



Acta Botánica Mexicana

ISSN: 0187-7151

rosamaria.murillo@inecol.edu.mx

Instituto de Ecología, A.C.

México

Equihua, Miguel

Análisis de la vegetación empleando la teoría de conjuntos difusos como base conceptual

Acta Botánica Mexicana, núm. 15, 1991, pp. 1-16

Instituto de Ecología, A.C.

Pátzcuaro, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=57401501>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## ANÁLISIS DE LA VEGETACIÓN EMPLEANDO LA TEORÍA DE CONJUNTOS DIFUSOS COMO BASE CONCEPTUAL

MIGUEL EQUIHUA<sup>1</sup>

Instituto de Ecología, A.C.

### RESUMEN

La mayoría de las técnicas de clasificación de muestras de comunidades ecológicas están orientadas a caracterizar grupos discretos. Sin embargo las comunidades ecológicas no están separadas por límites precisos, por el contrario existe mezcla entre ellas. Otra característica de este tipo de clasificaciones es que llevan a suponer que los elementos comprendidos dentro de una clase son todos equivalentes, es decir, todos ellos muestran las mismas características o tienen el mismo rango dentro del grupo. Esto es una simplificación excesiva para el caso de comunidades ecológicas, puesto que se ha demostrado que su estructura varía conforme las especies componentes responden de manera más o menos independiente a los gradientes ambientales. En consecuencia tanto el traslape como la heterogeneidad interna son características importantes de las comunidades ecológicas, atributos que no pueden incorporarse fácilmente con los enfoques convencionales de clasificación. En este artículo se propone que el uso de la teoría de conjuntos difusos, como una base conceptual alternativa, es apropiada para representar las comunidades ecológicas en forma más satisfactoria.

La técnica difusa de *k*-medias se empleó para clasificar una muestra de vegetación de la región de la Reserva de la Biósfera La Michililla, ubicada en el estado de Durango, México. Esta clasificación difusa se contrastó contra la de tipo convencional elaborada con el programa TWINSpan. Se consideró que cuatro grupos representan satisfactoriamente la vegetación del área. Se interpretó que dos de los grupos corresponden a mezclas entre selva baja caducifolia, matorral xerófilo, pastizal y bosque de encino-pino. Los otros dos son bosques de pino y encino. Todos los grupos se pueden ordenar sobre un gradiente principal de aridez.

Los resultados sugieren que una clasificación difusa de la vegetación es apropiada y útil. Se muestra que los grupos formados proporcionan una descripción coherente de las comunidades del área, en la que se conserva la información relativa a la variación natural y a la mezcla entre ellas. Se demuestra también que estos grupos pueden ser analizados en términos de su asociación con variables externas, lo que proporciona información sobre la ecología de las especies componentes y los factores que influyen sobre la estructura de cada comunidad.

### ABSTRACT

Most current techniques for classifying ecological community data are intended to typify discrete groups. However, ecological communities are not separated by distinct boundaries and some mixing between neighbouring communities occurs. Another aspect of these classifications is that they assume that the entities within each group are all equivalent, i.e. all of them will show the same characteristics or will have the same rank within the group. This is an over simplification since the structure of natural communities has been shown to vary as their component species respond more or less independently

---

<sup>1</sup> Dirección actual: University of York, York, YO1 5DD, Inglaterra

to environmental gradients; thus, both overlapping and internal heterogeneity are important features of ecological communities that cannot be incorporated easily in conventional classification approaches. In this paper, it is proposed that fuzzy set theory provides a conceptual basis which overcomes the limitations of conventional approaches.

The fuzzy *k*-means algorithm was used to classify a vegetation sample from the region of the biosphere reserve 'La Michilía', in the state of Durango, México. This fuzzy classification was compared with a conventional classification produced by the TWINSpan program. It was reckoned that four groups were a suitable representation of the vegetation of the area. It was interpreted that two of the groups were mixtures of deciduous tropical forest, xerophytic shrubland, grassland and oak-pine forest. The other two groups are oak-pine forests. All of the groups can be arranged along a main gradient of aridity.

The results suggest that a fuzzy set classification of vegetation data is appropriate and useful. It is shown that the groups formed give a sensible description of the ecological communities while also retaining the information on the natural variation and mixing between them. It is also demonstrated that these groups can be analysed in terms of their association with external variables, providing a good insight into the ecology of the component species and the factors influencing the structure of each community.

## INTRODUCCION

### Conjuntos difusos

Las ideas de la teoría de conjuntos difusos han sido discutidas en la literatura ecológica (por ejemplo Bosserman y Ragade, 1982; McBratney y Moore, 1985; Roberts, 1986 y Dayong, 1988); sin embargo no parecen haber alcanzado todavía una difusión muy amplia. Para los propósitos de este trabajo sólo se requieren algunos de los conceptos más básicos de esta teoría, los que a continuación se discuten. Como es sabido, un conjunto ordinario se define como una agrupación de elementos que comparten ciertos criterios. Se entiende que cualquier elemento puede ser miembro o no de un conjunto, según satisfaga su especificación. Un conjunto difuso se produce cuando los elementos pueden ser miembros *parciales* del mismo conjunto. Una forma de representar esto es definir una variable de *afiliación* que tome el valor '1' si el elemento pertenece al conjunto y '0' en caso contrario (de acuerdo con el concepto de un conjunto ordinario). Si esta idea se extiende para permitir que la variable de afiliación tome cualquier valor entre '0' y '1' el resultado es un conjunto difuso o nebuloso, porque sus fronteras no están claramente definidas. En el caso de los conjuntos ordinarios los criterios que se usan para definir un conjunto deben satisfacerse exactamente. En los conjuntos difusos puede pensarse que esos criterios caracterizan un elemento ideal o típico con el cual cualquier otro elemento tiene un cierto grado de similitud. Este grado de similitud determina su valor de pertenencia en el conjunto.

