.46886	-9.55324		3.28814
	Medio	-163.83521			46.46362	75.88604		-11.39672	-3.56493	-3.35786		3.62014

Como ya se especificó en capítulo correspondiente a los Materiales y Métodos, el descarte de las variables que no aportan información importante en la discriminación se realizó con la ayuda del programa *SAS versión 6.04*, entendiéndose por información importante, aquella aportada por aquellas variables que influyen directamente en la clasificación, es decir que no aumenten el error de clasificación errada. Este descarte se realizó siguiendo los siguientes criterios:

- Tomando en cuenta la información aportada por los resultados de las pruebas de medias poblacionales univariadas presentadas por la opción ANOVA del procedimiento DISCRIM, es decir se consideraron aquellas variables edáficas que presentan diferencias entre las medias poblacionales para las categorías de discriminación, esto como una garantía de que esas variables están asociados a diferentes grados de respuesta en rendimiento para esa variable en particular.
- Manteniendo el menor error de clasificación para el modelo y en especial para la categoría de rendimiento *Alto*.

Si bien Rao (1973), citado por Mardia *et al* (1979), propuso un estadístico basado en la distancia de Mahalanobis muestral para cada una de las variables considerada en el análisis discriminante, con la finalidad de crear una prueba para realizar el descarte de variables; para el desarrollo de esta metodología, no se pudo utilizar este estadístico por razones operativas, es decir, esta metodología no está incluida en los paquetes estadísticos utilizados para el desarrollo de esta propuesta. Así, el descarte de variables se llevó a cabo considerando los criterios expuestos anteriormente y constituye una aproximación bastante razonable a la prueba propuesta por Rao.

El primer criterio se plantea porque es recomendable verificar que existen diferencias entre las medias poblacionales de las variables consideradas (Mardia *et al*, 1979), esto con la finalidad de no engrosar el número de variables redundantes en el análisis. En el caso de la dosis 90-60-50, se observa que los F calculados para cada una de las variables presentes en la ecuación de discriminación, presentan diferencias a un nivel de significación α del 5%, es decir existen diferencias significativas entre las medias poblacionales para las variables presentes en la ecuación de discriminación, sin embargo, como ya se explicó anteriormente, esta diferencia sólo se presenta en las variables edáficas correspondientes a la categoría de rendimiento *Alto*, debido a la similitud que se presenta entre el valor de las medias muestrales de las categorías *Bajo* y *Medio*, esto explica la debilidad de estas ecuaciones para discriminar entre observaciones provenientes de las categorías *Bajo* y *Medio*. (Cuadro 32).

Cuadro 32. Resultados de las pruebas de medias univariadas para la dosis 90-60-50 en el estado Yaracuy.

<i>Variable</i>	<i>F</i>	<i>Prob. > F</i>
<i>P_{olsen}</i>	13.4815	0.0001
<i>K_{olsen}</i>	3.7277	0.0273
<i>pH</i>	5.3538	0.0061
<i>C.E.</i>	10.4317	0.0001
<i>Ca</i>	7.2293	0.0011
<i>Mg</i>	5.7811	0.0042
<i>a</i>	19.7284	0.0001
<i>A</i>	6.4787	0.0022

El segundo criterio proviene del hecho de que el método de resustitución subestima el valor de los errores de clasificación, (Mardia *et al*, 1979), por lo cual se trató de mantener el máximo número de variables que mantuvieran la estimación del error de clasificación al mínimo. En el caso de la dosis 180-60-50, se presenta una probabilidad de mala clasificación de cero, habiendo descartado seis (6) variables edáficas. (Cuadro 28).

Sin embargo, se incluyeron en varias ecuaciones algunas variables sin tomar en cuenta el primer criterio, ya que a pesar de tener medias poblacionales estadísticamente iguales, mostraron una influencia positiva en la discriminación. Se consideró que su influencia era positiva porque aportaron información que ayudó a clasificar algunas observaciones, es decir en caso de descartar alguna de estas variables los errores de mala clasificación hubieran aumentado. Es de esperarse que al aumentar el número de observaciones de las variables de suelo para la obtención de las ecuaciones discriminantes, estas variables sean descartadas, pero para la situación considerada mostraron una buena capacidad de discriminación. Un ejemplo de esta situación se presenta en la dosis 49-128-20, en donde se incluyeron en la ecuación de discriminación las variables edáficas K_{olsen} , $C.E.$, Mg y a , ya que a pesar de presentar medias poblacionales iguales a un nivel de significación α del 5%, si aportaron información en la discriminación. (Cuadro 33).

Cuadro 33. Resultados de las pruebas de medias univariadas para la dosis 49-128-20 en el estado Yaracuy.

<i>Variable</i>	<i>F</i>	<i>Prob. > F</i>
K_{olsen}	0.8919	0.4255
pH	13.2091	0.0002
$C.O.$	8.2985	0.0024
$C.E.$	1.2068	0.3200
$C.I.C.$	3.7762	0.0406
Ca	6.3581	0.0073
Mg	3.2956	0.0579
a	2.7616	0.0873
A	4.3036	0.0279

Al aplicar estos criterios a las dosis consideradas en el estado Yaracuy, se encontró que las variables discriminantes dependen de cada una de las dosis consideradas, es decir, no se puede establecer un patrón único sobre cual variable va a discriminar en una dosis dada, sin embargo, es de esperarse que con muestras mayores sólo se mantengan aquellas variables que aporten información importante en la ecuación de discriminación.

Cabe destacar que en caso del estado Yaracuy la mayor frecuencia de aparición de las variables edáficas estuvo altamente ligada con las variables texturales del suelo (% de arena y de arcilla), ya que en 25 de las 27 dosis estudiadas, estuvo presente por lo menos una de estas dos variables; se piensa que esta relación está en consonancia con la disponibilidad de agua por parte de la planta y obviamente con una mejor utilización de los nutrientes presentes y aportados al suelo. (Anexo 1).

Una alternativa a considerar en el caso de las dosis que presentaron una alta probabilidad de clasificación errada para la categoría de rendimientos *Bajo*, es la de fusionar dos categorías contiguas en una sola. Por ejemplo, para la dosis 90-60-50, se pudiera pensar en fusionar las categorías de rendimiento *Medio* y *Bajo* en una sola, y volver a calcular las ecuaciones de discriminación para la nueva situación. Sin embargo, en las localidades estudiadas del estado Yaracuy, se encontró buena discriminación entre las tres categorías; por lo tanto la propuesta anterior podría ser aplicable a sólo dos (2) de las dosis utilizadas: 90-60-50 y 90-60-0. En general, en las localidades se encontró buena discriminación entre las categorías para cada una de las dosis, y a pesar de que estas dos dosis mencionadas anteriormente presenten tal alta probabilidad de clasificación errada para la categoría de rendimiento *Bajo*, las ecuaciones podrían utilizarse tal y como están para realizar recomendaciones ya que la categoría a partir de la cual se derivan las conclusiones es la categoría de rendimiento *Alto*. (Cuadro 26 y Anexo 1).

No se descarta la posibilidad de utilizar estas ecuaciones para discriminar entre las categorías de rendimientos de muestras provenientes de otros estados o zonas productoras de maíz, ya que puede suponerse que si la muestra de suelo tiene características similares a las presentadas por las variables consideradas en este análisis (Cuadro 24), puede esperarse que la respuesta del cultivo de maíz se comporte de manera similar a la observada en las localidades

estudiadas. Debe recordarse que al discriminar entre las categorías de rendimiento de otras zonas, se dejan de considerar otros aspectos como clima, manejo, etc., que de una manera u otra pueden influir en los rendimientos y que pueden ser similares en el estado estudiado, por lo cual esta extrapolación será siempre una aproximación, además en este estudio no se tomaron en cuenta dichos aspectos debido a que no se llevaron registros por quienes realizaron los ensayos. Sin embargo, puede realizarse la extrapolación asumiendo que tal vez exista un aumento en la probabilidad de mala clasificación, pero puede constituir una buena estimación si no se tiene otra herramienta a la mano, o si no se cuenta con suficientes observaciones y combinaciones de dosis para el desarrollo de esta metodología. Las características que debe reunir un suelo proveniente de otra zona productiva, es que las variables edáficas deben presentar valores preferiblemente entre los rangos tomados por las variables en el estado Yaracuy, esto con la finalidad de evitar llegar a conclusiones absurdas y de incurrir en el mínimo de errores de clasificación.

Debe señalarse que el tamaño de muestra utilizado para la mayoría de las dosis fue pequeño en comparación con el utilizado por Fisher, (Fisher, 1936, citado por Mardia *et al*, 1979), y puede objetarse que este modelo es demasiado exigente en lo referente al tamaño de la muestra, debido a que se presentan errores debido a la subestimación de la probabilidad de mala clasificación.

Sin embargo, debe señalarse una gran diferencia que existe entre la aplicación que Fisher le dio a esta metodología y la utilizada en este trabajo, ya que fue aplicada por él en la clasificación taxonómica de distintas especies del género *Iris*, en donde la obtención de las observaciones resulta sumamente sencilla y poco costosa en comparación con la obtención de las observaciones en este caso, ya que para obtener las observaciones de rendimiento deben llevarse a cabo ensayos en diferentes localidades y esto conlleva a un seguimiento del ciclo de vida del cultivo estudiado, el costo que genera todo el personal empleado tanto en campo como en gabinete: investigadores, obreros, laboratoristas etc., además de otros costos como lo son: control de plagas, fertilización, control de competidores bióticos etc., por lo tanto la obtención de un número mayor de observaciones no es un proceso tan sencillo como el utilizado por Fisher y por otros investigadores que han trabajado con esta metodología como

el caso de Mata (1989) y Montiel (1986), en Venezuela, ya que estudiaron la clasificación de distintas especies de cangrejos del género *Mithrax* y la identificación y caracterización de paisajes de Mantecal, estado Apure, respectivamente.

Debe señalarse además, que la base de datos utilizada en este análisis posee observaciones provenientes de zonas con condiciones agroecológicas representativas de las zonas productoras y las observaciones provienen de zonas con características edáficas distintas (Cuadros 19 y 20), esto con la finalidad de contar con muestras de suelos con características asociadas a distintas categorías de rendimiento y algo muy importante, este análisis no requiere de un diseño experimental específico, basta con que se tengan las observaciones y las mediciones de las variables edáficas para que una observación se considere válida para realizar el análisis discriminante.

Otra alternativa para el problema que se presenta con las categorías con alta probabilidad de clasificación errada, es la de no tomar en cuenta esa dosis con miras a realizar recomendaciones por considerarla poco confiable o riesgosa, en el caso del estado Yaracuy se pudiera eliminar la siguiente dosis: 120-0-50, ya que presenta una probabilidad de clasificación errada de 0.75 para la categoría de rendimiento *Alto*, y esto atenta contra la intención de esta metodología, pues si se llegara a recomendar esta dosis, el rendimiento esperado estaría equivocado con una frecuencia de 75%, pueden considerarse en cambio, las dosis con alta probabilidad de clasificación errada para la categoría de rendimiento *Bajo*, ya que con fines de realizar recomendaciones, se consideran como primera opción aquellas dosis que estén asociadas a rendimientos altos. (Cuadro 28).

5.2.2. Estado Apure.

En el estado Apure se analizaron 14 combinaciones de nutrientes. Las dosis de nutrientes consideradas mostraron los siguientes rangos. N: desde 0 Kg./Ha. hasta 280 Kg./Ha., P: desde 0 Kg./Ha. hasta 60 Kg./Ha. y K: desde 0 Kg./Ha. hasta 40 Kg./Ha. (Cuadro 22).

A diferencia de lo ocurrido en las localidades del estado Yaracuy, los valores tomados por las variables edáficas de las localidades estudiadas en el estado Apure no variaron dentro de un espectro muy amplio, exceptuando las variables *C.O.* y *P_{olsen}* aunado además al hecho de que el número de variables tomadas en cuenta fue menor que las consideradas en el estado Yaracuy, esto generó que las ecuaciones discriminantes encontradas para cada una de las categorías de rendimientos dentro de cada dosis presentaran altas probabilidades de clasificación errónea para las categorías de rendimiento *Alto* y *Bajo*. (Cuadro 34 y Anexo 2).

Se pueden mencionar otras razones que ayudan a explicar por qué se presentó una clasificación deficiente de las ecuaciones discriminantes en el estado Apure.

- Las condiciones agroecológicas de las localidades son similares, ya que las observaciones de rendimiento y de las variables edáficas provienen en su mayoría de suelos de vega.
- Las observaciones de rendimiento y de las características edáficas sólo provienen de dos localidades: Queseras del Medio y Biruaca.
- El número de variables de suelos medido, fue considerablemente menor que las utilizadas en el estado Yaracuy, ya que no se consideraron las variables, *C.I.C.*, *Mg*, *a* y *A*, y se perdió la información que estas variables pudieran haberle aportado a las ecuaciones discriminantes. (Cuadros 8 y 12).

Cuadro 34. Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en el estado Apure.

VARIABLES DE SUELO MEDIDAS EN EL ESTADO APURE						
Valores	P_{olsen}	K_{olsen} ppm	Ca	C.O. %	C.E. dMhos/cm	pH 1:2.5
Mínimo	2	40	260	0.41	0.04	5.6
Máximo	39	240	1500	3.03	0.16	7.5

Todas las ecuaciones obtenidas en el estado Apure presentaron probabilidades altísimas de clasificación errónea para la categoría de rendimiento *Alto*, oscilando los valores entre 0.5 y 1.0, estos valores probabilísticos tan altos no permiten realizar predicciones acertadas, lo cual hace que el análisis discriminante sea muy deficiente en las localidades estudiadas del estado Apure. (Anexo 2).

Es lógico pensar que si las condiciones edáficas son similares para las distintas categorías de discriminación, los rendimientos obtenidos serán también similares y existirá poca o ninguna discriminación, es decir no se encuentran condiciones edáficas relacionadas con rendimientos altos y bajos en las localidades del estado Apure y la precisión de las ecuaciones de discriminación encontrada es baja. (Anexo 2).

Las medias de las variables edáficas medidas para los grupos *Medio* y *Alto*, se encuentran muy cercanas entre si, por lo cual, las observaciones de rendimiento cuyos valores de las variables edáficas se encuentran próximos a las medias muestrales de la categoría de rendimiento *Alto*, serán clasificadas dentro de la categoría de rendimiento *Medio*; debido a la tendencia del modelo de clasificar una observación dentro de la categoría de rendimiento con mayor probabilidad *a priori*, siendo esta precisamente la categoría de rendimiento *Medio*. (Anexo 2).

