



# **ASPECTOS MÉTRICOS DE LA BIOINDICACIÓN**

**Daniel Andrés Dos Santos  
María Celina Reynaga**

# ASPECTOS MÉTRICOS DE LA BIOINDICACIÓN

Daniel Andrés Dos Santos  
y María Celina Reynaga

## RESUMEN

El desvío de un sistema acuático respecto de un sistema de referencia se toma como indicio de impacto, y cuando dicho desvío implica una disminución de servicios, éste se traduce en deterioro. Se han concebido métricas basadas en bioindicadores como complemento a las medidas físicas y químicas para diagnosticar el estado de un sistema. Se introduce y ejemplifica con datos reales la metodología de curvas ROC (del inglés, Característica Operativa del Receptor) para medir la capacidad diagnóstica de los índices, establecer valores de corte y comparar su rendimiento entre sí. Se repasa brevemente el enfoque multimétrico y se ofrece una propuesta de agregación de métricas basada en lógica difusa.

**Palabras clave:** Calidad de agua, índice biótico, lógica difusa, curva ROC, bentos.

## ABSTRACT

*The deviation of an aquatic system with respect to the reference system is considered indicative of impact, and damage occurs whenever ecosystem services are impaired. Metrics based on bioindicators have been conceived as a complement to physical-chemical measures to diagnose the state of a system. ROC (Receiver Operating Characteristic) curve methodology is here introduced and explained through real data in order to measure the accuracy of an index, to establish cut-off values and to compare the diagnostic performance among them. The multimetric approach is briefly visited and a proposal of metric aggregation based on fuzzy logic is offered.*

**Keywords:** Water quality, biotic index, fuzzy logic, ROC curve, benthos.

## INTRODUCCIÓN

Los mineros británicos del carbón apelaban a una curiosa estratagema para evitar la intoxicación con monóxido de carbono. Llevaban consigo, hacia las entrañas del yacimiento donde se internaban, a un canario enjaulado. La suspensión de los gorjeos, ya fuere por muerte o enfermedad aguda del ave centinela, servía de señal para abandonar la mina antes que el gas venenoso hiciera mella en los mineros. Dicha práctica, que reconoce al fisiólogo escocés John Scott Haldane (1860-1936) como su propulsor, se extendió por casi ocho décadas a lo largo del siglo pasado. ¿Cuál es la lectura no trivial que podemos hacer de este relato que rezuma la esencia de la bioindicación? El estudio de los cambios de estado que involucran a seres vivos frente a presiones ambientales, puede proveernos de herramientas confiables, y no tan onerosas, para diseñar señales de alarma.

A la hora de tomar decisiones, adquiere especial relevancia la ponderación objetiva de las señales que se desprenden de un fenómeno; separar el ruido y las contingencias del azar del contenido relevante de información. En estas circunstancias, la mera opinión del experto, aunque valiosa, no escapa a la modulación que imprime la propia subjetividad en la interpretación de las señales. La implementación de procedimientos métricos contrastables y reproducibles ofrece una alternativa para: (i) eludir la ruta falaz del *argumentum ad verecundiam* o argumento de autoridad dado por un experto y (ii) revestir a la tarea de bioindicación con los venerables ropajes de la falsabilidad propios del sistema científico (sensu Popper 1935). Surgen así los números índices o medidas sintéticas de un conjunto de observaciones. Los índices de bioindicación se sustentan en respuestas observables de los organismos y sirven para medir impactos ambientales. Al revisar la literatura afín (e.g. Armitage *et al.*, 1983; Barbour *et al.*, 1995; Dos Santos *et al.*, 2011; Karr 1981; Prat *et al.*, 2009; Stribling & Dressing 2015), descubrimos una ingente cantidad de índices propuestos, que varían según el nivel de organización considerado (desde índices cifrados en marcadores moleculares hasta índices basados sobre aspectos estructurales de la comunidad biológica), el tipo de bioindicador y el espacio físico de interés (los hay marinos y continentales, de ríos y lagunas, de suelo y aire, etc.).

El monitoreo de los ecosistemas implica mediciones periódicas acerca de su integridad que suele estar expuesta a numerosos agentes de perturbación antrópica. En particular, el biomonitoreo representa una alternativa factible para tal fin que se focaliza en atributos mensurables del componente biótico existente. Los macroinvertebrados bentónicos (o bénticos) son buenos indicadores debido a su naturaleza sedentaria, ciclos de vida largos, amplia distribución y sensibilidad a los procesos de contaminación (Rosenberg & Resh 1993). Constituyen un conjunto de organismos macroscópicos que pueden habitar el fondo de los

ambientes acuáticos a través de actividades que van desde la búsqueda activa de microhábitats y alimento hasta el asentamiento sobre un sustrato físico. Incluyen numerosas especies repartidas en los distintos grupos de invertebrados: Arthropoda, Nematoda, Annelida y Mollusca entre otros. Están involucrados en procesos ecológicos relevantes y pueden influir en la calidad de vida de las poblaciones humanas. Ejemplos de tales procesos son la bioturbación, el ciclaje de nutrientes, la transferencia de recursos energéticos al medio aéreo circunvecino, la bioacumulación y la participación en redes ecológicas. Los matices diferentes de diversidad que están implícitos en lo expuesto anteriormente (o sea, diversidad taxonómica, filogenética y funcional), invitan a considerar a todo el conjunto de macroinvertebrados bentónicos como representante de la complejidad biótica que tienen los sistemas acuáticos. A partir del conocimiento de la estructura del bentos y de los atributos biológicos de sus miembros, es posible calcular medidas sintéticas de integridad ecológica. El biomonitorio organizado alrededor del conocimiento del bentos ha ganado aceptación dentro de las políticas ambientales de varios países así como en organismos internacionales. Por ejemplo: UK National Ecosystem Assessment (NEA), USA Environmental Protection Agency (EPA) y European Union Water Framework Directive (WFD) recomiendan el uso del bentos para elucidar el status de las aguas corrientes. En América Latina hay también países que han adoptado este enfoque en sus normativas directrices respecto a la gestión del agua (Braga *et al.*, 2006; Hantke-Domas 2011).

El manejo sustentable de los recursos dulceacuícolas ha ganado protagonismo tanto a escala regional como global (Gleick, 2003; Richter, *et al.*, 2003; Hart & Calhoun, 2010), y el denominado "manejo integral de los recursos acuáticos" se ha convertido en el correspondiente paradigma científico. Estos ecosistemas son importantes por los bienes y servicios de valor cuantificable que proporcionan a las poblaciones humanas como, por ejemplo, agua para consumo y recreación. Lo preocupante es que esos beneficios aportados y la biodiversidad asociada están expuestos a amenazas globales (Vörösmarty *et al.*, 2010). Teniendo en cuenta que esos servicios apuntalan los programas de calidad de vida y sustancian el desarrollo social sobre nuestro planeta (Díaz *et al.*, 2005), resulta clave la tarea de monitorear los cambios de los ecosistemas. Para el desarrollo de un índice de calidad de agua, el mayor desafío consiste en reducir la multiplicidad de descriptores de la realidad a un número único, o sea, expresar en forma sintética el estado de un sistema que está expuesto a un conjunto de variables ambientales capaces de influir en su respuesta (Lermontov *et al.*, 2011). La dificultad que entraña es poder integrar muchas variables en un único número, combinando información registrada en diferentes unidades de medición. La formulación de un índice de calidad de agua debe contemplar aquellas variables que se presumen críticas para un cuerpo de agua dado. Una de las ventajas que ofrece trabajar con índices numéricos es permitir la

comparación en el espacio y el tiempo, además de facilitar un medio de comunicación entre diferentes actores interesados por la gestión del agua.

En este capítulo explicaremos una técnica estadística para (i) evaluar la eficiencia de los índices en el diagnóstico de la calidad del agua, (ii) establecer umbrales o valores de corte para juzgar presencia/ausencia de un impacto y (iii) comparar la capacidad diagnóstica de índices entre sí. Repasaremos además la naturaleza de algunos índices empleados en el biomonitorio de aguas continentales superficiales y concluiremos con una propuesta superadora que remite a la lógica difusa para amalgamar diferentes estimaciones del estrés ambiental en un valor único que sustente la toma de decisiones.

## **SOBRE MÉTRICAS Y APLICACIÓN DE ÍNDICES**

Se denomina estresor de un ecosistema acuático a todo agente, natural o antrópico, que impacta en ciertos organismos reduciendo su capacidad de supervivencia y reproducción. En los miembros de una comunidad, las respuestas biológicas por exposición a estresores van desde mortalidad hasta cambios comportamentales (Wong & Candolin, 2015), y por concatenación de efectos pueden verse reflejadas en cambios estructurales de una comunidad (por ejemplo, abundancia relativa de organismos con respiración tegumentaria tales como los oligoquetos límnicos). La diversidad taxonómica, abundancia de taxones y características autoecológicas (e.g. tipos alimentarios, valores de tolerancia frente a estresores, hábito de locomoción y posicionamiento con respecto al sustrato) conllevan información potencialmente útil para detectar y comprender procesos de degradación ambiental (Stribling & Dressing, 2015). En un sentido amplio, las diferentes métricas propuestas para evaluar la integridad ecológica del ecosistema acuático pueden basarse en los siguientes conceptos (en forma independiente o combinada): 1) diversidad de las comunidades, 2) perfil de tolerancia frente a los estresores por parte de los taxones presentes, 3) prevalencia de clases o grupos funcionales de organismos y 4) completitud (*completeness*) que es el cociente entre número observado versus número esperado de taxones. Ya fuere en formato unimétrico o multimétrico, todos los índices colapsan en una magnitud unidimensional susceptible de ser expresada, por calibración, como valor normalizado en una escala acotada (e.g. escala 0-1).

Ejemplos de métricas influidas por el concepto de diversidad (número de entidades diferentes que conforman un sistema) son: riqueza taxonómica, equitatividad, EPT o número de especies pertenecientes a los órdenes Ephemeroptera-Plecoptera-Trichoptera, el popular índice de Shannon-Wiener ( $H' = - \sum_{i=1}^S p_i \ln p_i$ ), donde en la práctica  $S$  es el número de unidades sistemáticas reconocidas en la muestra y  $p_i$  la fracción representada por el taxón  $i$  en el total de

individuos colectados) y su emparentado, el número efectivo de especies calculado como el exponencial de la entropía de Shannon ( $= e^H$ ) (Jost 2006).

