



**UNIVERSIDAD
DE ANTIOQUIA**

**IMPLEMENTACIÓN DE UN MODELO DE USO
LIBRE PARA LA APLICACIÓN Y APOYO EN LA
INTERPRETACIÓN DEL ÍNDICE DE CALIDAD
ECOLÓGICA**

Autor

Elizabeth Flórez Córdoba

Universidad de Antioquia

Facultad de Ingeniería

Medellín, Colombia



Implementación de un modelo de uso libre para la aplicación y apoyo en la interpretación del índice de calidad ecológica

Elizabeth Flórez Córdoba

Trabajo de investigación presentado como requisito parcial para optar al título de:

Magister en Ingeniería Ambiental

Asesores:

PhD. Esnedy Hernández Atilano

PhD. Diana María Agudelo Echavarría

MSc. Alejandro Arenas Vasco

Línea de Investigación:

Recurso hídrico – Limnología

Grupo de Investigación:

GeoLimna

Universidad de Antioquia

Facultad de Ingeniería

Medellín, Colombia

2020

AGRADECIMIENTOS

Una meta más en mi vida se ve reflejada en este trabajo de investigación, como siempre más que agradecida con mi familia. Gracias a mis padres. Gracias a Marce por recordarme que todo lo que me proponga lo puedo lograr. Gracias a Caro por estar conmigo en los buenos y malos ratos en este viaje de mi vida. Gracias a Karen por su apoyo incondicional para creernos que esto sí era posible. Gracias a la profe Esneddy por confiar en mí y por motivarme a alcanzar esta meta en mi formación como profesional. Gracias a Diana y Alejandro por su tiempo, su dedicación y paciencia para sacar adelante este trabajo.

Y como olvidar lo que decía Jorge, “Vive cada día como si fuera el último. Estudia cada día como si fueras a vivir para siempre”.

CONTENIDO

RESUMEN	8
1 INTRODUCCIÓN	9
2 ANTECEDENTES	10
3 PLANTEAMIENTO DEL PROBLEMA	14
4 OBJETIVOS	16
4.1 OBJETIVO GENERAL	16
4.2 OBJETIVOS ESPECÍFICOS	16
5 MARCO TEÓRICO	17
6 METODOLOGÍA	22
6.1 MÉTODOS DISPONIBLES EN R PARA EL CÁLCULO DEL ICE	23
6.2 INTERPRETACIÓN DEL ICE	25
6.2.1 Detección de inconsistencias y corrección de datos	25
6.2.2 Gradiente ambiental	26
6.2.3 Análisis de ordenación	29
6.2.4 Reescalonamiento de scores	32
6.2.5 Valores óptimos y de tolerancia	33
6.2.6 Valor numérico del índice de calidad ecológica	35
6.2.7 Análisis de correlaciones y análisis de agrupamiento	36
6.3 VALIDACIÓN DEL MODELO DE CÁLCULO PROPUESTO EN R	36
7 RESULTADOS Y DISCUSIÓN	38

7.1	MODELO EN EL SOFTWARE R PARA LA APLICACIÓN DEL ICE Y CRITERIOS CLAVES PARA SU INTERPRETACIÓN	39
7.1.1	Depuración de la información.....	39
7.1.2	Lectura y estandarización de datos.....	41
7.1.3	Estimación longitud del gradiente ambiental.....	42
7.1.4	Ejecución del análisis de ordenación.....	43
7.1.5	Reescalamiento del ICE.....	45
7.1.6	Estimación de valores óptimos y de tolerancia.....	46
7.1.7	Valor de calidad ecológica.....	47
7.1.8	Análisis de correlaciones.....	48
7.1.9	Análisis de agrupamiento.....	49
7.2	VALIDACIÓN DEL MODELO DE CÁLCULO IMPLEMENTADO EN EL SOFTWARE R.....	49
7.2.1	Validación Río San Juan (Aguirre <i>et al.</i> , 2015).....	50
7.2.2	Validación Ciénaga de Ayapel (Zabala, 2016).....	53
7.2.3	Validación escenarios hídricos afectados por minería en el Chocó (Vargas, 2016).....	55
8	CONCLUSIONES Y RECOMENDACIONES	58
	REFERENCIAS BIBLIOGRÁFICAS.....	60

TABLAS

Tabla 1.	Métodos requeridos para el cálculo del ICE y su disponibilidad en softwares.....	23
Tabla 2.	Variables analizadas en los sistemas de interés y su respectiva comunidad hidrobiológica.....	37
Tabla 3.	Comparativo modelo de cálculo para el ICE – río San Juan.....	51
Tabla 4.	Resultados índices de calidad ecológica y RMSE – río San Juan.....	52

Tabla 5. Comparativo modelo de cálculo para el ICE – Ciénaga de Ayapel.....	53
Tabla 6. Resultados índices de calidad ecológica y RMSE – Ciénaga de Ayapel.....	54
Tabla 7. Comparativo modelo de cálculo para el ICE – Chocó.....	56
Tabla 8. Resultados índices de calidad ecológica y RMSE – Chocó.....	56

FIGURAS

Figura 1. Esquema de cálculo del ICE. MADS, 2018	22
Figura 2. Métodos estadísticos para el análisis de gradiente. Adaptado de Ter Braak, C. y Prentice, C., 1988; Lepš y Šmilauer, 2003	27
Figura 3. Modelo de respuesta de especies: (a) Lineal, donde a representa el intercepto y b el coeficiente de regresión; (b) Unimodal, donde μ se refiere al valor óptimo, t a la tolerancia y c al mayor valor de abundancia (Adaptado de Legendre & Birks, 2012)	29
Figura 4. Métodos de ordenación asociados con el análisis de gradiente ambiental: (a) Análisis de redundancia - RDA. Implica dos pasos: una regresión que produce valores ajustados (\mathbf{Y}) y residuos (\mathbf{Yres}), seguido por un análisis de componentes principales (PCA) de la matriz de valores ajustados (\mathbf{Y}). Adaptado de Legendre P. y Legendre L.F. 2012; (b) Análisis de correspondencia canónica - CCA. Los datos transformados (Chi-cuadrado) se someten a una regresión múltiple ponderada y los valores ajustados (\mathbf{Y}) se envían a un análisis de correspondencia mediante la descomposición de valores singulares (SVD). Legendre P. y Legendre L.F. 2012.....	31
Figura 5. Indicaciones reescalamiento de scores	33
Figura 6. Representación de los límites de tolerancia y rango óptimo de una especie bajo la influencia de un factor ambiental. Adaptado de Odum, E., 1972.....	34
Figura 7. Modelo de cálculo para el ICE en el software R.....	40

GRÁFICOS

Gráfico 1. Resultados ICE MADS, 2018 vs. Software R – Río San Juan.....	52
Gráfico 2. Resultados ICE MADS, 2018 vs. Software R – Ciénaga de Ayapel.....	55
Gráfico 3. Resultados ICE MADS, 2018 vs. Software R – fuentes hídricas en Chocó.....	57

LÍNEAS DE CÓDIGO

Líneas de código 1. Lectura y estandarización de datos	42
Líneas de código 2. Estimación longitud del gradiente ambiental.....	43
Líneas de código 3. Longitud del gradiente ambiental con respecto al primer eje de ordenación.....	43
Líneas de código 4. Ejecución del análisis de ordenación.....	44
Líneas de código 5. Análisis de significancia del modelo de ordenación.....	45
Líneas de código 6. Reescalamiento de scores.....	46
Líneas de código 7. Cálculo de valores óptimos y de tolerancia.....	47
Líneas de código 8. Estimación índice de calidad ecológica.....	48
Líneas de código 9. Análisis de correlaciones.....	49
Líneas de código 10. Análisis de agrupamiento.....	49

RESUMEN

En Colombia se cuentan con diferentes bases para el ordenamiento del recurso hídrico donde se integran mecanismos como el empleo de índices que permiten reflejar en una escala numérica el estado de la calidad del agua. A pesar de ello, el fortalecimiento de la gestión integral del recurso hídrico ha requerido complementar el uso de índices fisicoquímicos y biológicos por separado, teniendo en cuenta que los registros fisicoquímicos podrían reflejar condiciones instantáneas, mientras que los resultados biológicos pueden acumular información de diferentes periodos de tiempo. Con base en lo anterior, el Ministerio de Ambiente y Desarrollo Sostenible (MADS) estableció la evaluación del índice de calidad ecológica (ICE) como un enfoque integral del estado ecológico de los sistemas acuáticos al asociar aspectos fisicoquímicos y biológicos en un solo indicador.

Actualmente se cuenta con una metodología establecida por el MADS, donde se utilizan diferentes softwares licenciados para aplicar el ICE. El presente trabajo de investigación busca optimizar dicha aplicación por medio de la implementación de un modelo en el software R, donde se acoplen los diferentes métodos estadísticos multivariados requeridos para el cálculo del ICE. Además, se presentan criterios claves para la interpretación de los aspectos ecológicos asociados a los cálculos estadísticos y matemáticos. La metodología propuesta fue validada en tres casos de estudio mediante el cálculo del error cuadrático medio (RMSE), donde se evaluó el ajuste del modelo implementado en el software R con respecto a la metodología del MADS; obteniendo valores entre 0.00 y 0.03 lo que valida una estrecha relación entre ambas aplicaciones. El modelo de cálculo implementado en el software de uso libre resulta útil en la medida que optimiza una herramienta para potenciar los instrumentos utilizados en la gestión del recurso hídrico.

Palabras clave: índice calidad ecológica, estadística multivariada, R, modelo de uso libre.

1 INTRODUCCIÓN

El recurso hídrico presenta un deterioro ligado al crecimiento demográfico, el desarrollo económico y el cambio climático, solo por mencionar algunos factores que cada vez generan más complejidades a la hora de gestionar dicho recurso (GWP, 2009). Este tipo de afectaciones influyen sobre los ecosistemas acuáticos, por lo cual han desencadenado diversas discusiones desde los entes ambientales competentes con el propósito de establecer medidas relacionadas con la planificación y gestión ambiental de los territorios.

Desde un ámbito internacional, el enfoque de la gestión integral del recurso hídrico promueve el desarrollo y gestión coordinados del agua, la tierra y los recursos asociados, para maximizar el bienestar económico y social de manera equitativa sin comprometer la sostenibilidad de ecosistemas vitales, según lo establece la Asociación Mundial para el Agua (GWP TAC, 2000).

Con base en este tipo de orientaciones, Colombia ha definido un conjunto de criterios legislativos que fomentan la gobernanza del agua y la sostenibilidad del recurso hídrico, donde se incluye la Política Nacional para Gestión Integral del Recurso Hídrico (MAVDT, 2010), el Decreto 3930 de 2010; así como las guías para el Plan de Ordenación y Manejo de Cuencas Hidrográficas (POMCA) y el Plan de Ordenamiento del Recurso Hídrico (PORH) que incorporan un derrotero para la evaluación de la calidad del agua por medio de la aplicación de índices de calidad fisicoquímica e hidrobiológica (MADS, 2014, 2018).

En búsqueda de afianzar los análisis de la calidad del agua que normalmente se ejecutan durante los ejercicios de gestión del recurso hídrico, diferentes autores han tratado de relacionar la información fisicoquímica e hidrobiológica en un solo índice. De este modo, se destaca el ejercicio realizado por Haase & Nolte (2008) quienes desarrollaron el índice Invertebrate Species Index (ISI) para definir el nivel de detección del cambio biológico en la comunidad de macroinvertebrados acuáticos ante la afectación de las fuentes superficiales por nutrientes. Por otro lado, en 2011 Chalar *et al.* propusieron el índice biótico Trophic State Index for Benthic Invertebrates (TSI-BI) para establecer una clasificación del nivel trófico de los ecosistemas lóticos a partir de la relación entre la abundancia de macroinvertebrados acuáticos y las variables fisicoquímicas.

Por su parte, basados en las investigaciones anteriormente mencionadas en Colombia se formuló el índice de calidad ecológica (ICE_{RN-MAE}) a partir de la abundancia de los macroinvertebrados acuáticos y las

correlaciones entre las variables fisicoquímicas (Forero *et al.*, 2014). Como estrategia para la evaluación de los ecosistemas acuáticos en Colombia, el Ministerio de Ambiente y Desarrollo Sostenible (MADS) consideró la propuesta de Forero *et al.* (2014) en la Guía técnica para la formulación de PORH con el propósito de incluir una estimación del estado ecológico de los sistemas hídricos durante los ejercicios de gestión (MADS, 2018).

El índice de calidad ecológica (ICE) se estableció como un método basado en estadística multivariada que involucra el concepto de gradiente ambiental, el cual proporciona información sobre las distribuciones de los organismos en el espacio y el tiempo a partir del análisis de las abundancias de las especies en función de las variables ambientales (C. J. . Ter Braak & Prentice, 1988). De este modo, el ICE permite definir un gradiente ambiental de acuerdo con las particularidades fisicoquímicas que presenta cada cuenca y, a su vez, determinar valores óptimos y de tolerancia de cada organismo en función de dicho gradiente.

Dentro de la estructura hidrográfica definida en Colombia, se han considerado diferentes niveles de planificación que requieren la aplicación de los lineamientos descritos por parte de los entes gubernamentales. Frente a ello, las autoridades ambientales competentes se encuentran sujetos a recurrir a alternativas que implican supuestos que no cuentan con una explicación técnica adecuada, tan solo para dar cumplimiento a la normatividad. Por tanto, el presente trabajo de investigación tiene como propósito generar un modelo de cálculo abierto, que potencie el uso de herramientas estadísticas multivariadas libres y gratuitas aplicadas al campo de la ecología durante los ejercicios de gestión del recurso hídrico.

2 ANTECEDENTES

La presión ejercida sobre el recurso hídrico asociada a la contaminación directa dicho recurso, la ampliación de la frontera agrícola, su uso irrazonable, entre otros aspectos, sustenta la necesidad de definir enfoques para evaluar el estado ambiental de los ecosistemas acuáticos. Desde la década de 1980, diferentes autores han mencionado que es necesario promover dentro de las políticas buenas prácticas de manejo ambiental del recurso hídrico, para lo cual se han desarrollado diferentes tipos de índices cuantitativos que reúnen información fisicoquímica y biológica para caracterizar los ecosistemas acuáticos (Angermier & Karr., 1986; Angermier & Williams, 1993; Karr, 1981).

Es así como las condiciones de calidad del agua de una corriente superficial frecuentemente se evalúan a través del análisis de variables fisicoquímicas y de la implementación de índices de calidad, con los cuales se logran expresar los resultados en una escala numérica o en una gama de colores que se encuentran vinculados con una clasificación de la calidad del agua (Debels *et al.*, 2005; IDEAM, 2011).

Sin embargo, algunas limitaciones se han identificado para los índices de calidad fisicoquímica, dado que su característica de fácil entendimiento se contrasta con el hecho de que la calificación se establece a partir de un número, lo que puede generar pérdida de información y no representar la condición real de calidad en un cuerpo de agua (Castro-Roa & Pinilla-Agudelo, 2014); asimismo, se ha mencionado que el contar con una gran cantidad de datos para diferentes parámetros de calidad del agua en fuentes superficiales sin una interpretación adecuada no es útil para la gestión de la calidad del agua (Asadollahfardi, 2015).

De igual forma, los índices fisicoquímicos pueden estar fuertemente influenciados por las condiciones climáticas y se consideran como el reflejo del estado instantáneo de calidad del agua al momento de un monitoreo; mientras que los indicadores biológicos pueden acumular eventos de disturbio de varios lapsos de tiempo (Alba-Tercedor, 1996; Pinilla, 1998).

En cuanto a los aspectos biológicos de los sistemas acuáticos, es posible identificar las características de los organismos que habitan en estos sistemas y sus variaciones en el tiempo. Para ello, diferentes organismos han sido utilizados como indicadores de la calidad del agua, tales como los macroinvertebrados acuáticos y algas perifíticas como las diatomeas, quienes son sensibles a los cambios de las condiciones ambientales (Cairns & Pratt, 1993; Forero *et al.*, 2014; Giorgio *et al.*, 2016; Kwandrans *et al.*, 1998; McCormick & Cairns, 1994). Cabe considerar que se han propuesto diferentes grupos de organismos acuáticos como indicadores de los cambios ante los factores ambientales, pero ninguno ha sido seleccionado para analizar las variaciones en la composición de las comunidades asociados con la calidad del agua (Karr, 1981).

Además, los índices biológicos han generado controversias, dado que por lo general tienen en cuenta solo la presencia o ausencia de los organismos y el número de taxones en un área definida (Alba-Tercedor, 1996; Forero *et al.*, 2014; Ramos-Merchante & Prenda, 2017), y para gran parte de sus aplicaciones no se han establecido rangos de tolerancia a la contaminación dentro de una misma familia, asociado con la falta de conocimiento acerca de la taxonomía de los organismos y su respuesta a los contaminantes en las regiones de estudio (Forero *et al.*, 2014; Ríos-Touma *et al.*, 2014).

Cabe destacar que ciertos índices biológicos han sido desarrollado para zonas específicas, lo que podría ocasionar que los individuos caracterizados no sean comparables con las comunidades que se encuentran en el trópico (Forero, 2012; Moya *et al.*, 2007). No obstante, diferentes autores han tratado de relacionar enfoques multimétricos, multivariados y rasgos biológicos (Li *et al.*, 2010), lo que conlleva a obtener índices que conjugan variables fisicoquímicas y biológicas.

Karr (1981) realizó una evaluación de la integridad biótica mediante comunidades de peces como un esfuerzo para promover la protección efectiva de los recursos biológicos en comunidades de agua dulce. Para 1994, Kerans & Karr formularon el índice béntico de integridad biótica (B-IBI) en ríos del valle de Tennessee, el cual les permitió establecer el estado biológico de los cuerpos de agua en estudio.

También, se han propuesto índices con base en la información biológica que han permitido caracterizar el recurso hídrico y evaluar el cumplimiento de objetivos de calidad del agua establecidos. Algunos ejemplos corresponden al índice desarrollado originalmente por Karr, (1981) y modificado por otros (Agencia de Protección de Ohio, 1987; Leonard & Orth., 1986; Miller *et al.*, 1988; Steedman, 1988); que es considerado para analizar los ensambles de macroinvertebrados. De igual manera, se cuenta con los protocolos de bioevaluación rápida de la Agencia de Protección Ambiental de Estados Unidos (EPA US) (Plafkin *et al.*, 1989); el índice de bienestar - Index of Well-Being - (Gammon, 1976; Gammon *et al.*, 1981) y el índice de la comunidad de invertebrados (Agencia de Protección de Ohio, 1987; DeShon, 1995).

Norris & Norris (1995) mencionan que en países como Australia el uso de estadísticas multivariadas para producir modelos de predicción de calidad del agua podrían ser una herramienta útil. Reynoldson *et al.* (1997) consideraron que los métodos multivariados eran complejos, frente a lo cual se ha cambiado la percepción, dado que otros autores han indicado que este tipo de análisis posibilitan la evaluación de conjuntos de datos robustos de calidad del agua, para finalmente determinar los factores relevantes en la contaminación y los parámetros responsables de las variaciones espacio-temporales en el cuerpo de agua en estudio (Koklu *et al.*, 2010; Zhang *et al.*, 2009).

