

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

U.A.C.P. y P. DEL C.C.H.

I.I.M.A.S.

ANALISIS DE UN EXPERIMENTO ECOLOGICO

T E S I S

QUE PARA OBTENER EL DIPLOMA DE

ESPECIALIZACION EN ESTADISTICA APLICADA

PRESENTA

Hugo Riemann González

BIBLIOTECA
JUAN A. ESCALANTE H.
UNIDAD ACADÉMICA DE
LOS CICLOS PROFESIONAL
Y DE POSGRADO / ECH
UNAM

Junio, 1985



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

Resumen	4
Antecedentes	5
I Introducción	6
Descripción del Experimento	10
II Análisis	13
Análisis de Varianza	14
Análisis de Varianza Multivariado	23
Análisis de Conglomerados	26
Análisis Discriminante	30
III Resultados e Interpretación	33
IV Conclusión y Recomendaciones	66
V Bibliografía	69

RESUMEN

Se analiza un experimento Ecológico que presenta un diseño de bloques al azar en parcelas divididas, cuyo objetivo es evaluar la Adecuación de plántulas de especies arbóreas. Los factores de clasificación involucrados fueron: Luz, Densidad y Defoliación . Como indicadores de Adecuación se midieron las variables de respuesta, Sobrevivencia, Crecimiento y Ganancia de Hojas.

Se recurre al uso de las metodologías de Análisis de Varianza Univariado y Multivariado, para probar significancia de de las interacciones de interés. Para establecer diferencias entre las medias de las interacciones se recurre en el caso univariado a la modificación hecha por Sen a la prueba de Tukey y en el caso multivariado al análisis de variables canónicas .

Para definir grupos de Adecuación se recurre al análisis de conglomerados y para establecer cuales de las variables de respuesta están definiendo a estos grupos se hace uso del análisis discriminante.

ANTECEDENTES

Los datos para el presente estudio fueron proporcionados por el Dr. Rodolfo Dirzo, perteneciente al grupo de Ecología de Poblaciones del Instituto de Biología de la UNAM.

Los datos fueron recabados en la Estación de Biología Tropical "Los Tuxtlas", Veracruz, perteneciente a la UNAM durante los años 1981 - 1982 por la pasante en Ciencias Biomédicas Beatriz Zagorín Hass, a la memoria de quien se dedica este trabajo.

INTRODUCCION

En una ciencia como la Biología donde la variabilidad es una característica inherente a sus objetos de estudio, aun dentro de categorías bien definidas (llámese a estas categorías taxón, población, niveles de complejidad, etc.) las aproximaciones a la realidad sólo pueden ser llevadas a cabo por medio de modelos que contengan esta variabilidad en forma implícita, esto es, a través de modelos estocásticos.

La descripción de fenómenos aleatorios por medio de modelos estocásticos permite probar hipótesis estadísticas, entendidas estas como aquellas que se formulan con referencia a una población y que pueden especificar que una característica en particular (por ejemplo media o varianza) tendrá un valor dentro de un intervalo dado (Dolby, 1982). Una hipótesis estadística será aquella que esté enmarcada dentro del paradigma de la probabilidad y la estadística, a diferencia de una hipótesis científica que tendrá un carácter global y que no necesariamente involucra los conceptos de población y muestra. Por lo tanto una hipótesis estadística referida a una población tiene un carácter local.

El alcance de la inferencia estadística estará determinado por dos factores. La población conceptual o real sobre la que se hace la inferencia y el grado de generalización.

dad de los datos sobre los que esté basada ésta cuyos extremos serían aquellos constituidos por un solo experimento y los que están basados en una amplia gama de situaciones, esto último sólo le confiere un amplio grado de generalidad y no un carácter global a la hipótesis estadística, ya que, por conjetura global entendemos aquella que tiene un amplio significado explicativo dentro de la ciencia que le concierne. ¿Por qué son las conjeturas globales tan raras?, se pregunta Dolby (Opus cit.), a lo que responde con dos argumentos. El primero supone que el investigador tiene la creencia errónea de estar creando y formulando hipótesis globales cuando en realidad está probando hipótesis estadísticas locales, esto sólo se puede explicar a través de la ignorancia del investigador de los paradigmas tanto de su propia ciencia de interés como de la estadística. El segundo argumento es que existe una tendencia en la consultoría estadística a ser el consultado y no el consultante el que formula las conjeturas; de ser cierto este argumento, esto redundará en una proliferación de hipótesis estadísticas locales más que en imaginativas conjeturas globales (Opus. cit.). Un equilibrio en este punto sólo se podría lograr dando al investigador una formación estadística más sólida de tal manera que sea menos dependiente del asesor estadístico.

A pesar de que hace más de cuatro décadas que fueron desarrolladas las principales técnicas paramétricas de

análisis estadístico y que actualmente se cuenta en los centros de investigación y docencia con amplias facilidades de cómputo, su estudio está minimizado, cuando no ausente, en los programas de docencia de las escuelas de Biología del país, donde aún perdura una actitud de tipo descriptivo, la cual, si bien no debe ser menospreciada, debe ser precedida en lo posible por un criterio analítico por medio del cual se haga uso en forma conjunta tanto de las técnicas de análisis biológicas como estadísticas.

Es común que la actitud hacia la Estadística por parte de investigadores de la Biología caiga en dos extremos igualmente reprobables; aquellos que muestran una total reticencia hacia su uso y aquellos que suponen que su simple empleo da carácter de validez a cualquier investigación. El uso de procedimientos de análisis estadísticos no justifica una investigación, independientemente de que en esta se haya hecho o no uso de la Estadística en forma adecuada.

La validez de una investigación debe estar sustentada no sólo en criterios de tipo ético, también debe estar conceptualmente bien fundamentada de tal forma que las hipótesis a probar estén basadas en conceptos propios de la ciencia que le concierne.

El experimento que aquí se ha analizado forma parte de un programa de investigación que pretende evaluar el

impacto de la herbivoría en la dinámica poblacional de la vegetación .Los objetivos de este trabajo son establecer el efecto conjunto de los factores Luz, Densidad y Defoliación en la adecuación de plántulas de cuatro especies arbóreas, así como, reconocer grupos de adecuación respecto a esta combinación de factores .

Si bien el experimento fue planeado y llevado a cabo con todo rigor, este, no fue diseñado de antemano. Al desarrollar el modelo que describe el experimento se encontró que este tenía restricciones para el análisis en varios factores de interés para el investigador.

Dados los objetivos de este trabajo se han minimizado las interpretaciones y conclusiones biológicas, centrandó el interés en el aspecto estadístico del estudio.

DESCRIPCION DEL EXPERIMENTO

La investigación que aquí nos ocupa es el resultado de un experimento llevado a cabo con el fin de determinar el efecto de la luz, la defoliación y la densidad, sobre la adecuación en estado de plántula en individuos de cuatro especies arbóreas de la selva alta perennifolia del Estado de Veracruz; las especies utilizadas en el experimento fueron: *Omphalea cardiophylla*, *Nectandra ambigens*, *Psychotria chiapensis* y *Brosimum alicastrum*.

La adecuación de un individuo a su medio ambiente se determina por su capacidad de sobrevivencia y su éxito reproductivo, sin embargo, para el caso de las plántulas se midieron como variables de respuesta tres indicadores: crecimiento, sobrevivencia y ganancia de hojas (ver Tabla 1)

Para el desarrollo del experimento se utilizaron dos viveros, uno de ellos ubicado dentro de la selva y otro en una zona sin cobertura vegetal, (ver Figura 11.)¹¹ estos fueron considerados los tratamientos de sol y sombra. Dentro de cada vivero se dispuso de cuatro mesas destinadas cada una a una especie. Las plántulas fueron sembradas en charolas a dos densidades que designaremos como alta y baja. Las plántulas de cada charola se defoliaron en forma artificial a cuatro intensidades: cero (control), cinco, veinticinco y setenta y cinco por ciento del área foliar con el objeto de simular el daño por herbivoría de insectos. El

+Al presente la ubicación taxonómica está en duda con *Omphalea oleiphera* .

Variabes experimentales

o
como

Factores de clasificación

Categorías:

Indicadores de adecuación

Unidades

Indicadores de adecuación	Unidades	Factores de clasificación	Categorías:
Crecimiento	cm.	Luz	Sol Sombra
Sobrevivencia	%	Defoliación	0% de área foliar 5% de área foliar 25% de área foliar 75% de área foliar
Ganancia de hojas	%	Densidad	Alta Baja
		Especie	<i>O. cardiophylla</i> <i>N. ambigens</i> <i>B. alicastrum</i> <i>P. chiapensis</i>
		Tiempo	1 2 3 4

Tabla 1

Se representan aquí, las variables de respuesta medidas durante el experimento y los factores de clasificación con sus respectivos niveles

número de repeticiones de cada combinación de tratamientos varió de cinco a ocho. Las variables de respuesta fueron medidas en cuatro ocasiones para tres de las especies y só lo medidas en tres ocasiones para *B. alicastrum*.

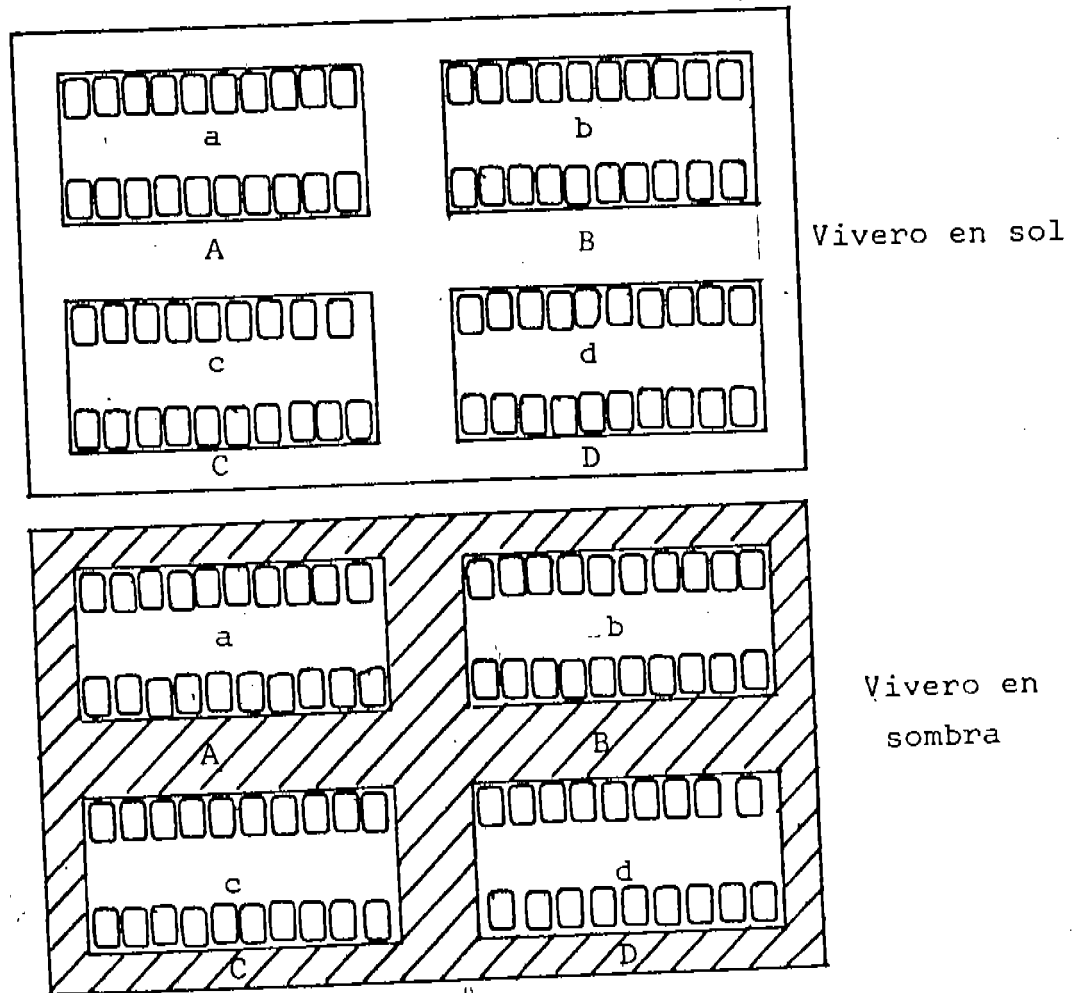


Figura No. 1

Representación del arreglo del experimento en el campo.

A, B, C y D representan las mesas donde fue hecho el experimento para cada una de las especies; a, b, c y d, representan las charolas donde se sembraron las plantas sometidas a las diferentes combinaciones de tratamientos.

ANALISIS

Para el análisis de los datos se siguieron dos vías independientes .

1) Analisis de varianza univariado para cada una de las tres variables de respuesta, haciendo uso de la modificación hecha por Sen (1969), a la prueba de Tukey para probar efectos de interacción. Se continuó con un análisis de varianza multivariado usando variables canónicas para prueba de medias.

2) En esta segunda vía se recurrió a procesar los datos mediante análisis de conglomerados con el objeto de definir grupos de adecuación . Se usaron dos procedimientos de análisis de grupos; jerárquico y de partición óptima. En forma subsecuente se hizo un análisis discriminante para cada uno de los dos procedimientos de agrupación, con el objeto de establecer cuales de las variables dependientes están definiendo los grupos de adecuación.