Como ejemplos de medidas construidas a partir de registros de tolerancia cabe citar a BMWP (Armitage *et al.*, 1983; Paisley *et al.*, 2013), FBI (Hilsenhoff, 1988) y ASPT (Walley & Hawkes, 1996; Sandin & Hering, 2004). Por su uso extendido en la región latinoamericana, describiremos sucintamente a la implementación clásica del índice BMWP concebida originalmente para alertar sobre la polución orgánica. Consiste en imputar *scores* o puntajes a taxones inclusivos, como son las familias de insectos acuáticos. Los puntajes admitidos son los números naturales entre 1 y 10, y están orientados a reflejar el grado de tolerancia a la polución en una relación inversa. Así, por ejemplo, las familias de plecópteros que son muy sensibles a la depleción de oxígeno suelen ser codificadas con puntaje 10, mientras que las familias de dípteros muy resistentes a condiciones anóxicas tienden a tener asignados puntajes bajos, incluso valor 1 para los quironómidos. El resultado final del BMWP para una muestra será la suma de los puntajes de tolerancia de los taxones presentes en la muestra. Mucho esfuerzo se ha destinado para calibrar o ajustar los puntajes a las particularidades de las áreas geográficas donde se procura aplicarlo.

Ejemplos ilustrativos que apelan a la tercera directriz de métricas (i.e. prevalencia de grupos específicos de organismos) son porcentaje de quironómidos y porcentaje de dípteros calculados a partir del total de individuos contados en una muestra. Finalmente, los valores cifrados en la noción de completitud se derivan al cotejar entre sí listas observadas (O) de taxones versus aquellas esperadas (E) (Wright, 1995; Hawkins, 2006). Cocientes O/E bajos se esperarían cuando las condiciones ambientales fuesen modificadas por algún estresor y la composición de la comunidad cambiase (por reemplazo o remoción de elementos indígenas). Un aspecto a favor de estas medidas es que brindan valores en una escala de cocientes que facilita comparaciones entre regiones e incluso ecosistemas. Un tema a considerar es la resolución taxonómica. Se acepta que el nivel de familia (o superior) es realizable a bajo costo y sin grandes esfuerzos de entrenamiento taxonómico. Es más, experiencias pasadas han mostrado la adecuación de este nivel taxonómico grueso para evaluación biológica del agua dulce utilizando macroinvertebrados (Chessman *et al.*, 2007, Prat *et al.*, 2009).

## METODOLOGÍA DE CURVAS ROC

Si bien la cuestión de la integridad ecológica admite matices o gradaciones y es multidimensional su espacio de variación, en la práctica, deseamos responder de forma fiable a un planteo de tipo binario: ¿está impactado o preservado el ecosistema? A tal fin acuden en nuestro auxilio

índices y umbrales de detección (como los 37.5 °C para juzgar la condición febril de un paciente) para diagnosticar el estado de conservación. Las exigencias que tenemos hacia los índices bióticos es que sean clasificadores precisos (minimizar la comisión de errores falsos positivos y falsos negativos), versátiles (responder a multiplicidad de agentes estresores) y operativos (baratos y rápidos para calcular). La metodología de curvas ROC (acrónimo de Receiver Operating Characteristic, o Característica Operativa del Receptor), desarrollada en el ámbito de la teoría de detección de señales, puede ayudarnos a medir la eficiencia diagnóstica de tests o clasificadores binarios (e.g. enfermo/sano, inocuo/inocuo, contaminado/preservado, etc.), como efectivamente lo son las métricas biológicas de contaminación en la praxis ambiental. Una curva ROC es una representación gráfica de la sensibilidad frente a la especificidad para un sistema clasificador binario según se varía el umbral de discriminación. En las siguientes líneas, iremos explicando y desagregando las partes de esta definición para una cabal comprensión de la misma. Dos Santos *et al.*, (2011) introducen esta herramienta estadístico-gráfica para evaluar la capacidad que tienen los índices bióticos basados en organismos bentónicos para discriminar, en ríos de Yungas, aquellos impactados antrópicamente versus otros de condición opuesta que discurren dentro de áreas conservadas. A los efectos de ilustrar los aspectos fundamentales de las curvas ROC, emplearemos un subconjunto al azar de los datos procesados en aquella publicación (Tabla 1).

Los datos de la Tabla 1 corresponden a dos índices calculados a partir de muestras cuali-cuantitativas del bentos en 35 ríos que atraviesan la selva montana de Yungas. Uno de los índices (i.e. #Familias) concierne al mero conteo de taxones diferentes o riqueza taxonómica a nivel de familia. El otro índice corresponde a BMWP', con scores de tolerancia adaptados para la región por Domínguez & Fernández (1998). Del total de ríos, 20 son *a priori* considerados de referencia o en buenas condiciones, por estar inmersos en áreas protegidas, mientras que el otro subconjunto de 15 ríos es asignado al lote de perturbados por registrarse en sus inmediaciones impactos producidos por la actividad humana (ganadería y agricultura, deforestación, uso recreativo, efluentes domésticos e industriales, alteración de cauces, extracción de áridos, etc.). Ambos índices están muy correlacionados (Fig. 1) y puede advertirse, a partir de las representaciones adjuntas de dispersión, que ambas tipologías de ríos se separan a lo largo del rango de valores de los índices, correspondiendo valores mayores para ríos de buenas condiciones. Con la metodología de curvas ROC, responderemos a los siguientes tres interrogantes: 1) ¿tienen los índices capacidad para discriminar sitios con distinto grado de impacto antrópico?, 2) ¿cuál es el valor de corte que deberíamos emplear con fines diagnósticos?, y 3) ¿cuál de las dos métricas (la riqueza taxonómicas a nivel de familia o el índice BMWP') tiene mejor rendimiento?

**Tabla 1.** Subconjunto de datos reales procesados en Dos Santos *et al.*, (2011). Valores calculados para un par de índices bióticos en ríos montaños de Las Yungas. Los sitios muestreados fueron clasificados en dos categorías según la presencia o no de impactos antrópicos.

Índice biótico			
#Familias		BMWP Yungas	
Tipo de sitio		Tipo de sitio	
Referencia	Impactado	Referencia	Impactado
6	0	39	0
8	0	43	0
9	2	57	8
10	3	59	12
11	4	61	19
12	6	71	33
14	6	87	36
14	6	90	37
15	9	92	39
15	10	93	53
15	10	96	67
17	11	96	67
17	12	101	70
18	13	113	77
18	14	115	80
18		121	
20		121	
23		146	
24		146	
27		156	

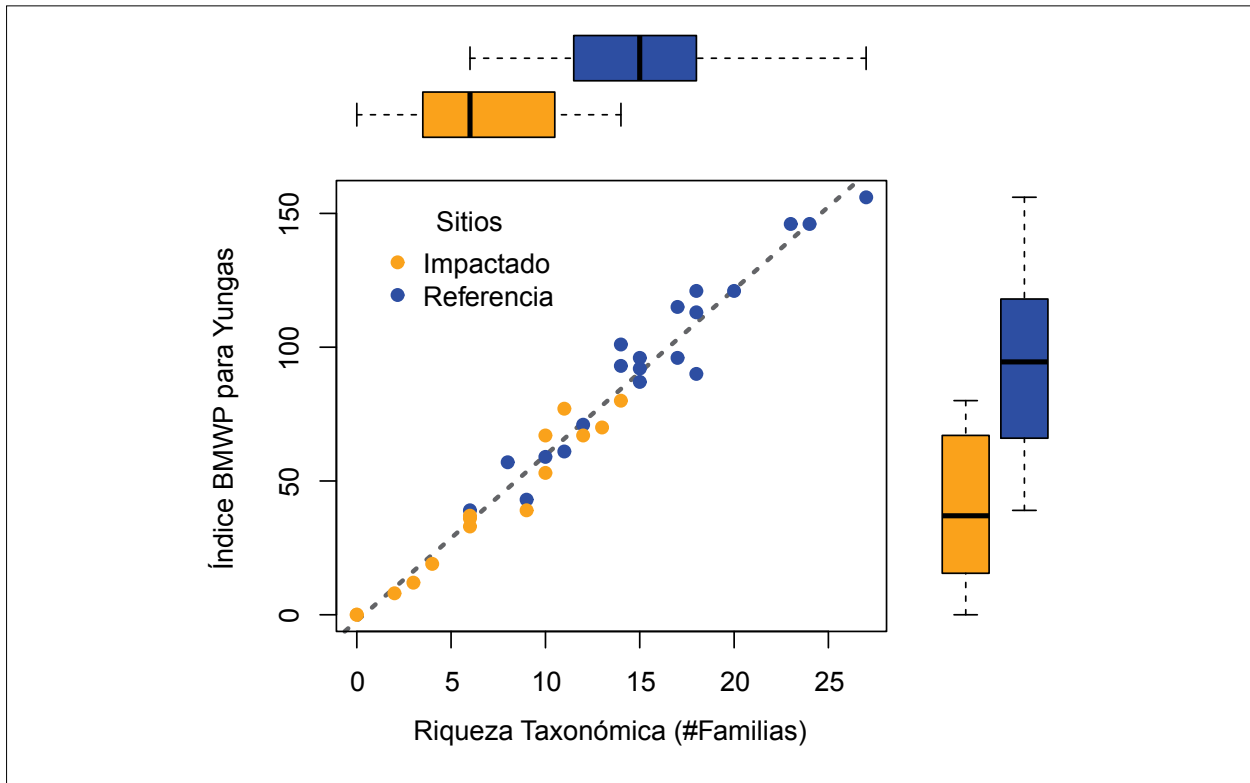
Especificidad y sensibilidad son los conceptos claves a partir de los cuales se construye la curva ROC. Supongamos que adoptamos como valor de corte #Familias = 9.5 (Tabla 2). En el contexto de este ejemplo, se considerarán casos "positivos" o "negativos" aquellos sitios cuya riqueza sea superior o inferior, respectivamente, al valor de corte. En la regla de decisión para el test, un resultado positivo se identifica así con la condición de preservado, mientras que se asimila como impactado al test negativo. Al contrastar la clasificación real de los ríos con la diagnosis resultante tras aplicar #Familias (umbral = 9.5), surge esta tetrada de alternativas diagnósticas: 1) verdaderos positivos (VP), son aquellos ríos reputados por el índice como saludables y que están efectivamente preservados, 2) falsos negativos (FN), son los ríos en buenas condiciones que dan un índice

bajo, 3) verdaderos negativos (VN), ríos en malas condiciones concordantes con el valor bajo del índice, 4) falsos positivos (FP), ríos expuestos a impacto antrópico pero clasificados erróneamente en estado conservado por el valor de la métrica empleada. Conociendo el número de instancias imputadas a cada una de estas alternativas, es posible calcular la sensibilidad y especificidad, para el punto de corte que les origina, de la siguiente forma:

$$\text{Sensibilidad} = \text{VP}/(\text{VP}+\text{FN})$$

$$\text{Especificidad} = \text{VN}/(\text{FP}+\text{VN})$$

Con otro valor de corte, la sensibilidad y especificidad variarán. En nuestro ejemplo, al fijar el umbral en #Familias



