



*Instituto de Investigaciones  
de la Amazonía Peruana*

**BIODAMAZ**

**Perú - Finlandia**

**GUÍA PARA ESTUDIAR PATRONES  
DE DISTRIBUCIÓN  
DE ESPECIES AMAZÓNICAS**



Documento  
Técnico  
Nº 06

**SERIE IIAP - BIODAMAZ**  
Iquitos - Perú

# GUIA PARA ESTUDIAR PATRONES DE DISTRIBUCIÓN DE ESPECIES AMAZÓNICAS

## DOCUMENTO TÉCNICO N° 06

Instituciones ejecutoras:



Instituto de  
Investigaciones de la  
Amazonía Peruana



Universidad de  
Turku, Finlandia



BIOTA BD

Biota BD Oy,  
Finlandia

Institución colaboradora:



UNAP  
IQUITOS  
PERU

Universidad Nacional de  
la Amazonía Peruana

BIODAMAZ, Perú - Finlandia  
Proyecto Diversidad Biológica de la Amazonía Peruana

**Comité Editorial**

Víctor Miyakawa Solís  
José Álvarez Alonso  
Filomeno Encarnación Cajañahupa  
Jorge Gasché  
Víctor Montreuil Frías  
ErasmO Otarola Acevedo

**Fotografías**

Instituto de Investigaciones de la Amazonía Peruana

El presente documento ha sido realizado con financiamiento del Ministerio de Relaciones Exteriores de Finlandia y del Gobierno del Perú, a través del Instituto de Investigaciones de la Amazonía Peruana - IIAP, en el marco del Convenio de Cooperación Técnica Internacional entre Perú y Finlandia: Proyecto Diversidad Biológica de la Amazonía Peruana - BIODAMAZ.

© 2004, Instituto de Investigaciones de la Amazonía Peruana - IIAP  
Proyecto Diversidad Biológica de la Amazonía Peruana - BIODAMAZ  
Av. Abelardo Quiñones km 2.5  
Apto. 784 - Teléfonos: (065) 264060 - 265515 - 265516 Fax: (065) 265527  
Iquitos - Perú  
Correo electrónico: [biodamaz@iiap.org.pe](mailto:biodamaz@iiap.org.pe)  
<http://www.iiap.org.pe/biodamaz>

**ISBN N°** 9972-667-13-8

**Hecho el depósito legal N°** 1501222005-0377

**Imprenta:**

Dominius Publicidad  
Telf.: (511) 4450735  
[dominius@infonegocio.net.pe](mailto:dominius@infonegocio.net.pe)

Los textos pueden ser utilizados total o parcialmente citando la fuente.  
Hecho en el Perú



## INDICE

Presentación.....	5
Resumen ejecutivo.....	7
Executive summary.....	9
I. INTRODUCCIÓN.....	11
II. TEORÍAS BIOGEOGRÁFICAS Y DISTRIBUCIÓN DE LAS ESPECIES	
1. Introducción.....	15
2. La Teoría de los Refugios Pleistocénicos.....	15
3. La Teoría de la Biogeografía de Islas.....	16
4. Las relaciones entre la abundancia y distribución de especies.....	17
Variación espacial en abundancia dentro de especies.....	17
Correlación entre la abundancia y distribución entre las Especies.....	17
La teoría.....	17
Variación espacial en la abundancia de una especie.....	17
5. Fuentes de información sobre la distribución de especies.....	18
Museos de Historia Natural.....	18
Bases de datos.....	18
Listados de especies.....	18
Encuestas.....	18
III. ESPECIES INDICADORAS	
1. Introducción.....	21
2. Seis grupos indicadores.....	22
IV. INTRODUCCIÓN A LA PERCEPCIÓN REMOTA	
1. Introducción.....	27
2. Imágenes Landsat.....	28
V. LOS PRINCIPIOS DE MUESTREO, ANÁLISIS DE DATOS E INTERPRETA- CIÓN DE RESULTADOS EN ESTUDIOS DE BIOGEOGRAFÍA Y DE ECOLO- GÍA DE COMUNIDADES	
1. Introducción a los análisis de patrones de comunidades de especies.....	33
2. ¿Qué son datos de biogeografía y ecología de comunidades?.....	33
Preguntas básicas en los estudios de patrones de comunidades.....	33
¿Qué tipo de datos se obtiene?.....	33
¿Cómo realizar el muestreo?.....	35
3. Autocorrelación espacial y sus consecuencias.....	36
4. Medir diferencias y semejanzas.....	37
Medidas simétricas y asimétricas de distancia.....	38
Índices binarios de similitud.....	38
Índices cuantitativos de similitud e índices de distancia.....	39
Transformaciones de los datos.....	43
Las matrices de similitud y distancia.....	44
5. Identificando los gradientes.....	45
Análisis directo de gradiente.....	45
Análisis indirecto de gradiente u ordenación.....	47
Explicando con variables medioambientales la variación en la composición de especies visible en las ordenaciones.....	52
6. Clasificación.....	56