Para ilustrar esta afirmación se presenta en el Cuadro 35 el caso de la dosis 140-30-20, esta dosis presentó probabilidades de mala clasificación para las categorías de rendimiento *Alto* y *Bajo* de 1.0, es decir, no clasificó correctamente ninguna de las observaciones

provenientes de esas dos categorías, sin embargo, al observar como fueron clasificadas esas observaciones vemos que todas fueron catalogadas como provenientes de la categoría de rendimiento *Medio*, es decir se confirma lo dicho en los resultados correspondientes al estado Yaracuy, que la inclusión de las probabilidades *a priori* beneficia a las ecuaciones de discriminación en la medida en que minimiza la posibilidad de clasificar una observación proveniente de la categoría *Bajo* como *Alto* y viceversa; nótese además que la categoría que presenta la mayor probabilidad *a posteriori* es la categoría de rendimiento *Medio*, lo cual ratifica lo expuesto anteriormente.

Cuadro 35. Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 140-30-20 en el estado Apure.

<i>Probabilidades a posteriori de clasificación</i>				
<i>Categoría de origen</i>	<i>Categoría de clasificación</i>	<i>Alto</i>	<i>Bajo</i>	<i>Medio</i>
ALTO	MEDIO *	0.1907	0.1712	0.6380
ALTO	MEDIO *	0.1907	0.1712	0.6380
ALTO	MEDIO *	0.1907	0.1712	0.6380
ALTO	MEDIO *	0.1907	0.1712	0.6380
ALTO	MEDIO *	0.1907	0.1712	0.6380
ALTO	MEDIO *	0.2067	0.1799	0.6134
ALTO	MEDIO *	0.2067	0.1799	0.6134
ALTO	MEDIO *	0.2067	0.1799	0.6134
ALTO	MEDIO *	0.2047	0.2557	0.5396
BAJO	MEDIO *	0.1485	0.2398	0.6117
BAJO	MEDIO *	0.1413	0.2547	0.6040
ALTO	MEDIO *	0.1413	0.2547	0.6040
BAJO	MEDIO *	0.1780	0.2134	0.6086
BAJO	MEDIO *	0.1780	0.2134	0.6086
BAJO	MEDIO *	0.1780	0.2134	0.6086
BAJO	MEDIO *	0.1780	0.2134	0.6086
BAJO	MEDIO *	0.0968	0.2468	0.6563
BAJO	MEDIO *	0.0577	0.0622	0.8801
BAJO	MEDIO *	0.1786	0.3114	0.5100
BAJO	MEDIO *	0.1786	0.3114	0.5100
BAJO	MEDIO *	0.1786	0.3114	0.5100
BAJO	MEDIO *	0.1786	0.3114	0.5100
BAJO	MEDIO *	0.1844	0.2723	0.5432
BAJO	MEDIO *	0.0937	0.2592	0.6471
BAJO	MEDIO *	0.0937	0.2592	0.6471
BAJO	MEDIO *	0.0937	0.2592	0.6471

Cuadro 35. Observaciones clasificadas de forma errónea y probabilidades a posteriori de clasificación para la dosis 140-30-20 en el estado Apure. (Continuación).

		<i>Probabilidades a posteriori de clasificación</i>		
<i>Categoría de origen</i>	<i>Categoría de clasificación</i>	<i>Alto</i>	<i>Bajo</i>	<i>Medio</i>
ALTO	MEDIO *	0.0937	0.2592	0.6471
ALTO	MEDIO *	0.1827	0.0616	0.7557
ALTO	MEDIO *	0.1827	0.0616	0.7557
ALTO	MEDIO *	0.1827	0.0616	0.7557
ALTO	MEDIO *	0.1827	0.0616	0.7557
ALTO	MEDIO *	0.1827	0.0616	0.7557
BAJO	MEDIO *	0.2738	0.0541	0.6721
BAJO	MEDIO *	0.1781	0.0602	0.7617
ALTO	MEDIO *	0.2495	0.0960	0.6545
ALTO	MEDIO *	0.2495	0.0960	0.6545
ALTO	MEDIO *	0.2495	0.0960	0.6545
ALTO	MEDIO *	0.2495	0.0960	0.6545
ALTO	MEDIO *	0.2495	0.0960	0.6545
BAJO	MEDIO *	0.3128	0.1153	0.5718
ALTO	MEDIO *	0.3913	0.0279	0.5808

Todas las variables de suelo consideradas en el análisis mostraron una marcada tendencia a tomar valores alrededor de las medias de las variables edáficas de la categoría de rendimiento *Medio*. Si bien es cierto que en algunas de las variables se presentaron diferencias significativas, como ocurre con la variable contenido de calcio (*Ca*), también es cierto que una sola variable no necesariamente va a aportar toda la información necesaria para realizar una buena discriminación. El resultado de esta situación se refleja en los coeficientes de las ecuaciones discriminantes, ya que éstos presentan valores similares para las tres categorías de rendimiento analizadas. Esta tendencia se mantuvo para todas las dosis estudiadas. (Cuadro 36 y Anexo 2).

Cuadro 36. Medias muestrales de las variables de suelo para las dosis 0-30-20, 57-12-8 y 140-30-20 en el estado Apure.

		MEDIAS MUESTRALES DE LAS VARIABLES DE SUELO					
<i>Dosis</i>	<i>Categorías</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i> <i>ppm</i>	<i>Ca</i>	<i>pH</i> <i>Sol 1:2.5</i>	<i>C.O.</i> <i>%</i>	<i>C.E.</i> <i>dMhos/cm</i>
0-30-20	<i>Alto</i>			612	5.96		
	<i>Bajo</i>			435	5.85		
	<i>Medio</i>			858	6.26		
	<i>Total muestra</i>			704.21	6.09		
57-12-8	<i>Alto</i>	18.33		986.67	6.53		
	<i>Bajo</i>	12.5		575	6.23		
	<i>Medio</i>	11.29		678.6	6.03		
	<i>Total muestra</i>	12.52		702.86	6.13		
140-30-20	<i>Alto</i>	13.32	122.73	677.27	6.02	1.56	0.08
	<i>Bajo</i>	11.37	103.8	577.9	6.06	0.93	0.07
	<i>Medio</i>	12.47	123.86	736.14	6.17	1.39	0.09
	<i>Total muestra</i>	12.45	120.58	701.45	6.13	1.35	0.08

El descarte de variables se realizó siguiendo los mismos criterios aplicados para el estado Yaracuy, pero en este caso se violó en repetidas ocasiones el primer criterio planteado, debido a la dificultad de encontrar variables que presentaran diferencias poblacionales en las medias de las características edáficas medidas para las categorías de rendimiento y a la vez mantuvieran el error de clasificación errada en su menor valor. La mayoría de las variables consideradas en las ecuaciones discriminantes no presentaron diferencias significativas a un nivel α del 5%, por lo cual, el descarte se realizó siguiendo el segundo criterio planteado en el estado Yaracuy.

Las variables presentes en las ecuaciones de discriminación variaron dependiendo de la dosis estudiada, es decir, en muchas ecuaciones discriminantes van a presentar variables distintas. (Cuadro 36).

La única variable que presentó una tendencia a aportar información a la ecuación discriminante, fue el contenido de calcio (Ca), esta variable estuvo presente en todas las ecuaciones discriminantes, sin embargo, como ya se explicó anteriormente la información aportada por esta variable es insuficiente para realizar una buena discriminación. (Anexo 2).

La aplicación del análisis discriminante en las localidades estudiadas en el estado Apure se considera inviable, porque sólo arroja estimaciones con una alta probabilidad de clasificar de forma errónea y no representa una herramienta confiable con la que se puedan realizar estimaciones de rendimientos.

Se debe destacar la importancia de contar con observaciones provenientes de localidades que presenten una alta variabilidad en las características edáficas para garantizar la obtención de una muestra en la cual se puedan encontrar asociaciones entre características edáficas y respuestas en forma de rendimiento. Desgraciadamente, la muestra estudiada en el estado Apure no presentó las características más deseables, ya que la homogeneidad presente en las características edáficas medidas atentó contra la precisión del modelo, pero, esto no indica que con un número mayor de observaciones que provengan de suelos con características físico-químicas distintas, se comiencen a observar asociaciones entre las características de suelo y las respuestas en rendimiento y se obtengan ecuaciones que presenten probabilidades de clasificación erradas menores, aumentando de esta manera su valor predictivo.

En el estado Yaracuy, aun presentando tamaños de muestra más pequeños que los obtenidos en Apure, se obtuvieron ecuaciones discriminantes con bajas probabilidades de clasificación errada y esto debido a que la variabilidad de características edáficas fue mucho mayor que la presentada por los datos provenientes del estado Apure.

No se descarta que en el futuro, si se obtienen observaciones de rendimiento y de las características edáficas de otras localidades ubicadas en el estado Apure, pueda ser aplicado este modelo obteniéndose resultados satisfactorios.

5.2.3. Estados Guárico y Táchira.

En los estados Guárico y Táchira se analizaron un total de 14 combinaciones de nutrientes. Dichas combinaciones de nutrientes presentaron los siguientes rangos: N: desde 0 Kg./Ha. hasta 230 Kg./Ha., P: desde 0 Kg./Ha. hasta 191 Kg./Ha. y K: desde 0 Kg./Ha hasta 120 Kg./Ha. (Cuadro 23).

Cuadro 37. Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en los estados Guárico y Táchira.

VARIABLES DE SUELO MEDIDAS EN LOS ESTADOS GUÁRICO Y TÁCHIRA						
<i>Valores</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i> <i>ppm</i>	<i>Ca</i>	<i>C.O.</i> <i>%</i>	<i>C.E.</i> <i>dMhos/cm</i>	<i>pH</i> <i>1:2.5</i>
<i>Mínimo</i>	1	40	12.5	0.45	0.02	5.2
<i>Máximo</i>	16	88	1500	2.8	0.77	7.7

A diferencia de lo ocurrido en las localidades del estado Yaracuy, y de una forma similar a la situación observada en las localidades estudiadas del estado Apure, las variables edáficas de las localidades estudiadas en los estados Guárico y Táchira no variaron dentro de un espectro muy amplio, presentando rangos de variación similares a los observados en el estado Apure y al igual que en el caso del estado apure se consideraron menos variables en el análisis, esto generó que las ecuaciones discriminantes encontradas presentaran en muchos casos altas probabilidades de clasificación errónea para las categorías de rendimiento *Alto* y *Bajo*. (Cuadro 37 y Anexo 3).

Varias de las ecuaciones obtenidas en los estados Guárico y Táchira presentaron altas probabilidades de clasificación errónea para la categoría de rendimiento *Alto*, siendo un total de ocho (8) dosis las que presentan probabilidades de clasificación errada cuyos valores oscilan entre 0.5 y 1.0, estas dosis son: 47-191-24, 47-49-24, 115-120-0, 115-120-120, 115-120-60, 183-191-24, 183-191-96 y 183-49-96, estos valores probabilísticos tan altos no permiten realizar predicciones acertadas para la categoría de rendimiento *Alto*, lo cual hace que el análisis discriminante sea muy deficiente o poco confiable para las dosis mencionadas. (Anexo 3).

La deficiente discriminación encontrada para las dosis mencionadas se puede explicar utilizando un argumento similar al planteado para el análisis realizado en las localidades del estado Apure, es decir, las medias de variables edáficas medidas para los grupos *Medio* y *Alto*, se encuentran muy cercanas, por lo cual, las observaciones de rendimiento cuyos valores de las variables edáficas se encuentran próximos a las medias muestrales de la categoría de rendimiento *Alto*, serán clasificadas dentro de la categoría de rendimiento *Medio*; debido a la tendencia del modelo de clasificar una observación dentro de la categoría de rendimiento con mayor probabilidad *a priori*, tal y como ocurrió en el estado Apure.

En el caso de la combinación de nutrientes 0-120-60, la cual presentó una probabilidad de clasificación errónea para la categoría de rendimiento *Alto* de cero, las medias de las variables de esta categoría de rendimiento se presentan bastante alejadas de los medias muestrales de la categoría de rendimiento *Medio*, en especial para la variable contenido de calcio (*Ca*), en cambio para la dosis 115-120-120 las medias muestrales de las variables edáficas presentes en la ecuación presentan valores muy similares a los de la categoría de rendimiento *Medio*; al igual que lo ocurrido en la dosis 0-120-60, la variable contenido de calcio (*Ca*), presenta valores muy alejados de la media general y de la media de la categoría de rendimiento *Medio*, sin embargo, como ya se explicó anteriormente, una sola variable no necesariamente puede aportar toda la información necesaria para obtener una buena discriminación, la consecuencia de este planteamiento se refleja en el valor de las probabilidades de mala clasificación para ambos casos, en el caso de la dosis 0-120-60 esta probabilidad tomó el valor de cero para la categoría de rendimiento *Alto*, y en el caso de la dosis 115-120-120, presentó el valor de 1.0 para la misma categoría de rendimiento. (Cuadro 38 y Anexo 3).

Cuadro 38. Medias muestrales de las variables edáficas de las combinaciones de nutrientes 0-120-60 y 115-120-120 en los estados Guárico y Táchira.

MEDIAS MUESTRALES DE LAS VARIABLES DE SUELO							
<i>Dosis</i>	<i>Categorías</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>
0-120-60	Alto			6.8	2.25	0.08	1053
	Bajo			6.33	1.37	0.05	330
	Medio			6.09	1.17	0.18	533.25
Total muestra				6.27	1.41	0.13	577.21
115-120-120	Alto			6.65		0.13	880
	Bajo			6.3		0.06	390
	Medio			6.21		0.14	559.43
Total muestra				6.27		0.13	577.21

Al igual que lo ocurrido en los estados Yaracuy y Apure, los coeficientes de las ecuaciones de discriminantes serán bastante diferentes en aquellas ecuaciones con probabilidades bajas o cero de mala clasificación y muy similares en las ecuaciones con altas probabilidades de clasificación errónea. Para las dosis analizadas anteriormente se puede observar la gran diferencia entre los valores en las ecuaciones de discriminación para las distintas categorías de las variables reacción del suelo (*pH*), contenido de carbono orgánico (*C.O.*), conductividad eléctrica (*C.E.*) y contenido de calcio (*Ca*) en el caso de la dosis 0-120-60 y la similitud entre los coeficientes de la ecuaciones de discriminación para la dosis 115-120-120. (Cuadro 39).

Cuadro 39. Coeficientes de discriminación correspondientes a las ecuaciones de las dosis 0-120-60 y 115-120-120 en los estados Guárico y Táchira.