### Diferencia entre probabilidad y nebulosidad

Es importante destacar la diferencia entre el concepto de probabilidad y la idea de nebulosidad. La probabilidad es una manera de expresar la frecuencia con la que se espera se presente un determinado evento cuando su ocurrencia es incierta, es decir es una medida de la esperanza que existe de que ocurra dicho evento. Por su parte la nebulosidad es una medida de similitud. En ciertas ocasiones, a pesar de esta diferencia, ambos

conceptos resultan más o menos equivalentes, pero en general no lo son. Existe también una semejanza superficial debida a que tanto las probabilidades como el grado de pertenencia se representan usualmente con valores en el intervalo  $[0, 1]$ , pero mientras que para las probabilidades se requiere que sumen 1 al considerar todos los posibles resultados, los valores de afiliación no tienen esta restricción (Kauffman, 1975). Los conjuntos difusos son de utilidad porque hacen posible representar y modelar la incertidumbre resultante de que los elementos de un conjunto puedan satisfacer sólo parcialmente su definición, la que generalmente es distinta de la incertidumbre asociada al proceso de observación, mejor representada en términos de probabilidad. De hecho puede verse que ambas ideas son complementarias.

#### Aplicabilidad en ecología

El concepto de ecosistema lo mismo que los de comunidad, asociación, nivel trófico, etc. pueden mencionarse como ejemplos de conjuntos difusos. En realidad, muchos conceptos y definiciones en Ecología dan lugar a conjuntos difusos al momento de aplicarlos a casos particulares. Por ejemplo, una muestra de bosques de pino no puede satisfacer exactamente una definición cualquiera de asociación. Puede verse que a la variabilidad resultante contribuye, además de la incertidumbre proveniente de las dificultades de muestrear los sistemas ecológicos (probabilística), la incertidumbre derivada de que las definiciones con que se trabaja son relativamente vagas o imposibles de aplicar en forma estricta. La teoría de los conjuntos difusos es útil para la descripción y el análisis bajo estas últimas circunstancias.

En la teoría de conjuntos ordinaria, el concepto de *partición* se aplica a un conjunto cuando es dividido en subconjuntos mutuamente excluyentes que no están vacíos. Esta idea es enteramente equivalente al concepto tradicional de un sistema de clasificación (tal tipo de partición se denomina *dura* en el presente trabajo). En la teoría de conjuntos difusos una partición no consiste necesariamente de subconjuntos mutuamente excluyentes; por el contrario, debe existir traslape al menos en un par de los subconjuntos resultantes (Bezdek, 1981). La indicada característica parece idónea para representar comunidades ecológicas, pues si bien la composición biótica varía más o menos en forma continua a lo largo de gradientes ambientales (Whittaker, 1975; Austin, 1985), usualmente puede reconocerse cierta estructura de comunidad. Esta estructura puede ser importante para explicar diversos fenómenos ecológicos (Roughgarden y Diamond, 1986). De tal manera, la teoría de conjuntos difusos parece capaz de dar cabida a la concepción de las comunidades ecológicas como entidades reconocibles, al mismo tiempo que acepta el hecho de la variación gradual de la composición biótica a lo largo de gradientes ambientales.

El objetivo de la presente contribución es evaluar la viabilidad del uso de conjuntos difusos para describir comunidades ecológicas. Los conjuntos difusos que se utilizan en este trabajo se definen en términos de composición específica (la cual se denomina *centroide*). Los valores de afiliación definen el grado en que cada sitio de muestreo corresponde a cada subconjunto. Bosserman y Ragade (1982) sugieren se recurra al "juicio de expertos" como una forma de estimar los valores de afiliación, que es una posibilidad frecuentemente empleada en otras aplicaciones de conjuntos difusos. En este trabajo, sin embargo, se optó por una forma más objetiva basada en el método de análisis de cúmulos

propuesto inicialmente por Dunn (1974) y más tarde elaborado por Bezdek (1974, 1981, 1987) y Bezdek et al. (1981a, 1981b). El procedimiento se basa en la idea de encontrar un número 'k' de centroides tal que las observaciones estén a la mínima distancia posible de ellos. El método es denominado *k*-medias difuso ("fuzzy *k*-means" o "fuzzy ISODATA" en inglés) y ha sido aplicado con éxito al análisis de vegetación (Equihua, 1990). Como una forma de contrastar la clasificación basada en conjuntos difusos con un enfoque de clasificación más convencional se empleó el programa TWINSpan (Hill, 1979) para producir una partición de tipo ordinario.

## MÉTODOS

### Obtención de una clasificación difusa

El método difuso de *k*-medias se basa en encontrar el mínimo de la siguiente ecuación

$$J(\mathbf{M}, \mathbf{C}, \mathbf{A}) = \sum_{i=1}^n \sum_{j=1}^k (m_{ij})^{\alpha} (d_{ij})^2, \quad (1)$$

en donde

$$(d_{ij})^2 = \|\mathbf{x}_i - \mathbf{c}_j\|_{\mathbf{A}}^2 = (\mathbf{x}_i - \mathbf{c}_j)^T \mathbf{A} (\mathbf{x}_i - \mathbf{c}_j)$$

es la medida de distancia.  $\mathbf{A}$  es la matriz que establece la norma en la que se basa la medida de distancia (esencialmente determina el peso que reciben las distintas variables), su dimensión es  $p$  que es el número de especies consideradas. El exponente  $\alpha$  es el parámetro de nebulosidad ( $1 \leq \alpha < \infty$ ),  $\mathbf{x}_i$  es el vector de abundancias de las especies en el sitio  $i$ , el vector  $\mathbf{c}_j$  es el centroide del grupo  $j$ ,  $m_{ij}$  es el valor de afiliación del sitio  $i$  en el grupo  $j$ .  $\mathbf{M}$  es la matriz de afiliaciones y  $\mathbf{C}$  la matriz de centroides ( $\mathbf{M}=\{m_{ij}\}$  y  $\mathbf{C}=\{\mathbf{c}_j\}$  respectivamente). Las constantes  $n$  y  $k$  son el tamaño de la muestra y el número de grupos a identificar respectivamente.

