

Experiencias en el Uso de Criterios de Similitud en el Estudio de Comunidades*

Experiences in the use of similarity criteria in community studies

FRANCISCO SAIZ

Laboratorio de Ecología, Universidad Católica de Valparaíso,
Casilla 4059, Valparaíso, Chile

(Recibido el 30 de enero de 1980)

SAIZ, F. Experiencias en el uso de criterios de similitud en el estudio de comunidades. (Experiences in the use of similarity criteria in community studies). *Arch. Biol. Med. Exp.* 13: 387-402, 1980.

This paper is a systematic presentation of the author's experience in the use of similarity criteria for the study of communities. Within a general framework, some of the similarity criteria are discussed. Some parameters to quantify them are given together with a sketch of the areas where they can be applied. Examples of their use and interpretation are given.

El presente trabajo tiene por finalidad exponer algunos criterios de similitud, usados en el análisis de comunidades, y la utilización de algunos índices matemáticos para cuantificar la información.

Junto con entregar el marco conceptual general bajo el cual se ha trabajado, se hará un breve análisis de los criterios de similitud, indicando algunos de sus principales campos de aplicación.

En cuanto a los instrumentos matemáticos para valorar la similitud, se entregará una selección de ellos, acompañados de algunos comentarios sobre su naturaleza y su uso.

El trabajo termina con consideraciones sobre la interpretación de la información y con ejemplos tomados de la experiencia personal del autor.

La intención de este trabajo es que tenga características de introductorio al análisis de similitudes y que sirva de estímulo a quienes lo comiencen a utilizar para que lo profundicen.

MARCO CONCEPTUAL GENERAL

Considerando una comunidad como un sistema compuesto por múltiples elementos bióticos, interdependientes, con una gama muy amplia de relaciones de calidades e intensidades muy variables y dependientes de las condiciones abióticas, podríamos aceptar, conceptualmente, que las comunidades están conformadas por un núcleo central altamente cohesionado y por esferas periféricas concéntricas con nexos cada vez más débiles respecto al centro.

El núcleo central estaría estructurado por los elementos más caracterizantes de la comunidad, en cuanto a otorgarle una individualidad definible, siendo los nexos entre sus constituyentes los más fuertes y tipificantes dentro de todo el complejo comunitario.

A medida que a este núcleo se asocian elementos menos intensamente relacionados, se van estructurando las esferas periféricas. Los elementos de estas esferas más externas tienen mayores posibilidades de pertenecer, simultáneamente, a esferas peri-

* Trabajo presentado en el Coloquio "Aplicación de Métodos Cuantitativos en Ecología de Poblaciones y Comunidades" (Santiago, 2 y 3 de agosto de 1979).

féricas de dos o más comunidades, sirviendo de unión entre ellas y pasando así a configurarse las meso y macrocomunidades, quedando las primeras como subsistemas unitarios de las últimas.

El modelo expuesto no se refiere solamente a una expresión espacial, temporal, trófica, etc., de las especies de la comunidad, sino a la resultante ecológica de todos los esquemas de organización.

Unidades de estudio con un núcleo difuso y gran concentración de elementos bióticos en la periferia no configurarían comunidades propiamente tales, sino muestras parciales, no representativas, o muestras de zonas de superposición de esferas periféricas de comunidades adyacentes.

El modelo expuesto implica un esquema de jerarquización de la importancia ecológica de las especies en cada comunidad, el que la individualiza estructural y funcionalmente, incluso en cualquier condición de estado dentro de su proceso sucesional. En este momento entramos en el campo de la tipificación de comunidades.

Caracterizar una comunidad es, entonces, detectar y caracterizar fundamentalmente su núcleo, siendo el análisis de su periferia necesario para establecer sus dependencias y su participación en unidades mayores.

En consecuencia, la individualización de una comunidad se da en razón inversa a la similitud con otras comunidades y directa a su diferenciación.

Como medida de la importancia ecológica de las especies se han considerado una serie de criterios parciales de valoración, suponiéndose que cada uno de ellos la estima con mayor o menor grado de precisión.

El más usado es el de abundancia de individuos por especie, el que se ha tratado de perfeccionar utilizando la biomasa. Otro, corresponde al uso de medidas de energía para ponderar dicha importancia. Personalmente, he utilizado también el criterio de constancia espacial o temporal de las especies, es decir, la frecuencia de ocurrencia en el espacio o en el tiempo.

CRITERIOS DE SIMILITUD Y ANÁLISIS COMUNITARIO

Es indudable que la densidad de una población es producto del grado de encaje entre su rango de tolerancia y el rango ambiental existente, incluyendo las limitantes bióticas de relaciones con otras especies o poblaciones (resistencia ambiental *versus* potencial biótico).

Dos poblaciones de una misma especie, analizadas en una misma época del año y que presentan diferentes densidades, reflejan condiciones ambientales diferentes, siendo distinta su posición jerárquica (en orden de importancia ecológica) en cada una de las situaciones consideradas. Ello conlleva a que serán poblaciones de otras especies las que asuman, en una de ellas, las posiciones prioritarias en la jerarquización numérica.

Desde este punto de vista dos grandes modelos pueden discutirse:

Modelo A. La presencia exclusiva de una especie en una comunidad en estudio refleja la existencia de un cierto rango ambiental diferente al de otra comunidad en la que la especie no existe, particularmente si las comunidades son espacialmente cercanas, ya que la posibilidad de llegada simultánea o de transferencia es alta, a diferencia de comunidades muy distantes en que la ausencia puede deberse a la baja probabilidad histórica de llegada simultánea o de transferencia actual.

Por lo tanto, un criterio de "presencia-ausencia" de especies que pondere similitudes, permitirá, en alguna medida, medir esquemas jerárquicos de importancia de especies, ya que en este caso se estaría aceptando la presencia como exponente del pleno encaje "rango de tolerancia/rango ambiental", y la ausencia como inexistencia del mismo.