A continuación se describen en forma intuitiva y haciendo caso omiso de los procedimientos matriciales las diferentes técnicas empleadas en el análisis.

ANALISIS DE VARIANZA

Si tenemos una muestra de individuos sometidos a diferentes combinaciones de factores de interés, la primera pregunta que surge es ¿ existe una diferencia significativa en la variable de respuesta de las diferentes combinaciones de tratamientos o los resultados sólo están reflejando la variabilidad aleatoria inherente a la población ?

El análisis de varianza parte del supuesto que el experimento toma muestras de poblaciones infinitas. Para que estas poblaciones sean comparables deben tener un mismo grado de generalidad, que está determinado por todos aquellos factores constantes comunes a las poblaciones a comparar.

Se ha comprobado empíricamente que poblaciones con un mismo grado de generalidad tienen varianzas iguales (Méndez 1976).

Si tenemos una población definida por A, B, C, D, constantes y por E, F, G, H, no constantes y la variable de respuesta Y tiene una distribución normal, la podemos representar por:

$$Y \sim N (\mu (A, B, C, D,) \sigma^2 (E, F, G, H,))$$

Si queremos comparar las poblaciones

$$Y_{1i} \sim N (\mu (A, B, C, D_1) \sigma^2 (E, F, G, H))$$

$$Y_{2j} \sim N (\mu (A, B, C, D_2) \sigma^2 (E, F, G, H))$$

representadas por el modelo:

$$Y_{ij} = \mu (A, B, C, D_i) + \epsilon_{ij}; \text{ con } \epsilon_{ij} \sim N(0, \sigma^2) \quad (1)$$

Estamos comparando los valores de $\mu (A, B, C, D_i)$ para $D_i = 1, 2$ donde se considera que los factores A, B, C, afectan por igual a los D_i poblaciones. Por lo que el modelo lo podemos escribir como:

$$Y_{ij} = \mu (A, B, C) + \mu (D_i) + \epsilon_{ij} \quad (2)$$

si consideramos a $\mu (D_i)$ como los i tratamientos el modelo anterior se representa por:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \text{ donde: } \epsilon_{ij} \sim N(0, \sigma^2) \quad (3)$$

$\mu =$ efecto de la media total

$\tau_i =$ efecto del i 'esimo tratamiento, $i = 1, 2$

y podemos expresar las comparaciones por:

$$H_0 : \tau_1 = \tau_2$$

Vs.

$$H_1 : \tau_1 \neq \tau_2$$

donde H_0 y H_1 representan la hipótesis nula y alternativa respectivamente para un nivel de significancia α dado, el cual representa la probabilidad de rechazar la hipótesis nula cuando esta es cierta.

Los procedimientos para la obtención de suma de cuadrados, grados de libertad y esperanzas de cuadrados medios necesarios para probar la significancia de tratamientos no se comentan aquí, recomendamos al lector interesado los textos de Montgomery (1976), Méndez (1981) y Anderson and McLean (1974).

Cuando se tienen grupos de unidades experimentales, de tal forma que cada grupo representa el nivel de uno de los factores a experimentar, cada uno de los conjuntos representa un bloque, y en este caso el efecto del bloque está confundido con el nivel del factor que representa, lo que significa que las unidades experimentales dentro de cada bloque están sujetas simultáneamente al efecto del nivel del factor a experimentar y a las condiciones que el propio bloque les imparte, por lo que no es posible probar el factor que representa el bloque, ya que, una diferencia significativa en bloques (niveles del factor que representan) no sería posible asociarla al nivel del factor o al efecto del bloque.

Esta restricción en el análisis se puede evitar al aleatorizar las u. e. de tal forma que no existan grupos de u. e. asociados a niveles de algún factor y que representen a este en forma única.

La introducción de bloques en el diseño representa como vemos una restricción en la aleatorización, y esto se representa en el modelo como un error de restricción (Anderson, 1970; Anderson and McLean, 1974; Méndez, 1981a; Méndez, 1983).

Dentro de los bloques se puede tener de nuevo restricciones en la aleatorización asociando grupos de u. e. más pequeños con otros factores, de esta forma se tiene un diseño en parcelas divididas. Si dentro de estas parcelas se tienen de nuevo restricciones, se tendrá un diseño en parcelas subdivididas y así sucesivamente.

El error de restricción no es estimable, ya que carece de grados de libertad. El objeto de incluirlo en el modelo es hacer patente que efectos es posible probar mediante pruebas de F.

Por lo que respecta a la definición del modelo del experimento es menester aclarar que los cuatro períodos de defoliación a que fueron sometidas las plántulas no estuvieron asociados con los cambios estacionales, ya que, el experimento no se corrió un año completo. Por lo que no es posible asociar cambios de tipo climático como temperatura y humedad ambiental con cada período de defoliación, asociado a esto, existe el hecho de que las plántulas fueron regadas en forma artificial y constante, no reflejándose en la adecuación el efecto de la precipitación natural.

Todo organismo pasa por cambios cíclicos anuales que son regulados por factores bióticos y abióticos del medio que los rodea. Los organismos del trópico cálido húmedo no escapan a este fenómeno aun en condiciones de relativa estabilidad ambiental como son los ecosistemas tropicales. Si cada uno de los períodos de de-

foliación y toma de datos se hubiesen llevado a cabo en fechas equidistantes y a lo largo del año, por ejemplo, durante los solsticios y equinoccios, hubiese sido posible medir la adecuación de las plántulas en función del factor Tiempo, donde este representaría tanto las diferentes condiciones abióticas (cambios en el clima y en la cantidad y calidad lumínicas), como bióticas, donde estos representan todos aquellos factores biológicos de tipo endógeno y cíclico que pudieran tener un efecto diferencial sobre la variable de respuesta.

El hecho de que el experimento no se haya seguido a lo largo de un año con el objeto de tener en cuenta estas variables hace innecesario llevar a cabo pruebas de significancia sobre el factor Tiempo, además de que su inclusión en el modelo no sólo complicaría excesivamente el análisis de los datos sino que podría llevar a conclusiones erróneas.

El modelo que representa este experimento es el siguiente :

$$\begin{aligned}
 Y_{ijklmqr} &= \mu + L_i + \delta_{1(i)} + E_k + \eta_q(ilk) + D_j + (LD)_{ij} \\
 &+ (DE)_{jk} + (LDE)_{ijk} + F_m + (FL)_{im} + (FE)_{km} + (FD)_{jm} \\
 &+ (LEDF)_{ijfm} + (LEF)_{ikm} + (LFD)_{imj} + (EDF)_{jkm} + \epsilon_{(ijklmq)r}
 \end{aligned}$$

donde:

μ = Media

L_i = Efecto del factor luz, $i = 1, 2$, fijo

$\delta_{1(i)}$ = Error de restricción asociado al factor luz.

$1 = 0$, aleatorio.

E_k = Efecto de especie, $k = 1, \dots, 4$ fijo

$\eta_{q(ilk)}$ = Error de restricción asociado a especie.

$q = 0$, aleatorio.

D_j = Efecto de densidad $j = 1, 2$, fijo

$(LE)_{ik}$ = Efecto de interacción luz*especie.

$(LD)_{ij}$ = Efecto de interacción luz*densidad.

$(DE)_{jk}$ = Efecto de interacción densidad*especie.

$(LDE)_{ijk}$ = Efecto de interacción luz*densidad*especie.

F_m = Efecto del factor defoliación $m = 1, \dots, 4$, fijo

$(FL)_{im}$ = Efecto de interacción defoliación*luz.

$(FE)_{km}$ = Efecto de interacción defoliación*especie.

$(FD)_{jm}$ = Efecto de interacción defoliación*densidad.

$(LEDF)_{ijkm}$ = Efecto de interacción defoliación*especie*
luz*densidad.

$(LEF)_{ikm}$ = Efecto de interacción luz*especie*defoliación.

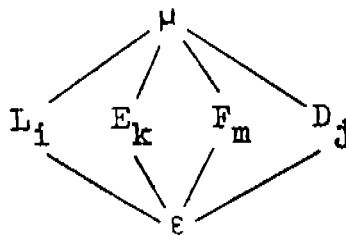
$(LFD)_{imj}$ = Efecto de interacción luz*defoliación*densidad.

$(EDF)_{ikm}$ = Efecto de interacción especie*densidad*defoliación

$\epsilon_{(ijklmq)r}$ = Error experimental.

Este modelo se define como un diseño en bloques al azar en parcelas divididas, donde los bloques están representados por los viveros, la parcela grande por la especie y la parcela chica por la charola.

La estructura factorial que representa este diseño es la siguiente :



Se ha hecho hincapié en que este diseño no es ortogonal, por lo que no es posible analizarlo a través del procedimiento ordinario, el proceso usual es pasar el modelo de ANDEVA a un modelo de regresión y resolverlo por mínimos cuadrados .

Para resolver un ANDEVA a través de regresión es necesario reparametrizar el modelo y esto se logra construyendo una matriz de variables de engaño que especifique los tratamientos y sus interacciones (matriz diseño) "X" tal que, $X'X$ tenga inversa única y de esta manera nos queda el modelo :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_h X_h + \varepsilon$$

el arreglo matricial es el siguiente:

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$$

donde \underline{Y} y $\underline{\varepsilon}$ son vectores aleatorios de $k \times 1$, X es una matriz de dimensiones $k \times h$ de cantidades conocidas y $\underline{\beta}$ es un vector de cantidades desconocidas de $h \times 1$ estimable a partir de la siguiente ecuación:

$$\underline{\beta} = (X'X)^{-1}X'\underline{Y}$$

La hipótesis nula queda representada por:

$$H_0 : \underline{\beta} = 0$$

La hipótesis alternativa es aquella en donde al menos uno de los elementos del vector $\underline{\beta}$ es diferente de cero .

Como el interés del investigador está centrado en analizar el efecto simultáneo de los factores involucrados en el experimento se decidió interpretar interacciones de segundo orden en las que estuviese implicada la defoliación .

Para la comparación de medias se usó la prueba de Tukey modificada" por Sen, según recomienda Vargas Chanes (1983) .

Dos medias se consideran diferentes para un nivel de significancia α , si para $m_1 > m_2$

$$m_1 - m_2 < Q^{\alpha}_{\gamma, v} \sqrt{\text{CME} (\gamma/t)}$$

donde:

γ = Producto de los niveles de los factores involucrados en el experimento.

$$t = r (\gamma - 1) + \lambda$$

CME = Cuadrado medio del error

r = Número de repeticiones de cada combinación de tratamiento.

λ = Número de bloques con igual número de tratamientos.

$Q^{\alpha}_{\gamma, v}$ = Cuantil de distribución de rango múltiple de la prueba de Duncan

$$v = b(\gamma - 1) - \gamma + 1$$

b = Número de bloques

ANÁLISIS DE VARIANZA MULTIVARIADO

Podemos considerar el análisis de varianza multivariado (MANOVA) como una extensión del caso univariado donde la variable de respuesta es un vector de dos o más elementos.

Una práctica común en la resolución de este problema es analizar en forma independiente cada una de las respuestas, pero ello trae como consecuencia que se ignoren las posibles correlaciones entre las variables dependientes .

El objeto primordial de MANOVA es determinar si existe una diferencia significativa entre los centroides de las diferentes combinaciones de tratamientos .

En el análisis de varianza univariado la prueba de significancia se lleva a cabo por medio del cociente de dos varianzas. En el caso multivariado se han desarrollado numerosos criterios de prueba, la mayoría de ellos basados en escalares como la traza o el determinante de las matrices de dispersión . Aquí sólo presentamos el criterio de máxima verosimilitud o lambda de Wilks.

$$\Lambda = \frac{|W|}{|T|}$$

donde $|W|$ y $|T|$,

representan los determinantes de las matrices de la suma de los cuadrados y productos cruzados dentro de tratamientos y Total respectivamente. El criterio de prueba es rechazar H_0 con valores pequeños de λ .

Una vez que hemos rechazado la hipótesis nula, nos enfrentamos al problema de determinar cual o cuales de las combinaciones de tratamientos son significativamente diferentes para lo cual podemos seguir los dos caminos existentes para ANDEVA: contrastes y comparaciones múltiples. Morrison (1976) recomienda la prueba de Roy, ya que, esta permite establecer intervalos de confianza tanto entre las combinaciones de tratamientos como entre las variables de respuesta. Pero esta prueba involucra el manejo de la matriz de diseño que es demasiado grande en el caso que nos ocupa, por lo que se optó por analizar los tratamientos a través de variables canónicas (Mardia, et al. 1979).

Si consideramos a μ_i como las coordenadas de un punto de p dimensiones donde $i = 1, \dots, k$. Cuando la hipótesis nula H_0 de tratamientos es cierta los vectores μ_1, \dots, μ_k son iguales por tanto, si r es la dimensión del hiperplano demarcado por μ_1, \dots, μ_k entonces, H_0 es equivalente a $r = 0$.

Las coordenadas canónicas de μ_i están definidas por

$$(q'_1 \bar{X}_i, \dots, q'_r \bar{X}_i) = \mu_i$$

donde: $i = 1, \dots, k$ definen las k medias canónicas

\bar{X}_i = vector de medias para las i 'ésimas combinaciones de tratamientos

$q_j = j$ 'ésimo eigenvector de la matriz $\Sigma^{-1} B$

para $\Sigma =$ matriz de varianza del error

$B =$ matriz de varianza entre tratamientos

Las variables canónicas están definidas así por la función lineal

$$Y_j = q_j'X$$

En las Gráficas 1 y 2 y Tablas 9 y 12 de variables canónicas se representan las K medias canónicas de las diferentes combinaciones, de tratamientos donde el radio de los círculos está definido en función del tamaño de muestra por:

$$\text{radio} = \frac{\chi^2_{r,\alpha}}{N_i}$$

donde: $r =$ número de variables canónicas a usar.