		<i>Coeficientes</i>						
<i>Dosis</i>	<i>Nivel</i>	<i>Constante</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>
0-120-60	<i>Alto</i>	-154.53177			57.98289	3.19433	-119.48914	-0.08269
	<i>Bajo</i>	-191.71081			66.49954	-0.0942	-138.81691	-0.09899
	<i>Medio</i>	-141.49768			57.01446	-0.05692	-113.96891	-0.08362
115-120-120	<i>Alto</i>	-80.82245			29.49407		-42.32421	-0.03781
	<i>Bajo</i>	-93.0943			32.01377		-49.30425	-0.04313
	<i>Medio</i>	-77.64762			29.39765		-42.24512	-0.03877

El descarte de variables se realizó siguiendo los mismos criterios aplicados para el estado Yaracuy y se presentaron los mismos problemas que en el estado Apure, es decir, en repetidas ocasiones no se utilizó el primer criterio planteado, debido a la dificultad de encontrar variables que presentaran diferencias poblacionales en las medias de las características edáficas medidas para las categorías de rendimiento y a la vez mantuvieran el error de clasificación errada en su menor valor. La mayoría de las variables consideradas en las ecuaciones discriminantes no presentaron diferencias significativas a un nivel α del 5%, por lo cual, el descarte se realizó siguiendo el segundo criterio planteado en el estado Yaracuy.

Las variables presentes en las ecuaciones de discriminación variaron dependiendo de la dosis estudiada, es muchas ecuaciones discriminantes van a presentar variables distintas. (Cuadro 39).

La única variable que presentó una tendencia a aportar información a la ecuación discriminante, fue el contenido de calcio (*Ca*), esta variable estuvo presente en trece (13) de las catorce (14) ecuaciones discriminantes, sin embargo, como ya se explicó anteriormente la información aportada por esta variable es insuficiente para realizar una buena discriminación. (Anexo 3).

La aplicación del análisis discriminante en las localidades estudiadas de los estados Guárico y Táchira se considera inviable, porque sólo arroja estimaciones con una alta

probabilidad de clasificar de forma errónea y no representa una herramienta con la que se puedan realizar estimaciones confiables.

Desgraciadamente y al igual que lo ocurrido en las localidades del estado Apure, la muestra estudiada no presentó las características más deseables, ya que la homogeneidad presente en las características edáficas medidas atentó contra la precisión del modelo.

Al igual que en la situación acaecida en Apure, se espera que al aumentar las observaciones y por ende los rangos de variación de las variables consideradas en este análisis, se encuentren asociaciones entre las variables edáficas y las respuestas en rendimientos y puedan obtenerse ecuaciones discriminantes que presenten probabilidades bajas de clasificación errónea, en especial para la categoría de rendimiento *Alto*, por ser esta la categoría considerada como más importante a la hora de realizar recomendaciones.

No debe olvidarse lo costoso que resulta la obtención de observaciones tanto edáficas como de rendimiento para la aplicación de este análisis, ya que se pueden considerar como pequeñas las muestras analizadas en estos estados, a pesar de ser de mayor tamaño que las analizadas en el estado Yaracuy, pero como ya se explicó anteriormente la homogeneidad de las características edáficas atentó contra la precisión del modelo planteado. (Cuadro 23).

5.2.4. Estado Monagas.

En el estado Monagas se analizaron un total de 15 combinaciones de nutrientes. Las dosis de nutrientes consideradas mostraron los siguientes rangos. N: desde 0 Kg./Ha. hasta 180 Kg./Ha., P: desde 0 Kg./Ha. hasta 200 Kg./Ha. y K: desde 0 Kg./Ha. hasta 140 Kg./Ha. (Cuadro 21).

Los rangos de variación para las variables edáficas consideradas para la realización del análisis discriminante fueron ligeramente mayores que los observados en los estados Apure, Guárico y Táchira, en especial para la variable reacción del suelo (pH) y potasio (K_{olsen}), pero, desgraciadamente este mayor rango de variación no repercutió en la obtención de ecuaciones de discriminación más precisas para las categorías de rendimientos *Alto* y *Bajo*. (Cuadro 40 y Anexo 4).

Cuadro 40. Rango de valores observados para las variables edáficas utilizadas en el análisis discriminante en el estado Monagas.

VARIABLES DE SUELO MEDIDAS EN EL ESTADO MONAGAS						
Valores	P_{olsen}	K_{olsen} ppm	Ca	C.O. %	C.E. dMhos/cm	pH 1:2.5
Mínimo	2	16	8	0.21	0.01	4
Máximo	63	116	1480	1.41	0.14	7.4

Un total de nueve (9) dosis presentaron altas probabilidades de clasificación errónea para la categoría de rendimiento *Alto*, los valores para estas probabilidades oscilan entre 0.5 y 0.6, dichas dosis son: 36-41-112, 90-0-70, 90-100-0, 90-100-70, 90-200-70, 144-159-112, 144-159-28, 144-41-112 y 144-41-28, estos valores probabilísticos tan altos no permiten realizar predicciones acertadas para la categoría de rendimiento *Alto*, lo cual hace que el análisis discriminante sea muy deficiente o poco confiable para las dosis mencionadas. (Anexo 4).

La deficiente discriminación encontrada para las dosis mencionadas se puede explicar utilizando un argumento similar al planteado para el análisis realizado en las localidades del estado Apure y las localidades de los estados Guárico y Táchira, es decir, las medias de

variables edáficas medidas para los grupos *Medio* y *Alto*, se encuentran cercanas entre sí, por lo cual, las observaciones de rendimiento cuyos valores de las variables edáficas se encuentran próximos a las medias muestrales de la categoría de rendimiento *Alto*, serán clasificadas dentro de la categoría de rendimiento *Medio*; debido a la tendencia del modelo de clasificar una observación dentro de la categoría de rendimiento con mayor probabilidad *a priori*, tal y como ocurrió en el estado Apure y en los estados Guárico y Táchira.

En el caso de la combinación de nutrientes 0-100-70, la cual presentó una probabilidad de clasificación errónea de cero para la categoría de rendimiento *Alto*, las medias de las variables de esta categoría de rendimiento se presentan bastante alejadas de los medias muestrales de la categoría de rendimiento *Medio*, en especial para la variable contenido de calcio (*Ca*), en cambio para la dosis 144-159-112, a pesar de que las medias muestrales de las variables edáficas presentes en la ecuación presentan valores diferentes a los de la categoría de rendimiento *Medio*, se presenta una alta probabilidad de mala clasificación, debido probablemente a que la distancia entre los dos grupos no es lo suficientemente grande como para obtener una buena discriminación; al igual que lo ocurrido en la dosis 0-100-70, la variable contenido de calcio (*Ca*), presenta valores muy alejados de la media general y de la media de la categoría de rendimiento *Medio*. (Cuadro 41).

La consecuencia de este hecho se refleja en el valor de las probabilidades de mala clasificación para ambos casos, en el caso de la dosis 0-100-70 esta probabilidad tomó el valor de cero para la categoría de rendimiento *Alto*, y en el caso de la dosis 144-119-112, presentó el valor de 0.6 para la misma categoría de rendimiento. (Anexo 4).

Cuadro 41. Medias muestrales de las variables edáficas de las combinaciones de nutrientes 0-100-70 y 144-159-112 en el estado Monagas.

MEDIAS MUESTRALES DE LAS VARIABLES DE SUELO							
<i>Dosis</i>	<i>Categorías</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>
0-100-70	Alto			6.27	1.00	0.09	1000
	Bajo			4.90	0.51	0.02	117
	Medio			4.68	0.42	0.04	44
Total muestra				4.90	0.50	0.04	170.17
144-159-112	Alto			5.82	0.89	0.09	579.2
	Bajo			4.60	0.48	0.022	73.60
	Medio			4.66	0.40	0.04	58
Total muestra				4.86	0.49	0.05	153.86

A diferencia de lo ocurrido en los estados Apure, Guárico y Táchira, ninguna de las ecuaciones discriminantes obtenidas para las dosis del estado Monagas presentó una probabilidad de clasificación equivocada de 1.0 para la categoría de rendimiento *Alto*, esto justifica el hecho de que los coeficientes de discriminación no sean tan cercanos como los observados en estos estados, y este comportamiento es debido a que el rango de las variables edáficas fue bastante más amplio que para los estados Apure, Guárico y Táchira. (Cuadro 42 y Anexo 4).

Cuadro 42. Coeficientes de discriminación correspondientes a las ecuaciones de las dosis 0-100-70 y 144-159-112 en el estado Monagas.

<i>Coeficientes</i>								
<i>Dosis</i>	<i>Nivel</i>	<i>Constante</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>
0-100-70	<i>Alto</i>	-57.5162			22.03124	43.87406	-362.38641	-0.0398
	<i>Bajo</i>	-90.32534			32.19903	76.54623	-452.28391	-0.11937
	<i>Medio</i>	-75.32165			29.21077	66.71156	-272.85698	-0.12274
144-159-112	<i>Alto</i>	-74.00868			24.80784	51.5534	-117.55234	-0.06102
	<i>Bajo</i>	-65.48706			24.46254	53.56159	-271.66626	-0.06138
	<i>Medio</i>	-58.46533			23.23473	44.98568	-159.56731	-0.05926

El descarte de variables se realizó siguiendo los mismos criterios aplicados para el estado Yaracuy y a diferencia de lo ocurrido para los estados Apure, Guárico y Táchira, la mayoría de las variables consideradas en las ecuaciones discriminantes presentaron diferencias significativas a un nivel α del 5%, por lo cual, el proceso de descarte de variables fue muy similar al llevado a cabo en el estado Yaracuy. En el caso de la dosis 144-159-112, a pesar de presentar diferencias significativas para las medias poblacionales de todas las variables presentes en las ecuaciones de discriminación, la probabilidad de mala clasificación fue alta, debido posiblemente a que las diferencias de las medias poblacionales se presenten con la categoría de redimiendo *Medio* y *Bajo*, en el caso de la dosis 36-41-112 se incluyeron las variables P_{olsen} y K_{olsen} , a pesar de no presentar diferencias significativas entre las medias poblacionales, sin embargo, aportaron información en la ecuación de discriminación. (Cuadros 43 y 44).

Cuadro 43. Resultados de las pruebas de medias univariadas para la dosis 144-159-112 en el estado Monagas.

<i>Variable</i>	<i>F</i>	<i>Prob. > F</i>
<i>pH</i>	5.4778	0.0106
<i>C.O.</i>	8.1175	0.0019
<i>C.E.</i>	12.6456	0.0002
<i>Ca</i>	6.7742	0.0045

Cuadro 44. Resultados de las pruebas de medias univariadas para la dosis 36-41-112 en el estado Monagas.

<i>Variable</i>	<i>F</i>	<i>Prob. > F</i>
<i>P_{olsen}</i>	2.3652	0.1155
<i>K_{olsen}</i>	0.4875	0.6201
<i>Ca</i>	5.4121	0.0115

Las variables presentes en las ecuaciones de discriminación variaron dependiendo de la dosis estudiada, es muchas ecuaciones discriminantes van a presentar variables distintas, a diferencia de las dosis presentadas como ejemplo. (Cuadro 42).

Todas las variable presentaron una tendencia a aportar información a la ecuación discriminante, sin embargo la más frecuente fue el contenido de calcio (*Ca*), esta variable estuvo presente todas las ecuaciones discriminantes, otra variable que presentó una tendencia muy marcada a aportar información fue la variable (*pH*), debido al amplio rango de variación presentado. (Anexo 4).

La aplicación del análisis discriminante en las localidades estudiadas en el estados Monagas se considera inviable, a pesar de tener una precisión mayor que la presentada en los estados Apure, Guárico y Táchira, debido a que arroja demasiadas estimaciones con una alta probabilidad de clasificar de forma errónea y no representa una herramienta con la que se puedan realizar estimaciones confiables.

A pesar de presentar un rango de variación más amplio que para los otros estados y de presentar probabilidades de mala clasificación menores, todavía el valor de estas probabilidades se mantiene muy alto como para considerarlas confiables.

Un común denominador de la situación presente en los estados Apure, Guárico, Monagas y Táchira, es que el número de variables utilizadas para la obtención de las ecuaciones discriminantes fue menor que el utilizado en el estado Yaracuy, en especial, la falta de las observaciones de las variables texturales, como lo son el contenido de arena (a) y el contenido de arcilla (A), que mostraron una frecuencia muy alta de aparición en las ecuaciones de discriminación y el hecho de estar estas variables relacionadas con la absorción de los nutrientes por parte de las plantas, bien pudieran haber influido negativamente en la precisión de las ecuaciones obtenidas en estos estados, por otro lado tampoco se consideró la Capacidad de Intercambio Catiónico ($C.I.C.$), y es bastante lamentable el haber perdido esta información, ya que en la *data* utilizada, para estos estados no se realizaron ni pruebas texturales, ni de capacidad de intercambio catiónico ni de contenido de magnesio (Mg). Pudiera hallarse en la falta de estas variables otra razón para encontrar una discriminación tan deficiente en estos estados, a diferencia de lo ocurrido en el estado Yaracuy.

No se descarta tampoco en este estado que un aumento en el número de observaciones tanto de rendimientos como de características edáficas favorezca a la obtención de ecuaciones discriminantes con menores probabilidades de clasificación errada.

5.3. OBTENCIÓN DE LAS DOSIS DE NUTRIENTES MÁS ADECUADAS.

A partir de las ecuaciones discriminantes de las localidades del estado Yaracuy, ya que fue este el único estado analizado que presentó un número de dosis suficientes para elaborar recomendaciones, se procede a la escogencia de las dosis más adecuadas para una situación particular.

El procedimiento se lleva a cabo tal y como se describe a continuación.

- Una vez que el productor traiga la muestra de suelo correspondiente a la zona de producción estudiada, que preferiblemente debe provenir del valle medio del estado Yaracuy, o de alguna otra zona que presente características edáficas y agroecológicas similares, se procede a realizársele a dicha muestra el análisis de suelo respectivo, midiendo cada una de las variables que requieran las ecuaciones discriminantes.
- Luego, dichos resultados se sustituyen en cada una de las ecuaciones discriminantes de cada categoría de rendimiento para cada una de las dosis estudiadas, esto con la finalidad de obtener la discriminación entre categorías de rendimiento para cada dosis, siguiendo los criterios de discriminación descritos en el capítulo de Materiales y métodos.
- Una vez obtenidos los resultados de cada una de las ecuaciones de discriminación, se escogen aquellas dosis para las cuales el análisis discriminante predice una categoría de rendimiento *Alto*, y se procede a conseguir la combinación de fertilizantes más económica (productos comerciales).