**Figura 1. Correlación entre los valores de dos índices bióticos implementados en ríos montanos de Yungas. Adviértase además la segregación de estos valores en función de la pertenencia a las clases de referencia e impacto. Línea punteada: recta de regresión lineal simple. Se emplean diagramas de caja marginales para representar la dispersión agrupada de los valores calculados para ambos índices. Los diagramas de caja permiten ver la dispersión de los puntos con respecto a la mediana (línea interna engrosada), los percentiles 25 y 75 y los valores máximos y mínimos.**

**Tabla 2.** Tabla de contingencia entre categorías diagnósticas obtenidas con la riqueza taxonómica (umbral 9.5) y la clasificación *a priori* producida por el *gold standard* (criterio certero o estándar de oro). VP, FP: verdaderos, falsos positivos. VN, FN: verdaderos, falsos negativos. TP, TN: totales de casos positivos, negativos.

Test: # Familias; umbral = 9.5				
SENSIBILIDAD	0.85 (= VP/TP)	Categorías predichas según test		TOTAL
		Preservado (+)	Impactado (-)	
ESPECIFICIDAD	0.60 (= VN/TN)			
Diagnóstico verdadero	Preservado (+)	17 (VP)	3 (FN)	20 (TP)
	Impactado (-)	6 (FP)	9 (VN)	15 (TN)
Regla de decisión: Preservado > Impactado				

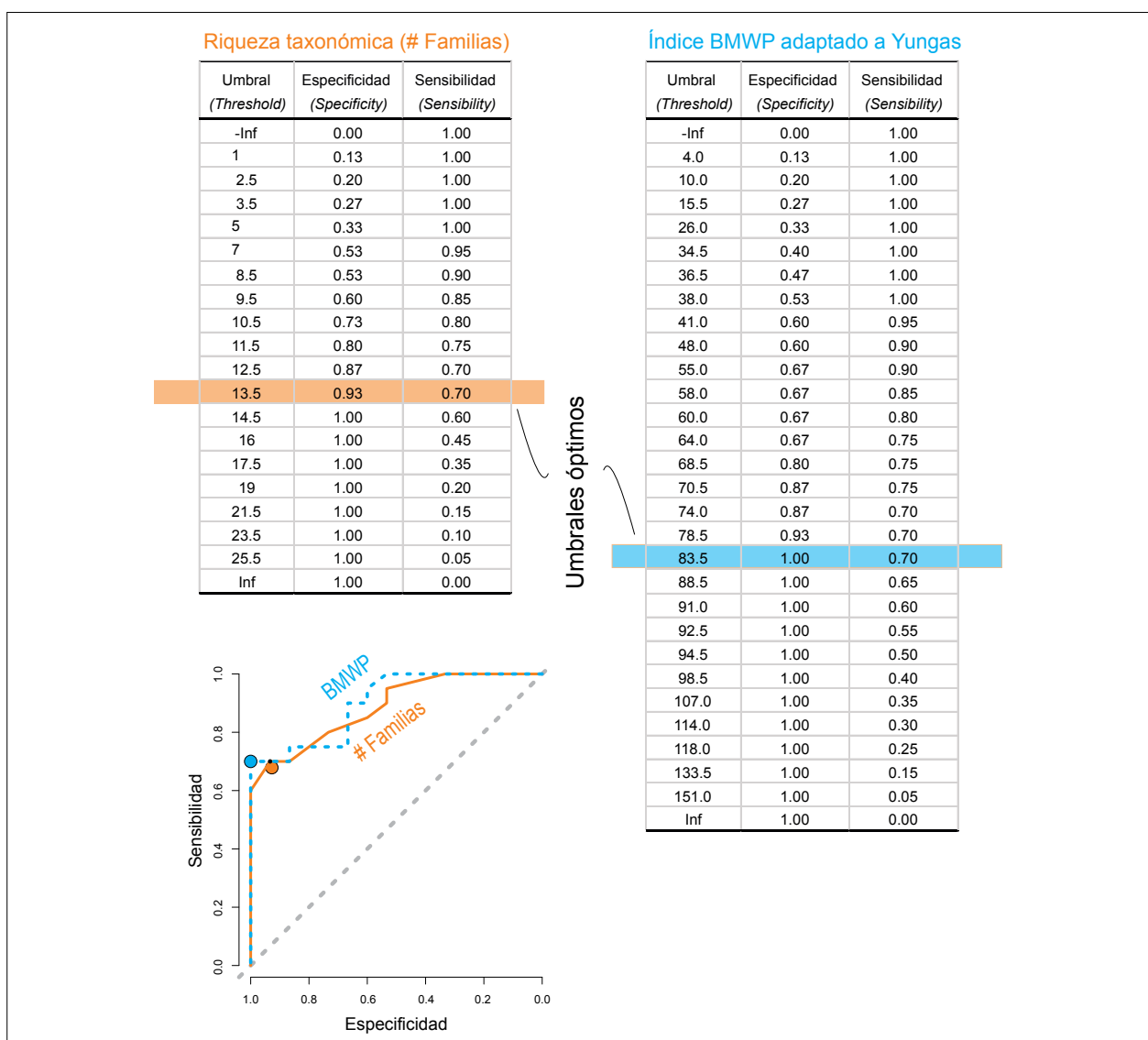
= 13.5 (Tabla 3), la especificidad trepa ostensiblemente de 60% a 93% respecto del umbral anterior, y la sensibilidad desciende en menor medida de 85% a 70%. Si recorremos todos los posibles valores de corte a lo largo de la escala métrica del índice, y unimos los respectivos puntos descritos por el par (1 - especificidad, sensibilidad), obtendremos nuestra curva ROC (Fig. 2). A título de referencia, en los gráficos de curvas ROC se traza también

una diagonal que va desde un extremo (100% especificidad, 0% sensibilidad) hasta el opuesto (0% especificidad, 100% sensibilidad). Esta diagonal se conoce como línea del azar o diagonal de la no-discriminación, puesto que representa al tipo de curva ROC que describiría un clasificador azaroso (e.g. arrojar una moneda al aire) para asignar las observaciones a las categorías de preservado o perturbado.

**Tabla 3.** Tabla de contingencia entre categorías diagnósticas obtenidas con la riqueza taxonómica (umbral 13.5) y la clasificación *a priori* producida por el *gold standard* (criterio certero o estándar de oro). VP, FP: verdaderos, falsos positivos. VN, FN: verdaderos, falsos negativos. TP, TN: totales de casos positivos, negativos.

Test: # Familias; umbral = 13.5				
SENSIBILIDAD	0.70 (= VP/TP)	Categorías predichas según test		TOTAL
		Preservado (+)	Impactado (-)	
ESPECIFICIDAD	0.93 (= VN/TN)			
Diagnóstico verdadero	Preservado (+)	14 (VP)	6 (FN)	20 (TP)
	Impactado (-)	1 (FP)	14 (VN)	15 (TN)

Regla de decisión: Preservado > Impactado



**Figura 2.** Gráfico de curvas ROC (del inglés Receiver Operating Characteristic, o Característica Operativa del Receptor) de dos métricas diagnósticas de calidad ecológica de ríos (#Familias y BMWP). La línea de no-discriminación es la diagonal punteada. El eje X tiene aquí la escala invertida, pero la misma curva suele representarse como 1 – especificidad (i.e. proporción de falsos positivos sobre el total de casos negativos) en dicho eje para quedar los valores en escala convencional ascendente hacia la derecha. Para cada curva ROC, el punto inflado cercano al vértice óptimo indica el punto de corte que determina la sensibilidad y especificidad conjuntas más alta. Los valores empleados para construir las curvas se detallan en las tablas anexas.



Entre todos los valores de corte explorados para construir la curva ROC, ¿cuál seleccionaríamos si el fin último es uno práctico para la toma de decisiones? Un criterio es seleccionar aquel punto de corte que maximice el índice de Youden (Cerdeira & Cifuentes 2012) calculado como especificidad + sensibilidad - 1. En términos gráficos, este punto es el más cercano al extremo superior izquierdo de la caja donde se inscribe la curva ROC. En la Fig. 2 pueden consultarse los puntos de corte seleccionados, 13.5 para #Familias y 83.5 para BMWP'. El lector debe percatarse de que el umbral de clasificación seleccionado según el índice de Youden es aquel punto de corte con la sensibilidad y especificidad más alta conjuntamente, lo cual no necesariamente significa que sea el punto donde el test alcance el máximo, tanto para la sensibilidad por un lado como para la especificidad por otro.