$N_i =$ número de observaciones

La intersección entre círculos se interpreta como similitud de tratamientos.

ANÁLISIS DE CONGLOMERADOS

Los conceptos de clasificación a identificación han estado estrechamente ligados al interés del hombre por conocer el universo que lo rodea, tales conceptos permiten la agrupación y reconocimiento de fenómenos, objetos, individuos, etc., que se distinguen por presentar características comunes, por lo tanto, es natural que estos hayan estado presentes en la Biología desde sus más remotos inicios, siendo en la taxonomía y la sistemática en donde han cristalizado en forma más evidente.

Si bien es frecuente que estos dos conceptos se confundan son **diametralmente** opuestos tanto en objetivos como en método, ya que, al clasificar procedemos en una forma inductiva separando en grupos a los objetos de interés, partiendo de sus características individuales hacia las más generales. Este es el concepto más estrechamente asociado a las técnicas conocidas como taxonomía numérica, análisis de conglomerados, agrupamiento, **reconocimiento de patrón**, tipología, etc.

El concepto de identificación es por lo contrario un procedimiento deductivo que tiene por objetivo asignar a alguno de los grupos previamente establecidos, el individuo de interés partiendo de características generales ha-

cia las individuales, este concepto se asocia más a los procedimientos de análisis discriminante que se tratan más adelante.

Las técnicas de agrupación pueden separarse en dos grupos:

- a) Técnicas de agrupamiento jerárquico
- b) Técnicas no jerárquicas o de partición óptima.

Las primeras a su vez se pueden separar en dos métodos: el agregativo, (aglomerativo), que parte del supuesto de considerar a cada caso como un grupo y va haciendo fusiones de estos formando cada vez menor número de grupos con mayor número de casos y el método divisivo que parte de considerar a todos los individuos como un solo grupo y va haciendo divisiones cada vez más finas; el procedimiento divisivo suele ser muy costoso a no ser que N sea muy pequeña ($N = 16$), ya que, hay $2^{n-1} - 1$ formas de hacer la primera división (Everitt, 1977), por lo que sólo haremos referencia al procedimiento agregativo. Las técnicas de tipo jerárquico forman g grupos donde $g = 1 \dots N$, y donde g y $g + 1$ tienen en común a $g - 1$.

Tanto las técnicas de partición óptima como las jerárquicas están basadas en el concepto de distancia o disimilitud entre los objetos a clasificar. A la fecha se han desarrollado numerosas medidas de distancia (Mardia et al., 1979).

El método jerárquico usado en este trabajo está basado en el procedimiento de Ward (1963). Este procedimiento se fundamenta en la idea de ir agrupando a los individuos de tal forma que la pérdida de información sea mínima, para lo cual se mide la suma de cuadrados total de las desviaciones de cada uno de los individuos al centro del grupo al que pertenece. En cada iteración del análisis se unen los conglomerados que minimicen el incremento de pérdida de información (minimizan la suma de cuadrados debida al error). En el estado inicial del análisis cada caso forma un grupo y la suma de cuadrados debida al error es igual a cero donde:

$$SCE = \sum_{i=1}^n X_i^2 - \frac{1}{n} \left(\sum_{i=1}^n X_i \right)^2$$

donde X_i es el valor del i 'ésimo individuo.

Las técnicas jerárquicas tienden a construir agrupaciones "naturales" que se representan en forma de diagrama de árbol o dendrograma.

Los métodos de partición óptima difieren de los jerárquicos en que permiten la reubicación de los casos y de esta manera optimizan en particiones sucesivas la ubicación de los casos formando grupos dentro de los cuales esté maximizada la similitud y entre los cuales se maximiza la disimilitud. Al igual que en caso de los métodos jerárquicos se han desarrollado numerosos procedimientos, aquí sólo comentaremos al algoritmo basado en el procedimiento del "Lider de Hartigan" (1975) y el de "k medidas de MacQueen" (1963). Se seleccionan un grupo de casos llamados puntos semilla, cada caso se asigna a la semilla más cercana para formar conglomerados temporales. Las semillas son reemplazadas por los centroides de los conglomerados temporales, el proceso se repite hasta que no haya cambios en los conglomerados.

Este procedimiento está basado en la distancia Euclidiana.

$$d_{ij}^2 = \sum_{k=1}^p (X_{ik} - X_{jk})^2$$

donde:

X_{ik} = i'ésima observación de la k'ésima columna de una matriz de datos X (n x p)

X_{jk} = j'ésima observación de la k'ésima columna de la matriz X.

ANÁLISIS DISCRIMINANTE

La consecuencia lógica de usar dos técnicas de análisis de grupos es evaluar cual de las dos es más eficiente y cuales son las variables que están jugando un papel más importante en la formación de los diferentes grupos en este caso de adecuación. Por lo cual se decidió aplicar un análisis discriminante a cada una de las dos técnicas de agrupación empleadas.

El análisis discriminante se emplea en aquellas situaciones en las que se tienen dos grupos de variables uno de ellos definido por variables de tipo categórico y otro por variables numéricas, entre las que existe una relación de dependencia.

Los objetivos de este tipo de análisis son:

- 1) Determinar si existen diferencias estadísticamente significativas entre los centroides de los dos o más grupos previamente establecidos por algún procedimiento de agrupación (dos de los cuales se presentaron en el punto anterior).

- 2) Establecer un procedimiento por el cual se puedan asignar individuos a los grupos previamente definidos.

3) Determinar cual o cuales de las variables independientes tienen un mayor poder discriminante.

Si se tienen dos poblaciones y dos variables, puede ocurrir que determinados valores puedan presentarse con alta probabilidad en las dos poblaciones.

En la Figura 2 se presenta en forma esquemática este problema, siendo A y B dos poblaciones, los puntos donde las elipses se intersectan definen una línea recta. Si sobre esta se traza una recta perpendicular D los casos de ambas poblaciones proyectados sobre esta línea llamada función discriminante tenderán a agruparse alrededor de su media y a maximizar la distancia entre sus centroides (se minimiza la varianza intra grupos y se maximiza entre grupos). Este concepto puede extenderse a tres o más poblaciones con tres o más variables, creándose según el caso, planos o hiperplanos discriminantes. Aunque en la Figura se han representado los supuestos del análisis discriminante, es decir, normalidad, homogeneidad de varianzas (igual tamaño de las elipses) e igualdad de covarianzas (misma orientación de los ejes de las elipses), en la práctica se ha visto que esta técnica no es muy sensible a desviaciones de los supuestos cuando estas no son muy marcadas (Hair et al., 1979).

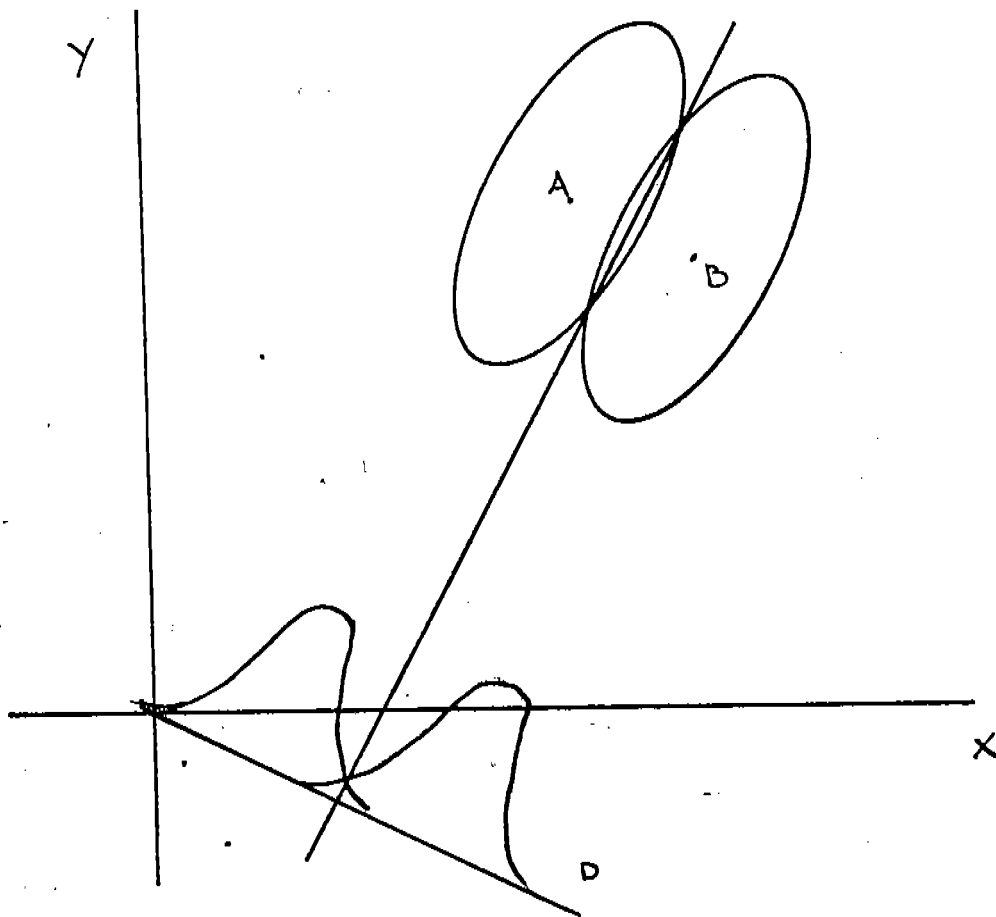


Figura 2

En esta figura se representan dos poblaciones (A y B) que proyectadas sobre los ejes X ó Y, presentan un grado de traslape tal que no permite establecer si existe diferencia significativa entre los centroides.

Al maximizar la varianza entre grupos es posible crear la función discriminante D, que permite establecer diferencias significativas entre los centroides y asignar nuevos casos a una u otra población.

RESULTADOS E INTERPRETACION

El proceso de los datos se llevó a cabo en el paquete estadístico SAS.

Análisis de Varianza Univariado

En la Tabla 2 se presentan los resultados del análisis de varianza para la variable Sobrevivencia, de este sólo se interpreta la interacción de segundo orden Luz Especie Defoliación, recuérdese que el factor Luz tiene un error de restricción asociado así como el factor Especie que además carece de interpretación biológica.

En la Tabla #3 se presentan los valores de las medias para las combinaciones de tratamientos de la interacción de segundo orden Luz*Especie*Defoliación para la variable de respuesta sobrevivencia. Los tratamientos que son significativamente diferentes a una α de 0.05 son los que se encuentran a los extremos de la Tabla, que corresponden a la especie *Omphalea cardiophylla* a los mismos niveles de defoliación pero en condiciones extremas de luz. La especie *Nectandra ambigens* también difiere significativamente en su sobrevivencia en condiciones extremas de luz, pero sólo para defoliaciones altas. Para defoliaciones 0, y 5% no existe una diferencia signifi-

Tabla 2

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: MS	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR > F	TYPE III SS	F VALUE	PR > F	R-SQUARE	C.V.
SOURCE	63	17515.22771359	278.01943752	10.26	0.0001	1990.04551027	73.44	0.0001	0.702299	5.4210
MODEL	274	7424.51704561	27.0714971			1532.01679364	18.85	0.0001		MS MEAN
ERROR	357	10090.71076798	28.26822780			2440.75762478	5.42	0.0014		96.02413708
CORRECTED TOTAL	357	21939.84076220				102.093300496	1.26	0.2893		
SOURCE	DF	TYPE I SS	F VALUE	PR > F	DF	TYPE III SS	F VALUE	PR > F		
LUZ	1	1931.02828731	71.26	0.0001	1	1990.04551027	73.44	0.0001		
ESP	1	1504.95667530	53.33	0.0001	1	1532.01679364	18.85	0.0001		
LUZ*ESP	1	2121.69238091	75.42	0.0001	1	2440.75762478	5.42	0.0014		
LUZ*DEF	1	114.74571961	4.11	0.0418	1	102.093300496	1.26	0.2893		
ESP*DEF	1	477.31752226	16.53	0.0001	1	453.74827095	1.68	0.0955		
LUZ*ESP*DEF	1	677.10944166	23.79	0.0001	1	653.34827095	2.33	0.0121		
DEN	1	2593.09211908	93.46	0.0001	1	2718.95673852	43.00	0.0001		
LUZ*DEN	1	1689.79213899	59.72	0.0001	1	1752.84672289	21.36	0.0001		
ESP*DEN	1	3198.02865718	112.12	0.0001	1	3214.73635447	39.34	0.0001		
LUZ*ESP*DEN	1	598.18831950	21.36	0.0001	1	619.35674427	7.56	0.0001		
LUZ*DEF*DEN	1	244.09739945	8.77	0.0032	1	249.95695320	3.78	0.0528		
ESP*DEF*DEN	1	430.9739945	15.41	0.0001	1	439.95695320	6.06	0.0138		
LUZ*ESP*DEF*DEN	1	359.60422719	12.77	0.0001	1	359.60422719	4.77	0.0288		

Análisis de Varianza para Variable Supervivencia.