Obviamente este planteamiento es válido sólo en el análisis comparativo en una unidad de tiempo. La proyección temporal del criterio significa valorar los cambios en el tiempo de los rangos ambientales mediante las modificaciones sufridas por la relación presencia-ausencia de especies, pudiendo detectarse desde la igualdad tempo-

ral de la unidad en estudio hasta la definición de períodos característicos, pasando por la posibilidad de secuencias graduales de paso en el tiempo.

Este criterio se basa en la isovaloración de las especies, no considerando su expresión cuantitativa.

Modelo B. Si al criterio anterior se le exige una ponderación de la importancia de las especies o poblaciones, se establecerán jerarquizaciones con una gama muy amplia de categorías, a diferencia del modelo A, en que las poblaciones se califican en dos condiciones extremas solamente y, por lo tanto, débil reflejo de una realidad compleja. Este nuevo criterio entrega un espectro más fino de ponderaciones de similitudes.

Si por ejemplo, al analizar dos unidades ecológicas se establece que el 40% de las especies son comunes, suponemos que el 40% de ellas tienen rangos de tolerancia que encajan en los rangos ambientales ofrecidos por ambas comunidades y que el 60% restante acoplaría en una u otra exclusivamente.

Estudiado más a fondo el 40% común, podemos detectar en él una gama de valores en la expresión del valor de importancia, desde la isovaloración en ambas unidades hasta la presencia mínima en una y muy abundante en la otra, tomando, la misma especie, distinta posición jerárquica en cada unidad en estudio, reflejo de rangos ambientales parcialmente diferentes.

Luego, la condición de "común" establecida con el criterio de "presencia-ausencia" puede desplazarse a la de casi no común con el que considera las frecuencias, o bien, acentuarse aún más por correspondencia de esquemas de jerarquización.

Al primer criterio lo he llamado "similitud taxonómica", ya que mide el grado de semejanza en función de la presencia o ausencia de especies, y al segundo "similitud biocenótica", ya que en alguna forma mide la similitud a través de valores de importancia ponderados bajo criterios más propios de la expresión de las biocenosis. Cancela da Fonseca (1966) engloba ambos criterios bajo el nombre de "afinidad cenótica". Sin embargo, considero más adecuado el uso del término similitud que el de

afinidad por la connotación de medida de relación directa que tiene el último.

El análisis precedente se ha hecho basándose en la importancia ecológica del número de individuos por especie. Iguales consideraciones generales son válidas si consideramos cualquier otro valor de importancia, como biomasa o energía, constancia, etc.

En una medida no despreciable, los criterios expuestos son reflejo, o se entroncan, con la diversidad de las unidades ecológicas tratadas. El criterio taxonómico con la variedad de especies y el biocenótico con los esquemas de diversidad de las especies.

En conclusión, al aplicar criterios de similitud se comparan esquemas jerárquicos de componentes de la organización. Ninguno de los criterios de similitud y de valoración de la importancia ecológica puede considerarse como globalizante integral en la comparación de esquemas de organización, recomendándose por ello el análisis conjunto de varios de ellos para obtener una aproximación más acertada de la realidad.

Análisis conjunto significa, en este caso, el uso combinado de criterios de similitudes taxonómica y biocenótica y de diferentes valores de importancia en la última.

Como resultante de un estudio de este tipo, se pueden obtener dos resultados extremos: a) coincidencia de todos los esquemas, algo muy improbable, pues raramente las unidades ecológicas tienen las mismas especies, y si llegaran a tenerlas, es difícil que se dé idéntica jerarquización de valores; b) no coincidencia de esquemas, algo también muy improbable, ya que en alguna medida existe relación entre biomasa, talla, número de individuos, energía en juego, constancia espacial, persistencia temporal, etc. Entre ambos extremos existe un amplio rango de coincidencias parciales de esquemas.

Es importante señalar que los criterios de similitud son complementarios con otros criterios de análisis de comunidades y que el grado de precisión alcanzado por un análisis de este tipo dependerá, entre otras variables, del valor del criterio de importancia utilizado, de la calidad de la informa-

ción base y de la información real que entreguen los instrumentos matemáticos utilizados. La interpretación de los mismos dependerá de la precisión anterior, de los conocimientos básicos y específicos del investigador y de la capacidad de interrelacionar de éste.

CAMPO DE APLICACION DE LOS CRITERIOS DE SIMILITUD

A modo ilustrativo se presentarán los campos o temas en los cuales es factible y útil el uso de criterios de similitud sin pretender agotar las posibilidades de aplicación en ecología.

Al utilizar los términos esquemas de organización u organización se estará pensando en "estructuras de especies", aunque lo que se diga es perfectamente aplicable a otros aspectos.

A. Sobre la organización de la unidad ecológica.

I. Dentro de una misma unidad ecológica.

Ia. Análisis espacial.

1. Definición de la unidad ecológica por análisis de múltiples muestras. Permite llegar a establecer, con cierta confiabilidad el núcleo fundamental de la unidad ecológica.
2. Relacionar asociaciones o subunidades ecológicas para vislumbrar la red organizativa a ese nivel. Por ejemplo, follaje, ramaje, inflorescencia, etc., consideradas, cada una de ellas, desde el punto de vista comunitario.
3. Detección de estratificaciones y zonaciones en las unidades ecológicas en estudio, al poder visualizar subunidades homogéneas espaciales (verticales u horizontales) y su heterogeneidad entre sí.

Ib. Análisis temporal.

1. Estudio de la secuencia de cambios de organización para detectar la unidad temporal de la comunidad o unidad ecológica,

es decir, definir el ciclo completo de sus fases fenológicas.

2. Detección y caracterización de períodos de organización gradual y secuencia. Por ejemplo, estudios fenológicos.

Ic. Análisis espacio-temporal

1. La utilización combinada de la información base permite este enfoque al integrar ambos tipos de análisis.

II. Entre unidades ecológicas

IIa. Análisis espacial

1. Evaluación de la composición, diferencial o no, de varias unidades ecológicas. Lleva a valorar el grado de dependencia entre unidades en comparación.
2. Relacionar esquemas de organización de asociaciones o subunidades particulares entre unidades ecológicas mayores; por ejemplo, subunidad follaje.
3. Establecimiento de gradientes de unidades ecológicas por comparación a una de ellas, significando una ordenación de esquemas de organización relativos a un patrón preestablecido.