El área bajo la curva (AUC, *area under the curve*) es una medida sintética de la capacidad diagnóstica de la métrica evaluada. Para el caso de un clasificador perfecto, o sea uno al cual pueda reconocerse un umbral que segregue en forma inequívoca a los dos casos de observaciones (todos los casos positivos agrupados hacia un lado del umbral y todos los casos negativos hacia el otro lado), resultará en una curva ROC cuyo AUC = 1. Por el contrario, un clasificador incapaz de discriminar ríos preservados versus dañados, se comportará de forma similar a la diagonal del azar y será su AUC ~ 0.5. Un test diagnóstico tendrá mayor capacidad discriminatoria en la medida que sus puntos de corte tracen una curva ROC alejada de la diagonal o línea de no-discriminación, cercana al extremo superior izquierdo donde sensibilidad y especificidad alcanzan el máximo. Existen tests concretos para evaluar la significancia del valor de AUC (DeLong *et al.*, 1988), y decidir en consecuencia si se rechaza o no lo que enuncia la hipótesis nula (i.e., AUC = 0.5). Si bien en teoría el valor de AUC está entre 0 y 1, ningún clasificador realista debería ser < 0.5; en caso contrario convendría estudiar mejor la orientación de nuestra regla de decisión. Para nuestro caso (Fig. 2), tanto #Familias (AUC = 0.8867; 95% IC: 0.7826-0.9908) como BMWP' (AUC = 0.9017; 95% IC: 0.8048-0.9985) poseen capacidad discriminatoria ya que sus valores de AUC son significativamente superiores a 0.5.

El valor AUC es una medida sintética de precisión clasificatoria independiente de la prevalencia de casos positivos y negativos en la muestra bajo estudio. Otra ventaja de este estadístico es su equivalencia con el test de rangos de Wilcoxon (Hanley & McNeil, 1982). Podemos interpretarlo como el valor promedio de sensibilidad para todos los valores posibles de especificidad, y viceversa. Corresponde también al valor de probabilidad que un ítem seleccionado al azar del conjunto de casos impactados, respecto de otro ítem extraído al azar del lote preservado, tenga un valor de índice coincidente con lo establecido por la regla de decisión. Así, para

BMWP', donde la regla subyacente de clasificación es Preservado > Perturbado, su valor calculado de AUC = 0.90 significa que para un par de sitios seleccionados al azar de sus respectivos conjuntos de conservación, el sitio preservado exhibirá un valor de BMWP' superior al del sitio perturbado en el 90% de las veces.

Además de responder a la pregunta si los diferentes índices bióticos son capaces de discriminar sitios en buen estado de otros en mal estado, podemos comparar diferentes curvas ROC construidas a partir de los mismos datos y testear si hay diferencias estadísticamente significativas entre sus valores de AUC (DeLong *et al.*, 1988). Para los detalles de cálculo, se sugiere leer Hanley & McNeil (1982) y Zweig & Campbell (1993). En nuestro ejemplo de estudio, no hay evidencia suficiente para rechazar la hipótesis nula de equivalencia entre los AUCs de #Familias y BMWP' ( $Z = -0.91737$ ,  $P = 0.3589$ ). Entre los recursos informáticos sugeridos para realizar gráficos y encarar los diferentes tests estadísticos, cabe mencionar al software *MedCalc* (disponible en <http://www.medcalc.org>) y el paquete *pROC* de R.

## ENFOQUE MULTIMÉTRICO CON PROPUESTA FUZZY DE AGREGACIÓN

### Índices multimétricos

Existe una suerte de premisa de precaución que ninguna iniciativa de bioindicación puede soslayar: hay múltiples tipos de estresores y los organismos pertenecientes a un taxón dado no responden unívocamente a los mismos, reconociéndose por lo tanto efectos diferentes con variaciones tanto en lo taxonómico como en lo geográfico. Así, por ejemplo, un estresor químico (ej. materia orgánica) diferirá de otro estresor físico (ej. rectificación del cauce), viéndose algunos organismos afectados por uno y beneficiados por otro; y dentro de un mismo estresor, puede no ser la respuesta del taxón uniforme a lo largo y ancho de su rango geográfico. Enunciado de esta forma, no se niega la validez de la bioindicación y se enfatiza su naturaleza compleja, pero hay un cuestionamiento implícito sobre la generalidad en los detalles, que invita al incesante ajuste *ad hoc* de los índices inspirados en biomonitores.

Dada la multiplicidad de causales y respuestas biológicas, una opción de resguardo frente a la incertidumbre y complejidad consiste en adoptar una estrategia multimétrica de evaluación. Según este enfoque, se calcula una medida única a partir de muchas y heterogéneas medidas individuales capaces de capturar distintas facetas de la estructura y función de las comunidades. De esta forma, cada métrica individual se vuelve un insumo numérico de una medida global orientada a inferir desviaciones de un sistema respecto de un estado referencial de integridad biótica. Daños severos en el sistema se

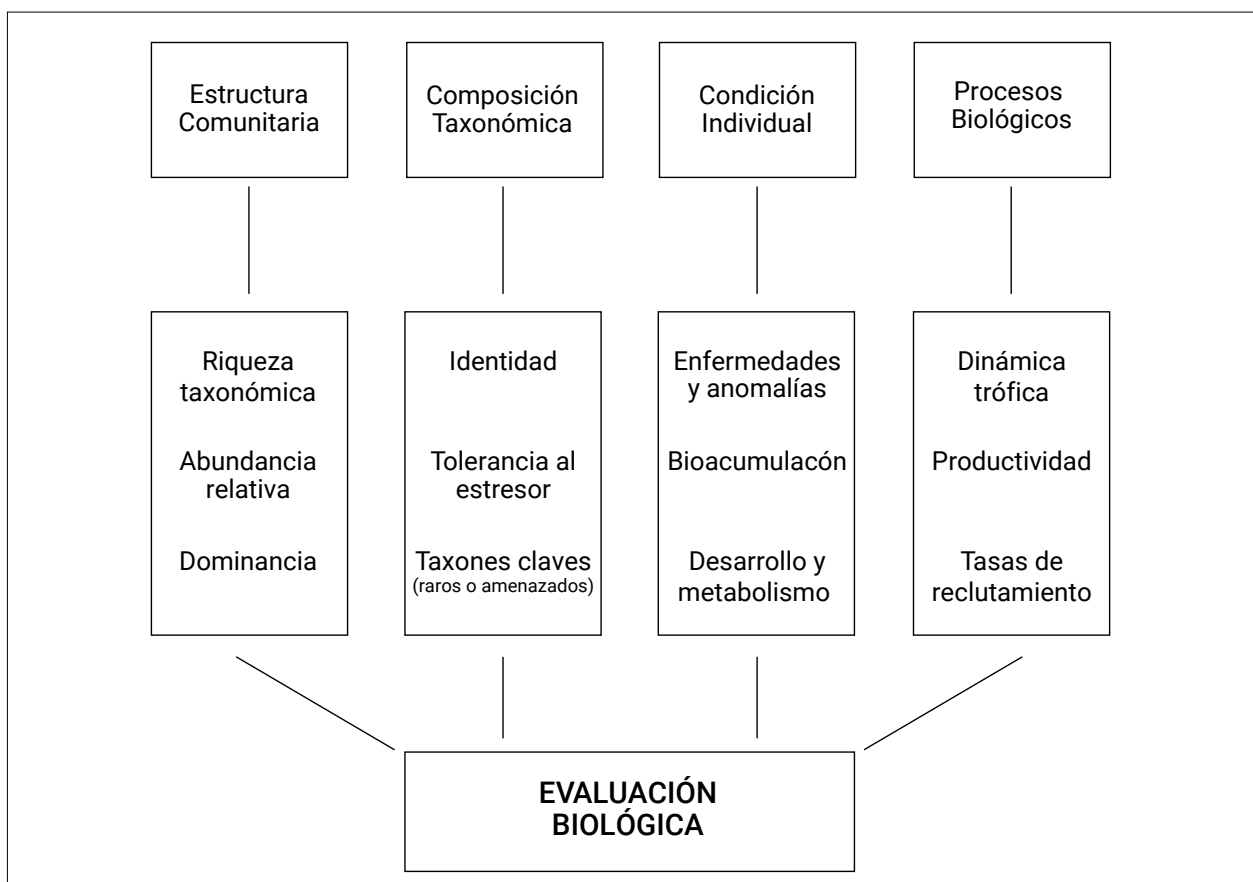


materializarán en valores anómalos a lo largo de la lista completa de métricas empleadas; impactos específicos y difusos pueden resultar indiferentes para la mayoría de las métricas y advertibles sólo por algunas pocas. El núcleo de este enfoque reside en el concepto de "integridad biológica" que ha sido definido como "la capacidad (de un cuerpo de agua) para sostener y mantener a una comunidad biológica balanceada con una composición específica, diversidad y organización funcional comparable con las presentes en hábitats naturales de la región" (Karr & Dudley, 1981; Schneider, 1992). A título de ejemplo de la importancia de estos conceptos, cabe mencionar que la EPA, agencia norteamericana para la protección del medioambiente, preconiza el uso de índices multimétricos de integridad biológica para monitoreo y diagnóstico de la condición de los ecosistemas acuáticos. Estos índices aditivos se llaman IBI (por sus siglas en inglés, índices de integridad biótica). El planteo original fue concebido por Karr (1981). Se procura que las métricas subsidiarias de un IBI cubran algunas de las 4 categorías conceptuales enunciadas por Barbour *et al.*, (1995) y expuestas en la Fig. 3. La Fig. 3 ilustra una estructura conceptual para los atributos de los ensamblajes que deberían ser considerados dentro de un programa de biomonitoreo sobre la condición general de la integridad ecológica.

El cálculo de un IBI implica seleccionar un conjunto de variables biológicas con potencial rol indicador, tales como números de especies en varias clases (nativos o intolerantes a la degradación). Sus valores son transformados a puntajes, por comparación con sitios de referencia. La escala discreta de normalización viene dada por los puntajes 1, 3 y 5. Se asigna el puntaje 5 a aquellos valores de las métricas dentro del rango esperado para sitios de índole prístina o conservada. Los valores esperables para sitios algo disturbados adquieren puntaje 3. El mínimo puntaje de 1 se da cuando la métrica asume valores concordantes con la expectativa para sitios dañados. El valor final de IBI proviene de la suma agregada de estos puntajes. Nótese que, si se empleasen 12 métricas individuales para componer el IBI, el mínimo sería 12 y el máximo 60. Aquí, nuevamente, la instancia de agregación es lo más polémico porque, al reducirse a un único guarismo, se desdibuja la información provista por todas las medidas individuales.