La ecuación (1) define toda una familia de criterios de agrupamiento como una función del parámetro de nebulosidad. Cuando  $\alpha=1$ , la solución de la ecuación (1) se asocia con una partición dura. Conforme  $\alpha$  crece los grupos se vuelven más difusos. Si bien una partición dura puede no ser enteramente satisfactoria por forzar una demarcación neta entre clases, una participación completamente difusa tampoco es útil (las afiliaciones son  $1/k$  en todos los casos). La experiencia con este método de análisis de conglomerados en otras aplicaciones ha mostrado que cuando  $\alpha=2$  se obtienen resultados satisfactorios

(Dunn, 1974; Bezdek, 1981; McBratney y Moore, 1985). En consecuencia este valor es el que se usó en este trabajo.

Para minimizar la ecuación (1) hay que resolver la siguiente pareja de expresiones

$$c_j = \frac{\sum_{i=1}^n (m_{ij})^2 x_i}{\sum_{i=1}^n (m_{ij})^2}, \quad 1 \leq j \leq k \quad (2)$$

$$m_{ij} = \left[ \sum_{h=1}^k \left[ \frac{(d_{ijA})^2}{(d_{hA})^2} \right]^{(1/\alpha-1)} \right]^{-1}, \quad 1 \leq j \leq k, 1 \leq i \leq n \quad (3)$$

En el procedimiento propuesto por Bezdek (1981) este sistema de ecuaciones es resuelto iterativamente para un número previamente establecido de grupos, una matriz **A** que en este trabajo fue la idéntica o la diagonal,  $\{1/s'_{ii}\}$  donde  $s'_{ii}$  es la varianza de la especie  $i$ , (según los grupos sean relativamente circulares o elipsoidales respectivamente), y una matriz inicial de valores de afiliación. La solución identifica un mínimo local de la ecuación (1). Debe notarse que las afiliaciones, para hacer posible la obtención de una partición difusa, deben sumar 1 para cualquier sitio. Cuando  $x_i = c_j$  (lo que implica  $d_{ijA} = 0$ ) el correspondiente valor de afiliación,  $m_{ij}$ , no puede calcularse con la ecuación (3), pero la afiliación del sitio  $i$  es obviamente 1 en el grupo  $j$  y 0 en los demás. Dada la propiedad de convergencia de este procedimiento a un mínimo local, es claro que diferentes valores iniciales de afiliación pueden desembocar en diferentes mínimos de la ecuación (1). Por lo tanto es importante especificar la estrategia empleada para generar dichos valores.

En muestreos de comunidades ecológicas es muy raro que un par de réplicas tengan similitudes altas, por el contrario, ésta se encuentra por lo general entre 90% y 50% (Gauch, 1982) debido a factores estocásticos, como por ejemplo, la dispersión y el establecimiento de los vegetales. Por supuesto que restricciones en el muestreo también contribuyen a limitar el grado de similitud medible. De acuerdo con Gauch (1982), las técnicas de ordenación recuperan selectivamente la información sobre los patrones ecológicos más relevantes para el conjunto de especies considerado. Al mismo tiempo relegan el ruido (o variabilidad no atribuible a algún factor ecológico en particular) a los ejes asociados con los eigenvalores más pequeños. Aprovechando esta propiedad de la ordenación para reducir el ruido en la muestra, los ejes de ordenación dominantes pueden usarse como las variables sobre las que se efectúa la clasificación. Tal procedimiento es el que se empleó en este trabajo, usando el método de promediación recíproca como técnica de ordenación (Hill, 1973).

De acuerdo con la ecuación (2) los centroides son un promedio ponderado donde las afiliaciones son los pesos. Esto sugiere que una vez obtenidos los valores de afiliación,



la misma ecuación (2) puede utilizarse para calcular los centroides de otras variables de interés asociadas con la muestra que se analiza. Esta circunstancia fue aprovechada para expresar los centroides en términos de especies, cuando la clasificación se llevó a cabo sobre ejes de ordenación.

Para contrastar las diferentes variantes del análisis se aplicó la técnica de correlación canónica con el propósito principal de determinar la proporción de la varianza del conjunto de especies que puede ser explicado con base en la clasificación. El índice de redundancia (Gittins, 1985) mide esta proporción y es el que se usó para efectuar las comparaciones. Tal parámetro estadístico se calcula como

$$Q = \sum_{i=1}^c (\|a_i\|^2/p) r_i^2$$

donde  $a_i$  es el vector  $i$  de correlaciones entre las variables y el correspondiente eje canónico,  $r_i$  es la correlación canónica respectiva,  $p$  es el número de variables en el conjunto considerado y  $c$  es el número de ejes canónicos estimados. Para el cálculo se consideró como uno de los conjuntos de variables a los conteos por especies y como el otro conjunto a las afiliaciones calculadas en el análisis de cúmulos (sólo se incluyeron  $k-1$  variables de afiliación puesto que se calculan con la restricción de que sumen uno). La misma forma de análisis se aplicó a las variables ambientales. También se empleó la correlación canónica para evaluar los resultados producidos por TWINSpan, en este caso la clasificación se representó mediante  $k$  variables indicadoras, una para cada grupo. Las variables indicadoras toman el valor uno para los individuos asignados al grupo correspondiente y cero para los restantes. Al igual que en el caso anterior sólo  $k-1$  variables indicadoras son necesarias para representar la clasificación completa.