IIb. Análisis temporal

1. Valoración de similitudes de esquemas temporales globales de organización.
2. Establecimiento de similitudes de secuencias temporales de organización. Por ejemplo, análisis comparado de secuencias de períodos o fases temporales entre dos o más unidades ecológicas. Por ejemplo, comparación de esquemas fenológicos.
3. Valoración de la dependencia temporal entre unidades ecológicas adyacentes o cercanas. Por ejemplo, interdependencia de esquemas fenológicos.

IIc. Análisis espacio-temporal.

1. Utilización combinada de la información base. Su campo de

aplicación es la combinación del de sus componentes.

B. Sobre las especies.

1. Valoración de la tendencia a la coexistencia espacial de especies. Definición de grupos por afinidades ecológicas.
2. Valoración de la tendencia a la coexistencia temporal de especies. Definición de grupos por afinidades ecológicas.
3. Valoración de la tendencia a la coexistencia espacio-temporal de especies. Definición de grupos por afinidades ecológicas.

Es indudable que la tendencia a la coexistencia de especies aumenta la probabilidad de relaciones entre ellas y, por lo tanto, puede interpretarse como una medida de centros de alta afinidad específica, defintorios de tipos de importancia ecológica.

La similitud medida entre especies en base a comparación de inventarios podría denominarse "similitud específica", para diferenciarla, por su intención, de las taxonómica y biocenótica, en que se estudian inventarios en base a especies.

INSTRUMENTOS MATEMATICOS SUSCEPTIBLES DE SER UTILIZADOS

A continuación se expondrán algunos de los instrumentos matemáticos que son usa-

dos para valorar la información base con criterios de similitud.

La enumeración no pretende ser exhaustiva ni ser una selección jerárquica de los mismos, ya que cada investigador debe seleccionarlos según la naturaleza del muestreo y según los objetivos concretos de su trabajo. Ningún instrumento matemático puede arreglar las deficiencias de origen de los datos ni las interpretaciones erradas de los resultados.

A. Instrumentos destinados a valorar la presencia-ausencia

Con el fin de facilitar la comparación de los índices que se expondrán a continuación, se ha utilizado para todos ellos la misma simbología. Llamaremos "elemento" a un inventario cuando se analiza la similitud entre especies, o a una especie cuando se analiza la similitud entre inventarios; y "condición" a cada especie o a cada inventario, cuya similitud se esté midiendo.

Simbología:

- a = elementos exclusivos de la condición A.
- b = elementos exclusivos de la condición B.
- c = elementos comunes por presencia a las condiciones A y B.
- d = elementos comunes por ausencia a las condiciones A y B.
- n = total de elementos considerados en la dupla A y B = a+b+c+d.

1. Indices de uso corriente (primera parte)*

1. Indice de Jaccard (1908)

$$S_J = \frac{c}{a + b + c}, \quad 0 \leq S_J \leq 1$$

2. Indice de Dice (1945)

$$S_D = \frac{2c}{a + b + 2c}, \quad 0 \leq S_D \leq 1$$

* Bibliografía original de índices puede ver en Cancela da Fonseca (2), Dagnelie (3) y Sokal y Sneath (12).

3. Índice de Sokal y Michener (1958).

$$S_{SM} = \frac{c+d}{a+b+c+d} = \frac{c+d}{n}, \quad 0 \leq S_{SM} \leq 1$$

4. Índice de Roger y Tanimoto (1960).

$$S_{RT} = \frac{c+d}{2(a+b)+c+d} = \frac{c+d}{n+a+b}, \quad 0 \leq S_{RT} \leq 1$$

5. Índice de Kulczynski (1927).

$$S_K = \frac{1}{2} \left(\frac{c}{a+c} + \frac{c}{b+c} \right), \quad 0 \leq S_K \leq 1$$

6. Índice de Ochiai (1957).

$$S_{0c} = \frac{c}{\sqrt{(a+c)(b+c)}}, \quad 0 \leq S_{0c} \leq 1$$

7. Índice de Odum (1950).

$$S_0 = \frac{(c+a) - (c+b)}{(c+a) + (c+b)}, \quad -1 \leq S_0 \leq 1$$

8. Índice de Forbes-Margalef (1958).

$$S_{FM} = \frac{(a+b+c+d)c}{(c+a)(c+b)} = \frac{nc}{(c+a)(c+b)}, \quad 0 \leq S_{FM} \leq n/c$$

9. Índice de Yule (1911).

$$S_Y = \frac{cd - ab}{cd + ab}, \quad -1 \leq S_Y \leq 1$$

10. Índice de Hamann (1961)

$$S_H = \frac{(c+d) - (a+b)}{a+b+c+d} = \frac{(c+d) - (a+b)}{n}, \quad -1 \leq S_H \leq 1$$

11. Índice phi.

$$S_\phi = \frac{cd - ab}{\sqrt{(c+a)(c+b)(d+a)(d+b)}}, \quad -1 \leq S_\phi \leq 1$$

2. Algunos comentarios sobre los índices descritos.

Se plantea como un inconveniente general para la interpretación de este tipo de índices, su dependencia del tamaño de número de elementos considerados en la dupla (n). Con n pequeños, la probabilidad de una estimación adecuada de la realidad subyacente a la valoración es baja, incrementándose a medida que n aumenta.

Dado que n depende del tamaño de la muestra, se recomienda utilizar una metodología que asegure un n relativamente grande y que, en lo posible, sea estándar para las situaciones a estudiar.

De los índices presentados he trabajado con mayor frecuencia los cuatro primeros índices de la nómina, habiéndose utilizado el resto en más de una oportunidad.