### Lógica difusa (fuzzy logic). Conceptos introductorios

El camino que conduce a la cuantificación de los impactos ecológicos por actividades humanas no está exento de



**Figura 3. Compartimentalización de los atributos que, en un escenario óptimo de trabajo, deberían ser incorporados en un enfoque multimétrico de medición dentro de un programa de biomonitoreo. Modificado de Barbour *et al.*, (1995).**

complicaciones: diferentes indicadores pueden entrar en conflicto, hay ausencia de algunas observaciones importantes o experimentamos incapacidad para incluir información potencialmente útil por ser su naturaleza cualitativa. Silvert (1997) advierte que abreviar en la Teoría de los Conjuntos Difusos (fuzzy set theory, Zadeh, 1965) es una buena alternativa para afrontar estas dificultades. Enunciados como "si el bosque de ribera está removido y la completitud taxonómica del bentos es pobre, entonces el estado ecológico de las aguas es malo", frecuentes en la praxis de la bioindicación y dotados de ambigüedad, tienen en lógica difusa un abordaje natural y pueden ser plasmados en los términos numéricos de un índice. La ambigüedad aducida viene a cuenta de que frente a un caso de estudio, siempre surgirán incertidumbres para juzgar de forma taxativa el contenido de los términos. Así, por ejemplo, ¿qué debo esperar para juzgar de forma inequívoca que la completitud del bentos es pobre? El núcleo de un sistema *fuzzy* viene dado por el enunciado de reglas inferenciales o cláusulas condicionales del tipo SI-ENTONCES. El procesamiento del sistema comporta evaluar la incumbencia del antecedente (o sea, determinar cuál es el grado de pertinencia en los términos de la parte SI de la cláusula) y llegar, por implicación, a un número indicador que nos ayudará a decidir la pertinencia de lo que se concluye.

La lógica difusa es una lógica alternativa a la lógica clásica donde se introduce un grado de vaguedad en las cosas que se clasifican. Un rasgo destacable de la lógica difusa es que provee un formalismo para trabajar con la ambigüedad y la ausencia de datos cuantitativos en un esquema de clasificación. Esto ciertamente no excluye la posibilidad de usar los conjuntos difusos para desprender graduaciones numéricas precisas en la toma de decisiones (Silvert, 1997). El proceso que se recorre para ir de conjuntos difusos hasta índices cuantitativos se llama "defuzzification" o decodificación, y debe ser considerada una cosa diferente del trasfondo conceptual del modelo construido a partir de reglas inferenciales. Un buen material de lectura, pionero en materia de bioindicación con bentos y lógica difusa, es el texto de un autor de nuestra región (Gutiérrez *et al.*, 2004). Adriaenssens *et al.*, (2004) revisan los modelos difusos en el contexto de manejo de ecosistemas. El texto de Lermontov *et al.*, (2011) será también de interés para los potenciales lectores de este libro. Ellos apelan a una arquitectura fuzzy de premisas que incluyen variables químicas para desarrollar un índice de calidad de agua.

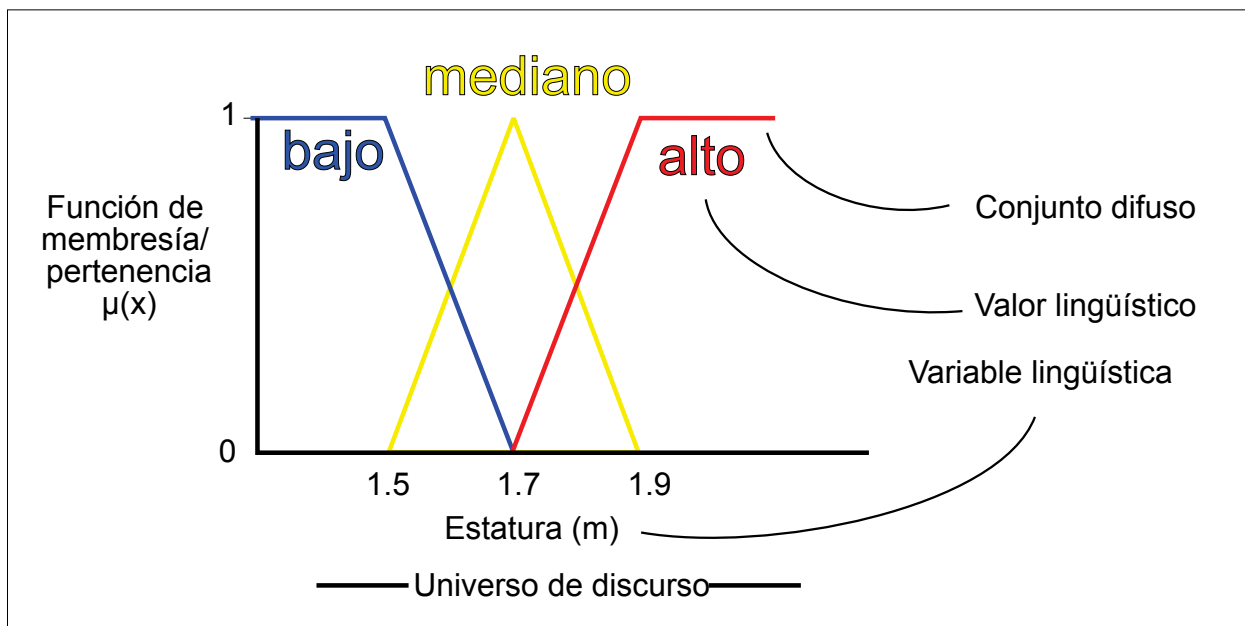
Para comprender un poco más las diferencias entre lógica difusa y lógica clásica, nos apoyaremos en el ejemplo tradicional de clasificar a una persona, tras medir su estatura, en una de estas tres etiquetas: "alta", "mediana" o "baja". En lógica clásica, los conjuntos son nítidos (*crisp sets*) y mutuamente excluyentes; una persona pertenece a una y solo una de las tres categorías. Para decidir sobre la pertenencia o no a una categoría,

se establecen valores de corte. Supongamos que el umbral de pertenencia para la clase "alta estatura" fuere  $> 1,80$  m y para la clase "baja estatura"  $< 1,50$  m. Según este proceder, una persona de 1,79 m sería considerada equivalente a otra de 1,51 m y diferente de otra la cual midiera 1,81 m. Pareciera que al haber optado por la delimitación de grupos nítidos, perdemos capacidad para captar graduaciones en las diferencias entre miembros de diferentes clases. Sin embargo, si dispusiéramos de una herramienta para caracterizar las alturas de forma que las transiciones fueran suaves, estaríamos reproduciendo la realidad clasificatoria mucho más fielmente. Y allí reside el atractivo de la lógica difusa.

Para representar los matices de pertenencia, definiremos a un conjunto (por ejemplo, "altos") tal que cada uno de sus elementos pertenezca a él con cierto grado (membresía). Formalmente, un conjunto difuso A se caracteriza por una función de pertenencia/membresía,  $\mu_A : U [0,1]$ , que asocia a cada elemento  $x$  de  $U$  un número  $\mu_A(x)$  del intervalo  $[0,1]$ , que representa el grado de pertenencia de  $x$  al conjunto difuso A. La Fig. 4 denota los valores de membresía para los conjuntos de altura en función de las posibles mediciones de estatura. Llamaremos *variable lingüística* a aquella noción o concepto que vamos a calificar de forma difusa (en nuestro ejemplo, la estatura). Por otro lado, el *universo de discurso* es el rango de valores que pueden tomar los elementos que poseen la propiedad expresada por la variable lingüística. Se entiende por *valor lingüístico* a las diferentes particiones que efectuamos sobre la variable lingüística (en nuestro caso, los valores alto, mediano y bajo). Por último, llamaremos *conjunto difuso* a la combinación de un valor lingüístico junto a su función de membresía. La función de membresía asocia a cada elemento del universo de discurso el grado con que pertenece al conjunto difuso. Si bien cualquier función podría emplearse para denotar conjuntos difusos, en la práctica se emplean funciones triangulares, trapezoidales, acampanadas y sigmoidales. En otras palabras, una función de membresía es una curva arbitraria cuya forma es definida por conveniencia.

### Razonamiento difuso. Implicación y decodificación

Se entiende por razonamiento difuso al proceso de realizar inferencias a partir de hechos y relaciones vagas (imprecisas). Las premisas refieren a observaciones con valor asignado en una variable lingüística y pueden agruparse en enunciados compuestos. Las operaciones sobre los conjuntos difusos son: unión ( $\cup$ ), intersección ( $\cap$ ) y complemento o negación ( $\neg$ ). Las premisas, articuladas en enunciados relacionales, dan lugar a reglas inferenciales del tipo SI-ENTONCES. El grado de pertenencia del consecuente que se postula en la regla es lo que se conoce como implicación. El proceso de decodificación consiste en agregar los diferentes resultados parciales de implicación



**Figura 4.** Ilustración de 3 conjuntos difusos asociados con la variable lingüística estatura, teniendo como valores lingüísticos a bajo, mediano y alto. Las funciones de pertenencia son de tipo L para bajo, Lambda o Triángulo para el mediano y Gamma para el alto. Se escogieron estas funciones porque trasladan o mapean nuestra subjetividad para juzgar a un valor dado de estatura en los conjuntos especificados. Así, una persona que mida 1.80 metros, para nuestro esquema de lógica difusa será 0.2 mediano y 0.8 alto, lo que en lenguaje coloquial sería alguien poco mediano y más bien alto.

y traducir el resultado global en un valor concreto para la variable de salida. La robustez de un sistema *fuzzy* depende del número y calidad de las reglas inferenciales. Inspirados en el caso didáctico presentado por Adriaenssens *et al.*, (2004), la Fig. 5 ejemplifica el funcionamiento interno de un proceso *fuzzy*. En este ejemplo, el sistema *fuzzy* diseñado consiste en estimar la aptitud del hábitat (variable output) a partir de los datos observados en heterogeneidad estructural y oxígeno disuelto (variables *inputs*). Durante el proceso se emplean cuatro reglas inferenciales, tres de ellas incorporan el conector lógico AND (Y en castellano), mientras que la restante incorpora el operador de disyunción OR (O en castellano). En la margen derecha de la Fig. 5 pueden leerse las reglas inferenciales. El ejemplo ilustra el funcionamiento de toda esta batería de reglas inferenciales para un caso hipotético donde las variables *inputs* tienen los siguientes valores: 46 para heterogeneidad estructural y 71 para oxígeno disuelto. Ingresando con estos valores, y atendiendo a los conjuntos difusos referidos en las reglas inferenciales, procedemos a interceptar las curvas de membresía respectivas. Para cada estrato de inferencia o regla actuante, entre los valores interceptados nos quedamos con la mínima membresía cuando hay un operador de conjunción (Y), mientras que retenemos la máxima membresía cuando se evalúa una disyunción (O). Finalmente, agregamos las implicancias para la variable *output*, uniendo todas las áreas parciales de inferencia. No debemos perder de vista que nuestro interés radica en estimar un valor numérico para nuestra variable *output* (en nuestro

ejemplo, aptitud del hábitat), y eso lo lograremos mediante la operación de decodificación. Para ello, calculamos el punto correspondiente al centroide del área total (polígono sombreado) que representa al conjunto difuso final. La coordenada *x* de dicho centroide representa el valor inferido para nuestra variable de salida.