De igual forma, estas técnicas estadísticas han permitido identificar semejanzas o diferencias entre puntos de monitoreo de calidad del agua y la influencia de las características naturales e intervenciones antrópicas en los parámetros medidos en una corriente superficial (Barakat *et al.*, 2016; Mitsch *et al.*, 2009; Shrestha & Kazama, 2007; Wang *et al.*, 2013). Algunos enfoques multivariados adoptan análisis estadísticos para predecir patrones de los organismos específicos en un sitio, y los organismos que se encontrarían en

ausencia de un estrés ambiental importante. De esta manera, se ha logrado demostrar que los análisis multivariados son una herramienta práctica para realizar biomonitoreos (Li *et al.*, 2010).

Con el propósito de analizar los puntos de muestreo de acuerdo con los gradientes de contaminación que se presentan en un río estudiado, se ha evaluado la asociación de técnicas multivariadas con organismos como las diatomeas epilíticas (Gómez, 1999). Otra propuesta ha sido evaluar el efecto de la perturbación en el ambiente sobre la composición de las especies de diatomeas con base en un análisis de gradiente indirecto, a partir del cual se propuso un índice multimétrico bajo la técnica de escalamiento multidimensional no métrico (Lane & Brown, 2007). Del mismo modo, se han realizado estudios donde los organismos indicadores son las algas perifíticas; tal es el caso del río Sava (Europa) en el que las diatomeas lograron manifestar la influencia de diversas actividades antrópicas sobre la calidad del agua (Vasiljević *et al.*, 2017).

En el río Potenza (Italia) se establecieron tres índices para expresar la sensibilidad de las diatomeas a los nutrientes (TI-D), materia orgánica (SI-D) y sales minerales (HI-D) con base en análisis de correlación, los cuales facilitaron la evaluación del efecto de la contaminación de tipo orgánico en los cuerpos de agua sobre las algas (Dell'Uomo & Torrisi, 2011). También, se identificó el Índice de Diatomeas Pampeano (IDP), mediante el cual los resultados de calidad biótica fueron correlacionados con las principales características químicas del agua (Gómez & Licursi, 2001).

Haase & Nolte (2008) desarrollaron el índice Invertebrate Species Index (ISI), teniendo en cuenta información de macroinvertebrados acuáticos hasta un nivel de especie adaptado para las comunidades específicas de los cuerpos de agua del sureste de Queensland (Australia), con la intención de brindar un mayor nivel de detección del cambio biológico ante la afectación de las corrientes por nutrientes. Por su parte, (Chalar *et al.*, 2011) propusieron en Uruguay el primer índice biótico para los ríos suramericanos, conocido como Trophic State Index for Benthic Invertebrates (TSI-BI), mediante el cual se estableció una clasificación del estado trófico de los sistemas lóticos de acuerdo con la abundancia de macroinvertebrados acuáticos y variables ambientales.

Con base en las anteriores investigaciones, Forero *et al.* (2014) plantearon una herramienta cuantitativa a partir de la abundancia de los macroinvertebrados acuáticos y las correlaciones entre las variables fisicoquímicas, teniendo en cuenta un gradiente ambiental para establecer el estado ecológico de la cuenca del río Negro (Colombia), y así formular el índice de calidad ecológica ICE_{RN-MAE} el cual se diferencia del enfoque cualitativo del BMWP-Col.

Este tipo de investigaciones han permitido fortalecer el uso de los índices ecológicos, los cuales se refieren a métodos de indicación que van desde el establecimiento de puntajes a taxones determinados, hasta modelos matemáticos complejos que estiman la respuesta de algunas especies acuáticas a las variaciones ambientales (Samboni Ruiz *et al.*, 2007). Es así como diferentes técnicas de estadística multivariada son utilizadas hoy en día para evaluar datos ecológicos y con base en ellas, relacionar la composición de especies con los gradientes ambientales (Greenacre & Primicerio, 2014; Leps & Smilauer, 2003).

3 PLANTEAMIENTO DEL PROBLEMA

Las afectaciones que se presentan sobre la calidad fisicoquímica e hidrobiológica en los cuerpos de agua pueden generar cambios en su estado ecológico y los niveles de organización de los ecosistemas acuáticos. Estas alteraciones se relacionan directamente con el aprovechamiento del agua superficial, por lo cual se hace necesario realizar seguimientos adecuados, así como el establecimiento de normatividad para potenciar el uso del recurso hídrico.

Tal como lo menciona Villamarín *et al.* (2013) “es fundamental implementar sistemas de evaluación de la calidad ecológica de los ríos, que puedan ser utilizados por las entidades gestoras del agua y que ayuden a la conservación del recurso hídrico”. De este modo, muchos países han mejorado sus leyes y políticas hídricas tales como los participantes en la Cumbre Mundial sobre el Desarrollo Sostenible celebrada en el año 2002 (GWP, 2009).

Países como Ghana, África del Sur y Malí adoptaron un enfoque integral en sus leyes y códigos del agua. España y Francia, han practicado la gestión del agua en cuencas durante décadas, y se han creado comisiones internacionales en Europa para los ríos Rin, Mosa, Scheldt, Mosela y Sarre, así como para el Lago Ginebra. Por su parte, Brasil promulgó su Ley Nacional de Aguas para promover la gestión del agua en forma integrada, descentralizada y participativa (GWP, 2009).

Colombia no ha sido ajeno a los enfoques sobre políticas hídricas, por lo cual se han desarrollado diferentes instrumentos orientados a la gestión integral del recurso hídrico que promueven la planeación del uso coordinado de los recursos, tal como se considera en el artículo 316 del Decreto Ley 2811 de 1974. También es el caso de la Política Nacional para Gestión Integral del Recurso Hídrico (PNGIRH) expedida en 2010 por el Ministerio de Ambiente, Vivienda y Desarrollo Territorial con el propósito de garantizar

la sostenibilidad del recurso hídrico y fortalecer el concepto de gobernanza del agua. En este mismo año, se emitió el Decreto 3930, donde se incluyen algunas disposiciones en cuanto a los usos y ordenamiento del recurso hídrico, así como ciertas consideraciones sobre los vertimientos líquidos.

Por su parte, el Plan Nacional de Desarrollo 2010-2014 expedido mediante la Ley 1450 de 2011 en su Capítulo VI sobre Sostenibilidad Ambiental, estableció que las cuencas hidrográficas serían el instrumento para avanzar en la planificación ambiental del territorio con base en la consolidación de pautas para la ordenación de las cuencas y de los acuíferos asociados; además de proponer la expedición de un manual enfocado en la gestión integral del recurso hídrico. A partir de estas y otras bases legislativas, se han desarrollado guías para planes como el Plan de Ordenación y Manejo de Cuencas Hidrográficas (POMCA) y el Plan de Ordenamiento del Recurso Hídrico (PORH) los cuales cuentan con fines diferentes, pero que son complementarios en el ejercicio de la gestión por la inclusión de enfoques para la evaluación de la calidad del agua a partir de la estimación de índices de calidad fisicoquímica e hidrobiológica (MADS, 2014, 2018)

En cumplimiento a lo establecido en instrumentos como la política nacional PNGIRH, y los planes de ordenamiento POMCA y PORH, las autoridades ambientales se enfrentan con impedimentos en su afán de tomar decisiones, puesto que deben recurrir a alternativas con alto grado de interpretación técnica o generar herramientas bajo supuestos que simplifican las condiciones del medio acuático en un alto porcentaje, lo que termina por conjugar una cantidad de parámetros en un solo valor.

Es así como se han desplegado diferentes estrategias para el análisis de los ecosistemas acuáticos a nivel de cuenca, con el propósito de conocer los procesos que se presentan en el cuerpo de agua principal, sus afluentes y la posible interacción entre las aguas subterráneas y superficiales. Actualmente, algunos de los lineamientos se encuentran plasmados en la Guía técnica para la formulación de PORH, donde se incluye la aplicación de índices de calidad a partir de variables fisicoquímicas, así como indicadores biológicos basados en comunidades hidrobiológicas que evalúan su estructura, diversidad, densidad, riqueza, dominancia y equidad (MADS, 2018).

Los índices propuestos en la guía para la formulación de PORH se basan generalmente en métodos adaptados para ciertas zonas del país, y se implementan con base en las características locales de los ecosistemas. Frente a ello, se ha propuesto evaluar el contexto ecológico de los ecosistemas acuáticos a partir del ICE, el cual se basa en métodos de estadística multivariada, que permiten conjugar la mayor

cantidad de variables fisicoquímicas e hidrobiológicas que se analizan durante los ejercicios de gestión y obtener de esta manera resultados de acuerdo con las características particulares del área de interés.

El presente trabajo de investigación pretende optimizar la interpretación ecológica de una herramienta numérica que abarca diferentes métodos multivariados para la evaluación de los ecosistemas acuáticos con base en información fisicoquímica y biológica. A partir de esta herramienta se espera conocer la respuesta de organismos hidrobiológicos frente a su exposición a diferentes condiciones ambientales en un cuerpo de agua y la interpretación asociada al comportamiento identificado.

Esta información servirá para potenciar los instrumentos existentes relacionados con la gestión del recurso hídrico, dado que se suministran nociones de una interpretación de aspectos ecológicos del ICE y se aborda la modelación matemática de uso libre como herramienta útil y eficiente para evaluar la condición ecológica de un cuerpo de agua bajo las características de la zona, lo que facilitará la toma de decisiones ajustadas a las particularidades que presentan las diferentes cuencas hidrográficas.

De este modo se plantea la pregunta:

¿Cómo optimizar el cálculo e interpretación del índice de calidad ecológica a partir de la implementación de un modelo matemático y estadístico?

4 OBJETIVOS

4.1 OBJETIVO GENERAL

Implementar un modelo en el software R para el cálculo del índice de calidad ecológica que optimice su aplicación y facilite su interpretación en el ámbito de la gestión integral del recurso hídrico.

4.2 OBJETIVOS ESPECÍFICOS

- Definir los métodos estadísticos y operaciones matemáticas apropiadas en el software R para el cálculo del índice de calidad ecológica.

- Establecer nociones de la interpretación del índice de calidad ecológica que abarque aspectos estadísticos y ecológicos enfocados a los ecosistemas acuáticos.
- Validar el funcionamiento del modelo de cálculo propuesto en el software R en tres casos de estudio.

5 MARCO TEÓRICO

El uso de índices ecológicos es una de las estrategias para apoyar el diagnóstico en diferentes enfoques, dado que el considerar especies indicadoras para caracterizar áreas ecológicas específicas ha sido de gran importancia en la biología de conservación y restauración (V. González *et al.*, 2011). Los indicadores ecológicos han sido utilizados con mayor frecuencia en comunidades terrestres, puesto que logran representar diferentes condiciones ambientales (Heino *et al.*, 2009; Noss, 1990; Spellerberg, 2005). También, se han utilizado los índices ecológicos con base en técnicas multivariadas para evaluar modelos de predicción paleoecológica (Spradley *et al.*, 2019) y se han evaluado condiciones específicas de cierta especie frente a un estado ecológico a partir de análisis estadísticos (McGeoch *et al.*, 2002).

En cuanto al estudio de la biota acuática, los reportes hidrobiológicos son funcionales en la evaluación de ecosistemas acuáticos, siendo algunos de los índices más utilizados el índice de diversidad de Shanon & Weaver (1949), el de dominancia de Simpson (Simpson, 1949), el índice de equidad de Pielou (1984), el índice BMWP adaptado para Colombia Roldán-Perez (2003) y el índice de valor de importancia (IVI) (Lamprecht, 1990), entre otros.

De igual manera, se ha identificado el uso de indicadores de integridad ecológica para ofrecer una evaluación de la condición de un ecosistema al integrar el papel de la biodiversidad en el mantenimiento de las propiedades de los procesos ecológicos (Mora, 2019). Así, en el contexto acuático los índices han buscado integrar el monitoreo de la calidad del agua y los indicadores ambientales con las prácticas de gestión del recurso (Pander & Geist, 2013).

Desde la perspectiva de la evaluación con índices, algunas estrategias han considerado esquemas con inclusión de estadística multivariada para los análisis ecológicos; una de estas propuestas se encuentra en la Guía técnica para la formulación de PORH coordinada por la Dirección de Gestión Integral del Recurso

Hídrico (DGIRH) del MADS, dado que se identificó que en el ejercicio de la gestión no se estaba incluyendo el estado ecológico de los sistemas hídricos (MADS, 2018).

Los lineamientos establecidos en la Guía para la formulación de PORH fueron aplicados por la DGIRH durante el año 2013 en varias regiones del país con condiciones diferentes desde el punto de vista climático, social, morfológico y de conflictos por uso, calidad y demanda del agua. En el marco de estos ejercicios, se implementó una alternativa para incluir los recursos hidrobiológicos como una valoración del estado ecológico de las corrientes con base en el ICE.

El ICE se describe como un método basado en estadística multivariada, mediante el cual se determinan valores óptimos y de tolerancia de un morfotipo en función de un gradiente ambiental. El ICE toma como base el monitoreo biológico, el cual considera los atributos estructurales y/o funcionales de las especies, las poblaciones, las comunidades y los ecosistemas; además de integrar las particularidades fisicoquímicas que presenta cada cuenca (Forero, 2012; Hart, 1994).

Con base en lo expuesto por Chalar *et al.* (2011); Haase & Nolte (2008) y la aplicación del ICE realizada en Colombia por Forero (2012) y Forero *et al.* (2014) a partir de macroinvertebrados acuáticos para la cuenca del río Negro, el ICE fue incluido en la Guía técnica para la formulación de PORH que fue adoptada por el MADS mediante la Resolución 751 de 2018. Para la ejecución del ICE se contemplan los pasos listados a continuación (MADS, 2018), de los cuales es posible extraer información relevante sobre el cuerpo de agua de interés:

1. Estandarizar los valores de los parámetros fisicoquímicos y bióticos, teniendo en cuenta el método propuesto por Guisande-González *et al.* (2006) considerado en la Ecuación 1.

$$VE = \frac{x - Min}{Max - Min}$$

Ecuación 1. Estandarización de datos

Donde VE se refiere al valor estandarizado; x es el registro para cada variable; Max y Min corresponden a los valores máximo y mínimo de cada variable. Con esta estandarización, los datos se establecen en un rango entre 0 y 1 sin perder la esencia de las distancias entre los registros.

2. Analizar las correlaciones entre los parámetros físicos, químicos, y la abundancia de los organismos. En este paso se establece el grado de correlación entre los parámetros mencionados

y los diferentes morfotipos, con el fin de eliminar del análisis posterior los parámetros altamente dependientes.

3. Estimar el gradiente ambiental en el cuerpo de agua a través de análisis multivariados mediante el empleo de programas estadísticos, donde las variables explicativas corresponden a los parámetros fisicoquímicos y las variables respuesta se refiere a la comunidad biótica (C. J. Ter Braak & Smilauer, 2002).
 - 3.1. Evaluar la heterogeneidad de la composición de la comunidad mediante un análisis de correspondencia sin tendencia (DCA por sus siglas en inglés) para definir el tipo de modelo de respuesta de los diferentes grupos taxonómicos a los gradientes ambientales. Cuando el gradiente es corto, se debe aplicar un una aproximación lineal y si es largo, se debe ejecutar un modelo unimodal (Leps & Smilauer, 2003).
 - 3.2. Calcular la significancia de la relación de los organismos con las variables ambientales a través de la prueba de permutación de Monte Carlo, donde se considera que la hipótesis nula del modelo es que la densidad de los organismos es independiente de las variables ambientales (Leps & Smilauer, 2003). A partir de ello es posible identificar las variables admitidas en el modelo.
 - 3.3. Verificar cuáles variables se encuentran altamente correlacionadas de acuerdo con los resultados del análisis de correlación. Se advierte que la selección de variables a incluir en el modelo es un paso que requiere de atención dado que es crítico para el análisis multivariado, por lo cual se recomienda tener en cuenta las variables significativas en la prueba de Monte Carlo; analizar el factor de inflación de la varianza para identificar la colinealidad en el modelo; revisar las variables con alta correlación entre ellas y con la abundancia de los organismos y considerar variables con significado ecológico importante para la comprensión de la composición y abundancia de los organismos en el sistema estudiado.
 - 3.4. Con las variables seleccionadas teniendo en cuenta las anteriores consideraciones, se debe ejecutar un análisis de correspondencia canónica (CCA) o análisis de redundancia (RDA), dado el caso, y verificar su significancia. A partir de los resultados de estos modelos, es

posible determinar la posición de las estaciones de monitoreo a lo largo del gradiente ambiental de acuerdo con la puntuación de la muestra en el eje.

- 3.5. Realizar un reescalamiento de los valores para facilitar la interpretación del índice. El máximo puntaje (mayor heterogeneidad) se equipará con un valor de 1 y el mínimo (menor heterogeneidad) con 10 y se hace una regresión. La ecuación resultante de la regresión permite calcular el resto de valores del eje. De esta forma, se conservan las diferencias entre los puntajes del gradiente y los resultados del índice se pueden expresar en una escala de 1 a 10.
4. Identificar los rangos de tolerancia y valores óptimos para cada taxón en función del gradiente ambiental del cuerpo de agua estudiado a través de un modelo de promedios ponderados, con el fin de establecer cuáles son los organismos indicadores.
 - 4.1. Determinar los organismos indicadores (baja tolerancia y alto óptimo) para el cuerpo de agua a partir del modelo de promedios ponderados. Tener en cuenta que un rango estrecho de tolerancia ambiental de los taxones es indicador de una condición ambiental concreta y, por tanto, responde a cualquier disturbio ambiental. Como en cualquier procedimiento en el cual se buscan organismos indicadores, el análisis puede verse afectado cuando los gradientes evaluados no son lo suficientemente amplios y cuando la mayoría de las estaciones tengan una misma condición ambiental.
5. Calcular el índice de calidad ecológica por estación mediante la aplicación de la fórmula propuesta por Haase & Nolte (2008) que se presenta en la Ecuación 2.

$$ICE_i = \frac{\sum_{i=1}^n (Tol_i * Op_i * Ab_i)}{\sum_{i=1}^n (Tol_i * Ab_i)}$$

Ecuación 2. Ecuación para el cálculo del índice de calidad ecológica

Donde Op_i corresponde al valor óptimo de cada taxón determinado a partir del promedio ponderado; Tol_i se refiere al valor de la tolerancia de cada taxón determinado a partir del promedio ponderado y Ab_i es el promedio de las abundancias de cada taxón por estación transformada a $\text{Log}_{10} + 1$ (promedio de la abundancia + 1).

- 5.1. Realizar una correlación entre los resultados del índice y los puntajes obtenidos en el análisis multivariado para verificar el cálculo del índice. Una buena relación entre las dos variables se estima con valores superiores a 0,7.
6. Realizar un análisis de correlación entre los valores del índice y los parámetros fisicoquímicos incluidos en el modelo de ordenación sin transformar, para luego analizar que altas correlaciones ($R^2 \geq 0,5$ o $R^2 \leq -0,5$) implican que dichos parámetros son importantes ecológicamente y que los cambios en estos alteran significativamente la distribución y abundancia de los organismos en el cuerpo de agua.
7. Ejecutar un análisis de agrupamiento (clúster) del índice por estación. Para este análisis se recomienda el método de agrupamiento de Ward y las distancias euclidianas. Se sugiere además la separación por tres conglomerados con el fin de generar tres categorías para facilitar la visualización espacial de los resultados.