El haber trabajado con mayor intensidad los índices mencionados se fundamenta en las siguientes cualidades importantes desde su punto de vista:

- a. Son de muy simple expresión.
- b. Son de interpretación fácil y directa.
- c. El rango de expresión entre 0 y 1 está más acorde con la idea de similitud, siendo por ello más fácil su interpretación ecológica.
- d. En conjunto engloban las siguientes posibilidades de análisis:
 - Consideración exclusiva de presencia de elementos. Índices S_J y S_D .
 - Consideración adicional de los elementos comunes por ausencia. Índices S_{SM} y S_{RT} .
 - No ponderación especial de ninguna categoría. Índices S_J y S_{SM} .
 - Ponderación especial de alguna categoría. Índices S_D y S_{RT} .

La naturaleza de los índices anteriores implica claras relaciones entre ellos. Desde

luego, todos tenderán a 1 a medida que $a + b$ tienda a 0. $S_{RT} \leq S_{SM}$, pudiendo decirse que S_{RT} es un S_{SM} disminuido por la sobreponderación de las presencias exclusivas. $S_J \leq S_D$, es decir, S_D es un S_J , aumentado por la sobreponderación de las presencias comunes. $S_J \leq S_{SM}$, luego S_{SM} es S_J enriquecido por la consideración de las ausencias comunes. La disminución o aumento del valor de un índice respecto al otro no es proporcional.

Los antecedentes anteriores y la experiencia me hacen recomendar los índices S_J y S_{SM} como básicos para iniciar este tipo de análisis, según se estime importante tomar en cuenta o no la condición de común por ausencia.

Si se considera que un conjunto de inventarios comprende la totalidad de las especies a tomar en cuenta, al analizar la similitud entre esos inventarios sería recomendable S_{SM} , para ponderar las ausencias comunes. Si, por el contrario, se estima que aún quedan especies no consideradas en el muestreo sería preferible S_J . Por otra parte, una proporción excesiva de coincidencias por ausencia puede hacer que el índice entregue una información difícil de interpretar, ya que la similitud se valoraría fundamentalmente por lo que no hay. Bajo esa condición es preferible valorar solamente lo que hay. Caracterizar condiciones ecológicas presentes es siempre preferible a hacerlo por condiciones ecológicas ausentes.

3. Índices de uso corriente (segunda parte).

Es de utilidad en el campo de las mediciones de la similitud el uso de Chi cuadrado (χ^2), cuya expresión, con la corrección de Yates incorporada y siguiendo la nomenclatura precedentemente expuesta, es:

$$12. \chi^2 = \frac{\left(n(|cd-ab| - \frac{n}{2}) \right)^2}{(a+c)(b+c)(a+d)(b+d)}, \quad 0 \leq \chi^2$$

El índice mide la magnitud de la diferencia entre los valores de presencia-ausencia

esperados bajo una distribución al azar, carente de asociación, y los valores observa-

dos en los inventarios o especies (según sea la naturaleza de la dupla). El valor de χ^2 será significativo si la diferencia anterior es tan alta que exista una muy baja probabilidad (del 50/o menos) de que los valores observados puedan corresponder a una distribución al azar carente de similitud. Si, para valores significativos de χ^2 , la diferencia $cd-ab$ es mayor que 0, podemos afirmar que hay similitud significativa entre los componentes de la dupla; pero si, $cd-ab$ es menor que 0, no hay similitud, y aún más, hay descarte significativo entre sus componentes, pues predomina la presencia exclusiva sobre la común.

Este índice presenta un problema cuando $d = 0$. En este caso $cd-ab$ será siempre menor que 0, por muy alto que sea el

número de presencias comunes. Para este caso sería aconsejable valorar $d = 1$ con el fin de dar importancia a las presencias comunes.

Algunos investigadores recomiendan calcular χ^2 dos veces para la dupla, una considerando las dobles ausencias y otra, sin hacerlo. La comparación de ambos resultados permitiría llegar a conclusiones.

Es conveniente recordar que el índice informa sobre la existencia de una correlación, o de su inexistencia, y no de su grado o naturaleza. El resultado es sensible al número de inventarios o especies considerados en el análisis (n), por lo que se han propuesto índices complementarios que consideran ese factor, por ejemplo, el descrito a continuación:

13. Coeficiente de correlación de punto o de Kendall.

$$\varphi = \sqrt{\chi^2 / n} = \frac{|cd - ab| - \frac{n}{2}}{\sqrt{(a+c)(b+c)(a+d)(b+d)}}, \quad -1 \leq \varphi \leq 1$$

Entre este tipo de índices destaca el que presento a continuación, que tiene la par-

ticularidad de modificar el denominador, de acuerdo a las magnitudes en juego.

14. Índice de Cole (1949). , $-1 \leq S_C \leq 1$

$$S_C = \frac{cd - ab}{(c+b)(b+d)}$$

Se usa cuando $cd \geq ab$ y $a > b$.

$$S_C = \frac{cd - ab}{(c+a)(a+d)}$$

Se usa cuando $cd \geq ab$ y $a < b$.

$$S_C = \frac{cd - ab}{(c+b)(c+a)}$$

Se usa cuando $cd < ab$ y $c \leq d$.

$$S_C = \frac{cd - ab}{(b+d)(a+d)}$$

Se usa cuando $cd < ab$ y $c > d$.

B. Instrumentos destinados a valorar la presencia y la frecuencia

A esta categoría pertenecen fundamen-

talmente los instrumentos que caen dentro del campo de los coeficientes de correlación.

1. Coeficiente de correlación de Pearson.

$$r = T = \frac{n \sum xy - \sum x \sum y}{\sqrt{[(n \sum x^2 - (\sum x)^2) [(n \sum y^2 - (\sum y)^2)]]}}, \quad -1 \leq r \leq 1$$

donde: n = número de elementos considerados en la dupla
 x = valor de importancia del elemento en la condición A
 y = valor de importancia del elemento en la condición B.

2. Coeficiente de correlación de Spearman.

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}, \quad -1 \leq r_s \leq 1$$

donde: n = número de elementos en la dupla
 D = diferencia entre rangos de valores asignados a x e y .

Estadísticamente, el coeficiente de Pearson se basa en supuestos de normalidad de la distribución estadística de las variables y de no correlación de errores (o independencia entre los elementos que aportaron las frecuencias x e y) y se refiere al grado de asociación lineal entre las dos series de valores (frecuencias).