### Formulación esquemática de un modelo difuso de bioindicación

Recapitemos, las 3 instancias involucradas en el proceso de ingeniería *fuzzy* de modelado para recuperar un valor indicador son:

1. Transformación difusa o *fuzzification* – Convertimos nuestros insumos (métricas basadas en bioindicadores) y resultado de salida (índice sintético de calidad ecológica) en valores lingüísticos, definiendo rangos y funciones de membresía.
2. Desarrollo de un esquema inferencial constituido por reglas SI-ENTONCES que se desprenden del conocimiento general que provee la investigación ecológica.
3. Decodificación – Convertimos nuestro resultado a una salida numérica concreta que permitirá interpretar el caso analizado. La técnica más frecuentemente usada es el método del centroide que es un promedio ponderado por la membresía.

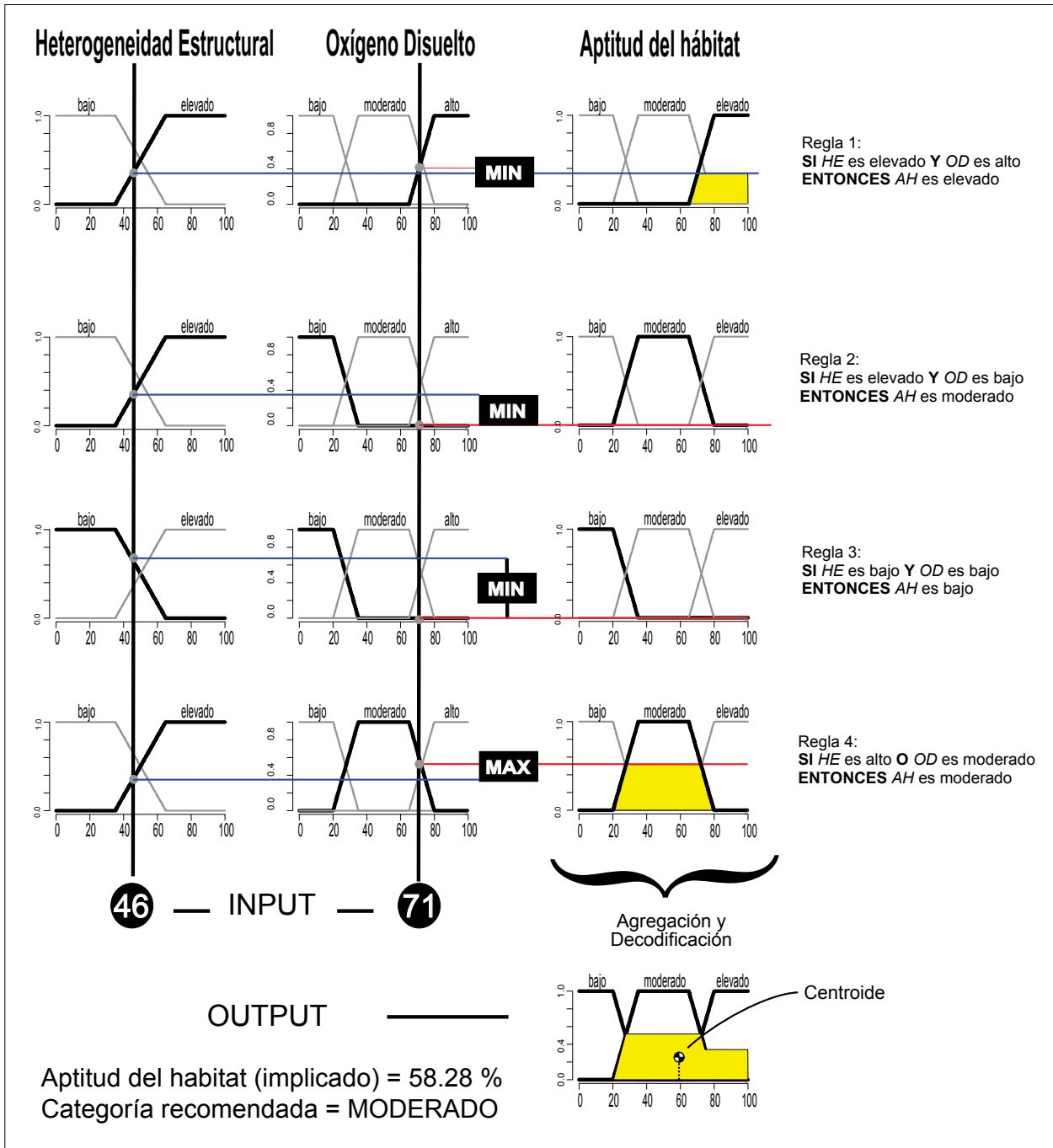


Figura 5. Ejemplo de proceso *fuzzy* para evaluación de calidad ambiental en ríos, con los valores normalizados de 46 y 71, respectivamente, para las variables de entrada (*inputs*) de heterogeneidad estructural (*HE*) y concentración de oxígeno disuelto (*OD*). El valor de la variable de salida (*output*) corresponde al estimador para la aptitud del hábitat que se obtiene después de la decodificación (centro de gravedad) del resultado global implicado a partir de las reglas inferenciales.

Tras revisar trabajos de distinta procedencia geográfica, pero con el interés común de diseñar y aplicar índices de integridad biótica calculados a partir del bentos de arroyos montañosos (e.g. USA: Stribling & Dressing, 2015; Hughes *et al.*, 2009; Alemania: Böhmer *et al.*, 2004; Korea: Jun *et al.*, 2012; Caribe: Black *et al.*, 2010), advertimos que ciertas métricas son seleccionadas con regularidad: riqueza taxonómica, porcentajes de algún grupo funcional

alimentario, abundancia relativa tanto de quironómidos como de miembros del grupo EPT. Es por ello, que en una formulación esquemática de sistema *fuzzy* hemos incluido 4 variables lingüísticas encarnadas por estas 4 directrices métricas recurrentes. Reservamos lugar para otra variable con un universo expandido en el rango [0, 100] y que reflejará el diagnóstico del estado ecológico y tendrá tres valores: bueno, regular y malo. Esta variable

representa nuestro índice sintético, que resulta de agregar las implicaciones de las reglas SI-ENTONCES que lo enlazan con las variables *inputs*. La Fig. 6 despliega el sistema *fuzzy* respectivo. Las variables *inputs* son:

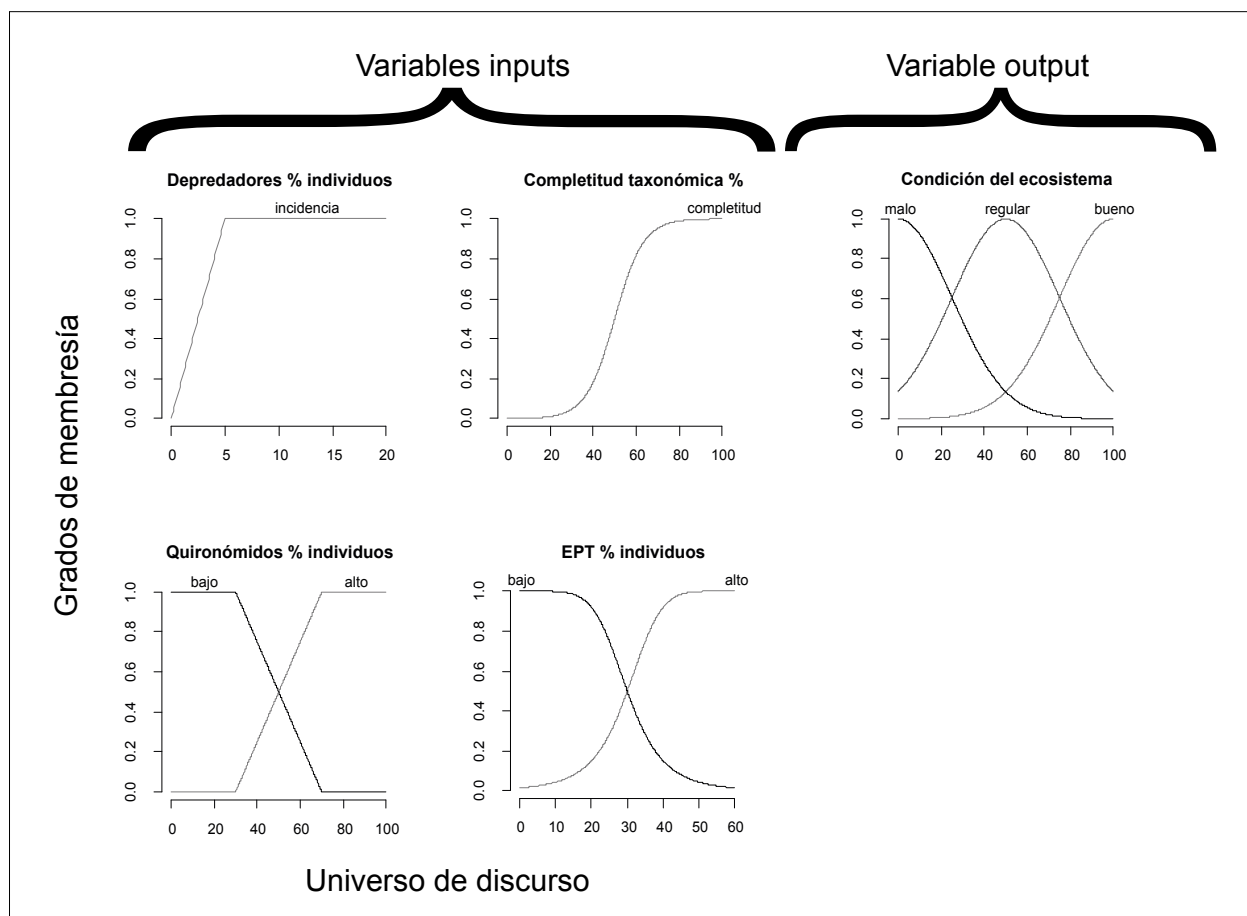
**Depredadores % individuos:** porcentaje del total de individuos colectados en una muestra que tienen un hábito alimentario depredador. Esta variable posee un único valor lingüístico, al cual hemos denominado incidencia, que denota el grado relativo de protagonismo que exhibe este grupo funcional en el seno de una comunidad. Este conjunto *fuzzy* está descrito por una fase lineal ascendente y otra constante a partir del 5% de fracción de depredadores.