Dentro de los softwares considerados por el MADS (2018) se encuentra CANOCO, el cual según Ter Braak (1989) fue diseñado para ofrecer técnicas estadísticas a los ecologistas que se habían interesado por conocer las respuestas de ciertas comunidades ante las variaciones en las condiciones ambientales. El programa CANOCO para Windows es un software comercial que requiere una licencia válida para su uso (Leps & Smilauer, 2003).

Para el análisis y visualización de datos ecológicos y paleoambientales se utiliza el programa de Microsoft Windows llamado C2. Este software realiza reconstrucciones ambientales a partir de un conjunto de datos secundarios mediante la aplicación de funciones de transferencia paleoecológica, o modelos de inferencia (Juggins, 2005). Se destaca que la versión gratuita de C2 se encuentra limitada a matrices de datos con menos de 75 muestras, tanto en sus filas como columnas; por lo cual se requiere de la obtención del licenciamiento del programa para eliminar la restricción.

Asimismo, son considerados programas estadísticos para análisis de datos como STATGRAPHICS Centurion XVI, el cual es una herramienta licenciada y compatible con Windows que permite la ejecución de análisis de correlaciones y de conglomerados, donde se combina la parte analítica con interfases gráficas. En la Figura 1 se esquematiza el cálculo del ICE a partir de lo contemplado en la Guía Técnica de PORH.

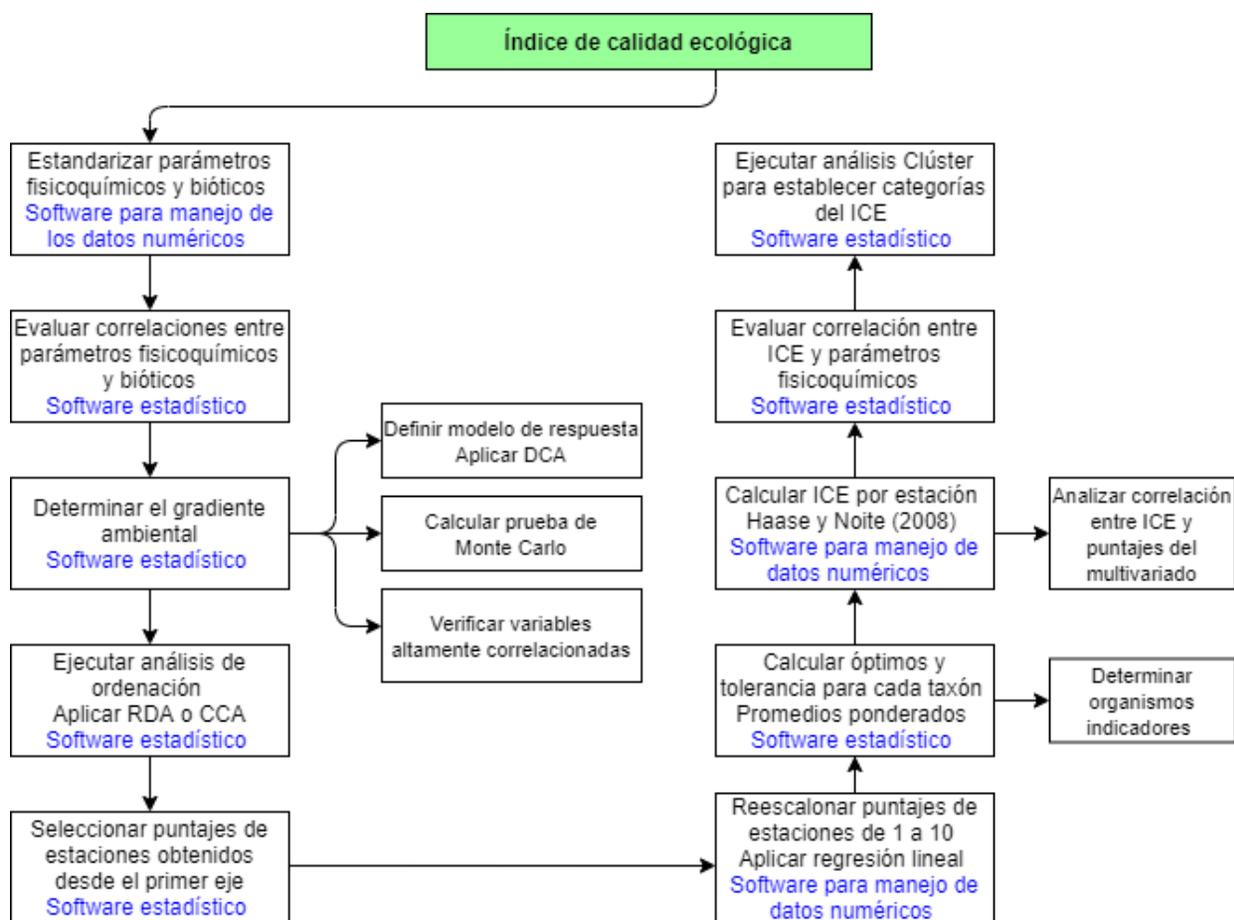


Figura 1. Esquema de cálculo del ICE. MADS, 2018

6 METODOLOGÍA

En los siguientes apartes se presentan los aspectos direccionados al cumplimiento de los objetivos específicos, donde se indaga sobre los métodos estadísticos y operaciones matemáticas para generar un

modelo de cálculo del ICE en un software de uso libre, así como las nociones que dan pie a la interpretación del índice. Finalmente, se describe como se realizará la validación del modelo de cálculo propuesto en tres casos de estudio.

6.1 MÉTODOS DISPONIBLES EN R PARA EL CÁLCULO DEL ICE

A partir de lo dispuesto por el MADS (2018) en cuanto al cálculo del ICE, se identificó que los programas estadísticos sugeridos son de uso restringido, dado que requieren de licencia para su implementación. Con el propósito de generar un modelo abierto en un software de uso gratuito, se consideró que R es un entorno y lenguaje de programación que proporciona gran cantidad de procedimientos estadísticos. Se destaca que R es un software gratuito y es distribuido bajo un *copyleft*, lo que indica que el programa y las mejoras de este son libres (Venables *et al.*, 2019).

Con base en lo anterior, se analizó la inclusión de los métodos estadísticos multivariados y análisis numéricos requeridos para el cálculo de dicho índice en el software estadístico R, y se evidenció que este programa implementa los métodos disponibles en CANOCO, C2 y STATGRAPHICS, tal como se resume en la Tabla 1.

Tabla 1. Métodos requeridos para el cálculo del ICE y su disponibilidad en softwares

Métodos/Análisis requeridos		Software de uso comercial			Software de uso gratuito
		CANOCO	C2	STATGRAPHICS	R
Estandarización de datos		x			x
Análisis de ordenación	CA, PCA, DCA, RDA, CCA	x			x
Funciones de transferencia	WA, PLS, WAPLS		x		x
Análisis de agrupamiento	Ward, centroide			x	x
Análisis de correlación	Spearman, Pearson			x	x

El lenguaje de programación R cuenta con paquetes y funciones que permiten el ensamblaje de un script operativo. En el caso del ICE, se analizaron diferentes métodos estadísticos multivariados y operaciones matemáticas encaminadas a generar un modelo de cálculo para el índice. En cuanto a la manipulación de datos, se presentan paquetes como ‘gdata’ que involucran funciones para leer archivos de Microsoft Excel

(Warnes *et al.*, 2017); asimismo, el paquete ‘readxl’, permite la lectura de archivos en formato .xls o .xlsx (Wickham *et al.*, 2019).

Por su parte, al considerar que los resultados de las mediciones fisicoquímicas generalmente presentan escalas numéricas diferentes al igual que en los registros hidrobiológicos, se requiere realizar una estandarización de la información, lo cual podría aplicarse mediante las funciones ‘decostand’ o ‘wisconsin’ incluidas en el paquete ‘vegan’ (Oksanen *et al.*, 2019).

Para la estimación de la longitud del gradiente ambiental, se consideró el análisis de correspondencia sin tendencia (DCA) integrado en el paquete ‘vegan’ de R bajo la función ‘decorana’ (Oksanen *et al.*, 2019). Con base en la información de las variables explicativas, se evalúa la longitud del gradiente ambiental para identificar el método de ordenación que se ajusta a la información recopilada del cuerpo de agua de interés.

De igual manera, el paquete ‘vegan’ proporciona diferentes métodos de ordenación para ser implementados en casos de ecología comunitaria (Oksanen *et al.*, 2019). Particularmente, las funciones ‘cca’ y ‘rda’ de este paquete ejecutan análisis de correspondencia canónico y análisis de redundancia de forma similar a los métodos del software propietario CANOCO (Ter Braak, 1989), pero la implementación actual en R se basa en lo descrito por Legendre y Legendre (2012). R también consolida el paquete ‘CCA’ (I. González & Déjean, 2015), donde funciones como ‘cancor’ y ‘cc’ permiten realizar asociaciones entre dos conjuntos de datos, solo que hacen referencia a un análisis de correlación canónica.

Con relación al análisis de óptimos y tolerancias, el paquete ‘rioja’ considera métodos numéricos para inferir el valor de una variable ambiental dado un conjunto de datos de especies y valores ambientales asociados (Juggins, 2018). Este paquete permite la ejecución de funciones de transferencia y otros métodos para analizar datos científicos del Cuaternario, como ‘WA’ y ‘WAPLS’.

En cuanto al análisis de agrupamiento, varias técnicas separan las observaciones en diferentes grupos; como el caso de la función ‘hclust’ del paquete ‘stats’, donde se cuenta con métodos como ‘ward’, ‘median’ y ‘centroid’. También, se puede implementar la función ‘agnes’ del paquete ‘clúster’ el cual presenta otros métodos para realizar análisis de agrupamiento. En cuanto a los análisis de correlaciones, el paquete ‘Hmisc’ calcula una matriz de coeficientes de correlación ya sea por el método de Pearson o de Spearman (Harrell, 2019).

6.2 INTERPRETACIÓN DEL ICE

Con el propósito de brindar nociones en cuanto a la interpretación del ICE para aspectos tanto ecológicos como estadísticos, se abordan ciertos criterios a considerar durante la ejecución del índice, así como también se presentan algunas definiciones que configuran el cálculo del ICE. Además, se analizan ciertos métodos estadísticos multivariados que son aplicados en temas de ecología comunitaria, tales como identificar los cambios en las comunidades bióticas a partir de las condiciones ambientales que presentan los ecosistemas, en este caso, los acuáticos.

6.2.1 Detección de inconsistencias y corrección de datos

La información base para la estimación del ICE parte de labores realizadas tanto en campo como en laboratorio. En el caso de presentarse valores erróneos en los resultados o fallas durante la toma de muestras fisicoquímicas e hidrobiológicas, las matrices de datos pueden presentar datos faltantes. Frente a lo anterior, se requiere la eliminación de la variable según el criterio del experto o realizar una imputación de datos, a través de métodos adecuados según la naturaleza de la información (CEPAL, 2007).

Con respecto a los métodos de imputación de datos, se pueden utilizar diversas técnicas para completar la información faltante sin tener que eliminar alguna variable. Por ejemplo, se puede emplear el método de medias no condicionadas, mediante el cual los datos omitidos se sustituyen por el promedio de los valores de la variable (CEPAL, 2007; Leps & Smilauer, 2003). También, se puede ejecutar una imputación a partir de medias agrupadas por datos condicionados para conformar categorías de covariables correlacionadas con la variable que se está analizando, y luego completar la información con datos de submuestras que presenten características similares. De igual manera, se podrían establecer variables latentes con base en un modelo de regresión con las muestras que se encuentren completas y así generar una variable indicadora para predecir las observaciones con datos faltantes (CEPAL, 2007).

Por otro lado, se recomienda realizar una estandarización de las matrices de datos explicativas y de respuesta, teniendo en cuenta que la intención de este procedimiento es suministrar el mismo peso a todas las variables fisicoquímicas y organismos considerados en cada uno de los cálculos. Este procedimiento influye en el efecto de los valores atípicos durante los análisis estadísticos cuando se trabaja con parámetros en diferentes unidades y escalas.

6.2.2 Gradiente ambiental

Términos como el gradiente ambiental son relevantes durante la evaluación de la calidad ecológica, dado que este permite interpretar la composición de una comunidad en términos de la respuesta de las especies a las variaciones naturales o antrópicas sobre las propiedades que caracterizan un ecosistema (Leps & Smilauer, 2003; Ter Braak & Prentice, 1988). No necesariamente los gradientes consideran la realidad física como continua en el espacio o el tiempo, aunque sí son un concepto útil para analizar las distribuciones de los organismos en el espacio y el tiempo a partir de una modelación de las abundancias de las especies en función de las variables ambientales (Ter Braak & Prentice, 1988).

En los métodos estadísticos aplicados a estudios ecológicos, el término análisis de gradiente se utiliza para cualquier método que pretenda relacionar la composición de especies con los gradientes ambientales, tales como los que se consideran en la Figura 2 (Leps & Smilauer, 2003; Lepš & Šmilauer, 1999). Generalmente, las técnicas adaptadas a este tipo de análisis reducen los datos a unos cuantos ejes de ordenación, los cuales se interpretan con base en el conocimiento de las especies y del ecosistema de interés (Ter Braak & Prentice, 1988).

En cuanto a la estimación del gradiente ambiental, se debe considerar que el análisis de la heterogeneidad en la estructura de una comunidad biológica se ha evaluado a partir de métodos estadísticos multivariados que permiten analizar la composición de las especies en un sitio en particular mediante la identificación de los organismos “cercaños” (Ter Braak & Prentice, 1988).

Tal es el caso de Opdam *et al.* (1984) quienes analizaron la variación en la composición de las especies de aves teniendo en cuenta los cambios en la vegetación en el valle del río Rhine, en la provincia de Utrecht (Países Bajos) a partir de un análisis de correspondencia sin tendencia (Detrended Correspondence Analysis - DCA). En este estudio se consideraron los puntos de las especies como óptimos en el espacio de ordenación y se analizó la composición de diferentes regiones del valle de acuerdo con las especies de aves más cercanas a la sección de interés. Además, interpretaron la longitud del primer eje del DCA como el gradiente ambiental en unidades de desviación estándar (σ).

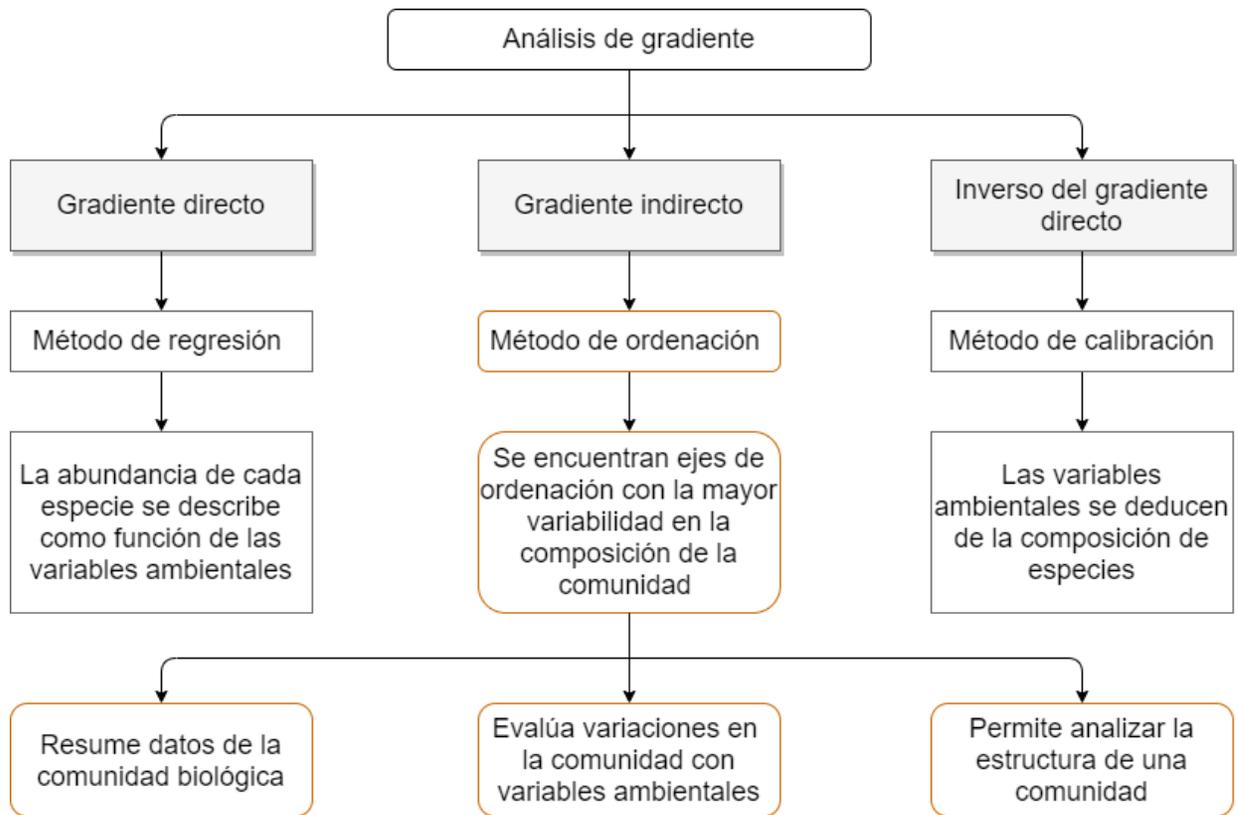


Figura 2. Métodos estadísticos para el análisis de gradiente. Adaptado de Ter Braak, C. y Prentice, C., 1988; Lepš y Šmilauer, 2003

El DCA surgió como una corrección al método del promedio recíproco o análisis de correspondencia, puesto que se presentaban dos fallas principalmente (Hill & Gauch, 1980): primero, el segundo eje canónico era una distorsión en forma de arco del primer eje, debido a que estos dos ejes se encontraban correlacionados, más no eran independientes; y segundo, no se conservaban las distancias ecológicas, dado que al ordenar las especies a partir de su posición con respecto al primer eje, la desviación estándar era más pequeña hacia los extremos del gradiente.

Mediante el DCA, se aseguró la ortogonalidad entre los ejes y se igualó la dispersión media dentro de la muestra de las puntuaciones de las especies a lo largo del gradiente; para finalmente obtener el perfil de abundancia de especies “promedio” en unidades de desviación estándar.

Para dar claridad al porqué es posible establecer que el gradiente corresponde a la longitud del primer eje del DCA en unidades de desviación estándar, se retoma el tema de gradiente ambiental, el cual permite conocer la composición de una comunidad al determinar la tasa de cambio de las especies. Por su parte,

la propiedad que permite medir la tasa de cambio de especies a lo largo de un gradiente es la diversidad beta y una unidad conveniente para expresar dicha tasa de cambio es el número de desviaciones estándar a lo largo del gradiente. De este modo, al implementar un DCA los ejes se escalan en unidades de diversidad beta (Gauch, 1982), por tanto, para las especies el escalamiento se realiza en unidades de desviación estándar (σ).