Para ejemplificar como se lleva a cabo este proceso se presenta a continuación un ejemplo ilustrativo.

Ejemplo 5.3.1:

Sean los datos presentados en el Cuadro 45, los resultados del análisis de suelo correspondiente a una muestra traída por un productor de maíz del valle medio del estado Yaracuy.

Cuadro 45. Resultados del análisis de suelo correspondiente a una muestra (ficticia) traída por un productor de maíz del valle medio del estado Yaracuy.

VARIABLES DE SUELO MEDIDAS EN EL ESTADO YARACUY										
	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>C.O.</i>	<i>a</i>	<i>A</i>
	<i>ppm</i>		<i>Sol 1:2.5</i>	<i>dMhos/cm</i>	<i>meq/100 gr. de suelo</i>				<i>%</i>	
Valores	8	65	5.7	0.65	11.3	8.4	0.1	1.78	18.5	36.7

Luego utilizando el programa *Microsoft Excel versión 5.0*, se realizó la rutina que se presenta a continuación y que presenta las siguientes características:

- Un cuadro inicial, llamado cuadro de entrada en el cual se presentan los valores tomados por cada una de las variables edáficas consideradas en la discriminación.
- Luego, para cada dosis, un cuadro en el cual se presentan los valores de las ecuaciones de discriminación para cada una de las categorías y los valores que toman las variables edáficas presentes en las ecuaciones de discriminación.
- Luego del cuadro anterior, aparecen los resultados de cada ecuación de discriminación, y la categoría en la cual la muestra de suelo fue clasificada, no se debe olvidar que estas ecuaciones clasifican dentro de la categoría para la cual se presente el valor mayor en la ecuación de discriminación, debido a que la ecuación discriminante se plantea sólo para una categoría a diferencia del criterio utilizado por Mardia *et al* (1979) de discriminar por categorías pareadas.
- Además se incluye los límites de las categorías de rendimiento *Alto* y *Bajo* para cada una de las dosis.
- Al final se presenta un cuadro resumen con los resultados de las discriminaciones para las dosis analizadas en el estado Yaracuy.

Como se observó anteriormente, las dosis para las cuales se espera un rendimiento *Alto* fueron las siguientes:

- 0-80-50, con un rendimiento mínimo esperado de 3881.9 Kg./Ha.
- 49-32-20, con un rendimiento mínimo esperado de 4641.42 Kg./Ha.
- 49-128-20, con un rendimiento mínimo esperado de 4443.96 Kg./Ha.
- 90-60-50, con un rendimiento mínimo esperado de 4836.58 Kg./Ha.
- 90-60-100, con un rendimiento mínimo esperado de 4898.63 Kg./Ha.
- 90-120-50, con un rendimiento mínimo esperado de 4813.39 Kg./Ha.
- 120-80-50, con un rendimiento mínimo esperado de 5253.53 Kg./Ha.
- 143-24-20, con un rendimiento mínimo esperado de 5589.82 Kg./Ha.
- 143-24-80, con un rendimiento mínimo esperado de 4981.77 Kg./Ha.
- 180-60-50, con un rendimiento mínimo esperado de 5643.63 Kg./Ha.
- 191-32-80, con un rendimiento mínimo esperado de 5901.83 Kg./Ha.

A diferencia de otros métodos para elaborar recomendaciones de dosis de nutrientes, esta metodología no arroja un resultado único como recomendación, sino que presenta un menú de opciones entre cuales escoger, por ello es que se hace necesario conocer los límites de cada categorías de rendimiento para cada dosis para de esta manera saber que rendimiento puede esperarse para esa dosis en particular.

Esta característica lo aleja mucho de lo que se obtendría utilizando la metodología de la superficie de respuesta, debido a que en esta metodología una vez estimada la función de producción se procede a optimizarla, es decir a obtener el máximo matemático y el óptimo económico de dicha función, arrojando así un sólo resultado como óptimo; en cambio el análisis discriminante presenta un abanico de opciones donde quien tendrá la última palabra sobre cual dosis utilizar será el usuario. Esto es debido a que como en el análisis discriminante los límites no son iguales para cada categoría en cada dosis, el productor podrá escoger entre

dosis que le produzcan rendimientos menores, pero con menos gasto de fertilizantes hasta dosis con rendimientos altos y consumos altos de fertilizantes.

A diferencia de los resultados que arroja el modelo discontinuo de respuesta lineal y de rendimiento máximo estable, los resultados obtenidos a través del análisis discriminante no son proporciones sino rendimientos expresados en términos de kg/ha, presentando además la ventaja de que pueden incluir en las dosis a analizar varios nutrientes al mismo tiempo, el análisis discriminante no sólo puede recomendar la dosis de nutrientes a aplicar, sino que además estima el valor del rendimiento mínimo esperado.

Como se observó en el ejemplo anterior, no es igual clasificar una observación como rendimiento *Alto* para la dosis 0-80-50, que clasificarla en la misma categoría para la dosis 191-32-80, es de esperarse que los rendimientos sean mayores en esta última dosis, debido a que las cantidades de nutrientes agregados, son mucho mayores que las aportadas en la primera.

Por otro lado, la recomendación de la dosis de nutriente más adecuada va a depender de lo que el productor quiera obtener, es decir, si el productor desea rendimientos altos se le podría recomendar la dosis 191-32-80, la cual presentó el rendimiento esperado más elevado; si por el contrario el productor no desea obtener rendimientos altos sobre la base de un gasto elevado en fertilizantes, se le podría recomendar aquella o aquellas dosis en las cuales se esperen rendimientos altos con el menor consumo de fertilizantes, como por ejemplo 49-32-20 u otra cualquiera.

Para nuestro ejemplo se escogió como criterio de recomendación el beneficio económico aportado por el cultivo considerando al factor fertilización como única fuente de costos, es decir, las dosis consideradas deben ser sometidas a un análisis posterior para poder recomendar la dosis más adecuada. El beneficio queda expresado por la siguiente ecuación:

$$\text{Beneficio (Bs/ ha)} = \text{Ingreso por venta del producto} - \text{Costos de producción}$$

en nuestro ejemplo el producto a vender es el grano de maíz cosechado y los costos de producción están representados por los costos de los fertilizantes.

Para el cálculo del ingreso por concepto de venta del producto se tomó como precio del maíz 147 Bs/kg, ya que fue este el precio fijado por el gobierno en el año 1997.

5.4. RECOMENDACIÓN DE LAS COMBINACIONES DE PRODUCTOS COMERCIALES MÁS ECONÓMICAS.

Como se explicó en el capítulo correspondiente a Materiales y métodos, la obtención de las dosis óptimas se realizó utilizando la metodología de la programación lineal, para lo cual se dispuso de un menú de diferentes productos comerciales con sus respectivos precios y contenido de nutrientes. (Cuadro 18).

A través programa *MPS versión 2.1*, se obtuvo que los productos comerciales que suplen los requerimientos de nutrientes al más bajo costo son: la urea perlada, como fuente de nitrógeno, el fosfato especial como fuente de fósforo y nitrógeno y el cloruro de potasio como fuente de potasio.

Los demás productos comerciales no fueron tomados en cuenta por el programa debido a que encarecían los costos debidos a la fertilización. Sin embargo no debe descartarse su uso si en el mercado no se encuentra alguno de los productos comerciales recomendados.

Debe aclararse que esta recomendación va a depender de las fluctuaciones de los precios de los productos comerciales, así como también que las recomendaciones son válidas para los precios de los productos comerciales a partir de 01/01/1998, además la recomendación sólo tendrá sentido si el productor no es un individuo precio aceptante, es decir, no muestra preferencia por algún producto comercial en especial, y además sea un individuo que busque la minimización de los costos y el maximización de los beneficios.

Las combinaciones óptimas para las dosis estudiadas aparecen en los anexos del 5 hasta el 8, es decir, las combinaciones de productos comerciales más económicas para todos los estados estudiados y sus respectivos beneficios, a pesar de que sólo se consideran factibles los resultados obtenidos en el estado Yaracuy. Se colocaron estos resultados ya que en el futuro pueden obtenerse los tamaño de muestra lo suficientemente grandes y representativos como para obtener las ecuaciones discriminantes y poder elaborar las recomendaciones más adecuadas para cada uno de los estados considerados en este trabajo.

Se recomienda para el ejemplo anterior la dosis 191-32-80, ya que fue esta la que generó el mayor beneficio, a pesar de ser una de las más costosas. Al aplicar esta dosis se espera un beneficio mínimo de 811647.19 Bs/ha, el cual es considerablemente mayor al beneficio aportado por las demás dosis. (Anexo 5).

6. CONCLUSIONES Y RECOMENDACIONES.

La metodología del análisis discriminante arrojó resultados satisfactorios en las localidades de el Estado Yaracuy, obteniéndose 27 ecuaciones de discriminación con probabilidades de clasificación errónea baja para la categoría de rendimiento *Alto*.

Estas ecuaciones permitirán realizar estimaciones de rendimientos para las localidades del Estado Yaracuy consideradas en este análisis, con un nivel de confiabilidad bastante elevado para la Categoría de rendimiento *Alto*, por lo cual las recomendaciones emanadas de la aplicación de la metodología del análisis discriminante se presentan como confiables, ya que los productores pueden esperar un rendimiento alto y en el peor de los casos obtendrían un rendimiento medio.

Para los otros Estados analizados en este trabajo no se encontró buena discriminación, debido a que la probabilidades de clasificación errónea para la categoría de rendimiento *Alto* fueron muy altas para la mayoría de las dosis consideradas en el análisis, por lo cual las ecuaciones obtenidas no son factibles de utilizar para realizar estimaciones de rendimiento.

Los cuadros en los cuales se presentan las ecuaciones de discriminación para el Estado Yaracuy permitirán aplicar este modelo a cualquier muestra proveniente de las localidades estudiadas.

Las combinaciones de fertilizantes óptimas, es decir las de menor costo, se presentaron como la combinación de los mismos tres productos comerciales (urea, fosfato especial y cloruro de potasio), para todas las localidades en los Estados analizados.

El problema que se presentó en los Estados Apure, Guárico, Táchira y Monagas, en los cuales las ecuaciones de discriminación presentaron altas probabilidades de clasificación errónea para la categoría de rendimientos *Alto*, podría estar relacionado con un tamaño de muestra pequeño, homogeneidad de las condiciones edáficas de las localidades estudiadas y que el número de variables edáficas consideradas en las ecuaciones discriminantes en estos Estados fue menor que las utilizadas en el Estado Yaracuy.

A pesar de lo costoso que resulta la realización de ensayos, es necesario obtener un número mayor de observaciones de suelos y de rendimientos para las dosis consideradas ya que de esta manera podría mejorarse la precisión de las ecuaciones de discriminación.

Es necesario además considerar un número mayor de variables y no descartar la información aportada por las variables edáficas mencionadas anteriormente, ya que como se observó en las localidades del Estado Yaracuy, con un mayor número de variables se dispone más información que luego puede ser utilizada en las ecuaciones de discriminación.

La extrapolación de la metodología hacia otras zonas productoras de maíz, sólo será recomendable si las condiciones edáficas son similares a las presentes en los Estados considerados, esto con la finalidad de no incurrir en errores mayores al estimar los rendimientos.

7. REFERENCIAS BIBLIOGRÁFICAS.

- ASCANIO, M. 1995. Comparación de diseños de superficie de respuesta en la experimentación con cultivo de tejidos vegetales. Universidad Central de Venezuela. Maracay. Venezuela. Trabajo de ascenso a la categoría de Profesor Asistente. 218 p.
- BANDES, L. y GARCÍA, M. 1991. Efecto de los diferentes niveles de nitrógeno, fósforo, potasio y abono orgánico en el cultivo de ajo porro (*Allium ampeloprasum* L.) bajo un diseño compuesto central rotatable y su validación comercial. Universidad Central de Venezuela. Maracay. Venezuela. Tesis de grado. 94 p.
- BASTIDAS, L.; GARCÍA, M.; GIRÓN, G.; GÜERE, M.; LOZADA, C.; MACHADO, J.; ORTEGA, M.; RÍOS, M.; RODRÍGUEZ, B.; RODRÍGUEZ, O.; SOLÓRZANO, D.; VILORIA, J.; ZAPATA, R. y ZOPPI, M. 1984. Los estudios de fertilidad en Venezuela. Evolución histórica, situación actual, proposiciones para el futuro. Universidad Central de Venezuela. Post-Grado en Ciencia del Suelo. Asignatura Fertilidad de Suelos. 23 p.
- CATE, R. and NELSON, L. 1971. A simple statistical procedure for partitioning soil test correlation data into two classes. SOIL SCI. AMER. PROC. Vol 35. pp. 658-660.
- CHACÍN, F. 1980. El diseño rotatable central compuesto y su utilización en ensayos de fertilización en maíz. Universidad Central de Venezuela. Maracay. Venezuela. Trabajo de ascenso a la categoría de Profesor Asistente. 80 p.
- CHACÍN, F. 1983. Comparación de algunos diseños de segundo orden en la exploración de superficie de respuesta en experimentos agrícolas. Universidad Central de Venezuela. Maracay. Venezuela. Trabajo de ascenso a la categoría de Profesor Agregado. 286 p.

- CHACÍN, F. 1993. Diseño y análisis de experimentos para generar superficies de respuesta. Volumen II. Universidad Central de Venezuela. Maracay. Venezuela. Trabajo de ascenso a la categoría de Profesor Titular. 485 p.
- COLWELL, J.; SUHET, A. and VAN RAIJ, B. 1988. Statistical procedures for developing general soil fertility models for variable regions. CSIRO. Australia. Division of soils. Divisional report N° 93. 68 p.
- GIRI, N. 1977. Multivariate statistical inference. Academic Press. 319 p.
- GONZÁLEZ, R. 1977a. Informe sobre calibración de los análisis de fósforo del suelo y abonamiento fosforado para el cultivo de maíz en el estado Portuguesa. 14 p.
- GONZÁLEZ, R. 1977b. Informe sobre calibración de los análisis de potasio del suelo y abonamiento potásico para el cultivo de maíz en el estado Portuguesa. 10 p.
- GONZÁLEZ, R.; CHIRINOS, A.; AVILÁN, L.; PÉREZ, R.; VILAIN, L.; BLANCO, F. y MARCOVIC, L. 1978a. Informe sobre el proyecto evaluación de métodos de laboratorio para el análisis de fósforo y potasio disponibles en el suelo. 92 p.
- GONZÁLEZ, R.; PÉREZ, R.; AVILÁN, L.; CHIRINOS, A. y BLANCO, F. 1978b. El fósforo disponible y el efecto del abono fosforado en maizales de Portuguesa, Aragua, Carabobo y Yaracuy. *Agronomía Tropical*. Volumen. 28. p. 31-54
- HELYAR, K. and GODDEN, D. 1977. The biology and modelling of fertilizer response. *The Journal of the Australian Institute of Agricultural Science*. March-June. pp. 22-30.
- HERNÁNDEZ, J. 1993. Estrategias metodológicas utilizadas en el desarrollo de modelos para realizar recomendaciones de fertilizantes y su evaluación agroeconómica. Universidad Central de Venezuela. Facultad de Agronomía. Maracay. Venezuela. 122 p.