**Complejidad taxonómica %:** Unidades sistemáticas presentes en una muestra respecto del total esperado. Por ejemplo, si en una muestra logran observarse sólo 5 familias de insectos acuáticos de un conjunto total esperado de 10 familias, entonces su complejidad en términos porcentuales será 50%. La razón o cociente entre lo observado (O) y esperado (E) se suele consignar en la literatura como cociente OE, y su versión porcentual se deriva

multiplicándolo por 100. Tiene un único valor denominado completitud y su valor de membresía es extraído a partir de una función sigmoidea.

**Quironómidos % individuos:** abundancia relativa de estos insectos con respecto a la abundancia total de individuos colectados en una muestra del bentos. En los cuerpos de agua muy poluidos, estos insectos predominan y pueden incluso ser el único tipo de macroinvertebrado presente en la muestra, o sea la métrica puede alcanzar el 100%. Los dos valores lingüísticos asignados son alto y bajo, quedando descritos sendos conjuntos difusos a través de funciones trapezoidales.

**EPT % individuos:** porcentaje de individuos del grupo EPT con respecto al total de individuos colectados en la muestra del bentos. Históricamente, se ha vinculado al grupo EPT (como ya fuere mencionado, conformado por insectos de los órdenes Ephemeroptera, Plecoptera y Trichoptera) con aguas corrientes bien oxigenadas, transparentes y en buen estado ecológico de conservación. Por ello, se esperan bajos valores en



**Figura 6.** Funciones de membresía que describen un sistema *fuzzy* esquemático. Este sistema sería aplicable para ríos montañosos y está destinado a integrar o agregar la información de múltiples fuentes (índices bióticos que se comportan como *inputs*) en un valor concreto que oficie de índice de calidad de agua expresable en la escala 0-100 (variable *output* particionada en 3 valores lingüísticos que denotan degradación ambiental).

esta métrica para sitios contaminados. Hemos establecidos dos valores contrastantes, etiquetados como bajo y alto, para esta variable. Las funciones de membresía son acampanadas y se cruzan en el valor 30% del universo de expresión.

En cuanto a las reglas inferenciales, nos restringiremos aquí a la siguiente dupla de reglas por razones de utilidad práctica y sencillez expositiva:

1) **SI** hay incidencia de depredadores **Y** se registra completitud composicional **ENTONCES** la calidad ecológica es buena.

2) **SI** el % de individuos pertenecientes a quironómidos es alto **Y** el % de individuos pertenecientes al grupo EPT es bajo **ENTONCES** la calidad ecológica es mala.

### Una aplicación con datos reales

Buscamos en la literatura científica producida fuera del ámbito de nuestro grupo de trabajo, alguna publicación sobre cálculo de índices bióticos en la Argentina, con macroinvertebrados bentónicos como bioindicadores, ríos montanos como objeto de estudio y con una tabla disponible de abundancia de organismos x sitios muestreados. Utilizamos el trabajo de Colla *et al.*, (2013) donde los autores muestrean el bentos del río El Tala (Catamarca) en

5 sitios distribuidos a lo largo de un gradiente altitudinal. Los sitios, numerados del 1 al 5, no evidencian en principio señales de impacto antrópico aunque sí se menciona que existen asentamientos humanos en las inmediaciones de los sitios 2 y 3. A partir de la tabla con datos crudos publicada por los autores, procedimos a calcular las métricas ligadas a las variables lingüísticas de nuestro esquemático sistema *fuzzy* presentado en el acápite anterior (Tabla 4).

Al alimentar nuestro sistema inferencial *fuzzy* con estos valores a título de *input* el *output* correspondiente por decodificación vía método del centroide, permite derivar la siguiente lista ordenada de sitios: Sitio II (= 32.25) < Sitio IV (= 40.15) < Sitio III (= 50.18) < Sitio I (= 57.83) < Sitio V (= 69.66). Nótese como el Sitio II adopta un valor muy presuntivo de impacto que contrasta con el resto. Este ejemplo ha sido corrido simplemente para demostrar cómo podemos hipotetizar sobre un caso de estudio en el que se desconoce la "verdad terrena" o las observaciones que los investigadores pueden haber realizado en su tarea de campo y que no están incluidas en el artículo publicado, incluso con una formulación intuitiva, no calibrada y esquemática de sistema *fuzzy* inferencial de calidad ecológica de agua. Con el afán de reforzar la comprensión de la técnica *fuzzy* para inferir un valor sintético sobre la calidad ecológica del ambiente, incluimos la Fig. 7 que detalla el cálculo del índice para el caso concreto del Sitio V.

**Tabla 4.** Valores calculados para las variables lingüísticas de un sistema *fuzzy* esquemático para diagnóstico de calidad ecológica del agua. Los datos originales proceden de Colla *et al.*, (2013) y corresponden a 5 sitios de muestreos distribuidos a lo largo del río El Tala (Catamarca). Todas las variables están expresadas porcentualmente. Riqueza % OE taxones: unidades sistemáticas presentes en la muestra respecto del total global identificado en el estudio y adoptado como valor esperado (= 25 familias de insectos). EPT % individuos: abundancia de insectos EPT con respecto a la abundancia total. Quironómidos % individuos: abundancia de estos insectos con respecto a la abundancia total. Depredadores % individuos: abundancia relativa del colectivo de taxones tradicionalmente reputados como depredadores. OE: cociente entre valores observados y esperados.

Variables	Sitio I	Sitio II	Sitio III	Sitio IV	Sitio V
RIQUEZA % OE taxones	64.0	60.0	56.0	60.0	68.0
EPT % individuos	30.5	14.4	21.3	13.8	42.7
Quironómidos % individuos	47.9	76.3	57.9	74.9	39.1
Depredadores % individuos	4.6	1.0	3.9	1.6	3.1

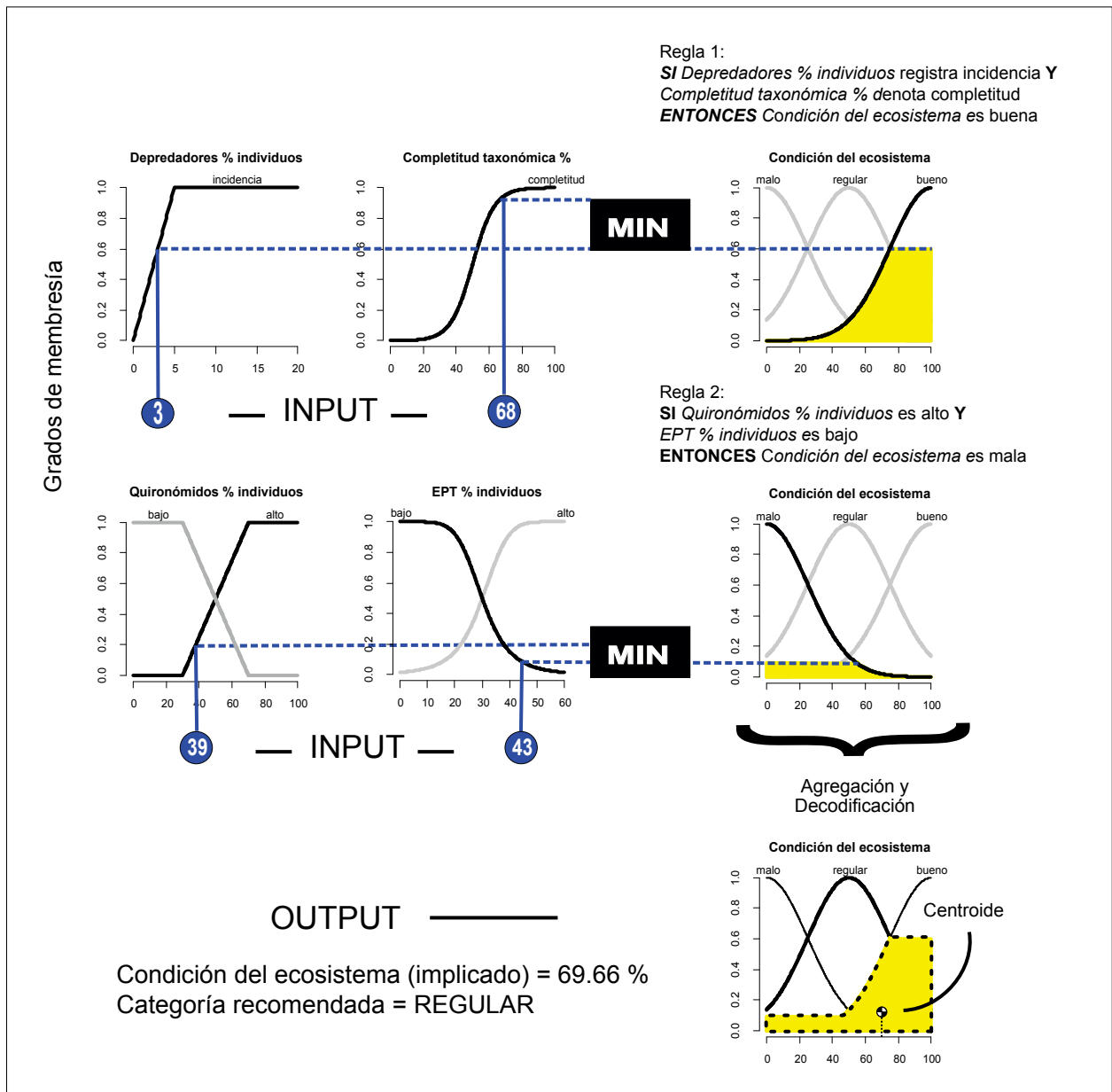


Figura 7. Implementación del sistema inferencial *fuzzy* presentado en la Figura 6. Los datos para las variables de entrada corresponden a las métricas individuales calculadas para el Sitio V (Tabla 4). Las conjuntos *fuzzy* parciales implicados (obtenidos tras aplicación de las respectivas reglas inferenciales) son unidos, y el centroide de dicho conjunto unión es calculado. La coordenada x del centroide corresponde al valor estimado para el índice en cuestión (medida sintética de la condición del ambiente).