Métodos diferentes de clasificación.....	56
TWINSPAN.....	57
Clasificación del vecino más cercano (nearest neighbour clustering, o single linkage clustering).....	57
Clasificación del vecino más distante (Furthest neighbour clustering, o complete linkage clustering).....	59
Clasificación del vecino relativo (intermediate linkage clustering, o proportional link linkage).....	59
Clasificación por el centroide (centroid clustering; UPGMC y WPGMC).....	62
Clasificación restringida a los vecinos (constrained clustering, chronological clustering).....	63
7. La prueba de Mantel: correlación entre matrices de distancia.....	63
 VI. METODOLOGÍAS DE MUESTREO PARA LOS GRUPOS INDICADORES	
1. Introducción.....	69
2. Criterios para ubicar los lugares de muestreo.....	69
3. Metodología para el inventario de plantas.....	70
Establecimiento del transecto de muestreo.....	70
El registro y el muestreo de plantas.....	72
El prensado de muestras.....	72
El secado de las muestras.....	73
Materiales específicos.....	74
4. Metodología para el inventario de la fauna	74
Generalidades.....	74
Hormigas.....	74
Anfibios anuros.....	76
Aves.....	77
 VII CONCLUSIONES.....	79
 VIII BIBLIOGRAFÍA.....	81
 ANEXO	
Ejemplo de una ficha de campo para el registro de plantas dentro de un transecto de 20 sub unidades.....	87
 EQUIPO TÉCNICO DEL PROYECTO.....	



## PRESENTACIÓN

Cada día es más evidente que el uso de los recursos biológicos excede en ciertas áreas la capacidad de regeneración natural, lo que resulta en la disminución de las poblaciones de plantas y animales aprovechadas y de aquellas que dependen de tales poblaciones. Para poder aprovechar sosteniblemente las poblaciones de la fauna y flora amazónica es importante conocerlas y conocer los requerimientos para que se mantengan viables. Estudios de campo son fundamentales para proporcionar dicha información básica.

Cuando se planifica el uso de la tierra es importante poder apoyarse con información actualizada y confiable sobre la distribución de especies, tanto de uso comercial como de subsistencia; igualmente importante es el conocimiento sobre especies raras y amenazadas para poder garantizar su conservación. Muchas veces es importante tener información sobre grandes extensiones, así como sobre tipos de ecosistemas y de bosques encontrados. Para tal fin es de mucha importancia conocer algunas especies que caracterizan estos ecosistemas y bosques; tales especies pueden indicar un cierto tipo de ecosistema cuando son encontradas en él.

El uso de especies indicadoras podría facilitar la clasificación de áreas extensas a través de estudios de distribución de especies indicadoras. Cuando se mapea las distribuciones de las especies en imágenes de satélite y mapas convencionales es posible estimar qué tipo de bosque o ecosistema hay en lugares donde se encuentran las especies indicadoras, dado que las especies indicadoras realmente reflejan los patrones de distribución de tipos de bosques y ecosistemas. Los inventarios biológicos, utilizando especies indicadoras y herramientas modernas de percepción remota, pueden de esa manera proporcionar información muy útil y necesaria para la planificación del uso de la tierra.