Ahora, para considerar que la longitud del primer eje del DCA se refiere al valor del gradiente ambiental Legendre, P. y Legendre, L. (1998) precisan en el libro *Numerical Ecology* que a lo largo de un gradiente regular aparece una especie, ésta se eleva a su mayor valor modal y desaparece a una distancia aproximada de 4σ . También mencionan que un cambio medio en la composición de especies ocurre entre 1σ y 1.4σ , por tanto, la longitud del primer eje en un DCA puede corresponder a una medida aproximada de la longitud del gradiente ambiental en unidades de rotación de especies, puesto que en este eje se logra almacenar los cambios en la composición de las mismas.

En síntesis, mediante un análisis de correspondencia sin tendencia o DCA se estima la longitud del gradiente ambiental con base en la información de variables biológicas, las cuales son medidas en unidades de desviación estándar (σ) para cada eje. Si la longitud del gradiente es corta (menor a 2.5σ), las especies presentan un comportamiento monótono y, por tanto, es apropiado utilizar un modelo lineal. Por su parte, al obtener una longitud del gradiente largo (mayor a 2.5σ), el uso de un modelo lineal no sería conveniente, dado que los datos son más heterogéneos y una gran cantidad de especies no se ajustarían al supuesto de respuesta lineal, por lo cual un modelo unimodal es el más adecuado (Legendre & Birks, 2012).

Al considerar la Figura 3 (a), el primer modelo representa la relación lineal entre el valor de abundancia (Y) de una especie y una variable ambiental (X) ajustada a datos artificiales identificados por puntos. Por su parte, la Figura 3 (b) se reproduce una relación unimodal entre el valor de abundancia (Y) de una especie y una variable ambiental (X).

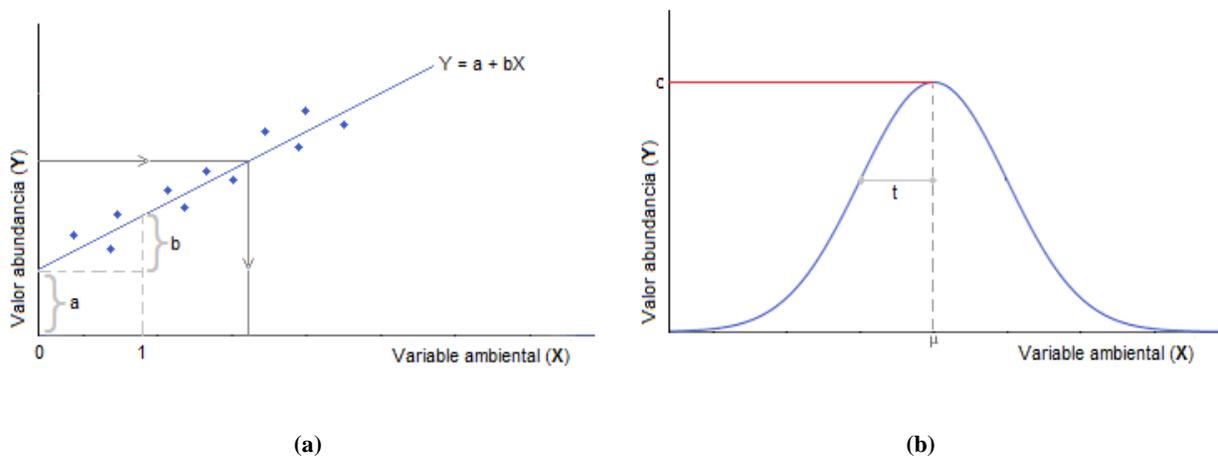


Figura 3. Modelo de respuesta de especies: (a) Lineal, donde a representa el intercepto y b el coeficiente de regresión; (b) Unimodal, donde μ se refiere al valor óptimo, t a la tolerancia y c al mayor valor de abundancia (Adaptado de Legendre & Birks, 2012)

6.2.3 Análisis de ordenación

Luego de estimar la longitud del gradiente ambiental y definir si la heterogeneidad en la estructura de la comunidad biológica corresponde a un modelo lineal o unimodal, se debe ejecutar un análisis de ordenación el cual se encuentra condicionado a obtener la máxima correlación posible entre las variables respuesta a partir de las variables explicativas. En el caso del modelo lineal, se hace referencia a un análisis de redundancia (RDA) y el unimodal se asocia con un análisis de correspondencia canónica (CCA).

Durante la verificación de los métodos, se evidenció que se pueden presentar confusiones al momento de seleccionar el análisis de ordenación, donde la abreviación CCA se utiliza tanto para el análisis de correspondencia canónica como para la correlación canónica; además, ambos métodos evalúan relaciones multivariantes entre dos conjuntos de datos (Ter Braak & Prentice, 1988).

Por otro lado, la correlación canónica también podría confundirse con el análisis RDA puesto que ambos asumen modelos lineales. Sin embargo, en el RDA se configuran dos conjuntos de datos, correspondientes a las variables independientes y variables dependientes, mientras que la correlación canónica considera ambos conjuntos de datos en igualdad de condiciones (Ter Braak & Prentice, 1988).

Con base en las anteriores consideraciones, se establece que para estudios ecológicos como la aplicación del ICE, se requieren de técnicas multivariadas que se asocien con el análisis de gradiente, donde los datos

de las especies u organismos se someten a un método de ordenación (Correspondence Analysis - CA, por ejemplo) para encontrar ejes principales de variación o variables latentes con un buen ajuste a los datos de las especies (Dargie, 1984; C. J. . Ter Braak & Prentice, 1988). Estos ejes se interpretan en términos de la variación conocida en el entorno mediante el uso de métodos gráficos (Gauch, 1982).

De este modo, se precisa que el análisis RDA conserva la distancia euclidiana entre los elementos de la matriz \hat{Y} de valores ajustados de Y por una regresión múltiple convencional de las variables explicativas X, como se presenta en la Figura 4 (a). Por su parte, el análisis CCA se refiere a una forma ponderada del análisis de redundancia, aplicado a la matriz dependiente \bar{Q} representado en la Figura 4 (b). En este método se realiza una aproximación a las distancias Chi-cuadrado entre los elementos de la matriz de datos dependiente bajo el supuesto de que los vectores de ordenación canónica se encuentran relacionados al máximo con las combinaciones lineales ponderadas de las variables explicativas (Legendre P. & Legendre L.F., 2012).

Con el propósito de disminuir la inestabilidad de los modelos de respuesta a utilizar para el cálculo del gradiente ambiental, se recomienda la identificación de posibles dependencias lineales fuertes entre las variables explicativas consideradas en el método de ordenación. Las dependencias lineales pueden explorarse calculando los factores de inflación de la varianza de las variables (VIF), que miden la proporción en que se infla la varianza de un coeficiente de regresión en presencia de otras variables explicativas (Borcard *et al.*, 2011).

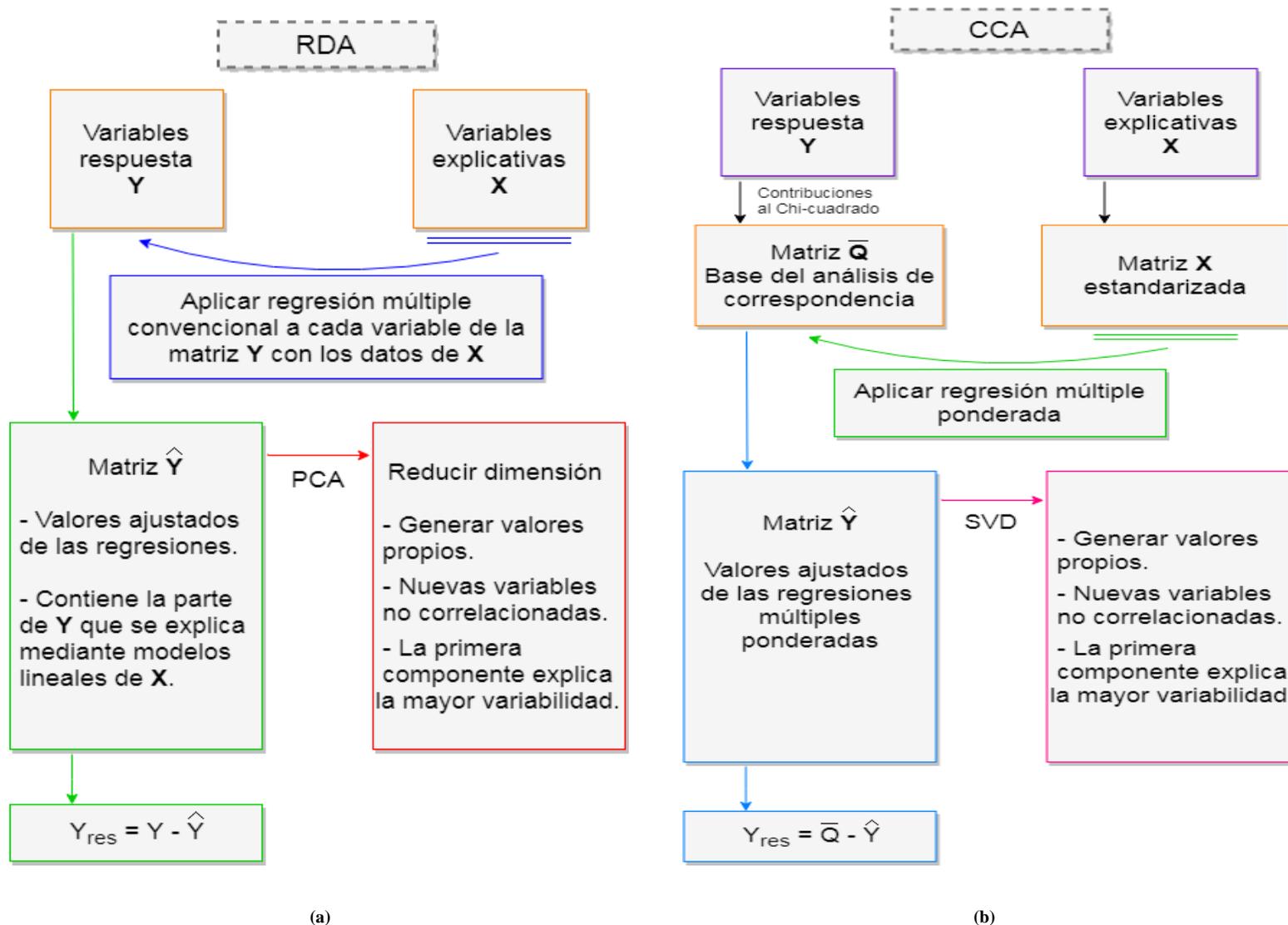


Figura 4. Métodos de ordenación asociados con el análisis de gradiente ambiental: (a) Análisis de redundancia - RDA. Implica dos pasos: una regresión que produce valores ajustados (\hat{Y}) y residuos (Y_{res}), seguido por un análisis de componentes principales (PCA) de la matriz de valores ajustados (\hat{Y}). Adaptado de Legendre P. y Legendre L.F. 2012; (b) Análisis de correspondencia canónica - CCA. Los datos transformados (Chi-cuadrado) se someten a una regresión múltiple ponderada y los valores ajustados (\hat{Y}) se envían a un análisis de correspondencia mediante la descomposición de valores singulares (SVD). Legendre P. y Legendre L.F. 2012

Luego de evaluar el VIF se genera un diagrama de ordenación conocido como triplot, donde se pueden representar las puntuaciones para las especies, las observaciones (sitios o puntos de monitoreo), y las variables ambientales. Algunas consideraciones para el análisis de este gráfico son (Leps & Smilauer, 2003):

- a. Identificar el ángulo entre los vectores de las variables ambientales (flechas), dado que este valor se puede aproximar a las correlaciones entre esas variables.
- b. Analizar en paralelo el triplot con los resultados de las puntuaciones obtenidos mediante el método de ordenación. En esta matriz de resultados se deben verificar los signos, los cuales indican el tipo de relación (directa o inversa) de las variables con respecto a los ejes de ordenación.
- c. Una observación (punto de monitoreo) o una especie se encuentra caracterizada por cierta variable ambiental si estos se localizan en el centroide de la variable.
- d. En el diagrama se pueden identificar datos atípicos al identificar los puntos más distantes sobre la línea de la variable
- e. La longitud del vector representado en el diagrama se relaciona con su variabilidad.

Por otro lado, a partir del modelo de ordenación lineal o unimodal se determinan las posiciones de los sitios o puntos de monitoreo con respecto al primer eje, según el caso. Dicha información se encuentra en los resultados bajo el término de *Site scores* (*Weighted averages of species scores*).

Los puntajes de interés para el cálculo del ICE obtenidos a partir de la aplicación del método de ordenación, se refiere a la posición de las unidades de muestreo individuales a lo largo del gradiente ambiental calculado (Leps & Smilauer, 2003). Los puntajes registrados como *Site scores* indican las coordenadas de los puntos de muestreo en el diagrama de ordenación (Borcard et al., 2011) e igualmente corresponde a las sumas ponderadas de las puntuaciones de las especies. Tener en cuenta que los scores con las puntuaciones más altas a lo largo del primer eje canónico presentan una mayor concentración de organismos en el ecosistema acuático de interés (Legendre P. y Legendre L.F., 2012).

6.2.4 Reescalonamiento de scores

El reescalonamiento de los scores se realiza con base en el estado trófico del cuerpo de agua de modo tal que a los puntos de monitoreo con mayor heterogeneidad en cuanto a su calidad fisicoquímica se les adjudica un valor de 1, dado que son más susceptibles a presentar cuadros de eutrofización asociados a altas concentraciones de nutrientes. Por su parte, los sitios con características oligotróficas cuentan con una

menor heterogeneidad y se asocian con un valor de 10 al momento de aplicar la regresión lineal (Chalar et al., 2011). En la Figura 5 se vinculan las consideraciones para realizar el reescalamiento.

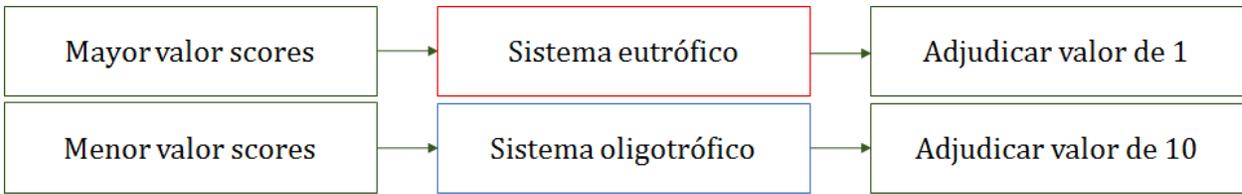


Figura 5. Indicaciones reescalamiento de scores

6.2.5 Valores óptimos y de tolerancia

Para introducir los términos de óptimo y tolerancia como aspectos relevantes dentro del cálculo del ICE, la discusión se enmarca en la Ley de Tolerancia de Shelford (1913) donde se indica que la existencia y prosperidad de un organismo en particular dependen de un conjunto de condiciones, y que la ausencia o el descenso del organismo pueden deberse a la deficiencia o al exceso cualitativo o cuantitativo con respecto a cualquiera de los diversos factores que se acercan a los límites de tolerancia del organismo en cuestión (Odum, 1972). Hesse (1924) estableció una ley más general, donde menciona que cada especie prospera mejor con un valor óptimo particular de una variable ambiental y no puede sobrevivir cuando el valor es demasiado bajo o alto.

Shelford (1913) mencionó que “el éxito de una especie, su abundancia, a veces su tamaño, están determinados en gran medida por el grado de desviación de un solo factor o factores respecto al rango óptimo de la especie”. Para apoyar el comentario de Shelford, Odum (1972) consideró que “los organismos no son simplemente esclavos del medio físico, sino que se adaptan y modifican el medio de modo que puedan reducir los factores que los limitan”.

La relación existente entre la abundancia de las especies y las variables medioambientales se conoce como coenoclina, que se refiere a la organización de las especies a lo largo de un gradiente ambiental y está basada en la respuesta fisiológica de los organismos ante los cambios del medio en el que se encuentran (Ramírez & Viña, 1998).

Gause (1930) declaró que "la Ley de Gauss es la base de las curvas ecológicas", pero también indicó que no se debe dejar a un lado que existen factores como la competencia, que producen cambios en diferentes

secciones de la curva de distribución. Si se analiza la influencia de un factor ambiental sobre cierto organismo tal como se presenta en la Figura 6, se identifica un valor mínimo y un valor máximo con un margen entre uno y otro, los cuales determinan los límites de tolerancia del organismo. Estos valores límites varían para cada especie de acuerdo con las adaptaciones y modificaciones que se presenten. En cuanto a la zona óptima del organismo, ésta se encuentra dentro de los límites de tolerancia identificados, pero en un rango menor.

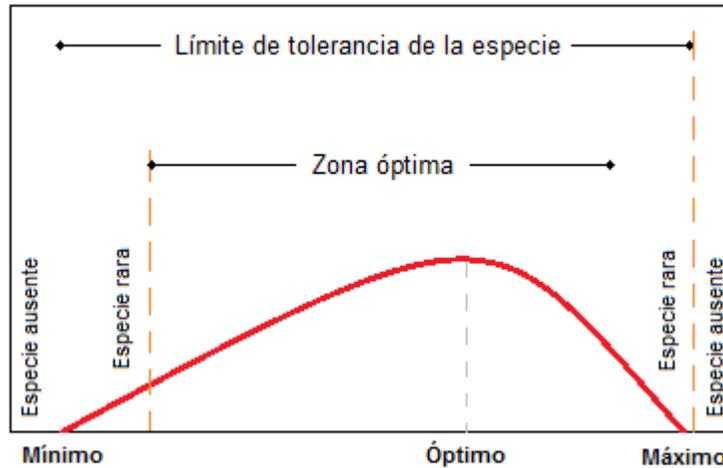


Figura 6. Representación de los límites de tolerancia y rango óptimo de una especie bajo la influencia de un factor ambiental. Adaptado de Odum, E., 1972

En el marco de las coenoclinas, se consideran individuos denominados eurítipicos que poseen un rango de tolerancia amplio, por lo que presentan una mayor dispersión espacio-temporal dentro de un sistema. Por otro lado, las especies estenotípicas presentan una tolerancia más restringida, lo que hace que se encuentren bajo condiciones particulares y que por ende su ubicación en el espacio y tiempo sea reducida. Estas últimas especies son buenas indicadores de las condiciones ambientales en las cuales se desarrollan (Ramírez & Viña, 1998).

Todo organismo es indicador de las condiciones del medio en el cual se desarrolla, puesto que de cualquier forma su existencia en un espacio y momento determinados responde a su capacidad de adaptarse a los distintos factores ambientales. Los límites de tolerancia son variables, y frente a una determinada alteración se encuentran organismos “sensibles” que no soportan las nuevas condiciones, comportándose como “intolerantes”, mientras que otros son “tolerantes” y no se ven afectados (Leps & Smilauer, 2003).

Al llegar la perturbación a un nivel letal para los intolerantes, estos mueren y su lugar es ocupado por organismos tolerantes. De igual manera, cuando la perturbación no sobrepasa el umbral letal, los organismos intolerantes abandonan la zona alterada, por lo cual generan espacio libre que puede ser colonizado por organismos tolerantes. Es así como las variaciones inesperadas en la composición y estructura de las comunidades de organismos vivos en los cuerpos de agua pueden interpretarse como signos evidentes de algún tipo de contaminación (Lynch & Gabriel, 1987).