#### Descripción de los datos empleados como ilustración

La zona muestreada se localiza en la porción sudoriental del estado de Durango (23° 15' a 23° 45' N y 104° a 104° 20' O). Incluye la totalidad de la Reserva de la Biósfera La Michilía. La topografía es irregular con altitudes que oscilan entre 1700 y 2900 m. Los climas van desde los secos esteparios ( $BS_0$ ) hasta los más secos de los subhúmedos ( $C(w_0)$ ), de acuerdo con la clasificación de Köppen modificada por García (1964). El régimen de lluvias es de verano. Predominan en la zona las rocas extrusivas ácidas. Los suelos son principalmente litosoles. La vegetación consta de bosques de encino y pino, matorral xerófilo, pastizal y bosque tropical caducifolio.

El muestreo se realizó dentro de dos transectos de 56 km de longitud y 10 km de ancho orientados de norte a sur. Su ubicación se planeó con la intención de captar la mayor variabilidad posible debida a la posición geográfica, de acuerdo con el esquema de muestreo en "gradsectos" propuesto por Gillison y Brewer (1985). Dentro de cada sitio de muestreo se ubicaron cuatro parcelas de 50 x 20 m, dispuestas con el eje mayor paralelo a la cota de nivel, una en cada una de las siguientes posiciones: norte, sur, intermedia (entre norte y sur), y sin pendiente apreciable. En cada parcela se contaron los fustes

de los árboles y arbustos más prominentes. La muestra consta en total de 203 parcelas (Austin, Becerra et al., 1984). Para cada parcela se estimaron, con base en información cartográfica, la temperatura promedio anual y la precipitación total anual. También se tomaron medidas de orientación y azimut de las elevaciones que definen el horizonte en torno a cada parcela, para calcular un índice local de la cantidad de radiación incidente relativo a la que recibiría una superficie plana (Austin, Cunningham y Fleming, 1984). Este índice, que tiene usualmente valores bajos en las laderas orientadas en dirección opuesta al Ecuador, se calculó como promedio anual y también para el mes de junio como representativo del verano.

## RESULTADOS

Se hicieron los análisis de conjuntos difusos y TWINSpan tanto sobre la matriz original de conteos como sobre su equivalente de presencias. En primer lugar se efectuaron ordenaciones para determinar el número de ejes a usar en la clasificación difusa. Se consideró que tres ejes son suficientes para resumir la información contenida en la muestra. Se elaboraron clasificaciones difusas basadas en la norma Euclídiana y la diagonal, que corresponden, en este caso, al uso de eigenvectores normalizados a longitud proporcional a sus eigenvalores y longitud unitaria respectivamente. En esta forma la norma Euclídiana implica que la clasificación da mayor peso a los ejes de acuerdo con su dominancia. El procedimiento de asignación inicial de afiliaciones que se empleó consistió en: 1) encontrar el eje de ordenación dominante en la muestra y dividirlo en  $k$  segmentos de igual longitud y 2) asignar a las observaciones encontradas en cada segmento una afiliación de 0.9 en el grupo que corresponde al segmento y  $0.1/(k-1)$  en cada uno de los grupos restantes. El método de ordenación empleado fue el de la promediación recíproca (Hill, 1973).

Como se mencionó en la sección anterior, la vegetación del área puede dividirse en cuatro grupos básicos. En el análisis de cúmulos se consideró que también cuatro grupos producen una partición razonable de la muestra. En el cuadro 1 puede verse que la clasificación que mayor redundancia muestra con el conjunto de especies, es la que se basó en los datos de conteos y utilizó la norma diagonal. Puede verse también que en general hay poca diferencia entre las clasificaciones basadas en conjuntos difusos y que cualquiera de ellas es superior a las clasificaciones duras producidas por TWINSpan. Los valores de redundancia de las especies sobre los grupos son bajos probablemente debido a que no todas las especies son igualmente valiosas para identificar los grupos. Seguramente es posible encontrar un subconjunto de especies que haga óptima la redundancia de las especies sobre los grupos, de modo similar a lo que se hizo con las variables ambientales. Sin embargo no se exploró esta posibilidad para las especies.

Para el análisis de la asociación entre variables ambientales y grupos, así como para comparar las clasificaciones, se seleccionó el subconjunto de variables que hizo óptima la redundancia. Este subconjunto consistió de la temperatura media anual ( $t$ ), la precipitación total anual ( $P$ ) y el cuadrado de ambas. Se incluyeron términos cuadráticos porque es de esperar que las afiliaciones tengan óptimos sobre estos gradientes. Los resultados del análisis de correlación canónica se resumen en el cuadro 2. Nuevamente los resultados para las clasificaciones difusas sugieren que se asocian mejor con las variables ambientales que las clasificaciones duras. También puede notarse que la



Cuadro 1. Resumen del análisis de correlación canónica entre las distintas clasificaciones y el conjunto de especies. Los valores de redundancia son una medida del grado en que la clasificación reproduce el comportamiento de las especies (penúltima columna) y a la inversa (última columna). Los valores de  $F$  corresponden a la aproximación de la  $T^2$  de Hotteling.

		ESPECIES					
Datos	Norma	Correlación Canónica			F (72/524)	Redundancia (%)	
		1	2	3		Grupos Especies	Especies Grupos
CONJUNTOS DIFUSOS							
P/A	Diagonal	89	85	80	20.02	71	20
P/A	Euclidiana	90	85	80	20.09	72	20
Conteos	Diagonal	93	88	82	28.21	78	20
Conteos	Euclidiana	93	88	82	27.92	76	20
TWINSpan							
P/A	_____	90	81	75	18.17	68	19
Conteos	_____	90	82	75	18.27	68	19

Cuadro 2. Resultados del análisis de correlación canónica entre las clasificaciones y las variables ambientales. Se seleccionaron las variables ambientales que hicieron máximos los valores de redundancia. Las variables consideradas son temperatura, precipitación, y el cuadrado de estos dos parámetros. La redundancia de grupos/variables es la proporción de la varianza de las variables ambientales explicable con la clasificación. La redundancia de variables/grupos indica el grado en que la clasificación puede ser descrita con base en las variables ambientales. Los valores de  $F$  corresponden a la aproximación de la  $T^2$  de Hotteling.