La interpretación de los ANDEVA se hizo con la suma de cuadrados tipo I. En este caso sólo se interpretó la interacción Luz*Especie*Defoliación que presentó un nivel de significancia de 0.0041.

Tabla 3

Comparaciones de las medias para la interacción Luz*Especie*Defoliación con variable de respuesta sobrevivencia

Luz	Esp.	Def.	sobrevivencia
1	1	1	100.0000
1	1	2	100.0000
1	1	3	100.0000
1	1	4	100.0000
1	2	2	100.0000
1	2	3	100.0000
1	4	1	100.0000
1	4	2	99.5772
1	2	1	99.3303
2	4	2	99.2225
1	2	4	99.0250
2	4	1	98.8000
2	4	3	98.2900
1	4	3	97.8150
2	2	1	97.4431
2	4	4	97.1975
1	4	4	97.1950
1	3	1	97.0000
2	2	2	96.8875
2	3	4	96.1666
2	3	1	95.8333
1	3	2	94.6666
1	3	3	94.3333
2	2	3	93.7500
2	3	3	93.6666
1	3	4	93.3333
2	2	4	92.6136
2	1	1	92.0454
2	3	2	88.8333
2	1	2	88.7625
2	1	3	87.5000
2	1	4	78.7500

Comparaciones múltiples para interacción de Segundo orden por medio de la modificación de Sen a la prueba de Tukey

Tratamientos.

Luz: 1 = sol

2 = sombra

Especies:

1 = *O. cardiophylla*

2 = *N. ambigens*

3 = *B. alicastrum*

4 = *P. chiapensis*

Defoliación:

1 = 0

2 = 5 por ciento

3 = 25 por ciento

4 = 75 por ciento

Los valores bajo la columna sobrevivencia representan las medias de las combinaciones de tratamientos.

Las líneas verticales unen a las medias que no son significativamente diferentes a una α de 0.05

cativa en la sobrevivencia en condiciones extremas de luz, por lo que podemos afirmar que existe un efecto sinérgico de los factores luz y defoliación sobre la sobrevivencia de las plántulas de estas dos especies, siendo el caso extremo, la especie *O. cardiophylla* para alta defoliación en condiciones de sombra, pues como se puede observar, este tratamiento difiere significativamente de los 31 restantes.

En la Tabla 4 se presentan los resultados del análisis de varianza para la variable Crecimiento, se interpreta de este la interacción de segundo orden Luz Especie Defoliación, que tuvo un nivel de significancia de 0.0001 .

En la Tabla 5 se presentan las comparaciones de medias para esta interacción . Existe aquí un efecto sinérgico de los factores Luz y Defoliación sobre el Crecimiento de las especies *O. cardiophylla* y *P. chiapensis*. Observese que el efecto simultáneo de de la defoliación y sombra producen un crecimiento significativamente menor en la especie O. cardiophylla respecto al tratamiento de sol en la misma especie . El mismo efecto se observa en la especie P. chiapensis .

En la Tabla 6 se presenta el resultado del análisis de varianza para la variable Ganancia de Hojas de este se interpretó la interacción de segundo orden Especie Defoliación Densidad .

Tabla 4

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: MCRE		SUM OF SQUARES		MEAN SQUARE	F VALUE	PR > F	TYPE III SS	R-SQUARE	C.V.
SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PR > F	TYPE III SS	R-SQUARE	C.V.	
MODEL	63	22994.10700268	364.98582544	49.41	0.0001	0.919095	48.7255		
ERROR	274	2624.09025488	7.38719071			ROOT MSE	5.57805720		
CORRECTED TOTAL	337	25018.19725756				2.71793869			
SOURCE	DF	TYPE I SS	F VALUE	PR > F	DF	TYPE III SS	F VALUE	PR > F	
LUZ	1	8706.47299964	907.85	0.0001	1	8554.83991123	887.33	0.0001	
ESP	1	9882.53691236	494.93	0.0001	1	9629.75525373	435.88	0.0001	
DEF	3	5309.42327110	239.58	0.0001	3	5254.61922461	277.02	0.0001	
LUZ*DEF	3	273.29325726	12.39	0.0001	3	261.86723011	11.82	0.0001	
ESP*DEF	3	67.07639136	1.02	0.4333	3	65.92202911	1.09	0.4333	
LUZ*ESP	3	26.42026037	0.66	0.6003	3	25.81963543	0.89	0.6003	
LUZ*DEF*ESP	1	0.47493913	0.02	0.8703	1	0.44210263	0.02	0.8703	
LUZ*DEF	1	2.36605533	0.32	0.5703	1	2.20339664	0.33	0.5703	
LUZ*ESP	1	0.61778284	0.08	0.8971	1	0.58302846	0.08	0.8971	
DEF*ESP	1	0.11923795	0.00	0.9304	1	0.10890264	0.00	0.9304	
LUZ*DEF*ESP	1	17.97424771	0.34	0.7936	1	17.72065899	0.35	0.7936	
LUZ*DEF*ESP*DEF	1	14.80586181	0.29	0.7337	1	14.70095899	0.29	0.7337	
LUZ*DEF*ESP*DEF*DEF	1	18.98170735	0.29	0.9775	1	18.98170735	0.29	0.9775	

Análisis de Varianza para Variable Crecimiento

De este análisis sólo se interpreta la interacción Luz*Especie*Defoliación que tuvo un nivel de significancia de 0.0001.

Comparaciones de medias para la interacción Luz*Especie*Defoliación con variable de respuesta Crecimiento

Luz	Esp	Def	Crecimiento
1	1	2	29.9525
1	1	1	28.4979
1	1	3	25.9191
1	1	4	19.9050
1	4	1	8.3488
1	4	2	6.9386
1	4	3	6.3200
1	2	1	6.0839
2	1	4	5.8350
1	2	2	5.3990
1	2	3	4.6750
1	3	2	3.7533
1	3	1	3.3133
2	1	3	3.0900
1	3	3	2.6733
2	1	2	2.6650
1	4	4	2.6400
2	1	1	2.4393
1	3	4	2.2500
1	2	4	1.8775
2	4	1	0.6853
2	4	2	0.5525
2	2	1	0.5310
2	2	2	0.4225
2	3	1	0.3033
2	3	4	0.2133
2	4	3	0.1425
2	2	2	0.1275
2	3	2	0.0866
2	2	4	0.0227
2	4	4	- 0.0850
2	3	3	- 0.2100

Comparaciones múltiples por medio de la prueba de Tukey modificada por Sen.

Los valores bajo la columna crecimiento representan las medias de las combinaciones de tratamiento.

Las líneas verticales unen a la medias que no son significativamente diferentes para una α de 0.05.

La simbología para Luz, Esp y Def corresponde a la de la Tabla 3

Tabla 6

GENERAL LINEAR MODELS PROCEDURE

DEPENDENT VARIABLE: MGT		SUM OF SQUARES		MEAN SQUARE	F VALUE	PR > F	TYPE III SS	F VALUE	PR > F	C.V.
SOURCE	DF	1681.09823230	513.15039357	26.63409893	14.25	0.0001	783.87851049	418.56	0.0001	150.8641
MODEL	63						107.06454150	19.86	0.0001	
ERROR	274						63.50134778	11.32	0.8554	
CORRECTED TOTAL	337	2194.24865387		1.7281166			1.45490390	0.53	0.6681	MGT MEAN
							28.593391722	1.97	0.0431	0.90651134
SOURCE	DF	TYPE I SS	F VALUE	PR > F	OF	TYPE III SS	F VALUE	PR > F		
LUZ	1	817.31707219	436.41	0.0001	1	783.87851049	418.56	0.0001		
ESP	3	109.87509921	19.26	0.0001	3	107.06454150	19.86	0.0001		
LUZ*ESP	3	63.15182915	11.21	0.0001	3	63.50134778	11.32	0.8554		
LUZ*DEF	3	0.81510999	0.15	0.9499	3	1.45490390	0.53	0.6681		
ESP*DEF	9	2.33831184	0.27	0.7451	9	28.593391722	1.97	0.0431		
LUZ*ESP*DEF	9	29.79478244	5.31	0.0256	9	31.14401972	1.97	0.0431		
DEN	1	36.12621766	6.55	0.0256	1	27.741725417	127.28	0.0001		
LUZ*DEN	1	24.0196415714	4.35	0.0416	1	27.741725417	127.28	0.0001		
LUZ*ESP*DEN	3	25.9197918259	4.67	0.0301	3	25.9197918259	44.66	0.0001		
LUZ*DEF*DEN	3	17.97791359	3.27	0.2522	3	17.363814556	1.31	0.2704		
LUZ*ESP*DEF*DEN	3	19.358101326	3.45	0.0171	3	20.01962403	3.56	0.0147		
ESP*DEF*DEN	3	0.39012151	0.07	0.9212	3	0.955326153	0.17	0.9147		
LUZ*ESP*DEF*DEN	9	61.26036327	11.03	0.0003	9	60.87627653	3.61	0.0003		
		22.32986577	4.03	0.0230	9	22.32986577	1.32	0.2230		

Análisis de Varianza para Variable Ganancia de Hojas

De este análisis se interpretó la interacción Especie*Defoliación*Densidad que presentó una significancia de 0.0003. Esta interacción no fue significativa para los ANDEVA con variable de respuesta, sobreviven-
cia y crecimiento, a su vez la interacción Luz*Especie*Defoliación con un nivel de significan-
cia de 0.0256 no se consideró significativa.

En la Tabla 7 se presentan las medias para la variable de respuesta ganancia de hojas, de la interacción de segundo orden especie*defoliación*densidad. Podemos concluir que sólo en la especie *O. cardiophylla* existe un efecto simultáneo de la defoliación y la densidad, que trae como consecuencia una ganancia diferencial de hojas, se puede observar en la parte inferior de la Tabla que para densidad alta se tiene pérdida de hojas. Es interesante notar que para alta densidad en la especie *O. cardiophylla* no existe una diferencia significativa en la ganancia de hojas en condiciones de cero y alta defoliación, por lo que es posible que el efecto de la densidad sea mejor explicado a través de la interacción de primer orden especie*densidad que también es significativa.

En conclusión para los tres ANDEVAS, la especie *O. cardiophylla* es la más sensible a condiciones de baja luminosidad, alta densidad y alta defoliación.

Si bien la significancia en las tres pruebas es muy alta, estas conclusiones deben tomarse con reserva para inferir sobre la adecuación de las plántulas, ya que, al hacer una prueba para cada variable independiente no se está considerando la correlación que pudiera existir entre las tres variables de respuesta.

Tabla 7

Comparaciones de medias para interacción Especie*Defoliación*Den-
sidad con Variable de Respuesta Ganancia de Hojas.

Esp	Def	Den	Ganancia de Hojas
1	4	2	3 2750
1	3	2	2.9000
2	2	2	2.4431
1	1	2	2.2916
2	1	2	2.1839
2	4	2	2.0677
2	3	2	1.7000
4	2	2	1.6875
4	1	2	1.5770
3	4	2	1.4916
2	1	1	1.3284
4	4	2	1.3062
3	2	2	1.3000
4	3	2	1.2540
3	3	2	1.2416
3	4	1	1.0903
4	1	1	1.0203
3	2	1	0.9700
3	3	1	0.9376
2	4	1	0.9275
2	2	1	0.8047
4	3	1	0.6735
2	3	1	0.6732
4	2	1	0.6325
4	4	1	0.6080
1	2	2	0.5500
3	1	2	0.2233
3	1	1	0.1856
1	2	1	- 1.3265
1	3	1	- 1.8960
1	1	1	- 2.1195
1	4	1	- 4.1722

Densidad:

1 = alta

2 = baja

Los valores bajo la columna ganancia de hojas representan las medias de las combinaciones de tratamientos.

Las líneas verticales unen a las medias que no son significativamente diferentes, para una $\alpha = 0.05$

ANÁLISIS DE VARIANZA MULTIVARIADO

En la Tabla 8 se tienen los resultados del análisis de varianza multivariado para la interacción de segundo orden Luz Especie Defoliación que arrojó una significancia de $P \ll 0.05$, en la Tabla 9 se presentan los valores de las coordenadas canónicas para las medias de esta interacción .

En la Gráfica 1 se presentan los resultados de la comparación de medias por medio de variables canónicas.