En este contexto, el Instituto de Investigaciones de la Amazonía Peruana IIAP, la empresa de consultorías ambientales, Biota BD Oy de Finlandia, y la Universidad de Turku de Finlandia, en el marco del Convenio Perú-Finlandia, Proyecto Diversidad Biológica de la Amazonía Peruana (BIODAMAZ), han contribuido con el documento "Guía para estudiar patrones de distribución de especies amazónicas". Este documento proporciona información sobre las metodologías para los inventarios biológicos, utilizando el concepto de especies indicadoras y el aporte de herramientas de percepción remota; en este caso sobre las imágenes de satélite y sobre el procesamiento de datos levantados mediante los estudios de campo. El propósito es que estas metodologías, cuando sean utilizadas para estudiar los patrones de distribución de especies amazónicas, puedan incrementar nuestro conocimiento sobre la diversidad biológica amazónica, de modo que esta información pueda ser aprovechada para alcanzar un uso racional y sostenible de la tierra y sus recursos, incluyendo conservación de especies amazónicas

**Dennis del Castillo Torres**  
*Presidente*  
*Instituto de Investigaciones de la*  
*Amazonía Peruana*

**Kimmo Pulkkinen**  
*Embajador de Finlandia*





## RESUMEN EJECUTIVO

La presente guía metodológica para estudiar patrones de distribución de especies amazónicas está orientada a jóvenes científicos, estudiantes y comunidad científica en general, y está basada en los estudios de campo ejecutados en el marco del proyecto Diversidad Biológica de la Amazonía Peruana, Perú-Finlandia (BIODAMAZ), en su Fase I. Dicho proyecto de cooperación técnica tiene su origen en un convenio entre los gobiernos del Perú y de Finlandia, cuya ejecución en su Fase I tuvo lugar del 1999 al 2002; la Fase II se encuentra en ejecución actualmente (2003-2007). El objetivo general del proyecto en la Fase I fue la conservación y uso sostenible de la diversidad biológica amazónica y su propósito fue brindar instrumentos de gestión y metodologías de investigación para el manejo sostenible de la diversidad biológica de la Amazonía Peruana. A través de estas acciones, el proyecto aspiró en última instancia a impulsar el desarrollo socioeconómico sostenible de la Amazonía Peruana. La contraparte nacional del proyecto BIODAMAZ es el Instituto de Investigaciones de la Amazonía Peruana (IIAP). La contraparte finlandesa en la ejecución del proyecto es un consorcio formado por la empresa finlandesa de consultorías ambientales, Biota BD Oy, y la Universidad de Turku. La Universidad Nacional de la Amazonía Peruana (UNAP) ha colaborado de muchas formas durante la ejecución del proyecto.

Es el propósito del proyecto BIODAMAZ que este documento sirva como una herramienta para realizar estudios biológicos de campo y para procesar adecuadamente la información. Se presentan secciones que ofrecen al lector la posibilidad de aprender los fundamentos teóricos para entender la distribución geográfica de especies, y también los principios de varios métodos numéricos comúnmente usados para analizar e interpretar las distribuciones y los factores que las controlan. Se espera que la parte teórica y la numérica sirvan para un público bastante amplio, por ejemplo estudiantes en las universidades peruanas. También se presentan para algunos grupos de organismos unos métodos estandarizados para documentación de composición de especies y otras características de comunidad en el campo. Esta parte probablemente despierte el interés de un público más selecto, pero por lo menos servirá como un ejemplo para ayudar a entender la calidad de trabajo que se requiere para los inventarios biológicos, que son parte de un conjunto que incluye, en teoría, inventarios de campo, análisis numéricos e interpretación de los resultados. Los inventarios biológicos constituyen los primeros estudios para futuros trabajos de investigación en ecología, sistemática o conservación, entre otros, y ofrecen resultados importantes para la toma de decisiones en políticas ambientales.

Los estudios de campo del proyecto BIODAMAZ fueron realizados para incrementar el conocimiento sobre la diversidad biológica y los patrones de distribución de la misma, para apoyar los procesos de planificación del uso de la tierra de modo que se logre un aprovechamiento sostenible de los recursos naturales, la ocupación ordenada de espacios amazónicos y la conservación de la diversidad biológica.