VARIABLES AMBIENTALES							
Datos	Norma	Correlación Canónica			F (12/584)	Redundancia (%)	
		1	2	3		Grupos Variables	Variables Grupos
CONJUNTOS DIFUSOS							
P/A	Diagonal	90	83	36	107.89	60	69
P/A	Euclidiana	90	84	36	112.54	62	70
Conteos	Diagonal	89	73	25	78.03	51	66
Conteos	Euclidiana	89	73	25	78.90	53	66
TWINSpan							
P/A	_____	89	83	32	100.24	46	68
Conteos	_____	89	84	34	102.00	46	67

clasificación que mayor redundancia mostró con las especies tiene un mínimo con las variables ambientales. Otro aspecto interesante es que los valores de redundancia de las variables ambientales sobre los grupos sugieren que es posible expresar las afiliaciones como una función de la temperatura y la precipitación.

A pesar de que hay un buen grado de asociación entre las variables ambientales y las clasificaciones difusas, puede considerarse que la redundancia con el conjunto de especies es una mejor indicación de la calidad de la clasificación, puesto que el grupo de variables ambientales elegido puede no incluir a los factores más importantes para explicar la variación de las especies y porque, por la forma como se calculan las estimaciones, están más sujetas a error. Esto también puede explicar la aparente relación inversa entre la redundancia de las especies con la de las variables ambientales.

Los centroides de la clasificación basada en conteos y norma diagonal se muestran en el cuadro 3. Para la interpretación de esta tabla una especie se puede considerar como elemento característico del grupo en el que muestra valores más altos. Su grado de asociación es proporcional al grado de contraste que muestre entre los diferentes grupos, es decir las especies más fuertemente asociadas con un grupo son aquellas que tienen valores altos en ese grupo y bajos en todos los demás. Este es el criterio que se usó para ordenar las especies por grupos en el cuadro 3. Las funciones de afiliación se graficaron como mapas del área en la figura 1.

El grupo 1 está caracterizado por *Calliandra eriophylla*, *Bursera fagaroides* y *Dasyliirion durangense*. Estas especies crecen en ambientes cálidos y secos, lo que está claramente indicado en este análisis, pues el grupo se asocia con la máxima temperatura y la mínima precipitación entre los grupos (Cuadro 4). *Acacia schaffneri* aparece asociada con la transición de la vegetación de este grupo hacia el grupo 2, lo mismo que *Opuntia leucotricha*, *Agave* sp. y *Mimosa biuncifera*. En esta forma el grupo 1 comprende comunidades transicionales entre el matorral xerófilo y la selva baja caducifolia.

El grupo 2 resulta definido por *Opuntia leucotricha*, *O. robusta*, *Mimosa biuncifera* y *Agave* sp., que son componentes típicos de los matorrales xerófilos y pastizales áridos del Altiplano Mexicano, y *Quercus grisea*. Este grupo está asociado con suelos someros (muestra la mínima frecuencia en suelo profundo en el cuadro 4). La transición entre matorral y bosque de pino y encino se caracteriza en el área por la presencia de *Q. grisea*, *Q. eduardii*, *Juniperus deppeana*, *Arctostaphylos pungens* y *Pinus cembroides*. Es interesante notar que el análisis sugiere que esta última especie se comporta de una manera bimodal, pues muestra dos máximos de abundancia uno en el grupo 2 y otro en el 4.

El grupo 3 está caracterizado por los encinos *Quercus durifolia*, *Q. rugosa*, *Q. sideroxyla* y *Q. eduardii* junto con *Pinus durangensis*, *P. arizonica* y *Juniperus deppeana*. En el estrato arbustivo están presentes *Arctostaphylos pungens* y *Pithecellobium leptophyllum*. Estas especies son comunes en los bosques de pino y encino, pero algunas de ellas también se encuentran en comunidades mezcladas con pastizal (Maysilles, 1959) o matorral xerófilo (Shelford, 1963; Rzedowski, 1978). De esta manera los elementos de este grupo pueden describirse como bosques de pino y encino de habitats secos.

Las especies asociadas con el grupo 4 son *Quercus urbanii*, *Pinus engelmannii*, *P. chihuahuana*, *P. cembroides*, *Arbutus xalapensis* y *A. glandulosa*. Tales árboles son

Cuadro 3. Promedio de abundancia de las especies en cada uno de los cuatro grupos difusos obtenidos de la clasificación. Las frecuencias se obtuvieron aplicando la ecuación (2) a la matriz de presencia/ausencia derivada de los datos originales. Las especies distintivas de un grupo son aquellas que muestran valores altos en tal campo y bajos en los demás. Este criterio se empleó para ordenar las especies en los cuatro grupos sugeridos.