En este caso los círculos 1 y 4 representan a la especie *O. cardiophylla* en condiciones de sol para los cuatro diferentes grados de defoliación, podemos considerar a estos tratamientos como los que representan la mejor adecuación, ya que, como se puede observar en la Tabla 9 las medias de estos tratamientos presentan los valores más altos de sobrevivencia y crecimiento. Los círculos 6, 7 y 8 representan los tratamientos de sombra para la especie *O. cardiophylla* en los tres niveles más bajos de defoliación, estos tratamientos podemos considerarlos semejantes entre sí, y de baja adecuación a causa de que las medias de estos tratamientos tienen valores negativos para la variable ganancia de hojas (en este caso pérdida) y la variable sobrevivencia presenta los valores más bajos (ver tabla 10)

Tabla 8

GENERAL LINEAR MODELS PROCEDURE

MANOVA TEST CRITERIA FOR THE HYPOTHESIS OF NO OVERALL LUZ*ESP*DEF EFFECT

DESIGN MATRIX FOR LUZ*ESP*DEF

H = TYPE III SS ECP MATRIX FOR LUZ*ESP*DEF
 E = ERROR SS MATRIX
 P = DEP. VARIABLES = 3
 Q = HYPOTHESIS DF = 27
 ME = DF = 273
 S = MINIP(Q) = 2.3
 N = 5(LAS(P-Q)-1) = 135.0
 W = 5(ME-P-1)

HOTELLING-LAWLEY TRACE = TR(E**I*H) = 0.27757236 (SEE PILLAI'S TABLE #3)
 F APPROXIMATION = 2(S*N+1)*TR(E**I*H)/(S*S*(2H+S+1)) WITH S(2H+S+1) AND 2(S*N+1) DF
 F(27,812) = 2.78 PROB > F = 0.0001

PILLAI'S TRACE = V = TR(H*INV(H+E)) = 0.23384657 (SEE PILLAI'S TABLE #2)
 F APPROXIMATION = (2H+S+1)/(2H+S+1) * V/(S-V)
 WITH S(2H+S+1) AND S(2H+S+1) DF
 F(27,822) = 2.57 PROB > F = 0.0001

WILKS' CRITERION = L = DET(E)/DET(H+E) = 0.77519936 (SEE RAD 1973 P 555)
 $U = \frac{(ME - S(P-Q)) * LN(L)}{S * (P-Q)}$ = 70.2996
 $Z = \frac{S * (P-Q) * U}{(P+Q-4) * (P+Q-5)}$ = 276.5000
 $W = \frac{(P+Q-2) * Z}{(P+Q-1)}$ = 2.9205
 F APPROXIMATION = (U*Z-2B)/(P+Q)*(1-L**1/Z)/L**1/Z WITH P*Q AND U*Z-2B DF
 F(27,795) = 2.68 PROB > F = 0.0001

ROY'S MAXIMUM ROOT CRITERION = 0.22852309 (SEE AMS VOL 31 P 625)
 FIRST CANONICAL VARIABLE YIELDS AN F UPPER BOUND
 F(9,274) = 6.96 (UPPER BOUND)

Análisis de Varianza Multivariado para Interacción de Segundo Orden. Luz*Especie*Defoliación
 Los resultados de las pruebas de Pillai, Hotelling-Lawley y Wilks, para probar efecto de esta interacción arrojaron un nivel de significancia de 0.0001 por lo que esta se considera altamente significativa.

Tabla 9

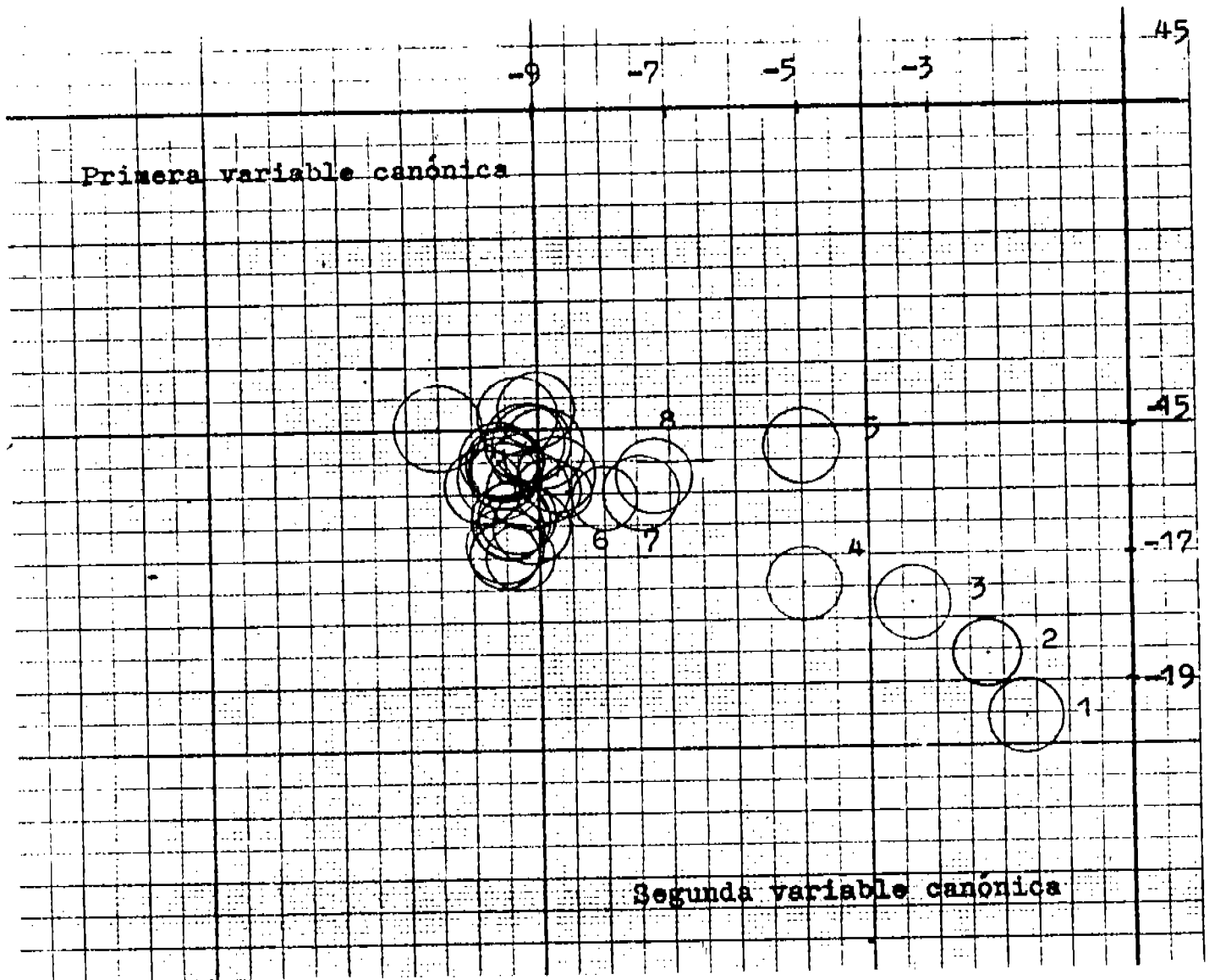
M	VC ₁	VC ₂	Radio
1	-2.24	-18.54	0.5
2	-1.64	-19.64	0.6
3	-3.35	-17.77	0.6
4	-4.97	-17.39	0.6
5	-9.30	-15.44	0.4
6	-9.56	-15.51	0.5
7	-9.55	-15.75	0.6
8	-10.52	-14.97	0.6
9	-9.11	-15.90	0.6
10	-8.93	-15.29	0.6
11	-9.19	-15.14	0.6
12	-9.32	-14.80	0.6
13	-8.61	-15.92	0.4
14	-8.84	-15.91	0.5
15	-8.75	-15.69	0.6
16	-9.57	-15.52	0.6
17	-7.99	-16.25	0.5
18	-7.42	-15.98	0.6
19	-7.24	-15.71	0.6
20	-5.04	-15.27	0.6
21	-9.40	-16.53	0.5
22	-9.38	-16.41	0.6
23	-9.10	-15.19	0.6
24	-9.05	-15.69	0.5
25	-9.12	-16.50	0.6
26	-8.94	-14.65	0.6
27	-9.50	-15.45	0.6
28	-9.75	-15.87	0.6
29	-9.36	-16.95	0.5
30	-9.48	-16.95	0.6
31	-9.50	-16.77	0.6
32	-9.64	-16.30	0.6

Valores de las variables canónicas para la interacción Luz* Especie* Defoliación.

VC₁ = Variable Canónica 1

VC₂ = Variable Canónica 2

Los números bajo la columna M se corresponden con los números de las medias de la Tabla 10



Gráfica # 1

Variables canónicas para interacción de segundo orden luz*especie*defoliación. Los círculos 1 a 4 representan la especie *O. cardiophylla* en condiciones de sol para defoliación de 0 a 75%. El círculo 5 representa a la misma especie en condiciones de sombra y 75% de defoliación, los círculos 6, 7 y 8 a *O. cardiophylla* con sombra y defoliaciones de 0, 5 y 25% respectivamente.

Tabla 10

	Luz	Esp	Def	M	MS	MCRE	MGT
①	1	1	1	1	100.0000000	28.4979167	1.51562500
②	1	1	2	2	100.0000000	29.9525000	1.01250000
③	1	1	3	3	100.0000000	25.9191667	2.60416667
④	1	1	4	4	100.0000000	19.9050000	2.27500000
	1	2	1	5	99.3303571	6.0839286	3.49553571
	1	2	2	6	100.0000000	5.3590909	3.46340909
	1	3	3	7	100.0000000	4.6750000	2.84625000
	1	2	4	8	99.0625000	1.8775000	3.60575000
	1	3	1	9	97.0000000	3.3133333	1.29366667
	1	3	2	10	94.6666667	3.7533333	1.81300000
	1	3	3	11	94.3333333	2.6733333	1.75033333
	1	3	4	12	93.3333333	2.2500000	2.07900000
	1	4	1	13	100.0000000	8.3488095	3.18398810
	1	4	2	14	99.5772727	6.9386364	2.79977273
	1	4	3	15	97.8150000	6.3200000	2.53250000
	1	4	4	16	97.1950000	2.6400000	1.99600000
⑥	2	1	1	17	92.0454545	2.4393939	-1.27295455
⑦	2	1	2	18	88.7625000	2.6650000	-1.78900000
⑧	2	1	3	19	87.5000000	3.0900000	-1.60025000
⑤	2	1	4	20	78.7500000	5.8350000	-3.17225000
	2	2	1	21	97.4431818	0.5310606	-0.34090909
	2	2	2	22	96.8875000	0.4225000	-0.31750000
	2	2	3	23	93.7500000	0.1275000	-0.47300000
	2	2	4	24	92.6136364	0.6227273	-0.36704545
	2	3	1	25	95.8333333	0.3033333	-0.88466667
	2	3	2	26	88.8333333	0.0866667	-0.45700000
	2	3	3	27	93.6666667	-0.2100000	0.42900000
	2	3	4	28	96.1666667	0.2133333	0.50300000
	2	4	1	29	98.8000000	0.6865385	-0.71019231
	2	4	2	30	99.2225000	0.5525000	-0.59100000
	2	4	3	31	98.2900000	0.1425000	-0.60500000
	2	4	4	32	97.1975000	-0.0850000	-0.08175000

Valores de medias para interacción de segundo orden Luz*Especie*Defoliación.

MS = Media sobrevivencia

MCRE = Media crecimiento

MGT = Media ganancia de hojas.

Los casos en los círculos 1 - 4 corresponden a *O. cardiophylla* en condiciones de sol y para defoliaciones de 0 a 75% respectivamente, el caso marcado con el círculo 5 corresponde a la misma especie en condiciones de sombra y máxima defoliación, los casos 6, 7 y 8 corresponden a la misma especie en sombra y defoliación de 0, 5 y 25% respectivamente.

El círculo 5 representa la misma especie en condiciones de sombra y defoliación extrema; claramente se diferencia de los tres anteriores y consideramos que es el caso que representa la más baja adecuación, ya que como se observa en la Tabla 10 las medias de las variables sobrevivencia y ganancia de hojas presentan los valores extremos más bajos. Podemos concluir que el resto de los tratamientos no se diferencian entre sí en forma significativa y que representan una adecuación intermedia, pues como se observa en la Gráfica 1 estos tratamientos forman un grupo compacto y se puede apreciar en la Tabla 10 que las medias de las variables de respuesta toman valores intermedios entre los casos extremos ya comentados.

En la Tabla 11 se presentan los resultados para la interacción de segundo orden significativa ($P \ll 0.05$) Especie Defoliación Densidad, en la Tabla 12 se presentan los valores de las coordenadas canónicas para las medias de esta interacción

En la Gráfica 2 se presentan los tratamientos de la interacción Especie*Defoliación*Densidad. El círculo 1 representa a la especie *O. cardiophylla* en condiciones de alta densidad y extrema defoliación. Esta combinación de factores generan una adecuación mínima extrema en esta especie, ya que, como se observa en la Tabla 13 los valores de las medias de esta combinación de tratamientos presentan los valores más bajos para las variables sobrevivencia y ganancia de hojas.

GENERAL LINEAR MODELS PROCEDURE

HANDVA TEST CRITERIA FOR THE HYPOTHESIS OF NO OVERALL ESP*DEF*DEN EFFECT

H = TYPE III SSECP MATRIX FOR: ESP*DEF*DEN

E = ERROR SSECP MATRIX

P = DEP. VARIABLES = 3

Q = HYPOTHESIS DF = 274

SE = DF OF E = 274

S = MIN(P,Q) = 3

M = 5(ABS(P-Q)-1) = 2.5

N = .5(NE-P-1) = 135.0

HOTELLING-LAWLEY TRACE = TR(E**I*H) = 0.19273600 (SEE PILLAI'S TABLE #3)

F APPROXIMATION = 2(S*N+1)*TR(E**I*H)/(S*S*(2M+S+1)) WITH S(2M+S+1) AND 2(S*M+1) DF

F(27,812) = 1.93 PROB > F = 0.0032

PILLAI'S TRACE = V = TR(H*INV(H+E)) = 0.17470424 (SEE PILLAI'S TABLE #2)

F APPROXIMATION = (2N+S+1)/(2M+S+1) * V/(S-V) WITH S(2M+S+1) AND S(2M+S+1) DF

F(27,822) = 1.88 PROB > F = 0.0045

WILKS' CRITERION L = DET(E)/DET(H+E) = 0.8244559 (SEE RAD 1973 P 555)

M = (S+1)*LN(L) = 50.7066

N = (P+1)*LN(L) = 176.9200

U = (P+1)*LN(L)/(P+Q+1) = 2.9265

Z = (P+2)/4 = 6.2500

F APPROXIMATION = (U*Z-2B)/(P+Q)*(1-L**1/Z)/L**1/Z WITH P*Q AND U*Z-2B DF

F(27,795) = 1.91 PROB > F = 0.0038

ROY'S MAXIMUM ROOT CRITERION = 0.13007245 (SEE AHS VOL 31 P 625)

FIRST CANONICAL VARIABLE YIELDS AN F UPPER BOUND

F(9,274) = 3.96 (UPPER BOUND)

Resultados de Análisis de Varianza Multivariado para Interacción Especie*Defoliación*Densidad
 Los resultados de las pruebas de Hotelling-Lawley, Pillai y Wilks para probar efecto de esta
 interacción arrojaron valores de .0032, .0045 y .0038 respectivamente por lo que se considera
 que esta interacción también es altamente significativa.