## CONCLUSIÓN

Los índices, al abstraer un conjunto de observaciones en un valor adimensional, reducen la complejidad del fenómeno estudiado pero facilitan la toma de decisiones, como realizar el diagnóstico de calidad ambiental o evaluar el cambio de un sistema como por ejemplo la respuesta a un proceso de restauración. No existe un índice universal, válido tanto para la estepa patagónica como para la nuboselva subtropical o los arroyos pampásicos por ejemplo, pero sí la posibilidad de enunciar reglas generales que permitan inferir el desvío de un ecosistema respecto de su condición natural por acción antrópica. Cambiará el elenco de organismos, cambiarán las sensibilidades de las métricas y los rangos de variación, demandando ello esfuerzos regionales de calibración; pero la necesidad de agregar las implicaciones de reglas inferenciales parece ser transversal a los diferentes grupos de investigación abocados a la bioindicación. Estimamos que el desafío a encarar en lo futuro es el diseño de módulos de premisas *fuzzy*, aptos para inferir calidad de agua y versátiles en lo que respecta a su adaptabilidad a diferentes escenarios de aplicación. Como Lotfi Zadeh, el fundador de la lógica *fuzzy* dijera en 1973, "debemos explorar nuestra tolerancia hacia la imprecisión".

## AGRADECIMIENTOS

Agradecemos a los editores por la confianza depositada en nosotros para la redacción de este capítulo.

## BIBLIOGRAFÍA

Adriaenssens, V., De Baets, B., Goethals, P. L. & N. De Pauw (2004). Fuzzy rule-based models for decision support in ecosystem management. *Science of the Total Environment* 319 (1-3), 1-12.

Armitage PD, Moss D, Wright JF. & MT. Furse. 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running water sites. *Water Res*, 17: 333 -347.

Barbour, M.T., J.B. Stribling & J.R. Karr. 1995. The Multimetric Approach for establishing Biocriteria and Measuring Biological Condition. Chapter 6. In: *Biological Assessment and criteria: Tools for Water Resource Planning and Decision Making*, ed. W.S. Davis and T.P. Simon, pp. 63-77. Lewis Publishers, Boca Raton, FL.

Black, A., Ford, M., Keough, K. & B. Richards. 2010. Development of a Benthic Index of Biotic Integrity. WPI IQP Proposal | D-2010 The Conservation Trust of Puerto Rico.

Böhmer, J., Rawer-Jost, C., & A. Zenker. 2004. Multimetric assessment of data provided by water managers from Germany: assessment of several different types of stressors with macrozoobenthos communities. In *Integrated Assessment of Running Waters in Europe* (pp. 215-228). Springer, Dordrecht.

Braga B, Flecha R, Pena SD & J. Felman. 2006. A reforma institucional do setor de recursos hídricos. In: Rebouças A, Braga B, Tundisi JG (eds) *Águas doces no Brasil: capital ecológico, uso e conservação*. Escrituras Editora, São Paulo, Brazil, pp. 639-673

Chessman B, Williams S, & C. Besley. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. *J North Am Benthol Soc*, 26: 546-565.

Cerda, J., & L. Cifuentes. 2012. Uso de curvas ROC en investigación clínica: Aspectos teórico-prácticos. *Revista chilena de infectología*, 29(2): 138-141.

DeLong, E.R., DeLong, D. & D.L. Clarke-Pearson. 1988. Comparing the area under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 44, 837-845.

Díaz S, Tilman, D. & J. Fargione. 2005. Biodiversity regulation of ecosystem services. In: Hassan R, Scholes R, Ash N (eds) *Ecosystems and human well-being: Current state and trends*. Island Press, Washington DC, USA, pp 297-329.

- Domínguez, E. & H.R. Fernández. 1998. Calidad de los ríos de la cuenca del Salí (Tucumán, Argentina) medida por un índice biótico. *Conservación de la Naturaleza* 12. Fundación Miguel Lillo, Tucumán.
- Dos Santos, D.A., Molineri, C., Reynaga, M.C. & C. Basualdo. 2011. Which index is the best to assess stream health? *Ecol Indic* 11: 582-589
- Gleick, P.H. 2003. Global freshwater resources: soft-path solutions for the 21st century. *Science*, 302: 1524-1528.
- Gutiérrez, J. D., Riss, W. & R. Ospina. 2004. Lógica difusa como herramienta para la bioindicación de la calidad del agua con macroinvertebrados acuáticos en la sabana de Bogotá-Colombia/Application of fuzzy logic as bioindication tool for the water quality with aquatic macroinvertebrates in the Sabana de Bogotá-Colombia. *Caldasia*, 161-172.
- Hanley, J. A. & B. J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143: 29-36.
- Hantke-Domas, M. 2011. Avances legislativos en gestión sostenible y descentralizada del agua en América Latina. Commission for Latin America and the Caribbean (ECLAC), LC/W.446, Santiago, Chile, <http://www.cepal.org/>
- Hart, D. D. & A. J. K. Calhoun. 2010. Rethinking the role of ecological research in the sustainable management of freshwater ecosystems. *Freshw Biol*, 55: 258-269.
- Hawkins, C.P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. *Ecol Appl*, 16: 1277-1294.
- Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. *J North Am Benthol Soc*, 7: 65-68.
- Hughes, D. L., Gore, J., Brossett, M. P., & J. R. Olson. 2009. *Rapid bioassessment of stream health*. CRC Press.
- Jost, L. 2006. Entropy and diversity. *Oikos*, 113(2): 363-375.
- Jun, Y. C., Won, D. H., Lee, S. H., Kong, D. S. & S. J. Hwang. 2012. A multimetric benthic macroinvertebrate index for the assessment of stream biotic integrity in Korea. *International Journal of Environmental Research and Public Health*, 9(10): 3599-3628.
- Karr, J.R. 1981. Assessment of biotic integrity using fish communities. *Fisheries*, 6: 21-27.
- Karr, J.R. & D. Dudley. 1981. Ecological perspective on water quality goals. *Environmental Management*, 5: 55-68.
- Lermontov, A., Yokoyama, L., Lermontov, M. & M. A. S. Machado. 2011. A fuzzy water quality index for watershed quality analysis and management. In *Environmental Management in Practice*. InTech.
- Paisley, M.F.D., Trigg, J. & W. J. Walley. 2013. Revision of the Biological Monitoring Working Party (BMWP) score system: derivation of present-only and abundance-related scores from field data. *River Res Applic*, 30: 887-904.
- Popper, K. R. 1935. *Logik der Forschung: zur Erkenntnistheorie der moderner Naturwissenschaft*. Verlag von Julius Springer.
- Prat, N., Ríos, B., Acosta, R. y M. Rieradevall. 2009. Los macroinvertebrados como indicadores de calidad de las aguas. En: Domínguez, E. y H.R. Fernández (Eds.): *Macroinvertebrados bentónicos sudamericanos. Sistemática y biología*, 631-654.
- Richter, B.D., Mathews, R., Harrison, D.L. & R. Wigington. 2003. Ecologically sustainable water management: Managing river flows for ecological integrity. *Ecol. Appl.* 13: 206-224.
- Rosenberg, D.M. & V.H. Resh. 1993. Introduction to freshwater biomonitoring and benthic macroinvertebrates. In: Rosenberg, D.M., Resh, V. (eds.) *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York, pp. 1-9.
- Sandin, L. & D. Hering. 2004. Comparing macroinvertebrate indices to detect organic pollution across Europe: a contribution to the EC Water Framework Directive intercalibration. *Hydrobiologia*, 516: 55-68.
- Schneider, E.D. 1992. Monitoring for Ecological Integrity. Chapter 79 in d. D.H. McKenzie, D.E. Hyatt, & V.J. McDonald, *Ecological Indicators*, Vol. 2, e, pp. 1403-1419. Elsevier Applied Science, New York.
- Silvert, W. 1997. Ecological impact classification with fuzzy sets. *Ecological Modelling*, 96(1-3): 1-10.
- Stribling, J.B. & S.A. Dressing. 2015. Technical Memorandum #4: Applying Benthic Macroinvertebrate Multimetric Indexes to Stream Condition Assessments, October 2015. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 14 p.
- Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O, et al., 2010. Global threats to human water security and river biodiversity. *Nature*, 467:555-561.
- Walley, W.J. & H.A. Hawkes. 1996. A computer-based reappraisal of the Biological Monitoring Working Party scores using data from the 1990 River Quality Survey of England and Wales. *Wat Res*, 30: 2086 -2094.

Wong, B. & U. Candolin. 2015. Behavioral responses to changing environments, *Behavioral Ecology*, 26 (3): 665-673.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. *Aust J Ecol*, 20:181-197.

Zweig, M.H. & G. Campbell. 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem*, 39, 561–577.

Zadeh, L.A. 1965. Fuzzy sets. *Inf. Contro* 8: 338-353.

Zadeh, L.A. 1973. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst., Man, Cybernetics, SMC*, 3: 28-44.