Para conocer la distribución de todas las especies, lo ideal sería inventariarlas en todos los rincones del planeta. Esto resulta prácticamente imposible de realizar, y por eso tenemos que suponer patrones de distribución con base en muestras, colecciones o investigaciones puntuales o focalizadas. Las suposiciones tienen que estar basadas en teorías o hipótesis de factores que controlan la distribución. En el documento se analiza tres teorías científicas (la teoría de los refugios pleistocénicos, la teoría de la biogeografía de islas y la teoría para explicar las relaciones entre la abundancia y distribución de especies basada en el concepto de nicho) y su significado en explicar la distribución de especies en la Amazonía. Se sigue discutiendo el rol de estas teorías y otras. La relevancia de las teorías científicas se evalúa a través de información levantada mediante especímenes en los museos de historia natural, la construcción de bases de datos y listados de especies y el uso de encuestas. La información tiene que respaldar las suposiciones de las teorías para que éstas sean aceptadas por la comunidad científica.

En este documento se discute sobre especies indicadoras de patrones de distribución de especies. La idea es buscar grupos de especies cuya composición varía entre localidades, paralelamente con la variación general de la composición florística y/o faunística entre estas localidades. Se puede, por ejemplo, clasificar los lugares de estudio según la composición de especies indicadoras, y si es que las indicadoras realmente sirven, esta



clasificación debe dar básicamente el mismo resultado que la clasificación de los lugares según la composición de especies de la flora y/o fauna más general.

Debido a la constante pérdida de hábitats y de biodiversidad en el mundo, existe una necesidad urgente de inventariar la composición de especies de plantas y animales para apoyar los planes de conservación. El objetivo principal de la conservación de la naturaleza es mantener poblaciones viables de todas las especies. Para poder asegurar la conservación de todas las especies, hay que conservar muestras de todos los tipos de hábitat que albergan especies especializadas. Para encontrar esos hábitats, prácticamente no hay otra alternativa que hacer muchos inventarios tratando de ubicar zonas que tengan una composición especial de especies. El uso de especies indicadoras de patrones de distribución de especies puede dar una solución para esta necesidad.

Bajo la hipótesis de que la distribución y abundancia de plantas está relacionada directamente con factores ambientales y formaciones geológicas, y la distribución y abundancia de animales con las formaciones vegetales y condiciones ambientales, el proyecto BIODAMAZ seleccionó seis grupos taxonómicos como potenciales indicadores para validar en el campo esta hipótesis: helechos (pteridofitas), melastomatáceas (la familia de plantas Melastomataceae), palmeras, insectos de hojarasca (hormigas), anfibios anuros y aves. Existen diferentes metodologías para el estudio de la flora y la fauna. Cada una de éstas tiene un fundamento y puede ser aplicada según el tipo de estudio que se quiera hacer, y según las condiciones para hacerlo. Para los estudios del proyecto BIODAMAZ se seleccionó una metodología para el inventario de plantas y las más adecuadas, según los propósitos del proyecto, para los inventarios de aves, anfibios anuros y hormigas. Estas metodologías se presentan en esta guía.

En este documento también se presentan diferentes métodos de análisis de datos; de cada uno de ellos se presenta la esencia, para permitir al lector la posibilidad de estimar su aplicabilidad al analizar sus propios datos, y para poder interpretar los resultados obtenidos con el método seleccionado. También se discuten los principios de muestreo.

Los bosques amazónicos cubren un área muy grande, y a menudo el acceso al bosque es difícil. Dadas estas condiciones, se hace necesario el uso de métodos de percepción remota para estudios a gran escala. La ventaja de la percepción remota está en su potencialidad para cubrir áreas grandes con un esfuerzo relativamente pequeño. Los métodos más utilizados en percepción remota son fotografías aéreas, imágenes de radar y de satélite. Las imágenes de satélite son una herramienta importante para biólogos que trabajan en los bosques amazónicos, porque ofrecen actualmente la fuente más segura de información sobre la variación geográfica de las características del paisaje.