En estudios paleoambientales, para obtener inferencias adecuadas de una variable ambiental se estiman los óptimos y tolerancias de los taxones incluidos en la matriz biológica (Smol, 2009; Vélez, 2019). Una buena aproximación de estos valores se logra mediante la aplicación de métodos unimodales, tales como el promedio ponderado (Weighted Averaging - WA) y el promedio ponderado por mínimos cuadrados parciales (Weighted Averaging Partial Least Square - WAPLS), los cuales son métodos numéricos que se basan en técnicas inversas para la estimación de la variable ambiental a reconstruir (Ter Braak & Prentice, 2004).

6.2.6 Valor numérico del índice de calidad ecológica

Con base en el desarrollo índice Invertebrate Species Index (ISI) propuesto por Haase y Nolte (2008), se estableció una ecuación para evaluar el cambio en la comunidad de macroinvertebrados acuáticos ante las variaciones de las corrientes superficiales generadas por nutrientes. Esta propuesta se adoptó para obtener el valor numérico del índice de calidad ecológica, donde se relacionan la abundancia de los organismos, así como los valores óptimos y de tolerancia de cada especie, tal como se indica en la Ecuación 2.

De esta manera, el valor ICE se estima mediante una calibración con respecto al estrés de los organismos de la comunidad hidrobiológica de interés, teniendo en cuenta los valores óptimos, de tolerancia y abundancia de especies sobre el producto de los registros de tolerancia por la abundancia de organismos. En otros términos, se evalúa la relación de la respuesta de los organismos a las condiciones ambientales en cuanto a sus óptimos, tolerancia y abundancia; entre la valoración del estrés de los organismos a partir de los rangos de tolerancia.

6.2.7 Análisis de correlaciones y análisis de agrupamiento

El análisis de correlaciones entre las variables fisicoquímicas y los resultados del índice de calidad ecológica permite identificar las variables con importancia ecológica en el cuerpo de agua de interés. Para establecer categorías del ICE, se realiza un análisis de agrupamiento por estación de monitoreo con el propósito de considerar la máxima homogeneidad en cada uno de los grupos.

Se recomiendan obtener tres agrupaciones, las cuales deben discriminarse en una clasificación de alta, media y baja calidad ecológica en el cuerpo de agua de en estudio. Tener en cuenta que dichas agrupaciones se encontrarán en una escala numérica entre 1 y 10, tal como se habría indicado desde el reescalamiento de los scores. Estas categorías son susceptibles a variaciones en sus rangos de acuerdo con el caso de estudio.

6.3 VALIDACIÓN DEL MODELO DE CÁLCULO PROPUESTO EN R

Dado que el índice de calidad ecológica no se encuentra condicionado a un sistema acuático ni a un grupo de organismos específico, la validación del modelo de cálculo propuesto para el ICE en el software estadístico R se realizó con base la información secundaria de diferentes investigaciones que permiten contrastar tres tipos de ecosistemas acuáticos donde se implementó este índice, a saber: río San Juan (Antioquia), Ciénaga de Ayapel (Córdoba) y fuentes de agua afectadas por minería en el corregimiento Jigualito y municipio de Cértegui (Chocó).

A continuación, se sintetizan los ejercicios utilizados para la validación y se remiten al estudio correspondiente:

- Con base en la publicación “*Calidad fisicoquímica e hidrobiológica del agua en el río San Juan, Andes, Antioquia*” (Aguirre *et al.*, 2015) se analizaron los cambios en la estructura del perifiton a partir del índice de calidad ecológica. El área de estudio se ubicó en la región Suroeste del departamento de Antioquia, correspondiente a la cuenca del río San Juan, el cual nace en los Farallones del Citará y desemboca en el río Cauca. El río San Juan tiene influencia en los municipios de Andes, Jardín, Bolívar, Salgar, Hispania y Betania, dado que sus cabeceras municipales se encuentran en su área vertiente.

- Como parte de las zonas inundables de la depresión Momposina, se consideró la evaluación del índice de calidad ecológica con base en el fitoplancton para diagnosticar el estado ambiental de la Ciénaga de Ayapel (Córdoba). Este ecosistema regula los regímenes hidrológicos de los ríos Cauca y San Jorge, así como de varios caños y quebradas; además de contar con una variedad de especies de flora y fauna. La información fue obtenida de la tesis de maestría “*Diagnóstico ambiental de la Ciénaga de Ayapel a través de la variación temporal de la morfología funcional y un indicador de calidad ecológica del fitoplancton*” (Zabala, 2016).
- Se tuvo en consideración el diagnóstico de la calidad ecológica de escenarios hídricos afectados por minería a cielo abierto en el corregimiento de Jigualito, perteneciente al municipio de Condoto, Chocó y en la cabecera municipal de Cértegui (Chocó, Colombia). En estas zonas se seleccionaron cuatro fuentes hídricas con distintos tiempos de cese del disturbio: La primera correspondió a la quebrada Pichirí, una fuente hídrica sin intervención minera; la segunda fue la quebrada Sabaleta que presentó entre 5 y 10 años de haber cesado la actividad minera; la tercera contó con más de 30 años de abandono de la actividad correspondiente a la quebrada Marcos Díaz y la como cuarta fuente hídrica se evaluó la quebrada Jorobibó que presentó recepción actual de vertimientos mineros. Para indagar sobre el tema, se remite al trabajo de Maestría titulado “*Diagnóstico de la calidad ecológica a través de algas perifíticas en escenarios hídricos afectados por minería a cielo abierto en diferente temporalidad, Jigualito (Condoto, Choco)*” (Vargas, 2016).

Para el procesamiento y análisis de la información secundaria, se utilizaron los resultados sistematizados de los parámetros fisicoquímicos e hidráulicos y la comunidad hidrobiológica considerada en las investigaciones citadas anteriormente, los cuales se listan en la Tabla 2. Los resultados obtenidos para el índice de calidad ecológica en cada estudio se confrontaron de acuerdo con la metodología propuesta en la Guía técnica para la formulación de planes de ordenamiento del recurso hídrico continental superficial (MADS, 2018) y el modelo de cálculo propuesto en R.

Tabla 2. Variables analizadas en los sistemas de interés y su respectiva comunidad hidrobiológica

Variables	Unidad	Sistema analizado		
		Río San Juan (Aguirre et al., 2015)	Ciénaga de Ayapel (Zabala, A., 2016)	Escenarios hídricos afectados por minería (Vargas, L., 2016)
Temperatura ambiente	°C	x		
Temperatura del agua	°C	x		x
Oxígeno disuelto	mg/L	x	x	x
Saturación oxígeno disuelto	%	x		

Variables	Unidad	Sistema analizado		
		Río San Juan (Aguirre et al., 2015)	Ciénaga de Ayapel (Zabala, A., 2016)	Escenarios hídricos afectados por minería (Vargas, L., 2016)
Conductividad eléctrica	μS/cm	x	x	x
pH del agua	U de pH	x	x	x
Transparencia	m		x	
Profundidad	m		x	
Velocidad	m/s			x
Caudal	m ³ /s			x
Turbiedad	UNT	x		x
Alcalinidad	mg/L	x		
Nitritos (NO ₂)	mg/L			x
Nitratos (NO ₃)	mg/L	x	x	x
Nitrógeno amoniacal (NH ₄)	mg/L	x	x	
Nitrógeno total	mg/L	x		
Fósforo total	mg/L	x		
Fosfatos (PO ₄)	mg/L	x	x	x
Dureza	mg/L	x	x	
Demanda química de oxígeno - DQO	mg/L	x		
Potencial REDOX	mV		x	
Sólidos disueltos totales	mg/L			x
Sólidos suspendidos	mg/L			x
Comunidad hidrobiológica		Algas perifíticas	Fitoplancton	Algas perifíticas

Para evaluar el ajuste del modelo implementado para el cálculo del índice de calidad ecológica en el software R con respecto a la metodología establecida en la guía de planes de ordenamiento del recurso hídrico, se analizó en cada uno de los casos el error cuadrático medio (RMSE por sus siglas en inglés), el cual permite conocer qué tan concentrados están los datos alrededor de la línea de mejor ajuste (Mayer & Butler, 1993).

7 RESULTADOS Y DISCUSIÓN

Diferentes paquetes y funciones integrados en el software R permiten la ejecución de los métodos requeridos para el cálculo del índice. Por tanto, se estableció un script en el lenguaje de programación R que consolida el modelo de cálculo para el ICE. Así mismo, dicho modelo se enlaza con criterios claves para la interpretación del índice y se valida su ejecución en casos de estudios con características

contrastantes en cuanto a la comunidad hidrobiológica considerada, la ubicación del cuerpo de agua de interés y afectaciones sobre el recurso hídrico.

7.1 MODELO EN EL SOFTWARE R PARA LA APLICACIÓN DEL ICE Y CRITERIOS CLAVES PARA SU INTERPRETACIÓN

Con base en el análisis de los métodos estadísticos y operaciones matemáticas disponibles en el software estadístico R, se generó un modelo de cálculo para el ICE mediante el cual se pretenden minimizar los errores en la estimación del valor del ICE asociados al uso de diferentes programas, y a su vez ofrecer un complemento para los insumos técnicos en información útil que facilite la toma de decisiones por parte de las autoridades competentes, mediante la transformación de los registros provenientes del monitoreo fisicoquímico e hidrobiológico sobre las condiciones ambientales de los ecosistemas acuáticos.

En la Figura 7 se esquematiza el ensamblaje general del modelo y se indican los paquetes, funciones y métodos apropiados para el cálculo del ICE en el software R. Igualmente, se presentan las subrutinas en R donde se separan por bloques cada uno de los pasos requeridos para la ejecución del índice, acompañadas por pautas que apoyan la interpretación de cada paso.

7.1.1 Depuración de la información

Inicialmente, se debe considerar que en los ejercicios de gestión del recurso hídrico como POMCA o PORH, las matrices de datos fisicoquímicos e hidrobiológicos incluyen una cantidad considerable de datos, a los cuales es necesario realizarles una depuración, con el propósito de identificar inconsistencias en las mediciones que puedan generar “ruido” durante el análisis de la información.

La aplicación del índice de calidad ecológica requiere como paso inicial la depuración de los datos, teniendo en cuenta aspectos como:

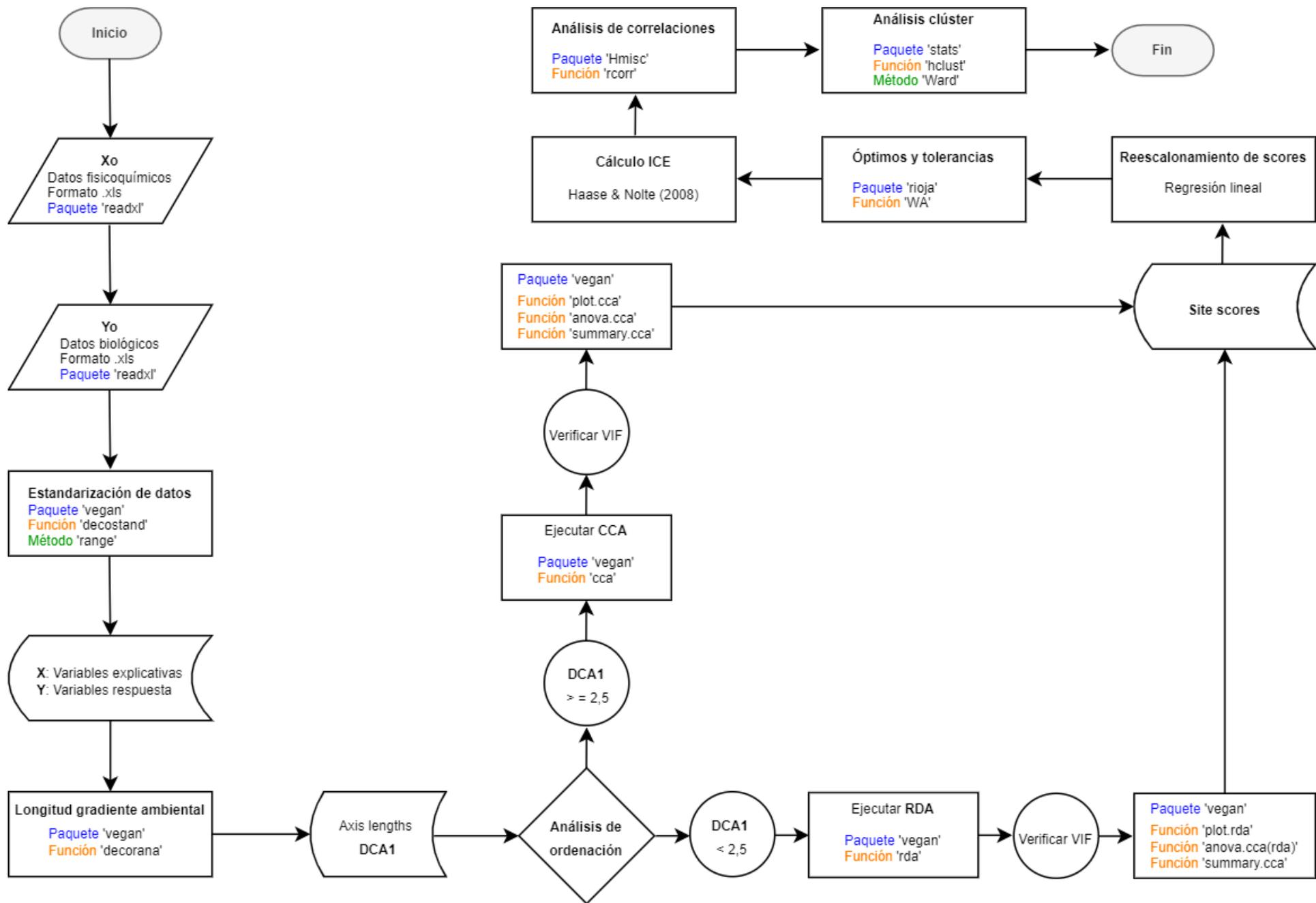


Figura 7. Modelo de cálculo para el ICE en el software R

- Identificar incoherencias en los registros y/o posibles errores en las mediciones en campo y en laboratorio.
- Evitar el uso de signos especiales como paréntesis, guiones, superíndices o subíndices en los nombres de las variables fisicoquímicas y los géneros de las comunidades hidrobiológicas de interés.
- Analizar la posible eliminación en los modelos estadísticos de las variables fisicoquímicas que no presentan variaciones en las diferentes estaciones de monitoreo.
- En la matriz de variables respuesta, las celdas sin registro de abundancia o densidad deben contener un valor de cero.
- Verificar que la matriz de variables explicativas y de variables respuesta contengan la misma cantidad de observaciones (sitios o puntos de monitoreo), con el propósito de contrastar ambas matrices y no presentar errores durante la ejecución de los modelos estadísticos de ordenación.

Además de las anteriores consideraciones, se requiere el reconocimiento de valores faltantes (*missing values*) en las matrices explicativas, puesto que habitualmente los paquetes estadísticos aplicados en el software R no se ejecutan con información incompleta.

Se destaca que desde el momento de la depuración de los datos se pueden identificar variables que no son fundamentales dentro del análisis de calidad ecológica en el cuerpo de agua de interés. También, se sugiere que los archivos tanto de variables explicativas como de variables respuesta presenten un formato con extensión .xls o .xlsx.

7.1.2 Lectura y estandarización de datos

Con relación a la lectura de las matrices con información fisicoquímica y biológica, se ha definido el paquete ‘readxl’, el cual permite la lectura de archivos en formato .xls o .xlsx. Por su parte, al considerar que los resultados de las mediciones fisicoquímicas generalmente presentan escalas numéricas diferentes al igual que en los registros hidrobiológicos, se requiere realizar una estandarización de la información. Las líneas de comando donde se ejecuta tanto la lectura como la estandarización de los datos en el software R se presentan en las Líneas de código 1.

Generalmente, la estandarización se refiere a la división de cada dato por la desviación estándar. También, los datos se pueden estandarizar mediante la resta de la media para que la variable resultante presente una

media de cero. En este caso, los datos han sido estandarizados con base en el método propuesto por Guisande-González *et al.* (2006) descrito en la Ecuación 1.

```
Lectura de datos _ Estandarización
library(readxl)

### Datos fisicoquímicos ###
Xo <- read_excel("Archivo de datos.xlsx",
                sheet = "Hoja de cálculo donde se encuentran los datos")
Xo <- as.data.frame(Xo)
rownames(Xo) <- as.vector(Xo[,1])
Xo <- Xo[,-1]
p1 <- readline(prompt="¿La matriz fisicoquímica se encuentra estandarizada?
                (1 = si, 0 = no) ")

if (p1 == 0){
  X <- decostand(Xo, method = "range")
} else {
  X <- Xo
}

### Datos biológicos ###
Yo <- read_excel("Archivo de datos.xlsx",
                sheet = "Hoja de cálculo donde se encuentran los datos")
Yo <- as.data.frame(Yo)
rownames(Yo) <- as.vector(Yo[,1])
Yo <- Yo[,-1]
p2 <- readline(prompt="¿La matriz biológica se encuentra estandarizada?
                (1 = si, 0 = no) ")

if (p2 == 0){
  Y <- decostand(Yo, method = "range")
} else {
  Y <- Yo
}
```

Líneas de código 1. Lectura y estandarización de datos

Al ejecutar los comandos de lectura de datos y estandarización de los mismos en el script, se obtiene la matriz estandarizada de variables explicativas (X) y la matriz de variables respuesta (Y). Tener en cuenta que se debe indicar en el código si los datos ingresados ya se encuentran estandarizados.

7.1.3 Estimación longitud del gradiente ambiental

Para evaluar la longitud del gradiente ambiental, se considera la heterogeneidad de los organismos de la comunidad de interés en el cuerpo de agua en estudio. De este modo, el método DCA debe implementarse

sobre la matriz biológica (Y), y así establecer las bases para evaluar el modelo de ordenación adecuado a partir de la composición de especies como respuesta a las variaciones en las condiciones ambientales del ecosistema acuático, tal como se describe en la Figura 2.