- LINARES, G. 1986. Estadística multivariada. Ciudad de La Habana. Cuba. Facultad de Matemática Cibernética.
- MACHADO, W. 1992. Comparación teórica y práctica de algunos diseños de tratamiento utilizados para estimar superficie de respuesta en ensayos de fertilización nitrogenada y densidad de siembra en el cultivo de maíz (*Zea mays* L.) y su validación comercial. Universidad Central de Venezuela. Maracay. Venezuela. Tesis de grado. 152 p.
- MARDIA, K. V.; KENT, J. T. and BIBBY, J. M. 1979. Multivariate Analysis. Academic Press. Londres. Inglaterra.
- MATA, M. 1989. Análisis discriminante y por conglomerados aplicados a la identificación de tres especies de cangrejos del género *Mithrax*. Universidad Central De Venezuela. Tesis de maestría. Maracay. Venezuela. 74 p.
- MONTIEL, E. 1986. Utilización de tres métodos (análisis de componentes principales, análisis de agrupamiento y análisis discriminante) para la identificación y caracterización de ambientes en Mantecal (Edo. Apure). Universidad Central De Venezuela. Tesis de maestría. Maracay. Venezuela. 250 p.
- MYERS, R. 1971. Surface Response. Allyn and Bacon, INC. Boston. Estados Unidos de América. 246 p.
- NATIONAL ACADEMY OF SCIENCES. 1961. Status and methods of research in economics and agronomics aspects of fertilizer response and use. National Academy of Sciences. National Research Council. Washintong D. C. Estados Unidos de América. Publication 918. 90 p.

- PFEIFFER, G. 1983. MPS-PC. Research Corporation/Research Software. Department of Agricultural Economics. Universidad de Nebraska-Lincoln. Estados Unidos de América.
- RAMÍREZ, R. 1989a. Estimación del requerimiento de fertilizante potásico por la yuca con base en el análisis de suelo. *Agronomía Tropical* Vol. 39. pp. 163-177.
- RAMÍREZ, R. 1989b. Relación entre el fósforo disponible y la respuesta de la yuca a la fertilización fosfatada. *Agronomía Tropical* Vol. 39. pp. 115-130.
- RAMÍREZ, R.; BEG, D.; COLMENARES, O.; MELÉNDEZ, E.; MARBAL, P.; BLANCO, F.; GUZMÁN, E.; HERNÁNDEZ, C. y CHIRINOS, A. 1988a. Relación entre el potasio disponible del suelo y el requerimiento de fertilizante potásico por el maíz. *Agronomía Tropical* Vol. 38. pp. 33-46.
- RAMÍREZ, R.; BEG, D.; COLMENARES, O.; MELÉNDEZ, E.; MARBAL, P.; BLANCO, F.; GUZMÁN, E.; HERNÁNDEZ, C. y CHIRINOS, A. 1988b. Relación entre la respuesta del maíz a la aplicación del fósforo y el P asimilable del suelo. *Agronomía Tropical* Vol. 38. pp. 5-19.
- RAMÍREZ, R.; DE SILVA, L.; TENÍAS, J.; RODRÍGUEZ, T. y CHIRINOS, A. 1989a. Calibración de cuatro métodos de análisis de suelo con la respuesta del maní a la fertilización con fósforo. *Agronomía Tropical* Vol. 39. pp. 5-21.
- RAMÍREZ, R.; DE SILVA, L.; TENÍAS, J.; RODRÍGUEZ, T. y CHIRINOS, A. 1989b. Relación entre el análisis de suelo y el requerimiento de fertilizante potásico por el maní. *Agronomía Tropical* Vol. 39. pp. 63-78.
- RAMÍREZ, R.; RODRÍGUEZ, T.; MILLÁN, A.; HERNÁNDEZ, C.; GUZMÁN, E. y TENÍAS, J. 1987. Relación del fósforo disponible del suelo con el requerimiento y respuesta del sorgo a la fertilización con este elemento. *Agronomía Tropical* Vol. 37. pp. 85-98.

- RAMÍREZ, R.; RODRÍGUEZ, T.; MILLÁN, A.; HERNÁNDEZ, C.; GUZMÁN, E. y TENÍAS, J. 1989c. Relación entre el requerimiento de fertilizante potásico por el sorgo y el potasio asimilable del suelo. *Agronomía Tropical* Vol. 39. pp. 179-193.
- RAMÍREZ, R. 1995. Impacto de la fertilidad de suelos en la producción agrícola del país y su posible contribución a una agricultura sostenible del futuro. Sociedad venezolana de la ciencia del suelo. Maracay. Venezuela. Boletín técnico especial. 40 años de la SVCS. pp. 134-146.
- ROUSE R. 1968. Soil test theory and calibration for cotton, corn and coastal bermudagrass. Agricultural experiment station. Auburn University. Alabama. Auburn. Alabama. Bulletin 375. 67 p.
- SAS INSTITUTE INC. 1989. SAS/STAT[®] user's guide. ANOVA-FREQ. Release 6. Fourth Edition. Volume 1. SAS INSTITUTE INC. Cary. North Carolina. USA. pp. 359-447.
- VILLASMIL, J.; MARTÍNEZ, E. y SEGURA, G. 1977. El diseño San Cristóbal y su uso en ensayos de fertilización en caña de azúcar. *Revista de la Facultad de Agronomía de la Universidad del Zulia*. Maracaibo. Venezuela. Volumen 3(4). pp. 7-25.
- WAUGH, D.; CATE, R. and NELSON, L. 1973. Modelos discontinuos para una rápida correlación, interpretación y utilización de los datos de análisis de suelos y las respuestas a los fertilizantes. Traducido al español por Sra. Isabel de Mendoza. Proyecto internacional de evaluación de la fertilidad del suelo del Departamento de Suelos, por la Estación experimental Agrícola de la Universidad del Estado de Carolina del Norte. Boletín técnico N° 7. 106 p.

8. ANEXOS

Anexo I. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Yaracuy.

Coefficientes

<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>	$\hat{P}_{ij(g)}$	\hat{P}_{ij}	Prob. a priori
90-60-50	<i>Alto</i>	-497.30866	-2.30263	-0.4819	60.00629		-134.14306		-0.73997	-73.3686	7.51376	21.00973	0.0952	0.2056	0.1963
	<i>Bajo</i>	-493.72696	-2.85577	-0.40003	62.78296		-147.92469		-1.03721	-64.92809	7.52997	20.154	0.8125		0.1495
	<i>Medio</i>	-484.5423	-2.66194	-0.38491	59.45386		-148.271		-0.68133	-64.54354	7.66467	20.24524	0.1		0.6542
180-60-50	<i>Alto</i>	-677.58477			83.61798		-415.88668				15.69668	23.0542	0.0	0.0	0.1667
	<i>Bajo</i>	-969.79584			99.30719		-516.31414				19.15584	27.74135	0.0		0.1667
	<i>Medio</i>	-851.35348			93.23267		-482.79062				17.90463	26.09555	0.0		0.6667
120-80-50	<i>Alto</i>	-38.05080				59.4683		-2.5496		-14.00635		1.27737	0.52	0.1449	0.1812
	<i>Bajo</i>	-20.83680				39.23211		-0.9877		-13.12037		0.86176	0.1034		0.2101
	<i>Medio</i>	-33.47475				64.73997		-3.04508		-9.94882		0.92848	0.0476		0.6087
120-80-0	<i>Alto</i>	-187.06121			51.10306	57.69898		-10.63795	-4.01416	-7.90208		4.24302	0.25	0.087	0.1739
	<i>Bajo</i>	-124.64403			43.3062	36.99299		-7.12691	-3.46886	-9.55324		3.28814	0.0		0.2174
	<i>Medio</i>	-163.83521			46.46362	75.88604		-11.39672	-3.56493	-3.35786		3.62014	0.0714		0.6087
90-60-100	<i>Alto</i>	-424.1551			19.91939				3.20307		8.66151	14.43535	0.0	0.0	0.2222
	<i>Bajo</i>	-354.46029			18.59489				2.75060		7.99027	13.06603	0.0		0.1111
	<i>Medio</i>	-449.16605			18.58338				3.41818		9.18512	15.00599	0.0		0.6667

Anexo I. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Yaracuy. (Continuación).

Coefficientes

<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
90-60-0	<i>Alto</i>	-21.09969	0.88321	-0.01329			34.58681		0.11187	-1.26799			0.3333	0.1667	0.1667
	<i>Bajo</i>	-9.21806	0.26363	-0.02799			25.65952		-0.04429	3.72823			0.6667		0.1667
	<i>Medio</i>	-7.8953	0.171193	0.01053			21.96203		0.08777	1.28045			0.0		0.6667
90-120-50	<i>Alto</i>	-27.1835	0.53029	-0.18333		18.36004	10.7376					1.02222	0.3333	0.1667	0.1667
	<i>Bajo</i>	-19.1656	0.30994	-0.16557		15.26483	7.34904					0.95074	1.0		0.1111
	<i>Medio</i>	-19.3199	0.17304	-0.137838		13.7838	15.39646					0.90781	0.0		0.7222
49-32-80	<i>Alto</i>	-582.13493		0.30422	153.57902	-105.74375	93.52282	-15.48075	-4.26979	-19.55742	3.54648	10.97433	0.0	0.0	0.15
	<i>Bajo</i>	-499.51945		0.25917	140.52815	-116.45828	92.0095	-7.84244	-4.15159	-28.38243	3.65096	8.74485	0.0		0.2
	<i>Medio</i>	-604.45312		0.31632	158.15336	-118.66775	103.51415	-12.64489	-4.42005	-26.65845	3.72938	10.11625	0.0		0.65
49-128-20	<i>Alto</i>	-1146		0.89366	290.89688	29.66767	198.6105	-50.922	-10.68513	-143.22314	2.31812	27.29148	0.0	0.0	0.1304
	<i>Bajo</i>	-790.78125		0.6633	236.07879	-3.10524	160.73788	-36.65248	-8.53166	-111.8438	2.78852	21.82723	0.0		0.2609
	<i>Medio</i>	-1025		0.78298	272.2577	23.68411	186.20548	-46.16262	-9.78913	-125.55441	2.73265	25.06916	0.0		0.6087
49-32-20	<i>Alto</i>	-48.81107	-0.66284	0.14852		62.18512		-5.6108	0.40743			1.49079	0.0	0.0	0.1
	<i>Bajo</i>	-15.09042	0.37759	-0.02898		56.9443		-5.43792	0.00721			0.83867	0.0		0.15
	<i>Medio</i>	-33.892	-0.033644	0.04357		75.62651		-7.12629	0.30484			1.40062	0.0		0.75

Anexo I. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Yaracuy. (Continuación).

Coefficientes

<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>	$\hat{P}_{ij(g)}$	\hat{P}_{ij}	Prob. a priori
37-96-80	<i>Alto</i>	-71.87583		0.58337			-41.78344	0.04954	1.85629		1.29734		0.0	0.0	0.1176
	<i>Bajo</i>	-39.89362		0.13482			1.78502	2.6192	0.52653		0.90336		0.0		0.1765
	<i>Medio</i>	-52.03601		0.36306			-20.95252	0.27875	1.30111		0.15752		0.0		0.7059
37-96-20	<i>Alto</i>	-88.22879		-0.06051	28.68889	7.48879			-1.99148	0.97912	-0.0779		0.0	0.0	0.1667
	<i>Bajo</i>	-94.99332		-0.19319	33.0868	9.50555			-2.7887	4.62728	-0.23106		0.0		0.1667
	<i>Medio</i>	-89.62998		0.04402	23.65892	3.07497			-1.45163	6.87881	0.53866		0.0		0.6667
37-24-80	<i>Alto</i>	-291.23773	14.03621	-1.12929			151.68171	-7.64227			4.56646	9.93065	0.0	0.0	0.1111
	<i>Bajo</i>	-135.54571	4.8142	-0.78492			87.22546	-1.85679			3.41296	7.6784	0.0		0.1667
	<i>Medio</i>	-170.24299	6.55711	-0.89045			114.96011	-3.50926			3.85307	8.74901	0.0		0.7222
37-24-20	<i>Alto</i>	-89.96069		-0.02024	27.70274	1.11456	15.883		-1.90049		0.0606		0.3333	0.0556	0.1667
	<i>Bajo</i>	-93.05722		-0.23384	32.4087	16.63013	14.46687		-2.94903		-0.33583		0.0		0.2222
	<i>Medio</i>	-85.81604		-0.08385	28.60837	8.60848	8.05661		-2.18242		-0.03877		0.0		0.6111
240-80-50	<i>Alto</i>	-581.78246			94.10768				-3.59191		6.6873	10.84902	0.25	0.0435	0.1739
	<i>Bajo</i>	-422.76609			80.10939				-3.08949		5.7356	9.20309	0.0		0.1739
	<i>Medio</i>	-534.73966			90.49596				-3.33767		6.43463	10.30832	0.0		0.6522

Anexo I. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Yaracuy. (Continuación).