En el proyecto BIODAMAZ se seleccionaron las áreas de muestreo de los grupos taxonómicos estudiados, con el criterio de que representen gradientes geológicos conocidos en condiciones climáticas algo diferentes, y que tengan acceso fácil. Dentro de estas áreas se ubicaron los puntos de muestreo en las imágenes de satélite, tratando de abarcar tanta variación en las imágenes como fuera posible. Al llegar en el terreno a una de las manchas elegidas para ser muestreada, se cuidó de no realizar los inventarios muy cerca de chacras ni de ríos, porque se trató de mantener siempre los inventarios dentro del bosque primario de tierra firme.

Esta guía pretende ser una herramienta para trabajar con metodologías de inventarios biológicos en el campo mediante el uso de especies indicadoras. Si bien es cierto que existen variados métodos de inventarios para los diferentes grupos taxonómicos, la selección de las metodologías presentadas aquí fue hecha con base en criterios científicos y en experiencias de investigadores con trabajos en la selva baja amazónica peruana. Además, metodologías de inventarios integrados con los objetivos establecidos son poco conocidas y escasas dentro de la literatura; por tanto, esta guía puede constituir el primer esfuerzo de aporte al conocimiento y de utilidad práctica para hacer evaluaciones de la diversidad biológica al nivel local o regional, basadas en la utilización de grupos indicadores e imágenes de satélite. Sin embargo, este trabajo no representa la última palabra en métodos de inventario y, desde luego, está sujeto a críticas, comentarios y sugerencias para su mejoramiento. Además, hay que notar que los métodos presentados aquí son diseñados para responder a ciertas interrogantes; principalmente para ver cómo cambian (o si cambian) la flora y la fauna de un lugar al otro, y si estos cambios tienen algo que ver con las condiciones medioambientales de los lugares estudiados. Otras interrogantes pueden requerir de metodologías algo diferentes.



## EXECUTIVE SUMMARY

This methodological guidebook to study distribution patterns of Amazonian species is oriented to young scientists, students and the scientific community in general. It is based on the field studies conducted in the framework of the project Biological Diversity of Peruvian Amazonia, Peru-Finland (BIODAMAZ), in its Phase I. The project is a technical cooperation project between the governments of Peru and Finland. Phase I of the project was implemented from 1999 to 2002; Phase II is currently being implemented (2003-2007). The overall objective of the project in Phase I was conservation and sustainable use of Amazonian biological diversity, and the project proposal was to develop environmental management and administration tools and research methodologies for sustainable management of Peruvian Amazonian biological diversity. Through these actions the project ultimately aimed to contribute to sustainable socioeconomic development in Peruvian Amazonia. The national counterpart of the project BIODAMAZ is the Research Institute of Peruvian Amazonia (*Instituto de Investigaciones de la Amazonía Peruana, IIAP*). *The Finnish counterpart is a consortium formed by the environmental consulting company, Biota BD Oy, and the University of Turku. The National University of Peruvian Amazonia (Universidad Nacional de la Amazonía Peruana, UNAP) has collaborated in many ways during the implementation of the project.*

The project BIODAMAZ aims that this document could serve as a guidebook in conducting biological field studies and correctly analysing the results. Separate sections treat the basic ideas and theories for understanding geographical distributions of species as well as the principles of various numerical methods commonly used for analysing species distributions and interpreting them in relation to factors that affect them. It is expected that theoretical and statistical sections can be of use to a wider audience, for example Peruvian university students. Some standardised sampling methods for studies of species composition and other community variables in the field are described for a limited number of groups of organisms. This section probably generates interest in a somewhat narrower group of readers; however, it can serve as an example that helps to understand the amount and quality of work required for biological inventories and to appreciate that a biological inventory is formed by a series of components, such as theory, field studies, numerical analysis and interpretation of results. Biological inventories are the basis for future studies in ecology, systematics or conservation, among others, and provide important results for decision-making in environmental policy issues.

The field studies of the project BIODAMAZ were conducted in order to increase the knowledge base on biological diversity and its distribution patterns with the aim of supporting the processes of land use planning for sustainable use of natural resources, organized occupation of Amazonian land, and conservation of biological diversity.