Así, con base en la información de las variables explicativas se evalúa la longitud del gradiente ambiental tal como se indica en las Líneas de código 2.

```
### Análisis longitud gradiente ambiental ###  
library(vegan)  
datos.dca <- decorana(Y)  
summary(datos.dca)  
temp <- datos.dca[["rproj"]]
```

Líneas de código 2. Estimación longitud del gradiente ambiental

Al ejecutar las líneas del script en R, se identifica la longitud del gradiente ambiental con base en el resultado asociado al primer eje en las Líneas de código 3. Se precisa que un cambio en la composición de organismos se logra almacenar en el resultado del primer eje del DCA; puesto que, según la literatura un cambio medio en la composición de especies ocurre entre 1σ y 1.4σ .

```
Call:  
decorana(veg = Y)  
  
Detrended correspondence analysis with 26 segments.  
Rescaling of axes with 4 iterations.  
  
Eigenvalues      DCA1   DCA2   DCA3   DCA4  
Decorana values  0.6696 0.3627 0.2557 0.2181  
Decorana values  0.8125 0.3521 0.2039 0.1322  
Axis lengths   5.7346 2.6169 2.7785 2.3363
```

Líneas de código 3. Longitud del gradiente ambiental con respecto al primer eje de ordenación

7.1.4 Ejecución del análisis de ordenación

Luego de estimar la longitud del gradiente ambiental se debe ejecutar un análisis de ordenación el cual se encuentra condicionado a obtener la máxima correlación posible entre las variables respuesta (Y) a partir de las variables explicativas (X).

```

### Análisis de ordenación###
library(xlsx)
gradiente <- max(temp[,1])
n <- 1
if (gradiente >= 2.5){
  while (n > 0){
    cca <- cca(Y~., data = X, scale = F)
    vif <- as.matrix(vif.cca(cca))
    vif <- cbind(vif,c(1:nrow(vif)))
    colnames(vif) <- c("vif","Número fila")
    print(vif)
    n <- readline(prompt="Ingrese número de fila a eliminar. Ingrese 0 para terminar ")
    n <- as.integer(n)
    if (n > 0){
      X <- X[,-n]
      Xo <- Xo[,-n]
    }
  }
  plot(cca)
  sum <- summary(cca)
  anova <- anova.cca(cca)
  scores <- as.data.frame(sum$sites)
  # Agregar Scores a Excel
  write.xlsx2(scores,file = "cca.xlsx",
             sheetName="scores",col.names=TRUE, row.names=TRUE, append=TRUE, showNA=TRUE)
}else{
  while (n > 0){
    rda <- rda(Y~., data = X, scale = F)
    vif <- as.matrix(vif.cca(rda))
    vif <- cbind(vif,c(1:nrow(vif)))
    colnames(vif) <- c("vif","Número fila")
    print(vif)
    n <- readline(prompt="Ingrese número de fila a eliminar. Ingrese 0 para terminar ")
    n <- as.integer(n)
    if (n > 0){
      X <- X[,-n]
      Xo <- Xo[,-n]
    }
  }
  plot(rda)
  sum <- summary(rda)
  anova <- anova.cca(rda)
  scores <- as.data.frame(sum$sites)
  # Agregar Scores a Excel
  write.xlsx2(scores,file = "rda.xlsx",
             sheetName="scores",col.names=TRUE, row.names=TRUE, append=TRUE, showNA=TRUE)
}
print(sum)
print(anova)

# Elegir solo el primer eje de los scores
(scores <- scores[,1])

```

Líneas de código 4. Ejecución del análisis de ordenación

Si se considera la Figura 3 (a), un modelo de respuesta lineal representa una composición datos homogéneos y debe ser aplicado como método de ordenación si se establecen valores menores a 2.5σ para la longitud del gradiente ambiental. Por su parte, al identificar un ajuste de los datos biológicos más heterogéneos, es

conveniente implementar un modelo unimodal, representado en la Figura 3 (b). En este caso, se deben obtener valores iguales o superiores a 2.5σ para el primer eje en el análisis DCA.

Tal como se observa en las Líneas de código 4, se ejecuta un análisis de correspondencia canónica (CCA) para valores del gradiente iguales o mayores a 2.5σ . Por su parte, si el gradiente ambiental indica una longitud menor a 2.5σ , el modelo de ordenación pertinente es un análisis de redundancia (RDA).

Durante la ejecución del modelo de ordenación, se debe evaluar la multicolinealidad entre las variables ambientales a partir del VIF. Frente a ello, se sugiere considerar valores para la varianza menores a 20, dado que registros por encima de este valor indican una colinealidad fuerte (Borcard *et al.*, 2011). Sin embargo, para el caso en el que el investigador considere que la presencia de la variable en el modelo de ordenación es imprescindible para la comprensión de la respuesta, es criterio de este determinar si la variable se mantiene a pesar de presentar un VIF elevado.

Otro aspecto relevante se relaciona con que los métodos multivariados de ordenación que se aplican para estimar el gradiente ambiental dividen la explicación de la varianza total en cuatro ejes, y el mayor porcentaje de varianza explicada se concentra en el primer eje. Es así como una particularidad del script del análisis de ordenación es la exportación de los resultados del primer eje de los *scores*, donde se determinan las posiciones de los sitios de monitoreo con respecto a este eje de ordenación.

```
Permutation test for cca under reduced model
Permutation: free
Number of permutations: 999

Model: cca(formula = Y ~ Ptot + pH + OD + Cond + PO4 + `Amonio NH4` + NO3 + Durtot + DS + Redox, data = X, scale = F)
      Df ChiSquare      F Pr(>F)
Model  10   2.1651 2.3014 0.001 ***
Residual 23   2.1638
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Líneas de código 5. Análisis de significancia del modelo de ordenación

Finalmente, se evalúa la significancia del modelo de ordenación, tal como se indica en las Líneas de código 5.

7.1.5 Reescalonamiento del ICE

Generalmente, los índices se presentan en una escala numérica para facilitar su comprensión y análisis. En el caso del ICE, los valores son reescalados entre 1 y 10 mediante una regresión lineal, dada la naturaleza de la relación entre las variables.

En este paso, se reescalonan los valores de los *scores* del primer eje obtenido a partir del análisis de ordenación como se indica en las Líneas de código 6. Cabe resaltar que en este punto del cálculo los valores de la varianza utilizados para el reescalonamiento consideran la información del modelo de respuesta de la relación entre las especies y las variables ambientales. Tal como se expresó en la Figura 5, el registro más bajo de los scores se escaló a un valor de 10, mientras que al score más alto se le adjudicó un valor de 1.

```

### Reescalonamiento de scores ###
scores.min <- scores[which(scores==max(scores))]
scores.max <- scores[which(scores==min(scores))]
m <- (9)/(scores.max-scores.min)
b <- 10 - m*scores.max
long.scores <- length(scores)
scores <- cbind(scores,rep(0,long.scores))
colnames(scores)[2] <- 'Puntuación'
for (i in 1:long.scores){
  scores[i,2] <- m*scores[i,1]+b
}
long.Yo <- length(Yo)
Yp <- scores[,2]
Yp1 <- Yp
for (i in 2:long.Yo){
  Yp <- cbind(Yp,Yp1)
}
colnames(Yp) <- colnames(Yo)
cca <- cca(Y~., data = X, scale = F)

```

Líneas de código 6. Reescalonamiento de scores

Al interpretar los resultados del reescalonamiento, se deben considerar con atención las características fisicoquímicas del cuerpo de agua de interés y las afectaciones que se presentan sobre el recurso hídrico, dado que es probable que los resultados del ICE sean mayores en los sitios de monitoreo que presentan alteraciones considerables en su calidad fisicoquímica.

7.1.6 Estimación de valores óptimos y de tolerancia

Con base en la Figura 6, cada organismo logra definir unos límites de tolerancia y un valor óptimo para su desarrollo de acuerdo con las condiciones ambientales. Es así como una especie tiende a presentar una mayor abundancia en torno a los valores ambientales cercanos a sus factores óptimos de crecimiento.

Dentro del contexto del índice de calidad ecológica, la tolerancia se relaciona con la adaptación del organismo al gradiente ambiental mediante cambios en su forma o el tamaño; mientras que el valor óptimo se vincula con la respuesta del organismo al gradiente ambiental. De este modo, la tolerancia se refiere al indicador numérico, con respecto a la abundancia de los organismos, de la habilidad de la especie para sobrevivir en un rango de condiciones ambientales. Por su parte, el valor óptimo puede considerarse como el indicador numérico, con respecto a la abundancia de los organismos, del buen desarrollo de una especie frente a ciertos factores ambientales. El cálculo de estos dos conceptos se presenta en las Líneas de código 7.

```
#### valores óptimos y de tolerancia ####  
library(rioja)  
# Promedio ponderado WA  
fit <- WA(Y, Yp1, tolDW=TRUE)  
# validación Cruzada  
rioja::crossval(fit)  
print(coefficients(fit))  
optytol <- t(coefficients(fit))
```

Líneas de código 7. Cálculo de valores óptimos y de tolerancia

En este punto del análisis, es posible identificar organismos indicadores para las condiciones ambientales al considerar especies sensibles, con límites de tolerancia bajos y valores óptimos altos. Se destaca también que cambios inesperados en la composición de las comunidades hidrobiológicas en los cuerpos de agua pueden evidenciar signos de algún tipo de contaminación

7.1.7 Valor de calidad ecológica

El valor ICE se estima mediante una calibración con respecto al estrés de los organismos a partir de la relación entre la sumatoria del producto de los valores óptimos, de tolerancia y abundancia de organismos sobre la sumatoria del producto de los valores de tolerancia por la abundancia de organismos.

En resumen, en el numerador se considera la respuesta de los organismos ante ciertas condiciones ambientales con base en sus óptimos, tolerancias y abundancia; mientras que en el denominador se valora el estrés de los organismos en función de su tolerancia.

El valor del ICE se obtiene para cada una de las estaciones de monitoreo incluidas en los análisis mediante la ejecución de las Líneas de código 8.

```
### cálculo índice de calidad ecológica ICE ###
# Transformación
Ytransf <- log10(Yo+1)
Ytransf2 <- Ytransf
Ytransf3 <- Ytransf
for (i in c(1:length(Ytransf2))){
  Ytransf2[,i] <- Ytransf2[,i]*optytol[1,i]*optytol[2,i]
  Ytransf3[,i] <- Ytransf3[,i]*optytol[2,i]
}
sumas <- matrix(0, nrow = nrow(Ytransf), ncol=3)
rownames(sumas) <- rownames(Ytransf)
colnames(sumas) <- c("Numerador", "Denominador", "ICE")
for (i in c(1:nrow(sumas))){
  sumas[i,1] <- sum(Ytransf2[i,])
  sumas[i,2] <- sum(Ytransf3[i,])
  sumas[i,3] <- sumas[i,1]/sumas[i,2]
}
print(sumas)
Xu <- cbind(Xo,sumas[,3],scores[,1])
colnames(Xu)[ncol(Xu)-1] <- "ICE"
colnames(Xu)[ncol(Xu)] <- "scores"
```

Líneas de código 8. Estimación índice de calidad ecológica

7.1.8 Análisis de correlaciones

Luego de establecer los valores del ICE en el cuerpo de agua, se deben identificar variables fisicoquímicas que presenten influencia sobre la calidad ecológica y puedan ser consideradas dentro de los programas de monitoreo y seguimiento del cuerpo de agua.

Para ello, se aplica un análisis de correlaciones para establecer relaciones inversas y directas con índice mediante la ejecución de las Líneas de código 9. Tener en cuenta que se consideran las variables ambientales que fueron incluidas en el modelo de ordenación sin estandarizar y los valores por estación para el ICE.

```

### Análisis de correlaciones ###
library(Hmisc)
co <- rcorr(as.matrix(xu), type = "spearman")
corr <- as.matrix(co$r)
corp <- as.matrix(co$p)
correlacion <- cbind(corr[,ncol(corr)-1],corp[,ncol(corr)-1])
colnames(correlacion) <- c("Corr ICE","p-value")
correlacion

```

Líneas de código 9. Análisis de correlaciones

7.1.9 Análisis de agrupamiento

Para categorizar la calidad ecológica en el cuerpo de agua que se encuentra en estudio, se ejecuta un análisis de agrupamiento mediante el método de Ward con base en la distancia euclidiana, tal como se encuentra descrito en las Líneas de código 10.

El análisis de conglomerados se encuentra restringido a establecer 3 grupos, con el propósito de evaluar ICE en una escala de baja, intermedia y alta calidad ecológica.