Los supuestos anteriores es más probable que se cumplan cuando se toman al azar números altos de elementos (n) de entre los que componen el universo.

El coeficiente de Spearman, al trabajar con rangos, utiliza información semicuantitativa, con menor contenido informativo que los datos originales, pero presenta la ventaja de no tener las restricciones de r para muestras pequeñas, puesto que no presupone distribuciones estadísticas especiales de las variables analizadas.

Un problema de estos índices, para el campo que estamos considerando, es la interpretación de valores 0 o inferiores. Habría que considerarlos a todos como expresión de carencia de similitud.

3. Índice de Winer

$$S_W = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}, \quad 0 \leq S_W \leq 1$$

donde: x e y son los valores de importancia de cada elemento en la condición A y B, respectivamente.

He seleccionado este índice en mis trabajos por su facilidad de cálculo, su rango de expresión entre 0 y 1, porque considera los datos cuantitativos con su carga informati-

va original, porque traduce a un rango de expresiones positivas toda la gama de valores entre -1 y 0 de los coeficientes de correlación anteriormente citados, y por-

que permite una más fácil comparación de los resultados obtenidos con los índices de presencia-ausencia seleccionados, cuyo rango también es de 0 a 1.

Puede decirse que es un índice de correlación de Pearson en que la medida de los elementos de cada condición tiene valor 0.

4. Índice de Odum (1950).

$$S_0' = \frac{\sum |x - y|}{\sum (x + y)}, \quad 0 \leq S_0' \leq 1$$

donde: x e y = valores de importancia de cada elemento en las condiciones A y B, respectivamente.

En realidad es un índice de disimilitud, donde 0 corresponde a la similitud máxi-

ma. Debería trabajarse como $1 - S_0'$, para hacerlo más comparable a los otros.

5. Índice de Agrell (1945)

$$S_A = \sum Q \text{ mínimos}, \quad 0 \leq S_A \leq 1$$

donde cada Q mín. se elige entre los valores de "importancia relativa", calculados para cada elemento en las condiciones A y B, respectivamente, de la siguiente forma:

$$Q_x = x/\sum x \quad Q_y = y/\sum y$$

El valor 0 estará dado cuando todos los Q mín. sean 0 y, por lo tanto, sólo haya presencias exclusivas. El valor 1 se obtendrá

cuando en las dos condiciones de la dupla haya presencia común de elementos con iguales valores de importancia relativa.

Puede ser útil donde los valores relativos de los datos base resulten aconsejables, por ejemplo, en algunos casos de análisis de duplas de especies en base a inventarios.

Personalmente no considero muy adecuado el uso del valor mínimo como aceptable de valoración ecológica dentro del contexto aquí tratado.

6. Índice de Kulczynski (1927).

$$S_K' = \frac{1}{2} \left(\frac{\sum F. \text{Mín.}}{\sum x} + \frac{\sum F. \text{Mín.}}{\sum y} \right) = \frac{(\sum F. \text{Mín.}) (\sum x + \sum y)}{2 \sum x \sum y},$$

$$0 \leq S_K' \leq 1$$

donde: F. Mín. = valor mínimo de importancia absoluta de cada elemento en la dupla considerada.

x e y = valores de importancia de cada elemento en las condiciones A y B, respectivamente.

Los valores extremos los toma en iguales condiciones que S_A , pero al contrario de éste, cuando los valores de importancia absoluta son muy distintos en ambas condiciones, entrega valores bajos de F . mín., los

que aún son rebajados por la ponderación especial del denominador.

El uso del valor mínimo me merece los reparos expuestos para S_A .

7. Índice de Elleberg (1956).

$$S_E = \frac{1/2 \sum c}{\sum a + \sum b + 1/2 \sum c} = \frac{\sum c}{2\sum a + 2\sum b + \sum c}, \quad 0 \leq S_E \leq 1$$

donde: c = suma de valores de importancia de un elemento en condiciones A y B, cuando ambos son mayores que 0.

a = valor de importancia de cada elemento exclusivo de la condición A.

b = valor de importancia de cada elemento exclusivo de la condición B.

La mayor ponderación de las presencias exclusivas es un factor que hay que considerar en la interpretación de los resultados, especialmente en el plano comparativo, al igual que para todos los índices que ponderan alguna categoría en especial.

Otro defecto es que valora la similitud máxima con cualquier distribución de valores de importancia bajo la condición de que todos los elementos sean comunes, no importando si existe concordancia o no en

los valores del elemento en las dos condiciones de la dupla.

C. Instrumentos matemáticos derivados de la teoría de la información

Para mis trabajos he seleccionado el índice de sobreposición (overlap) presentado por Horn (1966), el que trabaja directamente con la diversidad específica medida según la teoría de la información, usándose habitualmente el índice de Shannon.

1. Índice de sobreposición (Horn, 1966)

$$R_o = \frac{\text{Diversidad máxima} - \text{Diversidad observada}}{\text{Diversidad máxima} - \text{Diversidad mínima}}, \quad 0 \leq R_o \leq 1$$

Este mismo índice se utiliza para la medición de la diversidad específica relativa (medida de heterogeneidad) en el análisis de una comunidad o muestra.

Aquí nos interesa como una medida de similitud entre una dupla de comunidades o muestras, siendo la diversidad máxima aquella que se obtiene cuando los componentes de la dupla no tienen ningún ele-

mento común, y la diversidad mínima aquella que se obtiene cuando los componentes de la dupla tienen los mismos elementos y en las mismas proporciones. El índice mide grados de superposición de esquemas de jerarquización de especies entre esos extremos.

Por lo tanto, el valor 0 del índice corresponderá al análisis de una dupla en

que ningún elemento es común, y el valor 1 a la situación inversa, en que todos los elementos son comunes y mantienen la misma proporción entre ellos.