Coefficientes

<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
191-32-80	<i>Alto</i>	-56.85677					4.70082				1.21298	2.16633	0.5	0.1739	0.2609
	<i>Bajo</i>	-43.04536					4.26066				1.0713	1.82604	0.2		0.2174
	<i>Medio</i>	-51.82929					7.19377				0.17357	1.98664	0.0		0.5217
191-32-20	<i>Alto</i>	-167.35112	1.7452	-0.25581		70.56006	-20.44804		2.27564		3.72446	4.01489	0.2	0.0435	0.2174
	<i>Bajo</i>	-148.079	2.0252	-0.2842		60.28539	-21.9247		2.23861		3.61669	3.87488	0.0		0.1739
	<i>Medio</i>	-184.88333	1.92662	-0.2872		78.31987	-21.62245		2.44351		3.93128	4.06302	0.0		0.6087
191-128-80	<i>Alto</i>	-476.45938	-1.20029	0.5498	146.31508	-96.36225	121.01161	-12.71995	-8.17902	-87.95023		9.41812	0.0	0.0	0.1739
	<i>Bajo</i>	-341.53146	-0.55442	0.41467	123.50906	-71.87856	98.88687	-11.74793	-6.96241	-76.2601		8.15069	0.0		0.2174
	<i>Medio</i>	-407.58356	-0.36625	0.41489	130.63493	-20.31794	101.00492	-18.11239	-70.07714	-73.31094		9.04829	0.0		0.6087
191-128-20	<i>Alto</i>	-677.61916	1.1396	0.0627	141.90935	-129.37784	58.22293		-5.36934	-53.8144	6.48344	13.84368	0.0	0.0	0.2174
	<i>Bajo</i>	-540.57167	1.52229	-0.02019	123.34205	-110.61671	45.87358		-4.29412	-44.35898	6.14493	12.36869	0.0		0.1739
	<i>Medio</i>	-633.46232	1.47302	-0.04095	129.06648	-97.14505	46.68228		-4.28023	-40.02439	6.8958	12.89827	0.0		0.6087
143-96-20	<i>Alto</i>	-32.54691		-0.07648		17.56204	45.41332	2.50008		-0.186			0.0	0.0556	0.2222
	<i>Bajo</i>	-23.68006		-0.09752		18.22056	30.48757	2.03985		2.5129			0.0		0.1667
	<i>Medio</i>	-46.14048		-0.12146		24.9435	50.67805	2.5129		3.35024			0.0909		0.6111

Anexo I. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Yaracuy. (Continuación).

Coefficientes

<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>	$\hat{P}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
143-24-80	<i>Alto</i>	-38.03798			12.84054		-20.70713	-2.49401		6.45059		0.77111	0.0	0.0	0.2222
	<i>Bajo</i>	-90.54141			29.9981		-51.00271	-11.18169		47.50485		0.43816	0.0		0.1667
	<i>Medio</i>	-45.78899			18.96242		-29.77398	-6.10539		25.24787		0.5847	0.0		0.6111
120-0-50	<i>Alto</i>	-167.35918			45.01907	88.2127		-12.20972	-3.49713			3.25279	0.75	0.1739	0.1739
	<i>Bajo</i>	-112.61267			38.06775	62.64771		-9.30746	-3.05279			2.5774	0.0		0.2174
	<i>Medio</i>	-157.72283			42.99808	91.33161		-11.97695	-3.27113			3.08688	0.0714		0.6087
0-80-50	<i>Alto</i>	-230.1022						12.47201		-30.11256	4.1088	5.73799	0.0	0.0	0.1765
	<i>Bajo</i>	-230.83344						14.77706		-36.66079	4.18165	4.97007	0.0		0.1176
	<i>Medio</i>	-177.64032						12.09452		-28.16007	3.71904	4.595	0.0		0.7059
0-60-50	<i>Alto</i>	-71.76816	1.09118			-6.42882	95.51583	3.99405			0.87773		0.0	0.0	0.1667
	<i>Bajo</i>	-44.40431	-0.24252			6.74923	32.92041	4.01843			0.87104		0.0		0.1667
	<i>Medio</i>	-64.66935	-0.06159			7.0782	57.63783	4.68088			0.97581		0.0		0.6667
120-160-50	<i>Alto</i>	-68.3208				-45.61154	-5.66217	11.32326	0.5427	-5.66753	1.82393		0.5	0.1304	0.1739
	<i>Bajo</i>	-67.61896				-59.63751	-4.88803	11.81981	0.78218	-3.70401	1.98646		0.0		0.2174
	<i>Medio</i>	-70.23153				-35.28231	-1.80514	9.9018	0.6294	-4.19254	1.83648		0.0714		0.6087

Anexo I. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Yaracuy. (Continuación).

Coefficientes

<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>C.I.C.</i>	<i>Ca</i>	<i>Mg</i>	<i>a</i>	<i>A</i>	$\hat{P}_{ij(g)}$	\hat{P}_{ij}	Prob. a priori
120-80-100	<i>Alto</i>	-145.61679					-8.97492	6.46367	1.88112	1.30697	3.25527	2.77473	0.5	0.087	0.1739
	<i>Bajo</i>	-113.93305					-6.1614	5.83343	1.59522	0.48105	2.89984	2.35036	0.0		0.1739
	<i>Medio</i>	-148.45736					-9.50402	6.33727	2.03648	5.9879	3.3944	2.54325	0.0		0.6522
143-24-20	<i>Alto</i>	-358.38437			39.03574		-136.17602		0.74757	-43.03942	5.80481	12.90061	0.0	0.0556	0.1667
	<i>Bajo</i>	-364.22884			36.8132		-161.63427		1.24059	-34.32621	6.36314	12.91242	0.0		0.1111
	<i>Medio</i>	-341.39846			31.59971		-158.04429		1.55429	-33.69119	6.47188	13.15322	0.0769		0.72222

Anexo 2. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Apure.

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
0-30-20	<i>Alto</i>	-175.05463			64.63306			-0.06172	0.6	0.2632	0.2632
	<i>Bajo</i>	-179.4511			65.76074			-0.06647	0.25		0.2105
	<i>Medio</i>	-179.16205			64.87276			-0.05718	0.1		0.5263
140-0-20	<i>Alto</i>	-7.21722	-0.1343			1.19921	42.25986	0.01104	1.0	0.2	0.1
	<i>Bajo</i>	-5.07965	-0.08936			0.37625	62.70626	0.0063	0.0		0.2
	<i>Medio</i>	-9.0784	-0.20025			1.03699	65.91861	0.01531	0.1429		0.7
140-30-0	<i>Alto</i>	-175.89328			62.26637		369.89703	-0.09468	0.75	0.2381	0.1905
	<i>Bajo</i>	-178.64999			63.29995		375.70695	-0.10158	0.25		0.1905
	<i>Medio</i>	-170.88059			61.62232		348.16191	-0.09195	0.0769		0.619
140-30-20	<i>Alto</i>	-232.42938	1.3542	-0.06719	84.62815	8.55265	353.05247	-0.149	1.0	0.3306	0.17744
	<i>Bajo</i>	-240.33692	1.35705	-0.05738	86.2814	7.62919	352.14892	-0.15244	1.0		0.1532
	<i>Medio</i>	-233.10033	1.28953	-0.05973	84.88767	7.93248	356.16640	-0.14762	0.0		0.6694
140-30-40	<i>Alto</i>	-179.12466			65.09483	4.66962	390.99892	-0.10375	0.6667	0.1905	0.1429
	<i>Bajo</i>	-194.86602			67.95641	4.31909	448.55375	-0.11332	1.0		0.0952
	<i>Medio</i>	-190.83892			67.33413	4.72939	432.51497	-0.10855	0.0		0.7619

Anexo 2. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Apure. (Continuación).

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
140-60-20	<i>Alto</i>	-6.81553		0.01234		0.54584	57.86694	0.00295	1.0	0.2381	0.1429
	<i>Bajo</i>	-4.35952		0.01728		-0.13215	47.94766	0.000191	0.6667		0.1429
	<i>Medio</i>	-5.23404		0.00362		0.73174	39.57935	0.00643	0.0		0.7143
223-12-32	<i>Alto</i>	-7.13989	-0.17531				47.81673	0.01344	0.6	0.3684	0.2632
	<i>Bajo</i>	-4.75203	-0.0771				55.38663	0.00669	0.25		0.2105
	<i>Medio</i>	-8.40708	-0.14839				70.13613	0.01316	0.3		0.5263
223-12-8	<i>Alto</i>	-173.64765			61.67704	4.56459	366.1346	-0.09791	1.0	0.3	0.15
	<i>Bajo</i>	-171.7362			61.62841	4.02797	363.94474	-0.09897	1.0		0.15
	<i>Medio</i>	-175.87205			62.32821	3.87715	375.722	-0.09784	0.0		0.7
223-48-32	<i>Alto</i>	-11.70244	-0.32992				134.64672	0.00905	0.6667	0.1905	0.1429
	<i>Bajo</i>	-5.91642	0.02254				70.87788	0.00242	0.6667		0.1429
	<i>Medio</i>	-4.91104	-0.20781				67.05251	0.00914	0.0		0.7143
223-48-8	<i>Alto</i>	-8.73094		-0.02088		3.15337	67.741157	0.0057	0.5	0.1905	0.1905
	<i>Bajo</i>	-3.8408		0.01138		-0.06357	37.97861	0.00311	0.25		0.1905
	<i>Medio</i>	-6.93352		0.03664		0.31457	75.30257	0.0007699	0.0769		0.619

Anexo 2. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Apure. (Continuación).

Coefficientes											
Dosis	Nivel	Constant.	P_{olsen}	K_{olsen}	pH	C.O.	C.E.	Ca	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
280-30-20	Alto	-201.33483			72.05963		345.43652	-0.11465	1.0	0.2	0.1
	Bajo	-209.89532			73.66398		339.42335	-0.11561	0.6667		0.15
	Medio	-186.8314			69.33457		340.35581	-0.10767	0.0		0.75
57-12-8	Alto	-202.94807	1.72334		75.83417			-0.12673	0.6667	0.1905	0.1429
	Bajo	-223.23739	1.87457		80.47746			-0.1413	0.5		0.1905
	Medio	-190.70285	1.67289		74.35444			-0.12753	0.0		0.6667
57-48-32	Alto	-195.01216			70.15733		240.53992	-0.08012	1.0	0.3333	0.1667
	Bajo	-195.24174			70.59737		235.96922	-0.08333	1.0		0.1667
	Medio	-201.63704			71.64044		238.33403	-0.08126	0.0		0.6667
57-48-8	Alto	-239.99848			87.41531	12.9137	232.50448	-0.12377	1.0	0.2778	0.1667
	Bajo	-265.48329			92.11914	14.53796	212.60588	-0.12983	0.6667		0.1667
	Medio	-241.52404			88.00945	12.41279	217.48873	-0.12245	0.0		0.6667

Anexo 3. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en los estados Guárico y Táchira.

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
0-120-60	<i>Alto</i>	-154.53177			57.98289	3.19433	-119.48914	-0.08269	0.0	0.1176	0.1765
	<i>Bajo</i>	-191.71081			66.49954	-0.0942	-138.81691	-0.09899	0.25		0.2353
	<i>Medio</i>	-141.49768			57.01446	-0.05692	-113.96891	-0.08362	0.1		0.5882
115-0-60	<i>Alto</i>	-106.65475			39.34414	1.40772	-28.92107	-0.05458	0.3333	0.2353	0.1765
	<i>Bajo</i>	-132.20091			44.2266	2.1932	-30.37136	-0.06558	0.25		0.2353
	<i>Medio</i>	-103.21289			39.27967	1.7499	-31.45629	-0.05703	0.2		0.5882
115-120-0	<i>Alto</i>	-144.91253			51.18843	-3.57356	-46.19323	-0.06989	0.5	0.1176	0.1176
	<i>Bajo</i>	-152.60554			52.27114	-1.01896	-41.14054	-0.07457	0.25		0.2353
	<i>Medio</i>	-118.43657			46.27048	-1.24751	-40.5062	-0.06446	0.0		0.6471
115-120-120	<i>Alto</i>	-80.82245			29.49407		-42.32421	-0.03781	1.0	0.1765	0.1176
	<i>Bajo</i>	-93.0943			32.01377		-49.30425	-0.04313	0.5		0.1176
	<i>Medio</i>	-77.64762			29.39765		-42.24512	-0.03877	0.0		0.7647
115-120-60	<i>Alto</i>	-132.05072		0.35481	45.76255			-0.06784	0.7059	0.1667	0.1667
	<i>Bajo</i>	-179.52692		0.39519	53.83911			-0.08117	0.1818		0.2157
	<i>Medio</i>	-130.52014		0.28983	46.37691			-0.06786	0.0159		0.6176

**Anexo 3. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en los estados Guárico y Táchira.
(Continuación).**

<i>Coefficientes</i>											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	<i>Prob. a priori</i>
183-191-24	<i>Alto</i>	-126.15011			45.37473	1.9401	-84.15597	-0.06955	0.5	0.1176	0.1176
	<i>Bajo</i>	-140.12021			47.28161	4.80153	-87.6203	-0.07269	0.3333		0.1765
	<i>Medio</i>	-10613159			41.35643	3.25117	-72.61955	-0.06171	0.0		0.7059
183-191-96	<i>Alto</i>	-42.93704	2.54696		15.38025	-10.17331			1.0	0.2353	0.1765
	<i>Bajo</i>	-50.19359	2.79488		16.56608	-10.57471			0.0		0.1765
	<i>Medio</i>	-60.16804	3.12465		18.38788	-11.79652			0.0909		0.6471
183-49-24	<i>Alto</i>	-112.10173	2.23925	0.44305	38.97647	-10.57506	-69.63733	-0.05072	0.3333	0.1176	0.1765
	<i>Bajo</i>	-142.79	2.05711	0.47359	44.65375	-7.3477	-88.29966	-0.06458	0.3333		0.1765
	<i>Medio</i>	-118.38924	2.2381	0.44387	40.71729	-8.88788	-75.59831	-0.05689	0.0		0.6471
183-49-96	<i>Alto</i>	-283.70413	3.55126	1.24582	91.1215	-13.29069		-0.13459	0.6667	0.1176	0.1765
	<i>Bajo</i>	-392.96439	3.48954	1.40727	107.86531	-11.34979		-0.16302	0.0		0.1765
	<i>Medio</i>	-263.72893	3.29184	1.11672	88.90739	-11.47922		-0.13162	0.0		0.6471
230-120-60	<i>Alto</i>	-202.31581		0.90362	70.05697		-160.98723	-0.10292	0.25	0.1176	0.2353
	<i>Bajo</i>	-297.30453		1.13297	85.34313		-200.74695	-0.13075	0.0		0.1176
	<i>Medio</i>	-206.87917		0.92758	71.47947		-164.13826	-0.10887	0.0909		0.6471

**Anexo 3. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en los estados Guárico y Táchira.
(Continuación).**

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
47-191-24	<i>Alto</i>	-120.29103			44.47962	-0.0931	-42.40667	-0.07009	0.6667	0.1765	0.1765
	<i>Bajo</i>	-139.42344			47.45544	1.49028	-38.56345	-0.07498	0.25		0.2353
	<i>Medio</i>	-106.02991			41.38797	1.22484	-36.96696	-0.0631	0.0		0.5882
47-191-96	<i>Alto</i>	-153.26878			57.8438	3.02192	-97.19951	-0.08012	0.3333	0.1176	0.1765
	<i>Bajo</i>	-214.36009			68.90657	3.52498	-117.51166	-0.09862	0.0		0.1176
	<i>Medio</i>	-150.24891			57.8722	2.68041	-95.55749	-0.082	0.0833		0.7059
47-49-24	<i>Alto</i>	-202.29082		0.86216	69.93136		-159.49469	-0.10163	1.0	0.1176	0.1176
	<i>Bajo</i>	-293.71441		1.15095	83.73222		-198.57085	-0.1249	0.0		0.1176
	<i>Medio</i>	-203.65824		0.94604	69.95075		-162.11641	-0.10334	0.0		0.7647

Anexo 4. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Monagas.