Especies	CENTROIDES							
	Densidad (individuos/ha)				Frecuencia (%)			
	1	2	3	4	1	2	3	4
<i>Calliandra eriophylla</i> Benth.	443	0	2	1	86	0	3	0
<i>Bursera fagaroides</i> HBK.	71	2	1	0	77	7	1	1
<i>Dasyllirion durangense</i> Trel.	87	5	7	24	66	16	8	16
<i>Acacia schaffneri</i> (Wats.) Hermann	36	17	6	5	84	33	1	1
<i>Opuntia robusta</i> Wendl.	1	13	1	1	5	54	7	10
<i>Quercus grisea</i> Liebm.	19	147	24	18	16	95	27	36
<i>Mimosa biuncifera</i> Benth.	20	53	2	2	15	47	2	1
<i>Agave</i> sp.	23	39	7	2	49	38	5	2
<i>Opuntia leucotricha</i> DC.	35	40	1	2	78	62	4	6
<i>Pithecellobium leptophyllum</i> (Cav.) Dav.	0	0	24	1	0	0	26	1
<i>Quercus durifolia</i> von Seeman	0	1	17	0	0	3	16	0
<i>Pinus durangensis</i> Martínez	0	0	37	3	0	1	51	3
<i>Arctostaphylos pungens</i> HBK.	7	15	426	18	1	26	94	59
<i>Pinus arizonica</i> Engelm.	0	1	38	7	0	3	44	2
<i>Juniperus deppeana</i> Steud.	1	7	40	4	12	34	65	18
<i>Quercus rugosa</i> Née	1	1	70	27	0	0	65	56
<i>Quercus sideroxyla</i> Humb. et Bonpl.	1	0	97	45	0	0	36	34
<i>Quercus eduardii</i> Trel.	1	42	86	31	1	53	60	27
<i>Quercus urbanii</i> Trel.	1	0	9	211	1	2	42	85
<i>Pinus engelmannii</i> Car.	0	0	1	16	1	1	6	46
<i>Pinus chihuahuana</i> Engelm.	0	1	11	92	0	0	23	21
<i>Arbutus xalapensis</i> HBK.	0	1	17	43	0	3	42	82
<i>Arbutus glandulosa</i> Mart. et Gal.	0	0	10	23	0	0	30	56
<i>Pinus cembroides</i> Zucc.	3	26	2	66	12	36	3	55

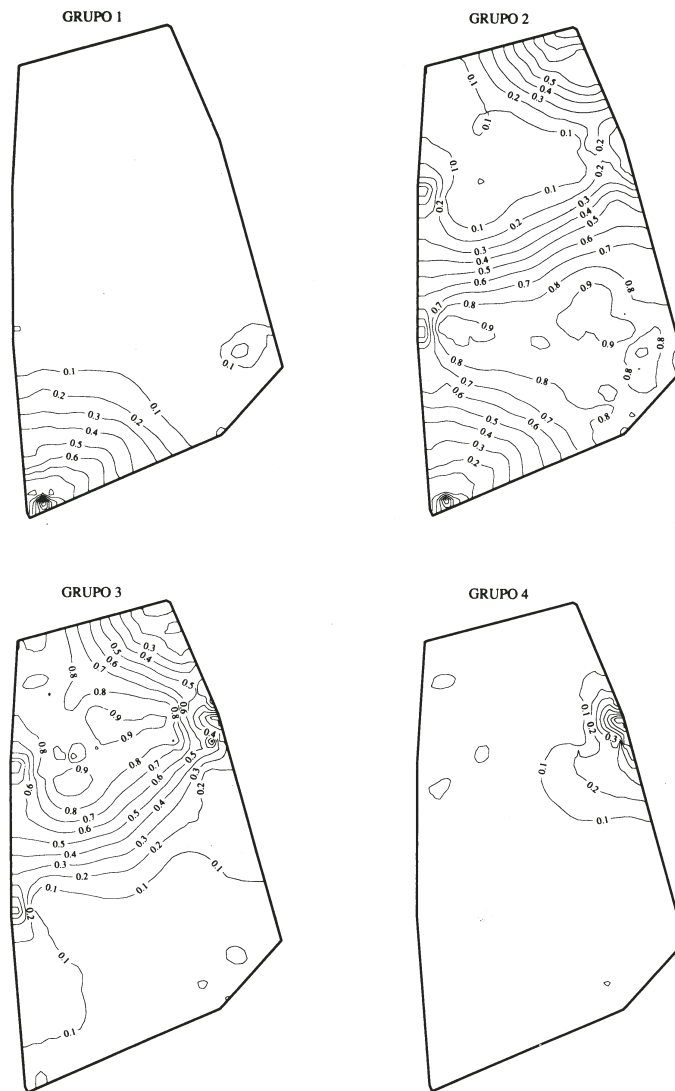


Fig. 1. Distribución geográfica de los grupos identificados con la clasificación difusa. Cada grupo se ilustra en un mapa. Las isolíneas separan niveles en los valores de afiliación dentro del grupo.

Cuadro 4. Promedios de las variables ambientales en cada uno de los grupos difusos. Los valores se obtuvieron aplicando la ecuación (2) a cada una de las variables indicadas.

Variables	CENTROIDES			
	1	2	3	4
Altitud (m)	1925	2169	2478	2627
Temperatura media (°C)	18.6	16.2	13.0	12.2
Precipitación total (mm)	412	500	591	637
Radiación de verano	0.93	0.96	0.94	0.92
Radiación anual	0.96	0.95	0.94	0.90
Frecuencia en suelo profundo (%)	46	23	46	36

componentes frecuentes de los bosques de las montañas de la Sierra Madre de Durango (Maysilles, 1959). De acuerdo con los resultados del análisis, las especies de este grupo crecen en las porciones más frescas y semihúmedas de la zona. El valor del índice de radiación anual sugiere (Cuadro 4) que este grupo prefiere las laderas protegidas. *P. cembroides* parece un poco fuera de lugar en el grupo. Era de esperar que esta especie fuera representante de los grupos 2 ó 3, sin embargo el análisis sugiere un comportamiento bimodal que es difícil de explicar sin información adicional. No obstante, este mismo tipo de patrón ha sido observado en otras ocasiones para *Dasyllirion* spp. (J. Rzedowski, comunicación personal) y es también aparente en los resultados de este análisis (Cuadro 3).

La distribución de los grupos de la clasificación de TWINSpan en el área se ilustra en la figura 2 junto con un mapa altimétrico de la Michilfa. Puede verse el alto grado de correspondencia entre TWINSpan y la clasificación difusa. Las isolíneas de 0.4 de afiliación coinciden aproximadamente con los límites entre clases producidos por TWINSpan. La correspondencia entre la distribución geográfica de los grupos y la topografía es notable como puede verse si se comparan los mapas de afiliación con el de altitud que se muestra.