Para facilitar su cálculo puede utilizarse la fórmula computacional:

$$R_o = \frac{\Sigma (x+y) \log_2 (x+y) - \Sigma x \log_2 x - \Sigma y \log_2 y}{(\Sigma x + \Sigma y) \log_2 (\Sigma x + \Sigma y) - \Sigma x \log_2 \Sigma x - \Sigma y \log_2 \Sigma y}$$

donde: x e y = valores de importancia de los elementos en las condiciones A y B, respectivamente.

Se usa \log_2 por correspondencia con la expresión en bits del índice de Shannon para la diversidad específica.

ANÁLISIS DE LA INFORMACION ENTREGADA POR LOS INDICES

La información obtenida por los índices de similitud puede ser utilizada directamente para el análisis de una dupla de unidades ecológicas, u organizada en esquemas globalizantes, cuando se trabaja simultáneamente con muchas unidades, entre las cuales se ha valorado la similitud por duplas.

Para estos casos he utilizado la elaboración de dendrogramas de relaciones de similitud entre unidades, siguiendo la técnica de "par de grupos ponderados" (weighted pair-group method), descrito en Sokal y Sneath (1963), donde se discute su uso junto al estudio crítico de otros métodos de agrupamiento de la información entregada por los índices.

Para estudios con enfoque temporal o fenológico, o para estudios entre zonas o estratos, es útil el análisis comparativo de esquemas generales (dendrogramas en nuestro ejemplo), confeccionados con valores obtenidos de índices de las dos categorías expuestas y con valores de importancia diferentes para los que consideran la frecuencia de elementos.

De tal análisis será fácil deducir un modelo tentativo general de lo que se está estudiando, ya que se tendrá una visión del comportamiento de varios parámetros simultáneamente.

Ejemplificación de algunos casos de aplicación

En Sáiz y Avendaño (1976) hay ejemplos de análisis con criterios de similitud de los siguientes aspectos:

1. Definición de núcleos de especies.
2. Constancia temporal de la estructura de la comunidad, a través de algunas taxocenosis particulares.
3. Grado de similitud entre taxocenosis de diferentes comunidades, valorada tanto por índices de similitud como de superposición.
4. Relaciones temporales de la composición específica por estrato entre comunidades hipogeas adyacentes.
5. Relaciones entre estratos de una comunidad hipogea, a través de algunas taxocenosis particulares.

Los ejemplos anteriores permiten, también, visualizar la diferente información presentada por índices de presencia-absencia y aquellos que consideran la frecuencia.

En Sáiz, Vásquez y Molina (1977) se utilizan los criterios expuestos para demostrar la diferencia entre la entomofauna asociada a la inflorescencia de *Acacia caven* en una condición de espinos solos (A) respecto a otra de espinos mezclados con árboles autóctonos (B), considerando insectos de los órdenes Díptera, Hymenóptera, Homóptera, Coleóptera y Heteróptera. Lo anterior dentro del objetivo general de demostrar la existencia o no de entomofauna asociada a la inflorescencia de *Acacia caven*. Los datos son los siguientes:

Indices utilizados	Espinos solos (A)	Espinos mezclados (B)
S_J		0.44
S_W		0.17
R_o		0.37
H'	4,36	5,22
J'	0.71	0.84

S_W ha sido calculado con el número de individuos por especie como valor de importancia. H' es el índice de diversidad específica según Shannon en bits/individuos y J' es el índice de heterogeneidad de Pielou.

Un ejemplo de esquemas globalizantes (dendrogramas), utilizados para el análisis de una situación, a la vez que se trabaja

simultáneamente con diferentes valores de importancia, se da en la Fig. 1 tomada de Vázquez y Sáiz (1978), donde se estudian las variaciones temporales temporales de la taxocenosis de coleópteros epigeos en los mismos ambientes del ejemplo anterior y durante el mismo período. Los valores de importancia utilizados son:

para S_J = presencia o ausencia de especies.

S_W = constancia espacial en cada recolección.

S_W' = abundancia de individuos por especie.

S_W'' = biomasa seca por especie.

Los números árabes de la Fig. 1 corresponden a las recolecciones (períodos de siete días cada uno) y su secuencia (1 a 7) a la cronológica.

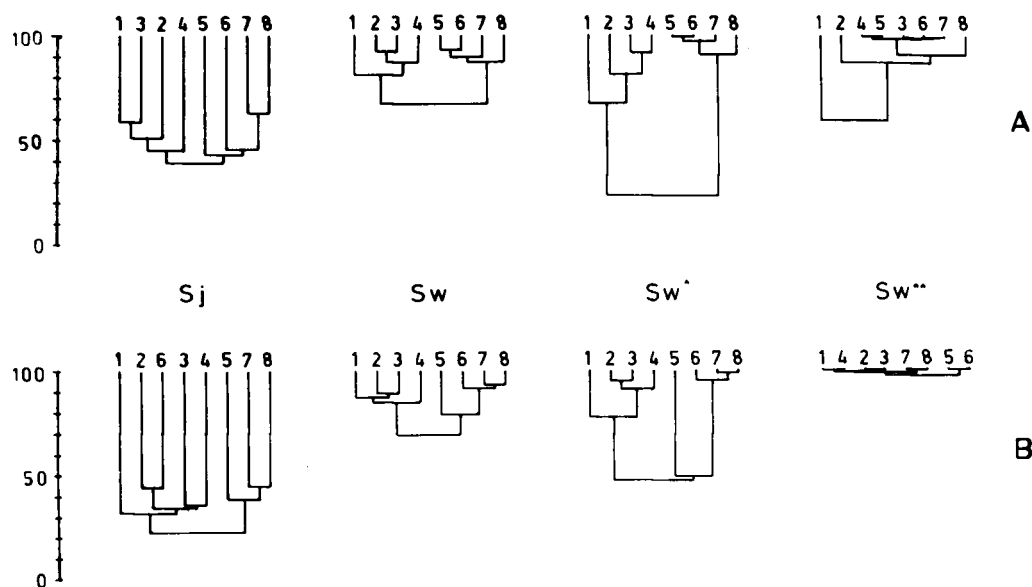


Fig. 1: Cambios temporales de la taxocenosis de coleópteros epigeos en estepas de *Acacia caven* en período de floración. A = espinos solos, B = espinos mezclados con otros árboles; S_J = similitud taxonómica, S_W - S_W' y S_W'' = similitud biocenótica, medida con valores de importancia de constancia espacial, abundancia de individuos y biomasa seca, respectivamente; Números árabes = recolecciones (Vázquez y Sáiz, 1978).