```

### Análisis clúster ###
distancia <- dist(scale(xu[, -ncol(xu)]))
fit <- hclust(distancia, method = "ward.D2")
plot(fit)
grupos <- cutree(fit,k=3)
rect.hclust(fit,k=3,border = "red")

```

Líneas de código 10. Análisis de agrupamiento

7.2 VALIDACIÓN DEL MODELO DE CÁLCULO IMPLEMENTADO EN EL SOFTWARE R

Los ecosistemas acuáticos cuentan con propiedades que los caracterizan, ya sean de tipo físico, químico, biológico, entre otros. Debido a los cambios naturales o antrópicos que se presentan sobre dichas propiedades, las especies de una comunidad biológica pueden presentar diferentes respuestas, como un

incremento en su población dado que las condiciones del ecosistema favorecen esta condición, o por el contrario, pueden propiciar la ausencia de una especie.

Dadas las particularidades que pueden presentarse entre las cuencas hidrográficas, la estimación del índice de calidad ecológica permite establecer valores óptimos y de tolerancia para un organismo de acuerdo con el gradiente ambiental que se determina en función de las variables fisicoquímicas propias del sistema que se analiza. Además, se recalca que el ICE se caracteriza porque no se encuentra condicionado a un ecosistema acuático.

Con base lo anterior, se tuvo en cuenta la información secundaria de tres investigaciones en diferentes sistemas acuáticos donde se ejecutó el cálculo del ICE a partir de la metodología descrita en la Guía técnica para la formulación de PORH (MADS, 2018), con el propósito de contrastar los resultados del ICE con los valores obtenidos a partir del modelo de cálculo implementado en el software estadístico R. En cada uno de los casos de estudio se analizó el error cuadrático medio para conocer el ajuste del modelo en R.

Para la ejecución de la metodología propuesta en la Guía de PORH (MADS, 2018), se utilizó la versión de prueba del software CANOCO 5 (Disponible en <http://www.canoco.com>), así como el Software for ecological and palaeoecological data analysis and visualisation C2 Versión 1.3 (Disponible en <https://www.staff.ncl.ac.uk/stephen.juggins/software/C2Home.htm>). También, se consideró la licencia académica del software estadístico STATGRAPHICS Centurion XVI. Por su parte, para el modelo en el software de uso libre se implementó la versión 3.6.1 del software R.

A continuación, se presenta la validación del modelo de cálculo en el software R para comparar los resultados con los valores del ICE obtenidos mediante el uso de herramientas de uso restringido, dado que requieren de licencia.

7.2.1 Validación Río San Juan (Aguirre *et al.*, 2015)

De acuerdo con la información secundaria obtenida de la publicación “*Calidad fisicoquímica e hidrobiológica del agua en el río San Juan, Andes, Antioquia*” (Aguirre *et al.*, 2015) se consideraron 6 estaciones de monitoreo durante 4 muestreos, para un total de 24 puntos de monitoreo incluidos en el análisis del índice de calidad ecológica. Además, se tuvieron en cuenta 15 variables fisicoquímicas y 35 especies de algas perifíticas que conformaron la matriz biológica.

Según Aguirre *et al.*, (2015), la aplicación del ICE en el río San Juan indicó que las relaciones entre las variables fisicoquímicas y la abundancia de las algas perifíticas correspondieron a un modelo unimodal, por lo cual se ejecutó un CCA como primer filtro del análisis. Lo anterior indica que al evaluar la heterogeneidad de la comunidad de algas perifíticas, la mayoría de especies no se ajustaron al supuesto de linealidad (Legendre & Birks, 2012).

También, se identificaron organismos que podrían ser utilizados como bioindicadores del gradiente ambiental en el sistema estudiado, tales como las algas *Chaetophora* sp., *Cymbella tumida*, *Gomphonema subclavatum*, *Tabellaria* sp. y *Zygnema* sp. La consideración de organismos bioindicadores se basa en la respuesta fisiológica identificada por parte de los organismos en cuanto a su alto valor óptimo ante las variaciones fisicoquímicas que se registra en el sistema acuático en estudio (Ramírez & Viña, 1998).

Por su parte, al considerar la clasificación del ICE, se evidenció que los valores altos de calidad ecológica se encontraron entre los 6.81 y 8.16, abarcando 17 de los 24 puntos de monitoreo analizados sobre el río San Juan. Además, la estación E2M3 registró un valor de 4.23 como calidad ecológica baja en el sistema analizado, sitio donde se identificaron organismos como *Oscillatoria* sp. que puede denominarse como eurítipico, puesto que poseen una alta tolerancia a la influencia de las condiciones ambientales.

Tal como se presenta en la Tabla 3, el método propuesto en la guía de planes de ordenamiento y el modelo en el software R no presentaron diferencias, tanto en el error del modelo de ordenación, como en la varianza explicada por los dos métodos.

Tabla 3. Comparativo modelo de cálculo para el ICE – río San Juan

Modelo de cálculo	Resultado DCA	Test de significancia de todos los ejes canónicos	Varianza explicada (%)
MADS, 2018	3.74	2.94	21.7
Software R	3.74	2.94	21.7

En cuanto a los resultados del ICE, en la Tabla 4 se observa el valor del índice para cada uno de los 24 puntos de monitoreo. De acuerdo con el análisis del error cuadrático medio, 23 resultados del ICE presentaron un error del 2%, a excepción de la estación E6M4 donde se obtuvo un RMSE del 3%. Sin embargo, en cada uno de los casos se evidenció un buen ajuste para el modelo de cálculo implementado en el software R, tal como se puede observar en el Gráfico 1.

Tabla 4. Resultados índices de calidad ecológica y RMSE – río San Juan

Sitio	Índice de calidad ecológica - ICE		RMSE
	MADS, 2018	Software R	
E1M1	8.16	8.17	0.02
E2M1	7.69	7.70	0.02
E3M1	6.81	6.82	0.02
E4M1	7.74	7.76	0.02
E5M1	6.03	6.04	0.02
E6M1	6.80	6.82	0.02
E1M2	8.00	8.02	0.02
E2M2	7.71	7.73	0.02
E3M2	7.88	7.91	0.02
E4M2	7.97	8.01	0.02
E5M2	6.21	6.22	0.02
E6M2	7.85	7.89	0.02
E1M3	8.14	8.16	0.02
E2M3	4.23	4.23	0.02
E3M3	7.76	7.78	0.02
E4M3	7.75	7.77	0.02
E5M3	7.66	7.68	0.02
E6M3	6.59	6.61	0.02
E1M4	7.90	7.91	0.02
E2M4	6.37	6.38	0.02
E3M4	7.96	7.99	0.02
E4M4	6.90	6.92	0.02
E5M4	6.54	6.56	0.02
E6M4	7.85	7.89	0.03

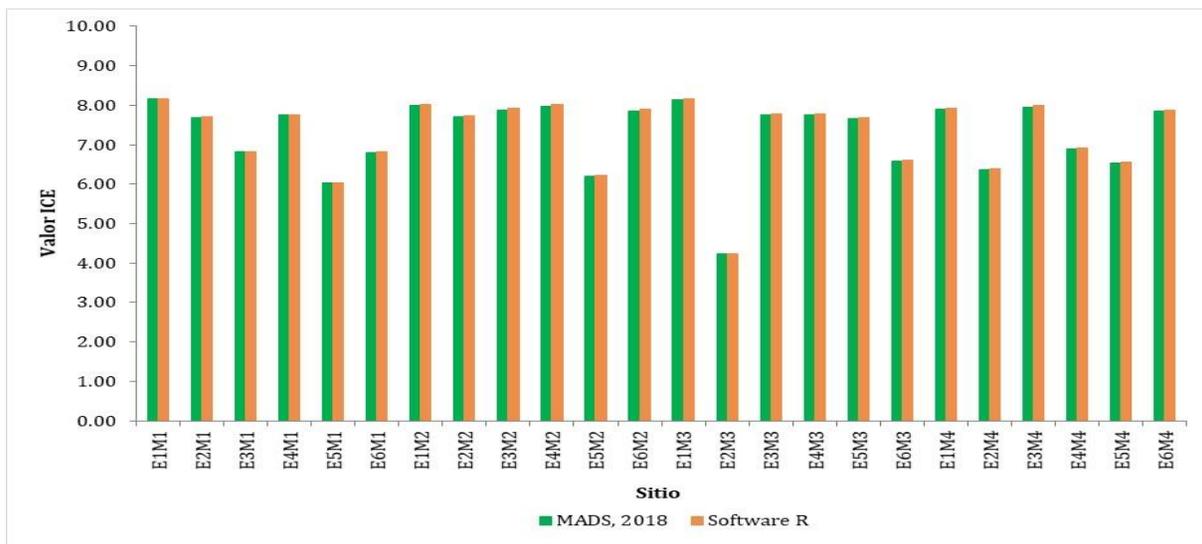


Gráfico 1. Resultados ICE MADS, 2018 vs. Software R – Río San Juan

7.2.2 Validación Ciénaga de Ayapel (Zabala, 2016)

A partir del trabajo de investigación “*Diagnóstico ambiental de la Ciénaga de Ayapel a través de la variación temporal de la morfología funcional y un indicador de calidad ecológica del fitoplancton*” (Zabala, 2016), se obtuvo una matriz con 34 puntos de monitoreo y 10 variables fisicoquímicas. Por su parte, la matriz biológica estuvo integrada por 79 especies pertenecientes a la comunidad del fitoplancton. Particularmente, se realizó una depuración de los datos de la matriz biológica donde se identificaron especies que presentaban una abundancia relativa menor al 2%, por lo cual no fueron incluidas en el cálculo del ICE.

En este diagnóstico, se tuvieron en cuenta estudios previos al año 2015 realizados en la ciénaga, los cuales fueron clasificados en cuatro épocas hidrológicas. Además, se ejecutó un monitoreo en el año 2015 como escenario ambiental actual que permitiera analizar las variaciones en el sistema acuático frente a los estudios anteriores.

Según el estudio realizado por Zabala (2016), la longitud del primer eje evaluado mediante el método DCA presentó un valor superior a 2.5σ y se implementó un modelo unimodal a partir de un CCA para evaluar el gradiente ambiental en la Ciénaga de Ayapel. Se destaca que la longitud del primer eje del DCA permite considerar los cambios en las especies de fitoplancton con base en las variaciones de calidad fisicoquímica que se registran en el cuerpo de agua de interés.

Al considerar el comparativo de la Tabla 5 entre los modelos de cálculo para el ICE a partir de la información recolectada en la Ciénaga de Ayapel, no se presentaron diferencias en cuanto a la varianza explicada por el modelo multivariado, así como en el error del método de ordenación ejecutado, correspondiente a un análisis de correspondencia canónica.

Tabla 5. Comparativo modelo de cálculo para el ICE – Ciénaga de Ayapel

Modelo de cálculo	Resultado DCA	Test de significancia de todos los ejes canónicos	Varianza explicada (%)
MADS, 2018	5.735	2.165	36.4
Software R	5.735	2.165	36.4

Por otro lado, se identificó que la mayoría de los taxones de la comunidad hidrobiológica fitoplancton obtuvieron un valor óptimo por encima de 5; mientras que algunos morfotipos como *Aphanocapsa aff. elachista*, *Euglena sp.*, *Plancktohyngbya sp.*, *Merismopedia sp.* y *Staurastrum cf.* presentaron valores por

debajo de 2, los cuales no representa una buena adaptación al sistema acuático analizado, dado que su abundancia y frecuencia en la ciénaga no fue representativa.

Referente a la clasificación de calidad ecológica para el desarrollo del fitoplancton, se registraron valores menores del ICE para las estaciones incluidas en el monitoreo del año 2015, considerado como escenario actual. En contraste, los resultados para los estudios previos fueron cercanos a un valor de 10, lo que permite indicar que la calidad ecológica presentaba una mejor clasificación bajo las condiciones ambientales registradas en la Ciénaga de Ayapel antes del 2015.

Con relación al ICE, en la Tabla 6 se presentan los resultados para las 34 estaciones de monitoreo diferenciadas de acuerdo con la nomenclatura asignada durante el ejercicio de investigación.

Tabla 6. Resultados índices de calidad ecológica y RMSE – Ciénaga de Ayapel

Sitio	Índice de calidad ecológica - ICE		RMSE
	MADS, 2018	Software R	
E1m2015	5.52	5.52	0.00
E2m2015	4.86	4.86	0.00
E3m2015	5.09	5.09	0.00
E5m2015	2.59	2.59	0.00
E6m2015	4.72	4.72	0.00
E7m2015	5.30	5.30	0.00
E8m2015	6.72	6.72	0.00
E1mLL	9.83	9.83	0.00
E2mLL	9.86	9.86	0.00
E3mLL	9.79	9.79	0.00
E5mLL	9.84	9.84	0.00
E6mLL	9.84	9.84	0.00
E7mLL	9.86	9.86	0.00
E8mLL	9.82	9.82	0.00
E1mS	9.71	9.71	0.00
E2mS	9.73	9.73	0.00
E3mS	9.51	9.51	0.00
E5mS	9.30	9.30	0.00
E6mS	9.83	9.83	0.00
E7mS	9.74	9.74	0.00
E8mS	9.80	9.80	0.00
E1mTT	9.72	9.72	0.00
E2mTT	9.81	9.81	0.00
E3mTT	9.75	9.75	0.00
E5mTT	9.81	9.81	0.00
E6mTT	9.76	9.76	0.00
E7mTT	9.66	9.66	0.00
E1mT	9.70	9.70	0.00
E2mT	9.64	9.64	0.00

Sitio	Índice de calidad ecológica - ICE		RMSE
	MADS, 2018	Software R	
E3mT	9.72	9.72	0.00
E5mT	9.27	9.27	0.00
E6mT	9.80	9.80	0.00
E7mT	9.60	9.60	0.00
E8mT	9.80	9.80	0.00

Según el error cuadrático medio, el modelo implementado en el software R se ajustó a la metodología propuesta en la guía de planes de ordenamiento del recurso hídrico, resultado que se puede observar en el Gráfico 2.

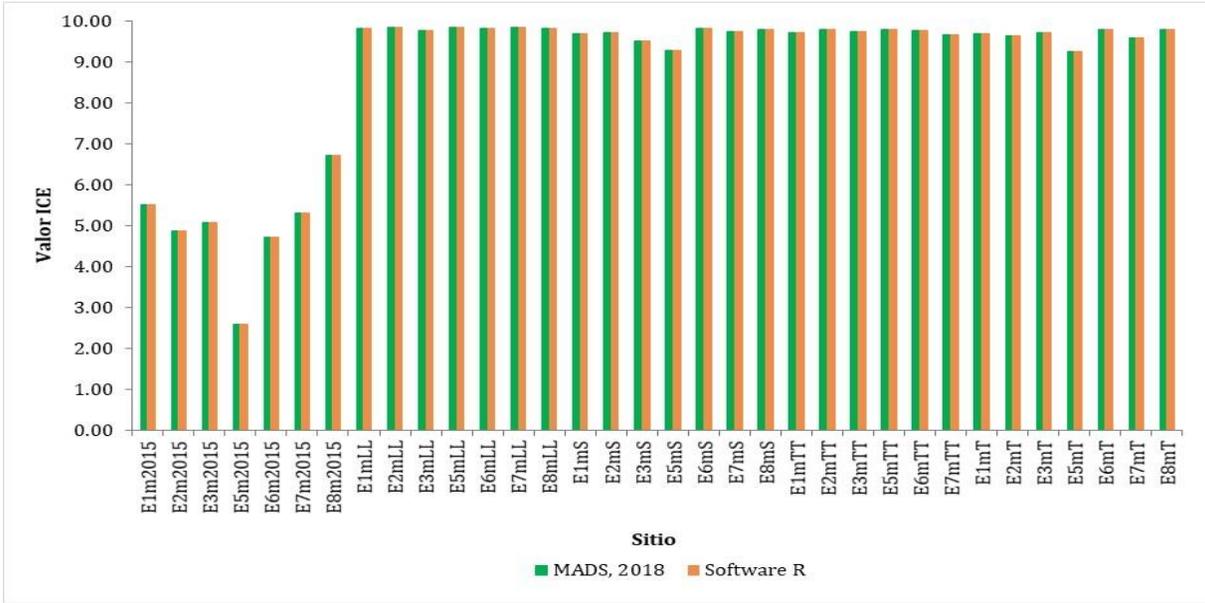


Gráfico 2. Resultados ICE MADS, 2018 vs. Software R – Ciénaga de Ayapel

7.2.3 Validación escenarios hídricos afectados por minería en el Chocó (Vargas, 2016)

En las diferentes fuentes hídricas consideradas en el trabajo de investigación “Diagnóstico de la calidad ecológica a través de algas perifíticas en escenarios hídricos afectados por minería a cielo abierto en diferente temporalidad, Jigualito (Condoto, Choco)” (Vargas, 2016), se consolidó una matriz fisicoquímica compuesta por 12 variables y 32 puntos de monitoreo. Además, se incluyeron 52 especies de algas perifíticas en la matriz biológica.

Vargas (2016) indicó que las relaciones entre las variables fisicoquímicas y la abundancia de las algas perifíticas correspondieron a un modelo unimodal. También, se estableció que la mayoría de los taxones considerados obtuvieron un valor óptimo por encima de 6, y que aquellos con un alto valor de tolerancia y un óptimo por debajo de 4 no fueron buenos indicadores de la calidad ecológica.

Al ejecutar el ICE, el método propuesto en la guía de planes de ordenamiento y el modelo en el software R no presentaron diferencias en cuanto a el error cuadrático del modelo de ordenación y varianza explicada por los dos métodos, información que se reporta en la Tabla 7.

Tabla 7. Comparativo modelo de cálculo para el ICE – Chocó

Modelo de cálculo	Resultado DCA	Test de significancia de todos los ejes canónicos	Varianza explicada (%)
MADS, 2018	2.594	3.317	16.6
Software R	2.594	3.317	16.6

Mediante la aplicación del ICE, se evidenció que las estaciones que no presentaban explotación minera registraron los mayores valores de calidad ecológica, por tanto, fueron los puntos referenciados para aplicar medidas de conservación en las cuencas. Estos resultados se contrastan con las fuentes de agua intervenidas con explotaciones mineras, donde las perturbaciones en términos ecológicos fueron más notorias. En la Tabla 8 se presentan los resultados del ICE bajo las dos metodologías consideradas.

Tabla 8. Resultados índices de calidad ecológica y RMSE – Chocó

Sitio	Índice de calidad ecológica - ICE		RMSE
	MADS, 2018	Software R	
M1E01	7.69	7.69	0.00
M1E02	6.92	6.92	0.00
M1E11	6.84	6.84	0.00
M1E12	7.03	7.03	0.00
M1E21	7.71	7.71	0.00
M1E22	7.60	7.60	0.00
M1E31	7.46	7.46	0.00
M1E32	7.46	7.46	0.00
M2E01	7.91	7.91	0.00
M2E02	8.24	8.24	0.00
M2E11	7.57	7.57	0.00
M2E12	7.01	7.01	0.00
M2E21	4.87	4.87	0.00
M2E22	6.80	6.80	0.00
M2E31	7.05	7.05	0.00

Sitio	Índice de calidad ecológica - ICE		RMSE
	MADS, 2018	Software R	
M2E32	8.59	8.59	0.00
M3E01	7.77	7.77	0.00
M3E02	8.24	8.24	0.00
M3E11	6.43	6.43	0.00
M3E12	6.55	6.55	0.00
M3E21	5.58	5.58	0.00
M3E22	5.64	5.64	0.00
M3E31	6.16	6.16	0.00
M3E32	6.13	6.13	0.00
M4E01	7.49	7.49	0.00
M4E02	7.88	7.88	0.00
M4E11	6.96	6.96	0.00
M4E12	6.90	6.90	0.00
M4E21	7.50	7.50	0.00
M4E22	6.84	6.84	0.00
M4E31	6.80	6.80	0.00
M4E32	6.23	6.23	0.00

De acuerdo con los resultados del RMSE, el error fue de cero para el caso de estudio de fuentes superficiales afectadas por minería en el Chocó, por tanto el ajuste modelo de cálculo para el índice de calidad ecológica en el software R no presentó variaciones con respecto a la metodología establecida en la guía de planes de ordenamiento del recurso hídrico, lo cual se puede observar en el Gráfico 3.

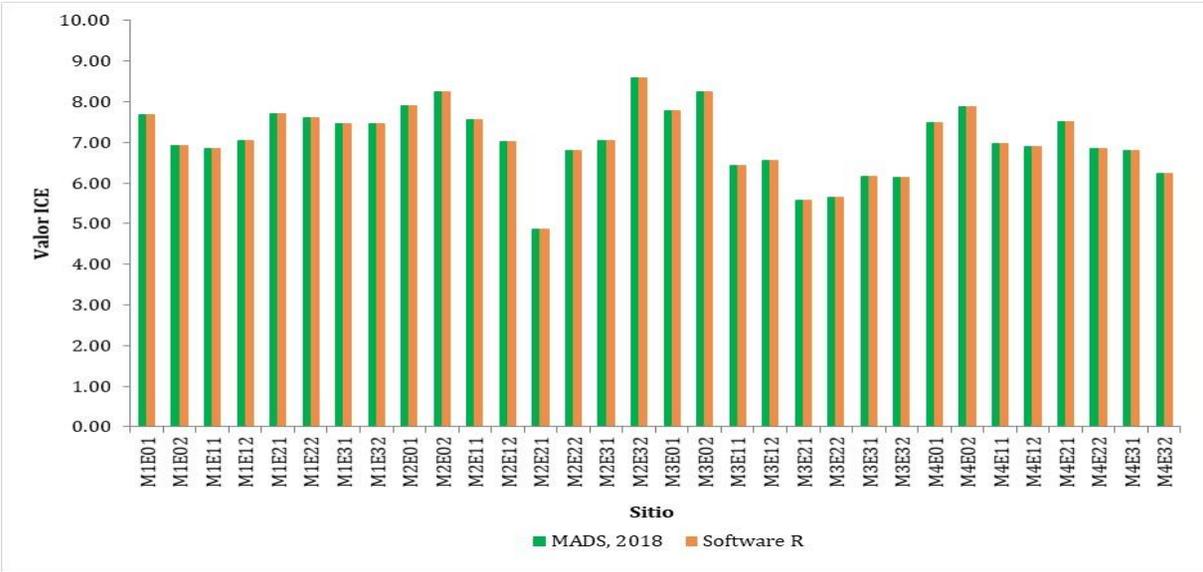


Gráfico 3. Resultados ICE MADS, 2018 vs. Software R – fuentes hídricas en Chocó

8 CONCLUSIONES Y RECOMENDACIONES

Como se estableció inicialmente en la metodología adoptada por el MADS en la guía para la formulación de PORH (MADS, 2018), diferentes softwares estadísticos permiten ejecutar los análisis requeridos para el cálculo del ICE. Frente a ello, se analizó el cómo minimizar los errores en la estimación del valor del índice ICE asociados al uso de diferentes programas estadísticos, definiendo el cálculo del índice en el software estadístico R. Este tipo de herramientas de uso libre fortalecen la aplicación de indicadores ecológicos como el ICE, los cuales se han basado en la implementación de modelos matemáticos complejos que estiman la respuesta de algunas especies acuáticas a las variaciones ambientales. De esta forma, se logró identificar que el software estadístico R permite la ejecución de los diferentes métodos estadísticos multivariados y cálculos matemáticos necesarios para el ICE según lo establecido en la guía para la formulación de PORH. Además, al R ser un programa de uso libre se promueve el uso de software legal para herramientas estadísticas multivariadas aplicadas al campo de la ecología.

Con base en los resultados de la validación del modelo de cálculo implementado en el software R, se evidenció una estrecha relación entre la metodología propuesta por el MADS para el cálculo del ICE y el modelo propuesto. Lo anterior se sustenta en el ajuste obtenido para el error cuadrático medio (RMSE), puesto que se observaron valores entre 0.00 y 0.03. Así mismo, se optimizó el tiempo de ejecución del cálculo del ICE con la implementación del modelo de cálculo.

El ejercicio realizado en el trabajo de investigación con base en enfoques ecológicos aplicados a los ecosistemas acuáticos que permiten el uso de diferentes análisis estadísticos multivariados, soporta el hecho de que estos análisis son una herramienta útil para predecir los cambios en las comunidades hidrobiológicas en un sitio en particular. Es así como la estructura de cálculo del ICE permite analizar las variaciones en la abundancia de diferentes grupos hidrobiológicos a partir de los cambios en las condiciones ambientales de un cuerpo de agua.

Por su parte, los índices de calidad fisicoquímica e hidrobiológica comúnmente utilizados se basan en métodos que generalmente se adaptan a las particularidades de una zona; mientras que el ICE proporciona información sobre las características locales del ecosistema acuático de interés y permite conjugar una mayor cantidad de variables fisicoquímicas e hidrobiológicas. Es por ello que la interpretación del ICE y la optimización del cálculo del mismo en el software R se orientó hacia el fortalecimiento de la gestión integral del recurso hídrico, dado que se presenta un complemento para los insumos técnicos a partir de la transformación de los registros provenientes del monitoreo fisicoquímico e hidrobiológico en un solo valor.

En cuanto a la interpretación del ICE, se obtienen criterios que facilitan el análisis de las dinámicas de las cuencas hidrográficas teniendo en cuenta las variaciones en el gradiente ambiental. Sin embargo, dadas las particularidades que pueden presentarse entre las cuencas hidrográficas, la estimación del ICE requiere de un análisis para cada caso, teniendo en cuenta que el gradiente ambiental se determina en función de las variables fisicoquímicas, y a su vez la evaluación de estas variables se encuentra ligadas a las afectaciones que se presente el cuerpo de agua de interés, tanto de tipo antrópico como natural. Luego, se deben establecer los valores óptimos y de tolerancia para los organismos del grupo biológico monitoreado en función del gradiente ambiental calculado.