## DISCUSION

El patrón de comunidades identificado tanto por TWINSpan como por la clasificación difusa, corresponde bien con lo que se puede esperar para la región (Gentry, 1957; Shelford, 1963; Miranda y Hernández, 1964; Rzedowski, 1978; González-Elizondo, 1983). En la cuenca del río Mezquital, que es la porción con menor altitud, se encuentran las comunidades termófilas (grupo 1). Conforme se gana altitud la vegetación se transforma en matorrales xerófilos y pastizales, que son los que dominan en los valles intermontanos (grupo 2). Los bosques de pino y encino se presentan en las partes altas (grupos 2 y 4), con su variante más húmeda principalmente en la Sierra de Urica (al noreste de la región). Ambos métodos de análisis identifican esencialmente el mismo patrón de

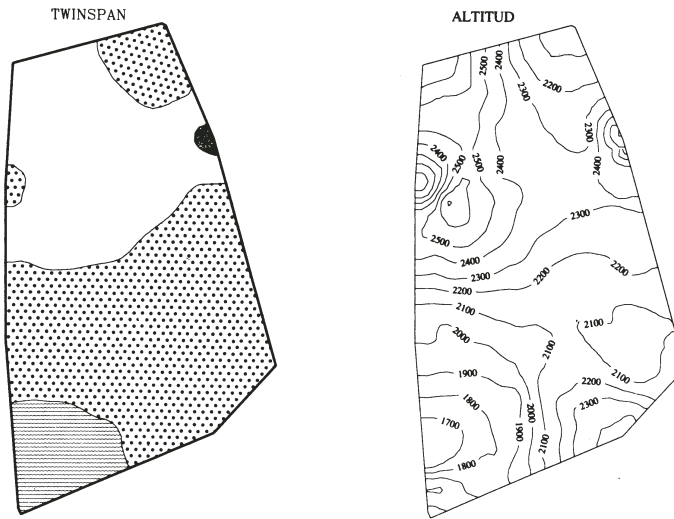


Fig. 2. En el mapa producido con base en la clasificación de TWINSPAN los grupos están indicados como sigue: 1) líneas onduladas, 2) punteado, 3) en blanco y 4) sombreado irregular. El mapa altimétrico se produjo con base en los datos de altitud registrados para cada una de las muestras.

comunidades vegetales. No obstante, la clasificación difusa permite visualizar la forma como varía la vegetación de acuerdo con los gradientes ambientales, muestra que dentro de cada grupo hay heterogeneidad y da oportunidad de analizar los patrones de traslape o mezcla entre comunidades. Todos estos aspectos se pierden en una clasificación dura.

Un ejemplo interesante que ilustra claramente las ventajas del enfoque de conjuntos difusos, es la distribución de los bosques semihúmedos de encino y pino (grupo 4). Puede verse fácilmente en el mapa correspondiente, que los valores de afiliación disminuyen más lentamente sobre las laderas sudoccidentales que sobre las nororientales, en la Sierra de Urica. Esto se puede explicar si se considera que las laderas sudorientales deben ser más húmedas, puesto que reciben los vientos cargados de humedad que vienen del Pacífico y que la Sierra de Urica es un obstáculo orográfico importante que puede provocar que ocurra condensación. Esto mismo produce una "sombra pluvial" en las laderas de la vertiente opuesta, por lo que rápidamente cambian las condiciones ambientales propicias para el bosque subhúmedo de encino y pino (de hecho la vegetación se convierte



rápida en matorrales xerófilos en esta dirección), lo que claramente se ve reflejado en los valores de afiliación.

En cierta forma la clasificación difusa tal como se aplicó en este trabajo reúne características tanto de ordenación como de clasificación, pero sobre todo extiende las posibilidades del análisis de gradientes convencional. Las técnicas de ordenación están orientadas preferentemente a la identificación de 'ejes' sobre los que la vegetación muestra mayor variación, lo que frecuentemente permite reexpresar la muestra en términos de un número reducido de variables compuestas. Estos ejes son técnicamente infinitos en longitud, pero obviamente, para fines prácticos, interesa solamente la región o regiones ocupadas por la muestra, por lo que el siguiente paso es identificar los patrones que pueden seguir subconjuntos de la muestra sobre esas nuevas variables. Una forma de abordar esto es localizar "centros de masa", lo que da cabida a la aplicación de técnicas de clasificación. Los métodos basados en conjuntos difusos son particularmente útiles en este sentido, debido a la importancia que tiene conservar la información relativa a los cambios de la vegetación en respuesta a los gradientes ambientales. Los resultados de este trabajo apoyan este punto de vista y sugieren que el método de *k*-medias difuso puede ser útil para el análisis de la vegetación.

Como se puede apreciar claramente en los resultados del análisis de la vegetación de la Michilía, el uso de conjuntos difusos permite concebir comunidades dentro de la noción de gradientes ambientales y la existencia de gradientes dentro de cada comunidad. Estos últimos, además, pueden ser distintos para cada una, en correspondencia con la riqueza de formas que la vegetación muestra en la realidad. Lo anterior permite pensar que las comunidades pueden formarse como respuesta independiente a los diferentes gradientes ambientales. El análisis de la asociación de las variables de afiliación con variables externas brinda la oportunidad de obtener información acerca de los factores que influyen sobre la estructura de las comunidades y la ecología de las especies que las constituyen. Esta información puede ser más detallada que la que se obtiene con las técnicas tradicionales de clasificación. Tal forma de representar (modelar) las comunidades ecológicas es útil y ofrece amplias posibilidades de desarrollo.