En la Fig. 1 se ejemplifica el caso de intensificación de la similitud taxonómica (S_J) al ser medida con S_W . Ello es visible por un nivel más alto de unión entre duplas y, en general, por un nivel más alto de cierre del conjunto.

La consideración de la constancia espacial (S_W) y la abundancia de individuos

(S_W') también acentúa la fractura que divide el tiempo total en dos períodos característicos. Finalmente, el esquema derivado de la utilización de la biomasa seca como valor de importancia (S_W'') entrega un modelo particular de relaciones.

S_J nos plantea un cambio bastante acentuado en la composición taxonómica de la

taxocenosis, a través del tiempo, con leve tendencia a definir dos períodos diferentes. S_W nos informa de dos períodos característicos, pero poco diferentes, resultantes de una alta concordancia de los esquemas jerárquicos de constancia espacial en el tiempo por períodos. Dicho de otra manera, las especies comunes entre recolecciones, en cada período, corresponden a las más constantes espacialmente.

S_W , además de la información entregada por S_W , nos dice que la fractura se debe fundamentalmente a un fuerte cambio en la densidad de las especies comunes, reflejo del cambio ambiental, según la interpretación expuesta en el capítulo referente al marco conceptual en que hemos usado estos índices.

El esquema de biomasa muestra, para espinos solos (A), una alta concordancia entre las especies comunes entre recolecciones intermedias, y una muy alta concordancia general para los espinos asociados a otros árboles autóctonos (B). Puede deducirse, por lo tanto, que las especies dominantes por biomasa, no corresponden, en general, a las dominantes por número de individuos.

En el mismo trabajo se usó R_0 , combinado con la diversidad específica (Shannon) para detectar también la fractura analizada.

He aquí los datos:

Recolecciones	Espinosa Solos (A)		Espinosa mezclados (B)	
	H'	R_0	H'	R_0
1	3.03		3.37	
2		0.95		0.69
3	3.53		3.41	
4		0.95		0.76
5	3.80		3.86	
6		0.91		0.84
7	3.50		3.45	
8		0.72		0.59
9	3.16		4.03	
10		0.93		0.74
11	3.29		3.79	
12		0.93		0.89
13	3.46		3.77	
14		0.91		0.91
15	3.07		3.38	

El R_0 fue calculado para duplas secuenciales de recolecciones en sentido cronológico.

El mismo índice lo he usado, junto al índice de diversidad específica (Shannon) y al de heterogeneidad (Pielou), para detectar la diversidad asintótica de un conjunto de muestras de musgos de una región antártica (Fig. 2, tomada de Pizarro y Sáiz, 1977).

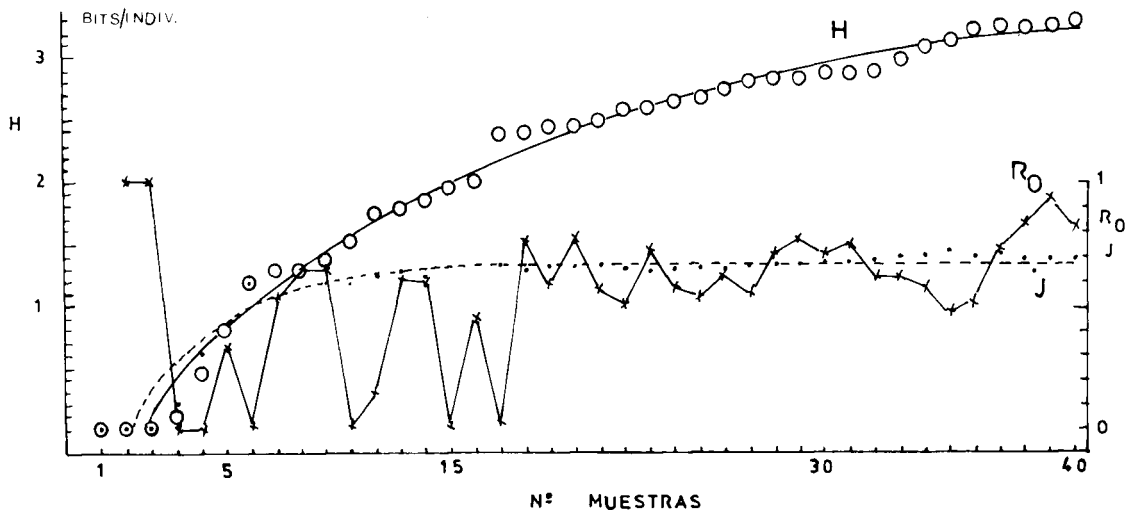


Fig. 2: Curva de valores acumulados de diversidad específica (H'), uniformidad (J) y sobreposición (R_0) en musgos antárticos (Pizarro y Sáiz, 1977).

En Palma, Sáiz y Pizarro (1978), S_J y S_W se ha utilizado para individualizar un foco de contaminación en el estero Lima-

che y para relacionar entre sí otros sectores del estero a través del estudio de sus hidrófitas (Fig. 3).