Por otro lado, la gestión integral del recurso hídrico es un tema que requiere de herramientas eficientes y que puedan ser aplicadas a diferentes zonas del país y considerar las diferentes afectaciones que presentan las cuencas hidrográficas. Cabe considerar que este tipo de herramientas pueden ser mejoradas con base en el conocimiento de la composición de los organismos en un cuerpo de agua en particular, por lo cual se recomienda integrar la evaluación de la calidad ecológica de los ecosistemas acuáticos con un análisis de la distribución del ensamblaje de los organismos que puedan ser considerados como indicadores de la calidad del agua.

Además, se recomienda continuar con el desarrollo y mejoramiento de índices que puedan ser adaptados a las características propias de las cuencas hidrográficas, dado que se podrían complementar los insumos técnicos de acuerdo con las dinámicas en los ecosistemas acuáticos relacionados con el incremento en el crecimiento poblacional, las afectaciones por el desarrollo económico y los impactos asociados al cambio climático sobre el recurso hídrico.

REFERENCIAS BIBLIOGRÁFICAS

- Agencia de Protección de Ohio. (1987). Users Manual for Biological Field Assessment of Ohio Surface Waters. Columbus, Ohio: OEPA, Division of Water Quality Monitoring and Assessment, Su. *Ohio Environmental Protection Agency (OEPA). Biological Criteria for the Protection of Aquatic Life, Vol. 2.*
- Aguirre, N. J., Jesús, V. F. de, Grajales, H., Hernández, E., Caicedo, O., Mesa, J., & Cortes, F. (2015). *Calidad fisicoquímica e hidrobiológica del agua en el río San Juan, Andes, Antioquia.*
- Alba-Tercedor, J. (1996). Macroinvertebrados acuáticos y calidad de las aguas de los ríos. *IV Simposio Del Agua En Andalucía, II*, 203–213.
- Angermier, P. L., & Karr., J. R. (1986). Applying an index of biotic integrity based on stream-fish communities: Considerations in sampling and interpretation. *North American Journal of Fisheries Management*, 6, 418–427.
- Angermier, P. L., & Williams, J. E. (1993). Conservation of imperiled species and reauthorization of the endangered species act of 1973. *North American Journal of Fisheries Management*, 18(7), 34–38.
- Asadollahfardi, G. (2015). *Water Quality Management*. <https://doi.org/10.1007/978-3-662-44725-3>
- Barakat, A., El, M., Rais, J., Aghezzaf, B., & Slassi, M. (2016). International Soil and Water Conservation Research Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques ☆. *International Soil and Water Conservation Research*, 4(4), 284–292. <https://doi.org/10.1016/j.iswcr.2016.11.002>
- Borcard, D., Gillet, F., & Legendre, P. (2011). *Numerical Ecology with R*. Springer-Verlag New York.
- Cairns, J., & Pratt, J. R. (1993). *A History of Biological Monitoring Using Benthic Macroinvertebrates*. 10–27.
- Castro-Roa, D., & Pinilla-agudelo, G. (2014). *Periphytic diatom index for assessing the ecological quality of the Colombian Andean urban wetlands of Bogotá*. 33(2), 297–312.
- CEPAL. (2007). Serie Estudios estadísticos y prospectivos No 54. Imputación de datos: teoría y práctica. *Comisión Económica Para América Latina y El Caribe*.

- Chalar, G., Arocena, R., Pacheco, J. P., & Fabián, D. (2011). Trophic assessment of streams in Uruguay : A Trophic State Index for Benthic (TSI-BI). *Ecological Indicators*, 11. <https://doi.org/10.1016/j.ecolind.2010.06.004>
- Dargie, T. C. D. (1984). On the integrated interpretation of indirect site ordinations: a case study using semi-arid vegetation in south-eastern Spain. *Vegetatio*, 55, 37–55.
- Debels, P., Figueroa, R., Urrutia, R., Barra, R., & Niell, X. (2005). Evaluation of water quality in the Chillian River (Central Chile) using physicochemical parameters and a modified water quality index. *Environmental Monitoring and Assessment*, 110, 301–322. <https://doi.org/10.1007/s10661-005-8064-1>
- Decreto 3930 de 2010. (2010). *República de Colombia*.
- Decreto ley 2811 de 1974. (1974). *República de Colombia*.
- Dell’Uomo, A., & Torrisi, M. (2011). The Eutrophication/Pollution Index-Diatom based (EPI-D) and three new related indices for monitoring rivers: The case study of the river Potenza (the Marches, Italy). *Plant Biosystems-An International Journal Dealing with All Aspects of Plant Biology*, 145(2), 331–341.
- DeShon, J. D. (1995). Development and application of the invertebrate community index (ICI). *Biological Assessment and Criteria: Tools for Risk-Based Planning and Decision Making*, W. S. Davis and T. Simon, Eds. Boca Raton, Fla.: Lewis Publishers.
- Forero. (2012). *Elaboración de un índice para la evaluación de la calidad ecológica de la cuenca del Río Negro (Antioquia) con base en los macroinvertebrados bentónicos*.
- Forero, L. C., Longo, M., Ramírez R., J. J., & Chalar, G. (2014). Índice de calidad ecológica con base en macroinvertebrados acuáticos para la cuenca del río Negro (ICE RN-MAE), Colombia. *Revista Biología Tropical*, 62(2), 233–247.
- Gammon, J. R. (1976). The fish populations of the middle 340 km of the Wabash River. Water Resources Research Center Technical Report 86. *West Lafayette, Ind.: Purdue University*.
- Gammon, J. R., Spacie, A., L., J., Hamelink, & Kaesler, R. L. (1981). Role of electrofishing in assessing environmental quality of the Wabash River. Pp. 307-324 in *Ecological Assessments of Effluent Impacts on Communities of Indigenous Aquatic Organisms*, J. M. Bates and C. I. Weber, eds. ASTM

- STP 730. Philadelphia, Pa.: Ame. *American Society for Testing and Materials*.
- Gauch, H. G. (1982). *Multivariate analysis in community ecology*. Cambridge University Press.
- Gause, G. F. (1930). Studies on the ecology of Orthoptera. *Ecology*, 11, 307–325.
- Giorgio, A., De Bonis, S., & Guida, M. (2016). Macroinvertebrate and diatom communities as indicators for the biological assessment of river Picentino (Campania, Italy). *Ecological Indicators*, 64, 85–91.
- Gómez, N., & Licursi, M. (2001). The Pampean Diatom Index (IDP) for assessment of rivers and streams in Argentina. *Aquatic Ecology*, 35, 173– 181.
- González, I., & Déjean, S. (2015). *Package ‘CCA.’*
- González, V., Caicedo, O., & Ramírez, N. (2011). *Aplicación de los índices de calidad de agua NSF, DINIUS y BMWP. 1*, 97–108.
- Greenacre, M., & Primicerio, R. (2014). *Multivariate analysis of ecological data*. Fundacion BBVA.
- Guisande-González, C., Barreiro-Felpeto, A., & Maneiro-Estraviz, I., Riveiro-Alarcón, I., Vergara-Castaño, A. Vaamonde-Liste, A. (2006). *Tratamiento de Datos*. Galicia, España: Ediciones Díaz de Santos.
- GWP, G. W. P. (2009). *Manual para la Gestión Integrada de Recursos Hídricos en Cuencas*. International Network of Basin Organizations - INBO. Global Water Partnership, Stockholm, Sweden. ISBN : 978-91-85321-78-0.
- GWP TAC, G. W. P. (2000). *Background Paper No. 4. Gestión Integrada de Recursos Hídricos*. Global Water Partnership, Stockholm, Sweden.
- Haase, R., & Nolte, U. (2008). *The invertebrate species index (ISI) for streams in southeast*. 8, 599–613. <https://doi.org/10.1016/j.ecolind.2007.08.005>
- Harrell, F. E. (2019). *Package ‘Hmisc.’*
- Hart, D. D. (1994). Building a Stronger Partnership between Ecological Research and Biological Monitoring. *Journal of the North American Benthological Society*, 13, 110–116.
- Heino, J., Virkkala, R., & Toivonen, H. (2009). Climate change and freshwater biodiversity: detected

- patterns, future trends and adaptations in northern regions. *Biological Reviews*, 84(1), 39–54.
- Hesse, R. (1924). *Tiergeographie auf ökologischer Grundlage*.
- Hill, M. O., & Gauch, H. G. (1980). Detrended correspondence analysis: an improved ordination technique. *Vegetatio*, 42, 47–58.
- IDEAM. (2011). *Formato Común Hoja Metodológica Índice de calidad del agua en corrientes superficiales (ICA)*. 1–10.
- Juggins. (2018). *Package ‘rioja’*.
- Juggins, S. (2005). C2 User guide. Software for ecological and palaeoecological data analysis and visualization. *University of Newcastle*.
- Karr, J. R. (1981). Assessment of Biotic Integrity Using Fish Communities. *Fisheries*.
[https://doi.org/10.1577/1548-8446\(1981\)006<0021](https://doi.org/10.1577/1548-8446(1981)006<0021)
- Kerans, B. L., & Karr, J. R. (1994). A Benthic Index of Biotic Integrity (B-IBI) for Rivers of the Tennessee Valley. *Ecological Applications*, 4(4), 768–785.
- Koklu, R., Sengorur, B., & Topal, B. (2010). (2010). Water quality assessment using multivariate statistical methods—a case study: Melen River System (Turkey). *Water Resources Management*, 24(5), 959–978.
- Kwandrans, J., Eloranta, P., Kawecka, B., & Wojtan, K. (1998). *Use of benthic diatom communities to evaluate water quality in rivers of southern Poland*. 193–194.
- Lamprecht, H. (1990). *Silvicultura en los Trópicos*. Eschborn: GTZ.
- Lane, C. R., & Brown, M. T. (2007). *Diatoms as indicators of isolated herbaceous wetland condition in Florida, USA*. 7, 521–540. <https://doi.org/10.1016/j.ecolind.2006.06.001>
- Legendre, P., & Birks, H. J. B. (2012). *From Classical to Canonical Ordination*.
<https://doi.org/10.1007/978-94-007-2745-8>
- Ley 1450 de 2011. *Plan Nacional de Desarrollo 2010-2014*.
- Leonard, P. M., & Orth, D. J. (1986). Application and testing of an index of biotic integrity in small,

- coolwater streams. *Transactions of the American Fisheries Society*, 115, 401–414.
- Leps, J., & Smilauer, P. (2003). *Multivariate Analysis of Ecological Data using CANOCO*. Cambridge University Press.
- Lepš, J., & Šmilauer, P. (1999). *Multivariate Analysis of Ecological Data*.
- Li, Y. I., Ogunnaike, B. A., & Roberts, C. J. (2010). Multi-variate approach to global protein aggregation behavior and kinetics: Effects of pH, NaCl, and temperature for α -chymotrypsinogen A. *Journal of Pharmaceutical Sciences*, 99(2), 645–662.
- Lynch, M., & Gabriel, W. (1987). Environmental tolerance. *The American Naturalist*, 129(2), 283–303.
- MADS, M. de A. y D. S. (2014). *Guía técnica para la formulación de planes de ordenamiento del recurso hídrico. Bogotá, D.C.: Colombia*.
- MADS, M. de A. y D. S. (2018). *Guía para el ordenamiento del recurso hídrico continental superficial. Bogotá, D.C.: Colombia*.
- Mayer, D. G., & Butler, D. G. (1993). Statistical validation. *Ecological Modelling*, 68, 21–32.
- McCormick, P. V., & Cairns, J. (1994). Algae as indicators of environmental change. *Journal of Applied Phycology*, 6(5), 509–526.
- McGeoch, M., Van Rensburg, B., & Botes, A. (2002). The verification and application of bioindicators: a case study of dung beetles in a savanna ecosystem. *Journal of Applied Ecology*, 39, 661–672.
- Miller, D. L., Leonard, P. M., Hughes, R. M., Karr, J. R., Moyle, P. B., Schrader, L. H., Thompson, B. A., Daniel, R. A., Fausch, K. D., Fitzhugh, G. A., Gammon, J. R., Halliwell, D. B., Angermeier, P. L., & Orth, D. J. (1988). Regional applications of an index of biotic integrity for use in water resource management. *Fisheries*, 13, 12–20.
- Mitsch, W. J., Gosselink, J. G., Zhang, L., & Anderson, C. J. (2009). *Wetland ecosystems. John Wiley & Sons*.
- Mora, F. (2019). The use of ecological integrity indicators within the natural capital index framework: The ecological and economic value of the remnant natural capital of México. *Journal for Nature Conservation*, 47, 77–92.

- Moya, N., Tomanova, S., & Oberdorff, T. (2007). Initial development of a multi-metric index based on aquatic macroinvertebrates to assess streams condition in the Upper Isiboro-Sécure Basin, Bolivian Amazon. *Hydrobiologia*, 589, 107–116.
- N. Gómez. (1999). Epipellic diatoms from the Matanza-Riachuelo river (Argentina), a highly polluted basin from the pampean plain: biotic indices and multivariate analysis. *Aquatic Ecosystem Health & Management*, 2(3), 301–309. <https://doi.org/DOI: 10.1080/14634989908656966>
- Norris, R. H., & Norris, K. R. (1995). *The need for biological assessment of water quality: Australian perspective*. 1–6.
- Noss, R. F. (1990). Indicators for monitoring biodiversity: a hierarchical approach. *Conserv. Biol.*, 4, 355–364.
- Odum, E. (1972). *Ecología*. (Tercera ed).
- Oksanen, J., Guillaume, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P., O’Hara, R. B., Simpson, G., Solymos, P., Stevens, M., Szoecs, E., & Wagner, H. (2019). *Package ‘vegan.’* <https://cran.r-project.org/web/packages/vegan/vegan.pdf>
- Opdam, P. F. ., Kalkhoven, J. T. ., & Philippona, J. (1984). *Verban tussen broedvogelgemeenschappel en begroeiing in een landschap bij Amerongen* (Issue 5).
- P, L., & LF, L. (2012). *Numerical ecology*.
- Pander, J., & Geist, J. (2013). Reviews Ecological indicators for stream restoration success. *Ecological Indicators*, 30, 106–118.
- Pielou, E. C. (1984). The interpretation of ecological data: a primer on classification and ordination. *John Wiley & Sons*.
- Pinilla, G. A. P. A. (1998). Indicadores biológicos en ecosistemas acuáticos continentales de Colombia: compilación bibliográfica. *Universidad Jorge Tadeo Lozano*.
- Plafkin, J. L., Barbour, M. T., Porter, K. D., Gross, S. K., & Hughes., R. M. (1989). Rapid Bioassessment Protocols for Use in Rivers and Streams: Benthic Macroinvertebrates and Fish. EPA/444/4-89001. Washington, D.C.: U.S. Environmental Protection Agency. *U.S. Environmental Protection Agency*.

- PNGIRH, P. N. para la G. I. del R. H.-. (2010). *República de Colombia*.
- Ministerio de Ambiente, Vivienda y Desarrollo Territorial - MAVDT- Bogotá, D.C. Colombia, (2010).
- Ramírez, A., & Viña, G. (1998). Limnología Colombiana. Aportes a su conocimiento y estadísticas de análisis. *Undación Universidad Jorge Tadeo Lozano*, 7, 70–72.
- Ramos-Merchante, A., & Prenda, J. (2017). Macroinvertebrate taxa richness uncertainty and kick sampling in the establishment of Mediterranean rivers ecological status. *Ecological Indicators*, 72, 1–12.
- Reynoldson, T. B., Norris, R. H., Resh, V. H., Day, K. E., & Rosenberg, D. M. (1997). The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society*, 16(4), 833–852.
- Ríos-Touma, B., Acosta, R., & Prat, N. (2014). The Andean Biotic Index (ABI): revised tolerance to pollution values for macroinvertebrate families and index performance evaluation. *Revista de Biología Tropical*, 62, 249–273.
- Roldán-Perez, G. (2003). Bioindicación de la calidad del agua en Colombia. Propuesta para el uso del método BMWP/Col. *Universidad de Antioquia*.
- Samboni Ruiz, N. E., Carvajal Escobar, Y., & Escobar, J. C. (2007). Revisión de parámetros fisicoquímicos como indicadores de calidad y contaminación del agua. *Ingeniería e Investigación*, 27(3), 172–181.
- Shanon, C. E., & Weaver, W. (1949). *The mathematical theory of information*.
- Shelford, V. . E. . (1913). Animal communities in a temperate America. *University of Chicago Press, Chicago*.
- Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques : A case study of the Fuji river basin , Japan. *Environmental Modelling & Software*, 22, 464–475. <https://doi.org/10.1016/j.envsoft.2006.02.001>
- Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163(4148), 688.
- Smol, J. P. (2009). Pollution of lakes and rivers: a paleoenvironmental perspective. *John Wiley & Sons*.
- Spellerberg, I. F. (2005). Monitoring ecological change. *Cambridge University Press*.

- Spradley, J., Glazer, B., & Kay, R. (2019). Mammalian faunas, ecological indices, and machine-learning regression for the purpose of paleoenvironment reconstruction in the Miocene of South America. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *518*, 155–171.
- Steedman, R. J. (1988). Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. *Canadian Journal of Fisheries and Aquatic Sciences*, *45*, 492–501.
- Ter Braak, C. J. (1989). CANOCO—an extension of DECORANA to analyze species-environment relationships. *Hydrobiologia*, *184*(3), 169–170.
- Ter Braak, C. J. ., & Prentice, I. C. (1988). *A Theory of Gradient Analysis*. *34*(03), 271–317. [https://doi.org/10.1016/S0065-2504\(03\)34003-6](https://doi.org/10.1016/S0065-2504(03)34003-6)
- Ter Braak, C. J., & Prentice, I. C. (2004). A theory of gradient analysis. *Advances in Ecological Research*, *34*, 235–282.
- Ter Braak, C. J., & Smilauer, P. (2002). *CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5)*.
- Vargas, Lady. (2016). *Diagnóstico de la calidad ecológica a través de algas perifíticas en escenarios hídricos afectados por minería a cielo abierto con diferente temporalidad, Jigualito (Condoto, Chocó)*.
- Vasiljević, B., Simić, S. B., Paunović, M., Zuliani, T., Krizmanić, J., Marković, V., & Tomović, J. (2017). Contribution to the improvement of diatom-based assessments of the ecological status of large rivers—The Sava River Case Study. *Science of the Total Environment*, *605*, 874–883.
- Vélez, C. A. (2019). *Distribución de diatomeas en sistemas esturiales y fluviales asociados del norte de la Patagonia: Implicancias en la reconstrucción cuantitativa de paleoambientes holocenos*.
- Venables, W. N., Smith, D. M., & Team, R. C. (2019). *An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics*.
- Villamarín, C., Paul, M. J., & Barbour, M. T. (2013). A tool to assess the ecological condition of tropical high Andean streams in Ecuador and Peru : The IMEERA index A tool to assess the ecological condition of tropical high Andean streams in Ecuador and Peru : The IMEERA index. *Ecological Indicators*, *29*(June), 79–92. <https://doi.org/10.1016/j.ecolind.2012.12.006>

- Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., & Shao, X. (2013). Assessment of surface water quality via multivariate statistical techniques : A case study of the Songhua River Harbin region , China. *Journal of Hydro-Environment Research*, 7(1), 30–40. <https://doi.org/10.1016/j.jher.2012.10.003>
- Warnes, G. R., Bolker, B., Gorjanc, G., Grothendieck, G., Korosec, A., Lumley, T., MacQueen, D., Magnusson, A., Rogers, J., & Others, A. (2017). *Package ‘gdata.’*
- Wickham, H., Bryan, J., Kalicinski, M., Valery, K., Leittenne, C., Colbert, B., Hoerl, D., & Miller, E. (2019). *Package ‘readxl.’*
- Zabala, A. (2016). *Diagnóstico ambiental de la ciénaga de Ayapel a través de la variación temporal de la morfología funcional y un indicador de calidad ecológica del fitoplancton.*
- Zhang, Q., Li, Z. W., Zeng, G. M., Li, J. B., Fang, Y., Yuan, Q. S., Wang, Y. M., & Ye, F. Y. (2009). *Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: a case study of Xiangjiang.*