M	VC ₁	VC ₂	Radio
1	3.19	19.73	0.5
2	6.37	20.21	0.5
3	3.65	18.40	0.6
4	5.19	20.64	0.6
5	2.95	18.66	0.6
6	6.69	20.02	0.6
7	0.73	17.28	0.6
8	6.95	19.93	0.6
9	4.21	19.13	0.5
10	4.97	19.63	0.4
11	3.80	19.29	0.6
12	5.06	19.53	0.5
13	3.53	18.65	0.6
14	4.49	19.68	0.6
15	3.55	18.23	0.6
16	4.48	19.26	0.6
17	3.19	19.06	0.6
18	3.28	19.49	0.6
19	3.71	18.68	0.6
20	3.81	17.58	0.6
21	3.50	17.89	0.6
22	3.94	19.22	0.6
23	3.65	18.15	0.6
24	4.11	19.16	0.6
25	4.23	19.72	0.5
26	4.62	19.81	0.4
27	3.87	19.76	0.6
28	4.60	19.74	0.5
29	3.79	19.56	0.6
30	4.17	19.44	0.6
31	3.41	18.75	0.6
32	4.05	19.70	0.6

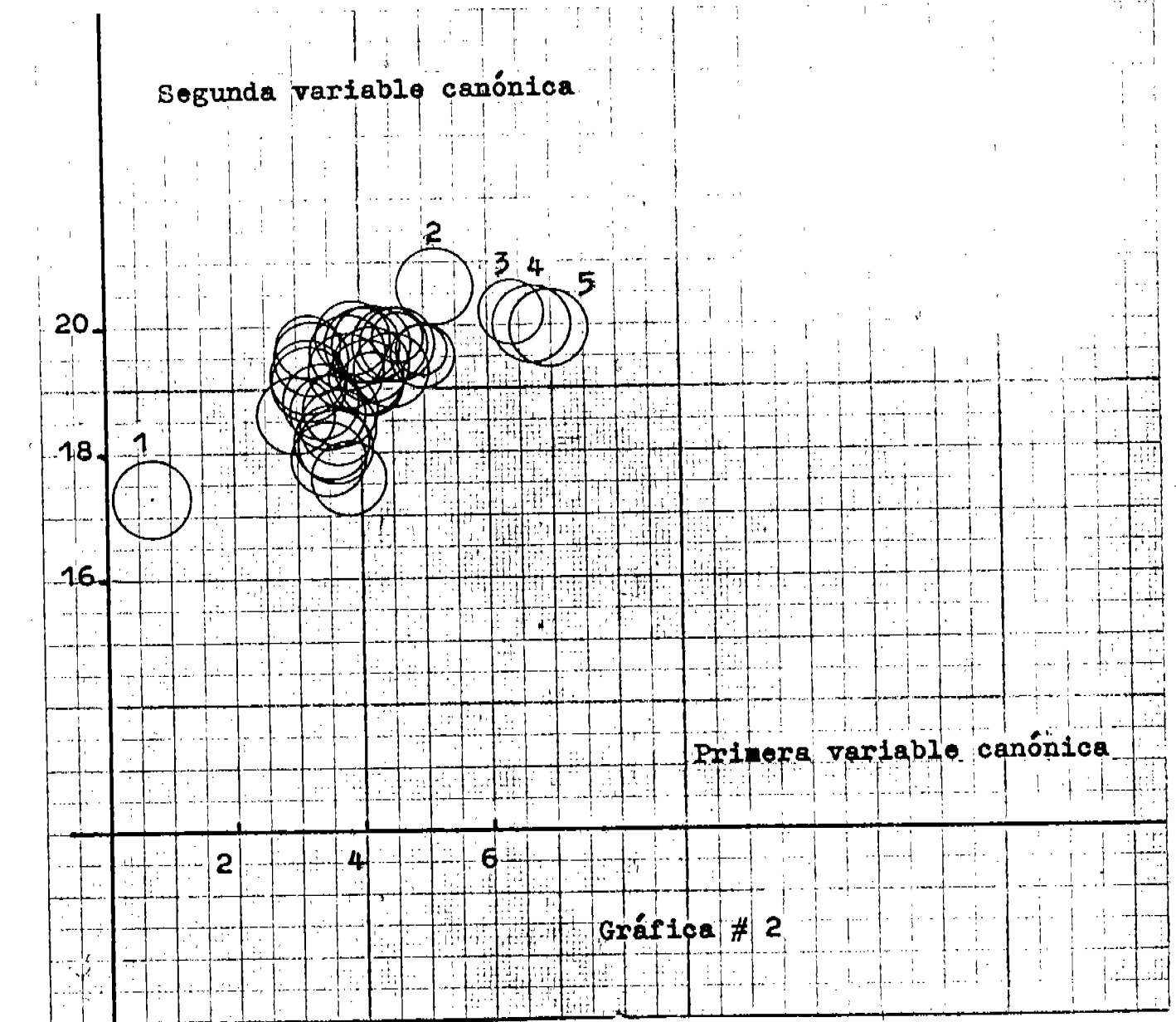
TABLA 12

Valores de las variables canónicas para la interacción Especie*Defoliación*Densidad.

VC₁ = Variable canónica 1

VC₂ = Variable canónica 2

Los números bajo la columna M se corresponden con los números de las medias de la Tabla 13 .



Variables canónicas para la interacción de segundo orden especie*defoliación*densidad. El círculo 1 representa a *O. cardiophylla* en condiciones de alta densidad y máxima defoliación. Los círculos 2 a 5 representan a la misma especie en condiciones de baja densidad y defoliaciones de 0, 5, 25 y 75 por ciento respectivamente.

	Luz	Esp	Def	M	MS	MCRE	MGT
	1	1	1	1	92.0454545	16.7863636	-2.11954545
②	1	1	2	2	100.0000000	15.3465278	2.29166667
	1	2	1	3	88.7625000	16.8350000	-1.32650000
③	1	2	2	4	100.0000000	15.7825000	0.55000000
	1	3	1	5	87.5000000	14.5516667	-1.89608333
④	1	3	2	6	100.0000000	14.4575000	2.90000000
①	1	4	1	7	78.7500000	11.3250000	-4.17225000
⑤	1	4	2	8	100.0000000	14.4150000	3.27500000
	2	1	1	9	96.5909091	3.3659091	1.32840909
	2	1	2	10	100.0000000	3.8565476	2.18392857
	2	2	1	11	96.8875000	2.9125000	0.80475000
	2	2	2	12	100.0000000	3.0954545	2.44318182
	2	3	1	13	93.7500000	2.1775000	0.67325000
	2	3	2	14	100.0000000	2.6250000	1.70000000
	2	4	1	15	92.1875000	1.2300000	0.92750000
	2	4	2	16	98.8636364	0.6113636	2.06772727
	3	1	1	17	95.3333333	1.8033333	0.18566667
	3	1	2	18	97.5000000	1.8133333	0.22333333
	3	2	1	19	94.3333333	1.8133333	0.97000000
	3	2	2	20	89.1666667	2.0266667	1.30000000
	3	3	1	21	90.5000000	1.1733333	0.93766667
	3	3	2	22	97.5000000	1.2900000	1.24166667
	3	4	1	23	92.0000000	1.1700000	1.09033333
	3	4	2	24	97.5000000	1.2933333	1.49166667
	4	1	1	25	98.8000000	4.8134615	1.02038462
	4	1	2	26	100.0000000	4.5166667	1.57702381
	4	2	1	27	98.7575000	4.1225000	0.63250000
	4	2	2	28	100.0000000	3.6931818	1.68750000
	4	3	1	29	97.9800000	3.3350000	0.67350000
	4	3	2	30	98.1250000	3.1275000	1.25400000
	4	4	1	31	94.3925000	1.3550000	0.60800000
	4	4	2	32	100.0000000	1.2000000	1.30625000

Tabla 13

Valores de las medias para la interacción de segundo orden
Especie*Defoliación*Densidad.

MS = Media sobrevivencia, MCRE = Media crecimiento

MGT = Media ganancia de hojas.

El caso marcado con el círculo 1 representa la especie *O. car-diophylla* en condiciones de alta densidad y extrema defoliación. Los casos marcados con los círculos 2, 3, 4 y 5 representan a la misma especie en condiciones de baja densidad y defoliación de 0, 5, 25 y 75% respectivamente.

El círculo 2 representa a la misma especie *O. cardiophylla* en condiciones de baja densidad y mínima defoliación, se considera a esta combinación de tratamientos como el que produjo la más alta adecuación, observese en la Tabla 13 que los valores de las medias de esta combinación de tratamientos presentan valores altos. Los círculos 3 a 5 representan a la misma especie en condiciones de baja densidad, pero con defoliación alta. Se considera que estos tratamientos producen también una alta adecuación pero sensiblemente menor al tratamiento con cero defoliación.

Análisis de Conglomerados.

Existen gran variedad de criterios para decidir el número de conglomerados en un análisis. La mayoría de las pruebas tienen distribuciones muestrales complejas o involucran a la hipótesis nula donde su rechazo no aporta información.

Arnold (citado en SAS 82) propone un criterio para definir el número de conglomerados basado en el determinante de la matriz de suma de cuadrados dentro de tratamientos $|W|$, aunque tiende a ser conservador con los procedimientos usados en este análisis y especialmente cuando se tiene más de una variable. Los procedimientos FASTCLUS y CLUSTER del paquete SAS tienden a minimizar más la traza que el determinante. Los algoritmos FASTCLUS Y CLUSTER están basados en el Criterio Cúbico de Conglomerados (CCC) que está fundamentado en el supuesto de que una distribución uniforme en un hiperrectángulo, será dividida en conglomerados a semejanza de hipercubos. Se pueden observar en la Tabla #143 bajo la columna "CUBIC CLUSTER CRITERIA" los valores para estos criterios de definición de número de grupos. Se observa que los valores aumentan a medida que se incrementa el número de conglomerados, cuando se esperaba tener un pico y una gradual declinación de los valores con mayor número de conglomerados: esto puede ser debido al excesivo redondeo de los datos provocados por el

WARD'S MINIMUM VARIANCE HIERARCHICAL CLUSTER ANALYSIS

EIGENVALUES OF THE CORRELATION MATRIX

EIGENVALUE	DIFFERENCE	PROPORTION	CUMULATIVE
1.600553	0.742999	0.533518	0.533518
1.857553	0.315660	0.285851	0.819369
2.541394	.	0.190631	1.000000

ROOT-MEAN-SQUARE TOTAL-SAMPLE STANDARD DEVIATION = 1.73205
 ROOT-MEAN-SQUARE DISTANCE BETWEEN OBSERVATIONS =

NUMBER OF CLUSTERS	FREQUENCY OF NEW CLUSTER	RMS STD OF NEW CLUSTER	SEMI-PARTIAL R-SQUARE	R-SQUARED	APPROXIMATE REPAIRED R-SQUARE	CUMULATIVE CLUSTER TIME CRITERION
32	1	0.352525	0.001324	0.971004	0.928187	22.2693
31	3	0.35468	0.001369	0.969655	0.926208	21.8760
30	4	0.155096	0.001372	0.968284	0.924127	21.5301
29	4	0.478568	0.001571	0.966713	0.921935	21.1908
28	4	0.173161	0.001694	0.965025	0.919621	20.8598
27	1	0.471027	0.001715	0.963310	0.917174	20.5334
26	1	0.315886	0.001795	0.961526	0.914582	20.2129
25	1	0.195815	0.001981	0.959783	0.911829	19.8978
24	1	0.355052	0.002181	0.957986	0.909074	19.5782
23	1	0.261161	0.002338	0.956222	0.906230	19.2639
22	1	0.372263	0.002654	0.954463	0.903387	18.9502
21	1	0.317011	0.003246	0.952717	0.900544	18.6376
20	1	0.426084	0.003430	0.950973	0.897701	18.3261
19	1	0.218173	0.003864	0.949241	0.894858	18.0156
18	4	0.237567	0.004233	0.947516	0.892015	17.7070
17	4	0.695695	0.004733	0.945798	0.889172	17.3993
16	1	0.375975	0.005115	0.944073	0.886329	17.0926
15	1	0.531987	0.005940	0.942347	0.883486	16.7869
14	1	0.518814	0.006532	0.940622	0.880643	16.4812
13	1	0.207058	0.006532	0.938897	0.877800	16.1755
12	2	0.388438	0.010702	0.937171	0.874957	15.8698
11	2	0.648938	0.015315	0.935445	0.872114	15.5641
10	2	0.315202	0.021847	0.933711	0.869271	15.2584
9	1	0.83515	0.031710	0.931978	0.866428	14.9527
8	1	0.536064	0.052979	0.930244	0.863585	14.6470
7	1	0.969897	0.062133	0.928509	0.860742	14.3413
6	1	1.035546	0.121658	0.926774	0.857899	14.0356
5	1	0.824351	0.269752	0.925039	0.855056	13.7299
4	1	0.824351	0.785372	0.923304	0.852213	13.4242

Tabla 14

Resultados de análisis de conglomerados para procedimiento jerárquico.

promedio de los valores para cada charola y el promedio de los datos por tiempo al simplificar el modelo; por lo que nos atenderemos a un criterio de tipo empírico buscando los extremos de adecuación, por lo cual se optó por un número de tres grupos que representaran adecuación alta, intermedia y baja. En la Tabla 15 se tienen los resultados para el procedimiento FASTCLUS (k - medias) de agrupación; bajo el renglón "DISTANCE BETWEEN CLUSTER MEANS" se observan las distancias entre las medias de conglomerados (centroides). Los grupos más distantes son 2 y 3 que representan los extremos de adecuación.