#### AGRADECIMIENTOS

El artículo es una versión revisada de una parte de mi tesis de maestría presentada en la Universidad de York, mientras recibía el apoyo económico del CONACYT y del British Council (1987- 1988). Al Dr. M.B. Usher mi agradecimiento por su apoyo y orientación durante el desarrollo de este proyecto. Los datos de la vegetación de La Michilía empleados para ilustrar este trabajo se tomaron del muestreo diseñado por el Dr. M.P. Austin durante su visita a México en 1984, con autorización del mismo; junto con él colaboramos en el trabajo de campo la Biól. J. Becerra, Biól. A. Carrillo, M. en C. M. Equihua, Dr. E. Ezcurra y Biól. J. López-Portillo. A la Biól. G. Benítez agradezco la lectura crítica del manuscrito y sus sugerencias para mejorarlo.

## LITERATURA CITADA

- Austin, M.P. 1985. Continuum concept, ordination methods, and niche theory. *Ann. Rev. Ecol. Syst.* 16:39-61.
- Austin, M.P., R.B. Cunningham y P.M. Fleming. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. *Vegetatio* 55: 11-27.
- Austin, M.P., J. Becerra, A. Carrillo, M. Equihua, E. Ezcurra y J. López-Portillo. 1984. Análisis de la respuesta funcional de especies arbóreas y arbustivas en la Reserva de la Biosfera de la Michilila, Dgo. y áreas aledañas, mediante el uso de métodos de análisis directo de gradientes. *Resúmenes IX Congreso Mexicano de Botánica*. México, D.F. p. 150.
- Bezdek, J.C. 1974. Numerical taxonomy with fuzzy sets. *J. Math. Biol.* 1: 57-71.
- Bezdek, J.C. 1981. Pattern recognition with fuzzy objective function algorithms. Plenum Press. New York. 256 pp.
- Bezdek, J.C. 1987. Some non-standard clustering algorithms. In: Legendre P. and Legendre L. (eds.). *Developments in numerical ecology (NATO ASI Series Vol. G14)*. Springer-Verlag. Berlin, pp: 225-287.
- Bezdek, J.C., C. Coray, R. Gunderson y J. Watson. 1981a. Detection and characterization of cluster substructure: I. Linear structure: Fuzzy c-lines. *SIAM J. Appl. Math.* 40: 339-357.
- Bezdek, J.C., C. Coray, R. Gunderson y J. Watson. 1981b. Detection and characterization of cluster substructure: II. Fuzzy c-varieties and convex combinations thereof. *SIAM J. Appl. Math.* 40: 358-372.
- Bosserman, R.W. y R.K. Ragade. 1982. Ecosystems analysis using fuzzy set theory. *Ecol. Model.* 16: 191-208.
- Dayong, Z. 1988. An index to measure the strength of relationship between community and site. *Ecol. Model.*, 40: 145-153.
- Dunn, J.C. 1974. A fuzzy relative of the ISODATA process and its use in detecting compact, well separated clusters. *J. Cybern.* 3: 32-57.
- Equihua, M. 1990. Fuzzy clustering of ecological data. *J. Ecol.*, 78: 519-534.
- García, E. 1964. Modificaciones al sistema de clasificación climática de Köppen (para adaptarlo a las condiciones de la República Mexicana). 2a. Ed. Universidad Nacional Autónoma de México. México, D.F. 246 pp.
- Gauch, H.G. 1982. Noise reduction by eigenvector ordinations. *Ecology* 63(6):1643-1649.
- Gentry, H.S. 1957. Los pastizales de Durango, estudio ecológico, fisiográfico y florístico. Ediciones del Instituto Mexicano de Recursos Naturales Renovables. México, D.F. 361 pp.
- Gillison, A.M. y K.R.W. Brewer. 1985. The use of gradient directed transect or gradsects in natural resource surveys. *J. Environ. Manag.* 20: 103-127.
- González-Elizondo, S. 1983. La vegetación de Durango, CIIIR-IPN-Unidad Durango (Cuadernos de Investigaciones Tecnológicas vol.1, no. 1), Vicente Guerrero, Dgo. 114 pp.
- Gittins, R. 1985. Canonical analysis. (Biomathematics Vol. 12). Springer-Verlag. Berlín. 351 pp.
- Hill, M.O. 1973. Reciprocal averaging: an eigenvector method of ordination. *J. Ecol.*, 61: 237-249.
- Hill, M.O. 1979. TWINSPLAN-a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. *Ecology and Systematics*. Cornell University. Ithaca, New York. 90 pp.
- Kauffmann, A. 1975. Introduction to the theory of fuzzy subsets. (Vol. I. Fundamental theoretical elements). Academic Press. New York. 416 pp.
- Maysilles, J.H. 1959. Floral relationships of the pine forests of western Durango, Mexico. Tesis Doctoral. University of Michigan. Ann Arbor. 177 pp.
- McBratney, A.B. y A.W. Moore. 1985. Application of fuzzy sets to climatic classification. *Agric. For. Meteorol.* 35: 165-185.
- Miranda, F. y E. Hernández X. 1964. Fisiografía y vegetación In: Beltrán, E. (ed.). *Recursos naturales de las zonas áridas del centro y noreste de México*. Instituto Mexicano de Recursos Naturales Renovables. México D.F. pp. 1-27.

- Roberts, D. W. 1986. Ordination on the basis of fuzzy set theory. *Vegetatio* 66: 123-131.
- Roughgarden, J. y J. Diamond. 1986. Overview: The role of species interactions in community ecology. In: Diamond, J. and T. J., Case (eds.). *Community ecology*. Harper and Row. New York. pp.333-343.
- Rzedowski, J. 1978. *Vegetación de México*. Limusa, México, D.F. 432 pp.
- Shelford, V.E. 1963. *The ecology of North America*. University of Illinois Press. Urbana. 610 pp.
- Whittaker, R.H. 1975. *Communities and ecosystems*. 2a. Ed. MacMillan. New York. 162 pp.