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
0-100-70	<i>Alto</i>	-57.5162			22.03124	43.87406	-362.38641	-0.0398	0.0	0.0417	0.125
	<i>Bajo</i>	-90.32534			32.19903	76.54623	-452.28391	-0.11937	0.0		0.0833
	<i>Medio</i>	-75.32165			29.21077	66.71156	-272.85698	-0.12274	0.0526		0.7917
144-159-112	<i>Alto</i>	-74.00868			24.80784	51.5534	-117.55234	-0.06102	0.6	0.1786	0.1786
	<i>Bajo</i>	-65.48706			24.46254	53.56159	-271.66626	-0.06138	0.4		0.1786
	<i>Medio</i>	-58.46533			23.23473	44.98568	-159.56731	-0.05926	0.0		0.6429
144-159-28	<i>Alto</i>	-60.61405			23.52819	51.92864	-291.03952	-0.05893	0.5	0.1786	0.1786
	<i>Bajo</i>	-74.29905			27.90603	61.91987	-472.34397	-0.07404	0.5		0.1786
	<i>Medio</i>	-60.67022			25.31264	49.75365	-333.30985	-0.06818	0.0		0.6429
144-41-112	<i>Alto</i>	-11.223				18.44132	34.58788	-0.00346	0.5	0.1429	0.1429
	<i>Bajo</i>	-6.2344				18.27847	-22.53587	-0.00802	0.5		0.1429
	<i>Medio</i>	-3.07964				8.72228	60.1612	-0.00743	0.0		0.7143
144-41-28	<i>Alto</i>	-10.98037	-0.04908				190.0187	0.00214	0.5	0.2143	0.1429
	<i>Bajo</i>	-3.39425	-0.01567				83.69148	0.00266	1.0		0.1429
	<i>Medio</i>	-2.15378	-0.00335				98.93573	0.00336	0.0		0.7143

Anexo 4. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Monagas.(Continuación).

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
180-100-70	<i>Alto</i>	-95.62728	0.18062	0.19492	32.28704	57.14817	-195.71197	-0.08334	0.25	0.2857	0.2857
	<i>Bajo</i>	-78.19028	0.25825	0.1271	30.07721	54.90778	-298.02335	-0.07674	0.5		0.2143
	<i>Medio</i>	-75.93602	0.22291	0.1341	29.46085	52.68012	-219.11633	-0.07741	0.2143		0.5
36-159-112	<i>Alto</i>	-58.06657			20.20647	40.20615	-145.71957	-0.04575	0.3333	0.1429	0.1071
	<i>Bajo</i>	-57.79642			22.77033	36.63524	-218.59402	-0.06183	1.0		0.1071
	<i>Medio</i>	-55.34534			22.28274	37.15627	-164.94588	-0.06301	0.0		0.7857
36-159-28	<i>Alto</i>	-121.70116	1.10147		38.93627	77.71199	-475.88742	-0.08595	0.25	0.0357	0.1429
	<i>Bajo</i>	-119.66979	1.02127		38.90056	93.46168	-575.58914	-0.1019	0.0		0.0714
	<i>Medio</i>	-86.07912	0.80354		33.95498	72.1025	-440.98198	-0.08499	0.0		0.7857
36-41-112	<i>Alto</i>	-4.24335	-0.03273	0.06462				0.00484	0.6	0.2222	0.1852
	<i>Bajo</i>	-4.27228	-0.07767	0.06148				-0.00222	0.75		0.1481
	<i>Medio</i>	-1.67322	-0.03594	0.07819				-0.00101	0.0		0.6667
36-41-28	<i>Alto</i>	-12.06264	0.21472				114.00537	0.00448	0.25	0.1071	0.1429
	<i>Bajo</i>	-3.22695	0.00717				50.2921	-0.0009367	1.0		0.0714
	<i>Medio</i>	-2.1612	0.03868				92.64882	-0.00373	0.0		0.7857

Anexo 4. Ecuaciones discriminantes correspondientes a la combinaciones de nutrientes en el estado Monagas. (Continuación).

Coefficientes											
<i>Dosis</i>	<i>Nivel</i>	<i>Constant.</i>	<i>P_{olsen}</i>	<i>K_{olsen}</i>	<i>pH</i>	<i>C.O.</i>	<i>C.E.</i>	<i>Ca</i>	$\hat{p}_{ij(g)}$	\hat{p}_{ij}	Prob. a priori
90-0-70	<i>Alto</i>	-38.68496			16.17526		61.47958	-0.03297	0.5	0.2308	0.1538
	<i>Bajo</i>	-40.64689			17.92017		7.44007	-0.04228	1.0		0.1538
	<i>Medio</i>	-46.05977			19.43511		10.46166	-0.04609	0.0		0.6923
90-100-0	<i>Alto</i>	-67.78923			24.7025	53.29161	-289.95003	-0.05734	0.6	0.1852	0.1852
	<i>Bajo</i>	-64.99956			24.93644	58.89134	-432.36289	-0.05931	0.4		0.1852
	<i>Medio</i>	-56.70562			23.09648	49.4241	-270.04793	-0.0583	0.0		0.6296
90-100-140	<i>Alto</i>	-41.8827		0.18389	18.33444		-125.27272	-0.03163	0.25	0.1071	0.1429
	<i>Bajo</i>	-69.56624		0.43769	26.11538		-452.56797	-0.05113	0.6667		0.1071
	<i>Medio</i>	-52.44227		0.31624	23.06614		-319.4165	-0.04606	0.0		0.75
90-100-70	<i>Alto</i>	-9.82209				10.76695	100.7452	-0.00321	0.5385	0.1758	0.1576
	<i>Bajo</i>	-4.73467				13.4469	4.73926	-0.00777	0.5		0.1697
	<i>Medio</i>	-3.5699				7.73526	86.12679	-0.00863	0.009		0.6727
90-200-70	<i>Alto</i>	-110.43068			36.51699	108.98227	-624.86475	-0.08082	0.5	0.08	0.16
	<i>Bajo</i>	-120.01171			38.8904	126.77803	-888.20181	-0.08968	0.0		0.12
	<i>Medio</i>	-79.87567			32.11738	92.20669	-559.81201	-0.07555	0.0		0.72

Anexo 5. Combinaciones óptimas de productos comerciales para las dosis de nutrientes en el estado Yaracuy.

<i>Dosis</i>	<i>Urea (kg/ha)</i>	<i>Fosfato especial (kg/ha)</i>	<i>Cloruro de Potasio (kg/ha)</i>	<i>Costo de fertilización (Bs/ha)</i>	<i>Rendimiento mínimo esperado (kg/ha)</i>	<i>Beneficio mínimo esperado (Bs/ha)</i>
0-60-50	0.0	142.857	83.333	25282.14	4274.64	603089.94
0-80-50	0.0	190.476	83.333	31234.52	3881.90	539404.78
37-24-20	60.559	57.143	33.333	15490.5	4537.90	651580.8
37-24-80	60.559	57.143	133.333	24400.5	4795.07	680474.79
37-96-20	0.932	228.571	33.333	31624.16	4274.64	596747.92
37-96-80	0.932	228.571	133.333	40534.16	4518.42	623673.58
49-128-20	0.518	304.762	33.333	41111.2	4443.96	612150.92
49-32-20	80.021	76.190	33.333	19599.65	3572.97	505626.94
49-32-80	80.021	76.190	133.333	28509.65	4641.42	653779.09
90-120-50	96.273	285.714	83.333	51688.35	4813.39	655879.98
90-60-0	145.963	142.857	0.0	30818.63	5499.90	777666.67
90-60-100	145.963	142.857	166.667	45668.63	4898.63	674429.98
90-60-50	145.963	142.857	83.333	38243.63	4836.58	642733.63

Anexo 5. Combinaciones óptimas de productos comerciales para las dosis de nutrientes en el estado Yaracuy. (Continuación).

<i>Dosis</i>	<i>Urea (kg/ha)</i>	<i>Fosfato especial (kg/ha)</i>	<i>Cloruro de Potasio (kg/ha)</i>	<i>Costo de fertilización (Bs/ha)</i>	<i>Rendimiento mínimo esperado (kg/ha)</i>	<i>Beneficio mínimo esperado (Bs/ha)</i>
120-0-50	260.87	0.0	83.333	30590.22	5262.11	742939.95
120-160-50	128.364	380.952	83.333	66442.81	5221.15	701066.24
120-80-0	194.617	190.476	0.0	41091.51	4763.08	659081.25
120-80-100	194.617	190.476	166.667	55941.51	5173.84	704612.97
120-80-50	194.617	190.476	83.333	48516.51	5253.53	723752.4
143-24-20	290.994	57.143	33.333	35953.11	5589.82	785750.43
143-24-80	290.994	57.143	133.333	44863.11	4981.77	687457.08
143-96-20	231.366	228.571	33.333	52086.77	5080.53	694751.14
180-60-50	341.615	142.857	83.333	55617.55	5643.63	773996.06
191-128-20	309.213	304.762	33.333	68523.37	5282.36	707983.55
191-128-80	309.213	304.762	133.333	77433.37	5887.75	788065.88
191-32-20	388.716	76.190	33.333	47011.82	5304.21	732707.05
191-32-80	388.716	76.190	133.333	55921.82	5901.83	811647.19
240-80-50	455.487	190.476	83.333	71681.73	5780.36	778031.19

Anexo 6. Combinaciones óptimas de productos comerciales para las dosis de nutrientes en el estado Apure.

<i>Dosis</i>	<i>Urea (kg/ha)</i>	<i>Fosfato especial (kg/ha)</i>	<i>Cloruro de Potasio (kg/ha)</i>	<i>Costo de fertilización (Bs/ha)</i>	<i>Rendimiento mínimo esperado (kg/ha)</i>	<i>Beneficio mínimo esperado (Bs/ha)</i>
0-30-20	0.0	71.429	33.333	11898.57	5149.05	745011.78
57-12-8	113.975	28.571	13.333	14880.42	4574.59	657584.31
57-48-32	84.161	114.286	53.333	26511.25	5364.11	762012.92
57-48-8	84.161	114.286	13.333	22947.25	5280.61	753302.42
140-0-20	304.348	0.0	33.333	29996.09	4975.06	701337.73
140-30-0	279.503	71.429	0.0	33748.45	5257.65	739126.1
140-30-20	279.503	71.429	33.333	36718.45	5391.54	755837.93
140-30-40	279.503	71.429	66.667	39688.45	5293.50	438456.05
140-60-20	254.658	142.857	33.333	43440.81	5032.34	696313.17
223-12-32	474.845	28.571	53.333	50489.64	5431.87	747995.25
223-12-8	474.845	28.571	13.333	46925.64	5149.40	710036.16
223-48-32	445.031	114.286	53.333	58556.47	5859.76	802828.25
223-48-8	445.031	114.286	13.333	54992.47	5338.60	729781.73
280-30-20	583.851	71.429	33.333	63744.53	5368.97	725494.06

Anexo 7. Combinaciones óptimas de productos comerciales para las dosis de nutrientes en los estados Guárico y Táchira.

<i>Dosis</i>	<i>Urea (kg/ha)</i>	<i>Fosfato especial (kg/ha)</i>	<i>Cloruro de Potasio (kg/ha)</i>	<i>Costo de fertilización (Bs/ha)</i>	<i>Rendimiento mínimo esperado (kg/ha)</i>	<i>Beneficio mínimo esperado (Bs/ha)</i>
0-120-60	0.0	285.714	100	44624.29	3278.02	437244.65
47-191-24	0.0	454.762	40	60409.24	3676.06	479971.58
47-191-96	0.0	454.762	160	71101.24	3795.66	486860.78
47-49-24	61.594	116.667	40	23616.9	3271.34	457270.08
115-0-60	250	0.0	100	31110	4554.30	638372.1
115-120-0	150.621	285.714	0.0	49089.44	4458.16	606260.08
115-120-120	150.621	285.714	200	66909.44	4695.70	623358.46
115-120-60	150.621	285.714	100	57999.44	4568.04	626732.44
183-191-24	239.648	454.762	40	81689.98	4614.70	600639.92
183-191-96	239.648	454.762	160	92381.98	5167.52	667243.46
183-49-24	357.246	116.667	40	49870.81	5119.18	702648.65
183-49-96	357.246	116.667	160	60562.81	4765.45	639958.34
230-120-60	400.621	285.714	100	80199.44	5162.06	678623.38

Anexo 8. Combinaciones óptimas de productos comerciales para las dosis de nutrientes en el estado Monagas.

<i>Dosis</i>	<i>Urea (kg/ha)</i>	<i>Fosfato especial (kg/ha)</i>	<i>Cloruro de Potasio (kg/ha)</i>	<i>Costo de fertilización (Bs/ha)</i>	<i>Rendimiento mínimo esperado (kg/ha)</i>	<i>Beneficio mínimo esperado (Bs/ha)</i>
0-100-70	0.0	238.095	116.667	40156.90	3508.15	475541.15
36-159-112	0.0	378.571	186.667	63953.43	4145.73	545468.88
36-159-28	0.0	378.571	46.667	51479.43	4273.97	576794.16
36-41-28	44.306	97.619	46.667	20294.79	4379.89	623549.04
36-41-112	44.306	97.619	186.667	32768.79	4409.15	615376.26
90-0-70	195.652	0.0	116.667	27768.91	4894.78	691763.75
90-100-0	112.836	238.095	0.0	39781.78	4685.13	648932.33
90-100-140	112.836	238.095	233.333	60571.78	4340.11	577424.39
90-100-70	112.836	238.095	116.667	50176.78	4657.22	634434.56
90-200-70	30.021	476.19	116.667	72584.65	5077.71	67383872
144-159-112	181.366	378.571	186.667	80058.77	4906.28	641164.39
144-159-28	181.366	378.571	46.667	67584.77	4446.03	585981.64
144-41-112	279.089	97.619	186.667	53617.49	4537.31	613367.08
144-41-28	279.089	97.619	46.667	41143.49	4645.36	641724.43
180-100-70	308.489	238.095	116.667	67550.7	5156.43	690444.51

Anexo 9. Ejemplo de una salida del computador del programa SAS versión 6.04 para el análisis discriminante.