Ambos grupos representan la adecuación extrema de la especie *O. cardiophylla* bajo condiciones de luz y sombra, el grupo intermedio está formado por las tres especies restantes. En el procedimiento de agrupación jerárquica "CLUSTER" se obtuvo para los tres grupos, una notable semejanza en cuanto a los casos agrupados con respecto al procedimiento no jerárquico. Coinciden en el grupo de alta adecuación la **gran mayoría** de los casos (ver Tablas 16 y 17). El grupo de baja adecuación "está formado por mayor número de casos que el correspondiente no jerárquico. Los individuos en los que difieren los dos procedimientos de agrupación están formando parte en el caso del procedimiento de partición óptima del grupo de adecuación intermedia y corresponden a aquellos casos con mayor distancia Euclidiana del centroide (ver Tablas 18 y 19).

11:36 FRIDAY,

FASTCLUS PROCEDURE

REPLACE = FULL MAXCLUSTERS = 3 MAXITER = 1
 SEED REPLACEMENTS: FIRST LEVEL = 7 SECOND LEVEL = 30
 INITIAL SEEDS

CLUSTER	MS	MCRE	MGT
1	81.2500	0.8000	-0.8775
2	100.0000	29.3000	8.7500
3	43.7500	1.0250	-7.9200

CLUSTER SUMMARY

CLUSTER	MEMBERS	RMS	ST DEV	MAX DISTANCE FROM SEED
1	278	3.512295		21.48942
2	43	4.621015		14.02415
3	17	6.298533		25.67144

STATISTICS FOR VARIABLES

VARIABLE	TOTAL STD	WITHIN STD	R-SQUARED	VAR RATIO
MS	8.60264623	5.09721362	0.65100722	1.86538875
MCRE	8.64017012	3.61074034	0.82639784	4.76029706
MGT	2.55139017	2.30699369	0.18720792	0.23046319
OVER-ALL	7.19026055	3.84474461	0.71577776	2.51337346

APPROXIMATE EXPECTED OVER-ALL R-SQUARED = 0.5426 CUBIC CLUSTERING CRITERION = 7.15052
 WARNING: THESE VALUES ARE INVALID IF VARIABLES ARE CORRELATED

CLUSTER MEANS

CLUSTER	MS	MCRE	MGT
1	97.2542	2.5691	1.0836
2	99.8927	2.8117	1.6351
3	66.1471	3.5677	-3.8173

CLUSTER STANDARD DEVIATIONS

CLUSTER	MS	MCRE	MGT
1	5.01299	2.86448	1.91654
2	0.76349	6.91330	3.96058
3	10.36534	2.22594	2.58092

11:36 FRIDAY,

FASTCLUS PROCEDURE

DISTANCES BETWEEN CLUSTER MEANS

CLUSTER	1	2	3
1			
2	23.70534		31.50663
3	31.50563	40.94599	

Tabla 15

Resultados de análisis de conglomerados para el procedimiento de partición óptima de 3 grupos

Tabla 16

Grupo de alta adecuación usando procedimiento de partición óptima.

CASO	distancia al centroide
38	0.1555
23	1.4077
42	2.0277
9	2.0175
16	3.2700
26	3.2754
31	3.3311
24	3.3766
29	3.5388
13	3.6688
19	4.4354
7	4.5152
2	4.5899
27	4.9541
3	5.3553
6	5.4194
11	5.4921
35	5.5587
34	6.2690
37	6.3760
211	6.4394
36	7.0209
10	7.3759
5	7.5186
1	7.7076
12	8.7235
22	8.7574
8	5.8832
20	9.1637
15	10.1219
41	10.2771
118	10.2802
25	10.7465
39	10.7505
4	11.8793
40	12.0574
14	12.2445
17	12.5819
21 ⁿ	13.3158
18	13.3223
28	13.5721
32	15.0242

Tabla 17

CASO

CASO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53			
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

NUMERO DE CONGLOMERADOS

Grupo de alta adecuación formado por el procedimiento jerárquico.

Compárense los casos agrupados con la tabla 16

Tabla 18

Grupo de baja adecuación usando el procedimiento de partición óptima.

CASO	distancia al centroide
207	2.3630
208	4.2222
197	4.3341
206	4.5079
188	8.4209
135	9.4554
205	10.8574
272	14.3600
246	18.6112
111	20.4404
174	20.8641
173	20.9544
196	21.0572
195	21.1549
204	21.2102
184	21.6488
176	25.6714

ESTADÍSTICA
JUAN A. ESCALANTE H.
 UNIDAD ACADÉMICA DE
 LOS CICLOS PROFESIONAL
 Y DE POSGRADO / CCH
 U N A M

Tabla 19

CASO	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53
NUMERO DE CONGLOMERADOS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53

Grupo de Baja Adecuación formado por medio del Procedimiento Jerárquico.

Comparar los casos agrupados con la tabla 18

Análisis Discriminante.

El análisis discriminante se procesó en el paquete SPSS, Versión 6. En la Tabla 20, se tienen los resultados para el procedimiento de agrupación jerárquico (WARD).

Las dos funciones discriminantes tienen un alto poder para separar grupos como se puede ver de los coeficientes de las raíces características o eigenvalores y de la correlación canónica entre las variables y la función discriminante, en ambos casos los valores son altos .92 y .82.

La primera función explica el 73.2% de la discriminación y la segunda del 26.8%.

El mapa territorial se interpreta con los coeficientes de mayor valor de las funciones discriminantes estandarizadas que en este caso corresponden a los valores de las variables crecimiento y sobrevivencia para la primera y segunda función respectivamente.

En el mapa territorial 1 el grupo 2 se interpreta como el de más alta adecuación, ya que está definido por el mayor crecimiento y la mayor sobrevivencia. Estas plántulas corresponden a la especie *O. cardiophylla* que fueron sometidas al tratamiento de luz y baja densidad.

Tabla 20

ANALISIS DISCRIMINANTE
 FILE NAME (CREATION DATE = 02/09/84) 02/08/84 PAGE 1
 DISCRIMINANT ANALYSIS
 ANALYSIS NUMBER 1
 SOLUTION METHOD - DIRECT
 PRIOR PROBABILITIES - EQUAL
 GROUP 1 GROUP 2 GROUP 3
 .33333 .33333 .33333

NUMBER REMOVED	EIGENVALUE	CANONICAL CORRELATION	PERCENT OF TRACE	MILKS LAISUDA	CHI-SQUARE	D.F.	SIGNIFICANCE
0	5.70305	.92240	73.2	.04824	1015.54076	6	0
1	2.02553	.82254	26.8	.32335	378.22245	2	0

2 FUNCTIONS WILL BE USED IN REMAINING ANALYSES

STANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	1	2
V1	.30458	1.71462
V2	2.81224	-.37333
V3	-.54191	.10652

UNSTANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	1	2
V1	.35405E-01	.201678
V2	.302053	-.43251E-01
V3	-.212293	.41720E-01
CONSTANT	-.490425	-.19.1575

CENTROS OF GROUPS IN REDUCED SPACE

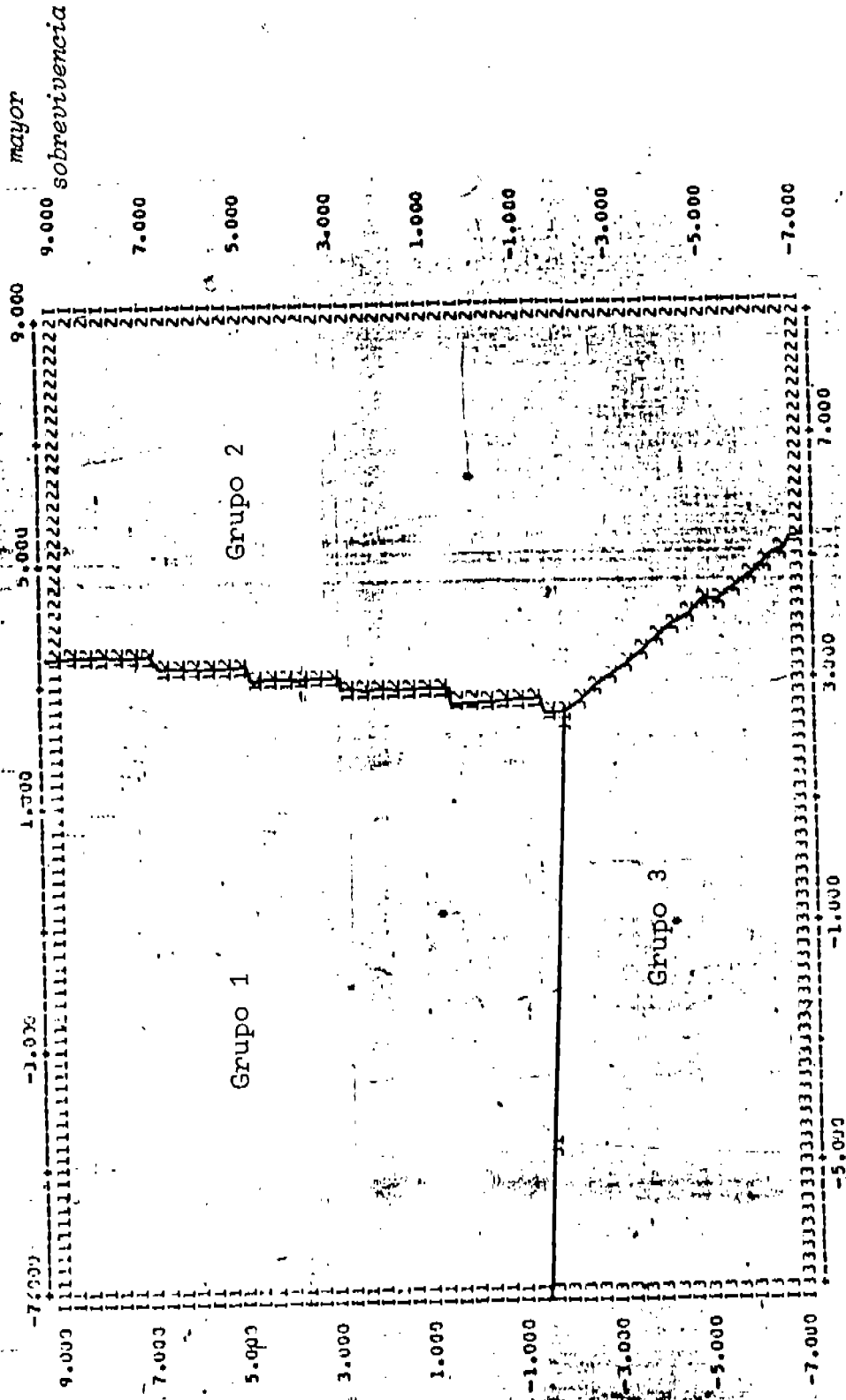
GROUP	1	2	3
GROUP 1	.85665	.59761	
GROUP 2	6.37861	-.11189	
GROUP 3	-1.38861	-4.25573	

Resultados de Análisis Discriminante para el Procedimiento de Agrupación Jerárquica.

v1 = Supervivencia, v2 = Crecimiento y v3 = Ganancia de hojas

Los coeficientes de las funciones discriminantes estandarizados estan representados en los ejes vertical y horizontal del mapa territorial 1.

TERRITORIAL MAP OF DISCRIMINANT SCORE 1 (HORIZONTAL) VS. DISCRIMINANT SCORE 2 (VERTICAL). * INDICATES A GROUP CENTER.



Mapa territorial del análisis discriminante subsecuente al análisis de conglomerados tipo jerárquico (ward).
 El eje vertical corresponde a la primera función discriminante.
 El eje horizontal corresponde a la segunda función discriminante.
 Los casos de los grupos 2 y 3 están representados en las tablas 17 y 19 respectivamente.

El grupo 3 se caracteriza por su bajo crecimiento y baja sobrevivencia, estas plántulas corresponden a la misma especie pero en condiciones de sombra y alta densidad. El grupo 1 se caracteriza por un bajo crecimiento y una alta sobrevivencia y corresponde a las tres especies restantes.