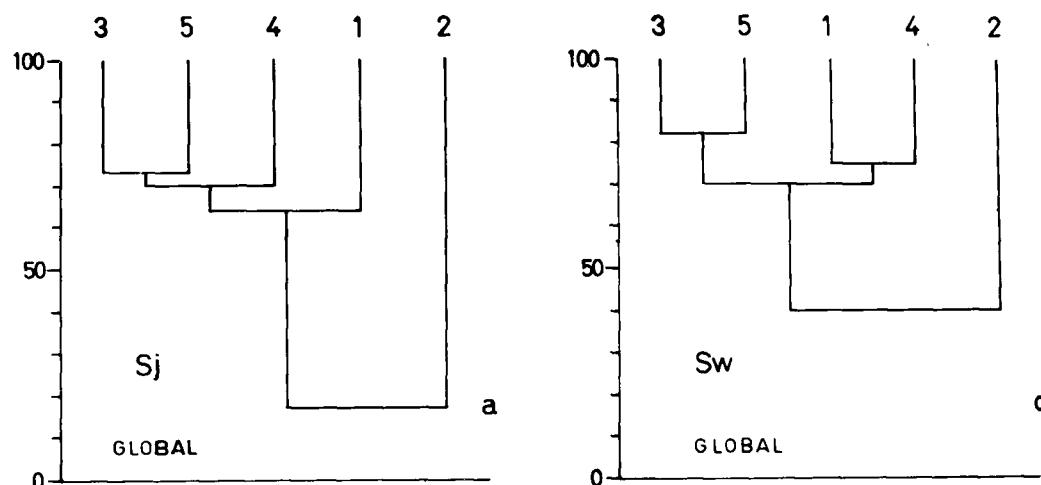


Fig. 3: Relaciones de similitud de la taxocenosis de hidrófitas entre sectores del estero Limache en función de grados de contaminación. $-S_J$ = similitud taxonómica; S_W = similitud biocenótica medida con valores de importancia de constancia espacial (Palma, Sáiz y Pizarro, 1978).

La secuencia numérica corresponde a la secuencia normal del curso del estero. El número 2 identifica al foco de contaminación. La asociación de 1 y 4 y de 3 y 5 define dos grados menores de contaminación, considerándose el grupo 1-4 como el más puro.

En el campo limnológico (Avendaño y Sáiz, 1977) se utilizaron los criterios de similitud para definir la unidad general de la masa de agua de la laguna El Plateado, respecto a rotíferos. Para ello se midió la concordancia en profundidad de la presencia y ausencia temporal de las especies. Los resultados son los siguientes:

Profundidad (m)	Presencia-ausencia simultánea (o/o)
0-4-8	90,74
0 - 8	91,92
0 - 4	91,04
4 - 8	94,14

También se calculó esta relación disminuyendo el universo de posibles coincidencias, ya sea en las presencias comunes, ya sea en las ausencias comunes.

En el mismo trabajo del análisis comparado de los dendrogramas derivados de S_J y S_W (medido éste con abundancia de individuos) se definen tres períodos diferenciales en el ciclo anual de la laguna (Fig. 4).

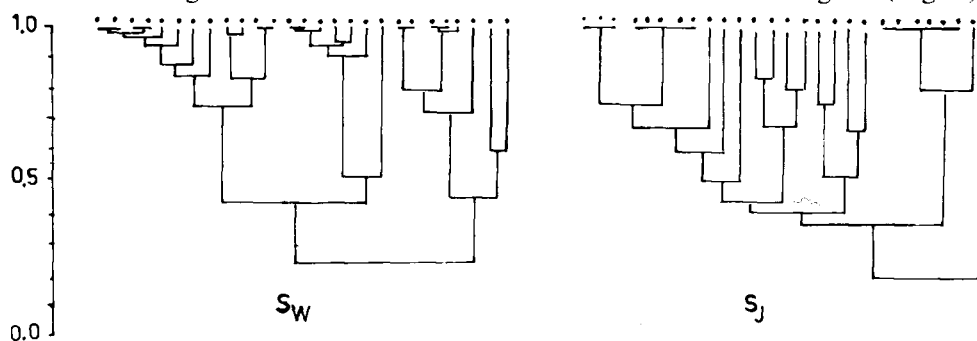


Fig. 4: Relaciones de similitud temporales en la taxocenosis de rotíferos de la laguna El Plateado. $-S_J$ = similitud taxonómica; S_W = similitud biocenótica, medida con valores de abundancia de individuos. Los puntos reemplazan a los números de las recolecciones (Avendaño y Sáiz, 1977).

En el gráfico se ha omitido el número de las muestras en la parte superior por no tener importancia relevante en esta presentación.

Más ejemplos del uso de criterios de similitud pueden verse en Sáiz (8), Sáiz y Calderón (9) y Etchegaray, Sáiz y Hajek (4).

AGRADECIMIENTOS

Al Sr. Dunny Casanova Z., por su valiosa ayuda en aspectos estadísticos y al Sr. Patricio Domínguez T., por la revisión general del manuscrito.

REFERENCIAS

1. AVENDAÑO, V., SAIZ, F., *An. Mus. Hist. Nat., Valparaíso*, 10: 121-133, 1977.
2. CANCELA DA FONCECA, J-P., *Rev. Ecol. Biol. Sol.*, 5: 381-407, 1966.
3. DAGNELIE, P., *Bull. Ser. Carte phytogeogr.*, 8: 5-77, 1960.
4. ETCHEGARAY, J., SAIZ, F., HAJEK, E., *Ser. Cient. Inst. Antárt. Chileno*, 5: 35-44, 1977.
5. HORN, H., *Amer. Nat.*, 100: 419-424, 1966.
6. PALMA, B., SAIZ, F., PIZARRO, C., *An. Mus. Hist. Nat., Valparaíso*, 11: 61-67, 1978.
7. PIZARRO, C., SAIZ, F., *Ser. Cient. Inst. Antárt. Chileno*, 5: 81-96, 1977.
8. SAIZ, F., *An. Mus. Hist. Nat., Valparaíso*, 8: 29-50, 1975.
9. SAIZ, F., CALDERON, R., *An. Mus. Hist. Nat., Valparaíso*, 9: 65-72, 1976.
10. SAIZ, F., AVENDAÑO, V., *An. Mus. Hist. Nat., Valparaíso*, 9: 89-104, 1976.
11. SAIZ, F., VASQUEZ, E., MOLINA, H., *An. Mus. Hist. Nat., Valparaíso*, 10: 135-152.
12. SOKAL, R.R., SNEATH, R., *Principles of numerical taxonomy*. Freeman & Co., 1963.
13. VASQUEZ, E., SAIZ, F., *An. Mus. Hist. Nat., Valparaíso*, 11: 69-80, 1978.