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis

107 Observations 106 DF Total
8 Variables 104 DF Within Classes
3 Classes 2 DF Between Classes

Class Level Information

CATEG	Frequency	Weight	Proportion	Prior Probability
ALTO	21	21.0000	0.196262	0.196262
BAJO	16	16.0000	0.149533	0.149533
MEDIO	70	70.0000	0.654206	0.654206

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis Simple Statistics

Total-Sample

Variable	N	Sum	Mean
POLSEN	107	855.00000	7.99065
KOLSEN	107	9658	90.26168
PH11	107	751.10000	7.01963
CA	107	1828	17.08037
CE	107	52.91000	0.49449
MG	107	97.20000	0.90841
TOTAL	107	4073	38.06355
A	107	2436	22.76262

Total-Sample

Variable	Variance	Std Dev	Label
POLSEN	47.25463	6.87420	P EN PPM
KOLSEN	2636	51.34352	K EN PPM
PH11	1.00914	1.00456	PH EN SOLUCION 1:1
CA	158.15989	12.57616	CA EN PPM (SOLUCION)
CE	0.02013	0.14188	COND. ELECT. EM dMHOS/CM
MG	0.30870	0.55561	MG EN PPM (SOLUCION)
TOTAL	156.02894	12.49115	ARENA EN %
A	51.83255	7.19948	ARCILLA EN %

CATEG = ALTO

Variable	N	Sum	Mean
POLSEN	21	293.00000	13.95238
KOLSEN	21	2280	108.57143
PH11	21	160.20000	7.62857
CA	21	544.40000	25.92381
CE	21	12.70000	0.60476
MG	21	11.70000	0.55714
TOTAL	21	525.90000	25.04286
A	21	570.80000	27.18095

CATEG = ALTO

Variable	Variance	Std Dev	Label
POLSEN	102.14762	10.10681	P EN PPM
KOLSEN	4476	66.90259	K EN PPM
PH11	0.77914	0.88269	PH EN SOLUCION 1:1
CA	264.29290	16.25709	CA EN PPM (SOLUCION)
CE	0.01019	0.10093	COND. ELECT. EM dMHOS/CM
MG	0.00557	0.07464	MG EN PPM (SOLUCION)
TOTAL	49.05057	7.00361	ARENA EN %
A	13.55062	3.68112	ARCILLA EN %

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ

ESTADO YARACUY

USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis Simple Statistics

CATEG = BAJO

Variable	N	Sum	Mean
POLSEN	16	68.00000	4.25000
KOLSEN	16	1013	63.31250
PH11	16	111.90000	6.99375
CA	16	232.60000	14.53750
CE	16	6.77000	0.42312
MG	16	15.10000	0.94375
TOTAL	16	699.30000	43.70625
A	16	311.10000	19.44375

CATEG = BAJO

Variable	Variance	Std Dev	Label
POLSEN	5.53333	2.35230	P EN PPM
KOLSEN	1275	35.71315	K EN PPM
PH11	1.06729	1.03310	PH EN SOLUCION 1:1
CA	107.68383	10.37708	CA EN PPM (SOLUCION)
CE	0.02684	0.16382	COND. ELECT. EM dMHOS/CM
MG	0.27062	0.52022	MG EN PPM (SOLUCION)
TOTAL	101.52596	10.07601	ARENA EN %
A	38.61996	6.21450	ARCILLA EN %

CATEG = MEDIO

Variable	N	Sum	Mean
POLSEN	70	494.00000	7.05714
KOLSEN	70	6365	90.92857
PH11	70	479.00000	6.84286
CA	70	1051	15.00857
CE	70	33.44000	0.47771
MG	70	70.40000	1.00571
TOTAL	70	2848	40.68000
A	70	1554	22.19571

CATEG = MEDIO

Variable	Variance	Std Dev	Label
POLSEN	26.83727	5.18047	P EN PPM
KOLSEN	2204	46.94904	K EN PPM
PH11	0.94770	0.97350	PH EN SOLUCION 1:1
CA	113.29819	10.64416	CA EN PPM (SOLUCION)
CE	0.01697	0.13027	COND. ELECT. EM dMHOS/CM
MG	0.36634	0.60526	MG EN PPM (SOLUCION)
TOTAL	137.48133	11.72524	ARENA EN %
A	58.48187	7.64734	ARCILLA EN %

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis

Total-Sample Standardized Class Means

Variable	ALTO	BAJO	
POLSEN	0.867261143	-0.544158459	P EN PPM
KOLSEN	0.356612587	-0.524879887	K EN PPM
PH11	0.606181469	-0.025758725	PH EN SOLUCION 1:1
CA	0.703190248	-0.202197895	CA EN PPM (SOLUCION)
CE	0.777234420	-0.502958481	COND. ELECT. EM dMHOS/CM
MG	-0.632221397	0.063603611	MG EN PPM (SOLUCION)
TOTAL	-1.042393176	0.451735552	ARENA EN %
A	0.613701807	-0.460986844	ARCILLA EN %

Variable MEDIO

POLSEN	-0.135799266	P EN PPM
KOLSEN	0.012988770	K EN PPM
PH11	-0.175966732	PH EN SOLUCION 1:1
CA	-0.164740413	CA EN PPM (SOLUCION)
CE	-0.118208388	COND. ELECT. EM dMHOS/CM
MG	0.175128451	MG EN PPM (SOLUCION)
TOTAL	0.209464112	ARENA EN %
A	-0.078742121	ARCILLA EN %

Pooled Within-Class Standardized Class Means

Variable	ALTO	BAJO	
POLSEN	0.963987547	-0.604848935	P EN PPM
KOLSEN	0.365674055	-0.538217001	K EN PPM
PH11	0.630587907	-0.026795838	PH EN SOLUCION 1:1
CA	0.743366706	-0.213750380	CA EN PPM (SOLUCION)
CE	0.843561552	-0.545879628	COND. ELECT. EM dMHOS/CM
MG	-0.660121654	0.066410471	MG EN PPM (SOLUCION)
TOTAL	-1.212660346	0.525523194	ARENA EN %
A	0.644641364	-0.484227331	ARCILLA EN %

Variable	MEDIO	
POLSEN	-0.150945079	P EN PPM
KOLSEN	0.013318812	K EN PPM
PH11	-0.183051609	PH EN SOLUCION 1:1
CA	-0.174152782	CA EN PPM (SOLUCION)
CE	-0.128295979	COND. ELECT. EM dMHOS/CM
MG	0.182856960	MG EN PPM (SOLUCION)
TOTAL	0.243678517	ARENA EN %
A	-0.082711876	ARCILLA EN %

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis	Pooled Covariance Matrix Information
Covariance Matrix Rank	Natural Log of the Determinant of the Covariance Matrix
8	10.1545058

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis	Pairwise Squared Distances Between Groups
-----------------------	---

$$D(i|j) = (\bar{X}_i - \bar{X}_j)' \text{COV}^{-1} (\bar{X}_i - \bar{X}_j)$$

Squared Distance to CATEG			
From CATEG	ALTO	BAJO	MEDIO
ALTO	0	15.97342	12.48501
BAJO	15.97342	0	1.47998
MEDIO	12.48501	1.47998	0

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis

Pairwise Generalized Squared Distances Between Groups

$$D^2(i|j) = (\bar{X}_i - \bar{X}_j)' \text{COV}^{-1} (\bar{X}_i - \bar{X}_j) - 2 \ln \text{PRIOR}_j$$

Generalized Squared Distance to CATEG

From CATEG	ALTO	BAJO	MEDIO
ALTO	3.25661	19.77390	13.33367
BAJO	19.23003	3.80048	2.32864
MEDIO	15.74162	5.28046	0.84867

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis

Univariate Test Statistics

F Statistics, Num DF= 2 Den DF= 104

Variable	Total STD	Pooled STD	Between STD	R-Squared	RSQ/(1-RSQ)
POLSEN	6.8742	6.1844	3.8022	0.205883	0.2593
KOLSEN	51.3435	50.0712	16.1873	0.066891	0.0717
PH11	1.0046	0.9657	0.3741	0.093346	0.1030
CA	12.5762	11.8965	5.3559	0.122056	0.1390
CE	0.1419	0.1307	0.0707	0.167090	0.2006
MG	0.5556	0.5321	0.2142	0.100051	0.1112
TOTAL	12.4912	10.7373	7.9856	0.275043	0.3794
A	7.1995	6.8539	2.9211	0.110787	0.1246

Univariate Test Statistics

Variable	F	Pr > F	Label
POLSEN	13.4815	0.0001	P EN PPM
KOLSEN	3.7277	0.0273	K EN PPM
PH11	5.3538	0.0061	PH EN SOLUCION 1:1
CA	7.2293	0.0011	CA EN PPM (SOLUCION)
CE	10.4317	0.0001	COND. ELECT. EM dmHOS/CM
MG	5.7811	0.0042	MG EN PPM (SOLUCION)
TOTAL	19.7284	0.0001	ARENA EN %
A	6.4787	0.0022	ARCILLA EN %

Average R-Squared: Unweighted = 0.1426433
Weighted by Variance = 0.0833078

Multivariate Statistics and F Approximations

S=2 M=2.5 N=47.5

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.27506226	10.9939	16	194	0.0001
Pillai's Trace	0.82259112	8.5584	16	196	0.0001
Hotelling-Lawley Trace	2.28051776	13.6831	16	192	0.0001

Roy's Greatest Root 2.11245614 25.8776 8 98 0.0001

NOTE: F Statistic for Roy's Greatest Root is an upper bound.
 NOTE: F Statistic for Wilks' Lambda is exact.

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
 ESTADO YARACUY
 USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis Linear Discriminant Function

$$\text{Constant} = -0.5 \sum_j \bar{X}'_j \text{COV}^{-1}_j \bar{X}_j + \ln \text{PRIOR}_j \quad \text{Coefficient Vector} = \text{COV}^{-1}_j \bar{X}_j$$

CATEG

	ALTO	BAJO	MEDIO	Label
CONSTANT	-497.30866	-493.72696	-484.54230	
POLSEN	-2.30263	-2.85577	-2.66194	P EN PPM
KOLSEN	-0.48190	-0.40003	-0.38491	K EN PPM
PH11	60.00629	62.78269	59.45386	PH EN SOLUCION 1:1
CA	-0.73997	-1.03721	-0.68133	CA EN PPM (SOLUCION)
CE	-134.14306	-147.92469	-148.27100	COND. ELECT. EM dmHOS/CM
MG	-73.36860	-64.92809	-64.54354	MG EN PPM (SOLUCION)
TOTAL	7.51376	7.52997	7.66467	ARENA EN %
A	21.00973	20.15400	20.24524	ARCILLA EN %

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
 ESTADO YARACUY
 USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis

Classification Results for Calibration Data: WORK.CORN

Resubstitution Results using Linear Discriminant Function

Generalized Squared Distance Function:

$$D(X) = \sum_j (X - \bar{X}_j)' \text{COV}^{-1}_j (X - \bar{X}_j) - 2 \ln \text{PRIOR}_j$$

Posterior Probability of Membership in each CATEG:

$$\text{Pr}(j|X) = \frac{\exp(-0.5 D(X))}{\sum_k \exp(-0.5 D(X))}$$

Posterior Probability of Membership in CATEG:

Obs	From CATEG	Classified into CATEG	ALTO	BAJO	MEDIO
5	BAJO	MEDIO *	0.0001	0.0867	0.9131
7	BAJO	MEDIO *	0.0004	0.2761	0.7234
12	BAJO	MEDIO *	0.0004	0.2761	0.7234
13	BAJO	MEDIO *	0.0009	0.2855	0.7136

17	BAJO	MEDIO *	0.0009	0.2855	0.7136
18	BAJO	MEDIO *	0.0009	0.2855	0.7136
21	ALTO	MEDIO *	0.0434	0.0660	0.8906
26	ALTO	MEDIO *	0.0001	0.0608	0.9391
31	MEDIO	ALTO *	0.9951	0.0001	0.0049
33	MEDIO	ALTO *	0.9951	0.0001	0.0049
38	MEDIO	ALTO *	0.9898	0.0007	0.0095
49	BAJO	MEDIO *	0.0002	0.0514	0.9484
55	MEDIO	BAJO *	0.0000	0.8256	0.1743
57	MEDIO	BAJO *	0.0000	0.8256	0.1743
59	MEDIO	BAJO *	0.0000	0.8256	0.1743
60	BAJO	MEDIO *	0.0000	0.1482	0.8517
64	BAJO	MEDIO *	0.0000	0.1482	0.8517
67	BAJO	MEDIO *	0.0001	0.4653	0.5347
68	BAJO	MEDIO *	0.0001	0.4653	0.5347
69	BAJO	MEDIO *	0.0001	0.4653	0.5347
96	MEDIO	ALTO *	0.9990	0.0001	0.0010
104	BAJO	MEDIO *	0.0000	0.0702	0.9297

* Misclassified observation

ANALISIS DISCRIMINANTE 90-60-50 EN MAIZ
ESTADO YARACUY
USING LINEAR DISCRIMINANT FUNTION

Discriminant Analysis

Classification Summary for Calibration Data: WORK.CORN

Resubstitution Summary using Linear Discriminant Function

Generalized Squared Distance Function:

$$D_j(X) = \frac{1}{2} (X - \bar{X}_j)' \text{COV}_j^{-1} (X - \bar{X}_j) - 2 \ln \text{PRIOR}_j$$

Posterior Probability of Membership in each CATEG:

$$\text{Pr}(j|X) = \frac{\exp(-.5 D_j(X))}{\sum_k \exp(-.5 D_k(X))}$$

Number of Observations and Percent Classified into CATEG:

From CATEG	ALTO	BAJO	MEDIO	Total
ALTO	19 90.48	0 0.00	2 9.52	21 100.00
BAJO	0 0.00	3 18.75	13 81.25	16 100.00
MEDIO	4 5.71	3 4.29	63 90.00	70 100.00
Total	23	6	78	107
Percent	21.50	5.61	72.90	100.00

Priors	0.1963	0.1495	0.6542
--------	--------	--------	--------

Error Count Estimates for CATEG:

	ALTO	BAJO	MEDIO	Total
Rate	0.0952	0.8125	0.1000	0.2056
Priors	0.1963	0.1495	0.6542	

Anexo 10. Ejemplo de una salida del computador para el programa MPS versión 2.1, para realizar cálculos de programación lineal.