El procedimiento discriminante aplicado al análisis de conglomerados no jerárquico produjo, como era de esperarse, resultados muy semejantes al anterior dada la similitud de los resultados entre los análisis de conglomerados. En este caso también ambas funciones tienen un alto poder discriminante aunque la segunda con un menor poder que la correspondiente al caso anterior.

Los eigenvalores (Tabla 21), 5.51 y 1.89 son muy semejantes al caso anterior, los valores de la correlación canónica son también altos en ambas funciones .92 y .80 para la primera y segunda función respectivamente. El porcentaje de explicación es de 74.4 y 25.6. En general en este segundo caso los valores de la segunda función discriminante son ligeramente más bajos. Los coeficientes de las funciones discriminantes corresponden a las mismas variables del caso anterior, por lo que el mapa territorial de este análisis discriminante se interpreta en la misma forma (ver mapa 2).

ANALISIS DISCRIMINANTE
 FILE: HONAME (CREATION DATE = 02/09/84)
 ANALYSIS NUMBER 1
 SOLUTION METHOD - DIRECT.
 PRIOR PROBABILITIES - EQUAL

GROUP	1	GROUP 2	GROUP 3
	.3333	.3333	.3333

DISCRIMINANT ANALYSIS

NUMBER REMOVED	EIGENVALUE	CANONICAL CORRELATION	PERCENT OF TRACE	MILKS LAMBDA	CHI-SQUARE	D.F.	SIGNIFICANCE
0	5.51567	.92007	74.4	.02276	984.27235	4	0
1	1.39770	.30926	25.6	.34510	356.41174	2	0

2 FUNCTIONS WILL BE USED IN RETAINING ANALYSES

STANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	1	2
V1	-.10864	1.62599
V2	-.11607	-.18186
V3	-.44930	

UNSTANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	1	2
V1	-.126288E-01	.188905
V2	.304191	-.134467E-01
V3	-.176298	.712771E-01
CONSTANT	-.333092	-.181287

CENTROIDS OF GROUPS IN REDUCED SPACE

GROUP	1	2	3
GROUP 1	-.97082	.28578	
GROUP 2	0.96322	.50513	
GROUP 3	.53437	-.595103	

Análisis Discriminante para el Procedimiento de Agrupación de Partición Optima.

v1 = Sobrevivencia, v2 = Crecimiento y v3 = Ganancia de hojas.

Las funciones discriminantes estandarizadas estan representadas en los ejes horizontal y vertical del mapa territorial 2.

02/08/84 PAGE

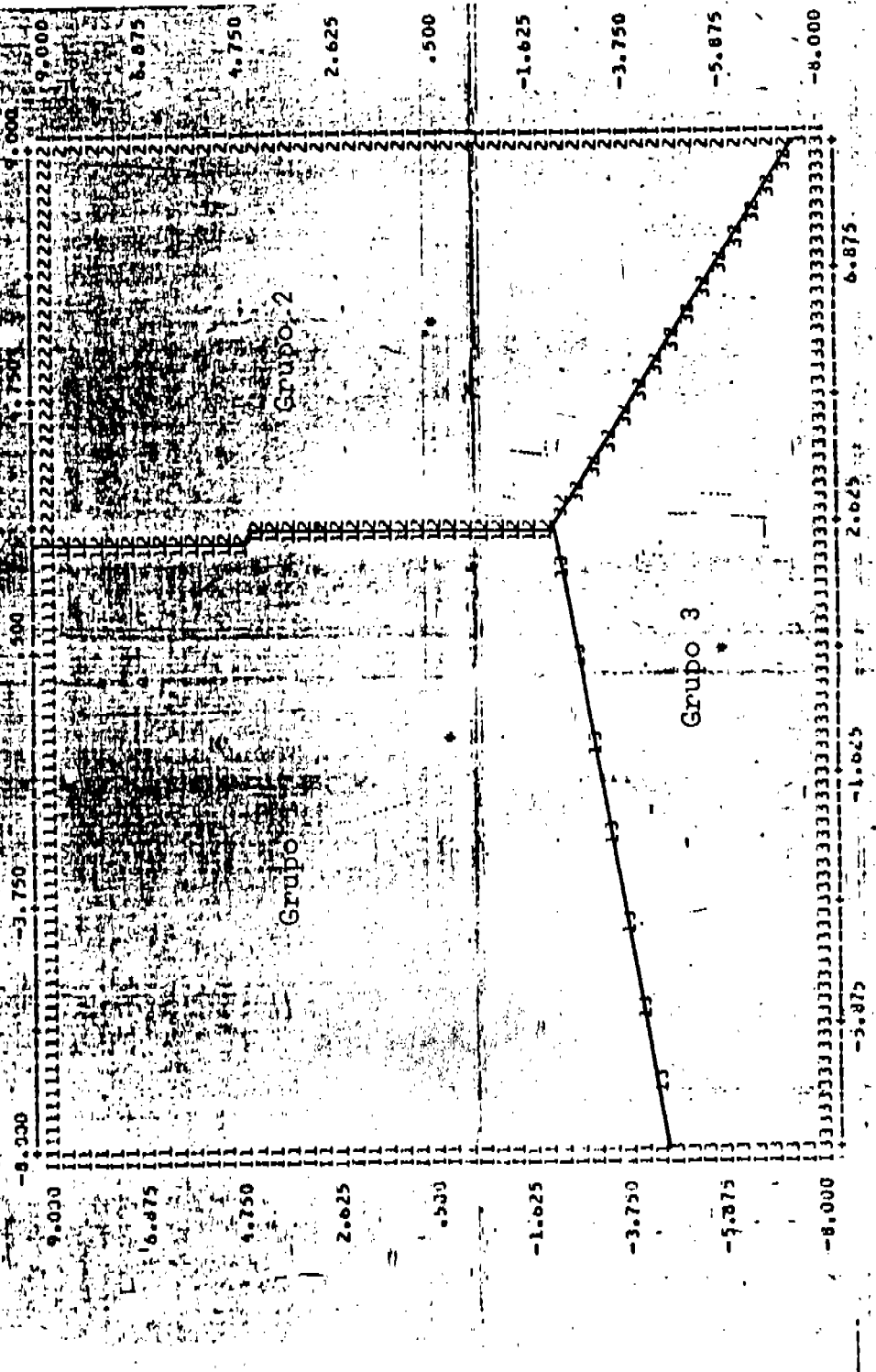
Mapa Territorial 2

ANALISIS DISCRIMINANTE

TERRITORIAL ASP OF DISCRIMINANT SCORE 1 (HORIZONTAL) VS. DISCRIMINANT SCORE 2 (VERTICAL) * INDICATES A GROUP CENTROID.

Mayor

Sobrevivencia



Mayor crecimiento

Mapa territorial del análisis discriminante subsecuente al análisis de conglomerados de partición óptima. Los ejes vertical y horizontal corresponden respectivamente a la primera y segunda función discriminantes. Los casos de los grupos 2 y 3 están representados en las Tablas 16 y 18 respectivamente.

CONCLUSIONES Y RECOMENDACIONES

Si bien los análisis de varianza univariado y multivariado arrojan resultados similares, el ANDEVA como se asentó anteriormente no está considerando la correlación que existe entre las variables de respuesta y el hecho de que ambos tipos de análisis presenten una estrecha coincidencia en sus resultados puede deberse más a la bondad de los datos analizados que a la bondad de las técnicas empleadas. Por ello se recomienda que cuando se tenga más de una variable de respuesta se emplee análisis multivariado que permite una aproximación holística a los resultados, al analizar en forma conjunta las variables considerando sus posibles correlaciones .

El hecho de que la aplicación de dos técnicas de análisis de conglomerados tan dispares como la jerárquica y la de partición óptima produzca resultados con una semejanza tan estrecha, no hace sino confirmar lo que se deja entrever del análisis de variables canónicas; la presencia de tres grupos de adecuación claramente identificables.

No es extraño que de los resultados tan semejantes de los análisis de conglomerados se produzcan de estos, dos análisis discriminantes igualmente similares, que permiten afirmar que los grupos de adecuación extrema están for-

mados por la especie *O. cardiophylla* y el grupo de adecuación intermedia por las tres especies restantes.

Los análisis discriminantes permiten concluir que sólo dos variables son necesarias para establecer la adecuación de las plántulas, pudiéndose omitir la variable Ganancia de Hojas, ya que los coeficientes de esta variable en las funciones discriminantes toman valores muy cercanos a cero .

El diseño del experimento presenta varias complicaciones innecesarias que dificultan y hacen poco confiable su análisis .

1) La presencia de errores de restricción que no permiten probar factores de interés.

2) El número diferencial de combinaciones de tratamientos con la consecuente pérdida de ortogonalidad del experimento que sólo puede ser analizado a través de procedimientos que requieren de una alta cantidad de memoria en el proceso de cómputo.

3) La asignación de factores e interacciones de interés a la parcela grande, cuando en un análisis de parcelas divididas debieron estar asignadas a la parcela chica, que permitiría una prueba de F más confiable al llevar a cabo ésta con el error experimental.

Es por ello que es importante planear previamente el experimento a realizar y de esta manera saber de antemano cuales son los factores que se pueden probar y confundir, aquellas interacciones de alto orden que no sean de interés para el investigador. Al mismo tiempo es recomendable no involucrar en el experimento un número excesivo de **factores** con gran número de niveles. Siendo aconsejable hacer en un principio un experimento exploratorio con sólo dos niveles extremos en cada factor, ya que, un experimento sencillo es más económico, de análisis más simple y de más fácil interpretación.

BIBLIOGRAFIA

- Anderson, V. L., (1970). Restriction Errors for Linear Models
(an aid to develop models for
designed experiments. Biometrics 25:
255-268.
- Anderson, V.L., and McLean, A., (1974). Design of Experiments.
A Realistic Approach.
Marcel Dekker, Inc. New York.
- Cohen, J., (1968). Multiple Regression as General Data-Analytic
System. Psychological Bulletin 70 (6):
426-443.
- Cooley, W. W. and Lohnes, R. P., (1971). Multivariate Data
Analysis. John Wiley and Sons. New
York.
- Dolby, G. G., (1982). The Role of Statistics in the Methodology
of the Life Sciences Biometrics. 38:
1069-1083.
- Elston, R. C. and Bush, N., (1964). The Hypothesis that can Be
Tested When There are Interactions In
Analysis of Variance Model. Biometrics.
20: 687-692.

- Everitt, B. S. (1976). Cluster Analysis. Heineman Educational Books. Ltd. London.
- Gabriel, K. R. (1978). A Simple Method of Multiple Comparisons of Means. Journal of American Statistical Association. 73: 724-729.
- Goodnight, J. H. et al. A Users Guide to SAS-82. SAS Institute.
- Hair, J. F. et al., (1979). Multivariate Data Analysis. Petroleum Pub. Co. Tulsa.
- Hartigan, J. A. (1975). Clustering Algorithms. John Wiley and Sons. New York.
- MacQueen, J. B., (1967). Some Methods for Classification and Analysis of Multivariate Observations Proceedings of the Fifth Berkley Symposium on Mathematical Statistics and Probability. 1: 281-297.
- Mardia, K. V, Kent, J. T. and J. M. Bibby. (1979). Multivariate Analysis. Academic Press. London.
- Marriot, F. H. C., (1974). The Interpretation of Multiple Observations. Academic Press. London.

- Méndez, R. I., (1976). Lineamientos Generales para la Planeación de Experimentos. Comunicaciones Técnicas Serie Azul No. 15 IIMAS-UNAM. México.
- Méndez, R. I., (1981). Modelos Estadísticos Lineales, Interpretación y Aplicaciones. CONACyT. México.
- Méndez, R. I., (1981a). El Error de Restricción en los Diseños con Bloques. Comunicaciones Técnicas. Serie Azul No. 21. IIMAS-UNAM. México.
- Méndez, R. I., (1983). Comentarios Sobre el Diseño y Análisis de Experimentos con Animales. Comunicaciones Técnicas. Serie Azul No. 67, IIMAS-UNAM. México.
- Méndez, R. I. y Márquez, J. A. (1976). Las Hipótesis que se Prueban en Modelos con Dos Criterios de Clasificación. Comunicaciones Técnicas. Serie Azul No. 19. IIMAS-UNAM. México, D.F.

- Montgomery, C. D. (1976). Design and Analysis of Experiments.
John Wiley and Sons. New York.
- Morrison, F. D. (1976). Multivariate Statistical Methods.
McGraw Hill. Tokyo.
- Nie, H. N., et al. (1975). Statistical Package for the Social
Sciences. New York.
- Overrall, E. J. and Spiegel, K. D. (1969). Concerning Least
Squares Analysis of Experimental Data.
Psychological Bulletin. 72: (5): 311-
322.
- Sen, K. P. (1969). "A Generalization of the T-method of Multiple
Comparisons for Interactions Journal of
American Statistical Association, 64:
290-295.
- Vargas, Ch. D. (1983). El Estado Actual de las Comparaciones
Múltiples de Medias. Tesis Profesional,
UNAM, México.
- Ward, H. J. (1963). Hierarchical Grouping to Optimize an Objective
Function. Journal of American
Statistical Association. 58: 236-244.

Weber, C. D. El Análisis de Diseños Experimentales Utilizando las Terminales de Tiempo Compartido de la Computadoras. Trabajo presentado en las Reuniones de Estadística Efectuadas en Fort Collins. Agosto 23-26 de 1971. Traducción. I. Méndez.