In order to know the distribution of all the species, the ideal would be to make an inventory of all the species all over the world. This is practically impossible to do, so we need to base our suppositions on species distribution patterns on samples or focalized and punctual research and collections. These suppositions need to be based on theories or hypothesis on factors that control the distribution. In this document three scientific theories (Theory of Pleistocene Refuges, Theory of Island Biogeography, theory that explains the relations between the abundance and distribution of species based on the concept of niche) are analysed, together with their significance in explaining species distributions in Amazonia. The role of these theories and others are still in debate. The relevance of scientific theories is evaluated by information gathering through data obtained from natural history museums, database and species list construction, and inquiries to knowledgeable people. The information gathered has to back up the suppositions based on the theories in order for these to be accepted by the scientific community.

In this document indicators of species distributions are discussed. The idea is to find species groups whose species composition varies between localities as the general composition of floral and faunal species varies. Study sites can be classified using the composition of indicator species, and if these really work, this classification can give the same result as the classification of sites using the general composition of flora and fauna.

Due to the constant loss of habitats and biodiversity in the world, there is an urgent need to

make inventories of species compositions of plants and animals to support conservation plans. The principal objective of nature conservation is to maintain viable populations of all species. In order to secure that all species are preserved, samples of all habitat types that harbour specialised species need to be preserved. In order to find these habitats there is really no alternative than to make many inventories. The use of indicator species of species distribution patterns can give a solution to this.

Based on the hypothesis that the distribution and abundance of plants are linked directly to environmental conditions and geological formations, and the distribution and abundance of animals to vegetation formations and environmental conditions, the project BIODAMAZ selected six taxonomic groups as potential indicators to validate this hypothesis in the field. These groups were: ferns (Pteridophyta), plants of family Melastomataceae, palms, insects of leaf litter (ants), anuran amphibians and birds. There are different study methodologies for flora and fauna. Each of these has a theory behind it and can be used according to the objectives of the study and conditions for conducting it. For the studies of the project BIODAMAZ, one methodology for plants was selected and for birds, anuran amphibians and ants the most suitable methodologies were selected. These methods are presented in the guidebook.

In this document there are also presented different methods of data analysis, of each method its theory is presented in order to facilitate the lector's evaluation of the applicability of each method for his/her own data analyses, and to help in the interpretation of the results obtained with the method selected. Also, the principles of sampling are discussed.

Amazonian forests cover a very large area, and often the access to them is difficult. Given these conditions, it is necessary to use remote sensing methods at large scale studies. The advantage of remote sensing is in its potential to cover large areas with a relative small effort. The most used methods in remote sensing are aerial photographs and radar and satellite images. Satellite images are an important tool to biologists working in Amazonian forests, because at the moment they offer the most secure source of information on geographical variation of landscape characteristics.

In the project BIODAMAZ the sampling sites for the taxonomic groups studied were selected so that they represent known geological gradients and different climatic conditions, and are easily accessible. Within these areas sampling points were located in satellite images so that the greatest portion of variation in the images was represented in the points. At the sampling point researchers took care not to conduct the inventories close to neither cultivations nor rivers as the goal was to make the inventories in not inundated forests.

The purpose of this guidebook is to present methods for biological inventories using indicator species. The methods presented here were selected on scientific criteria and personal experience that the researchers have of inventories in lowland Amazonian rain forest. Furthermore, integrated inventory methods are not well known and are scarce in literature; for these reasons this guidebook can provide practical information on making biological inventories on local and regional scales, based on indicator species and use of satellite images. However, this work is not, by all means, the last word in inventories, and is subject to critique, comments and suggestions to make it more useful. Also, it is important to note that the methods presented here are designed to respond to certain questions especially to see how the flora and fauna change (and if they do) from one place to another, and if these changes have anything to do with environmental conditions in the study sites. Other questions can require the use of other methods.



## INTRODUCCIÓN

La presente guía metodológica para estudiar patrones de distribución de especies amazónicas pretende ser un documento que sirva como una herramienta para realizar estudios biológicos de campo y para procesar adecuadamente la información así levantada. Los inventarios biológicos constituyen los primeros estudios para futuros trabajos de investigación en ecología, sistemática o conservación, entre otros, y ofrecen resultados importantes para la toma de decisiones en políticas ambientales. La guía está orientada a jóvenes científicos, estudiantes y comunidad científica en general.

La guía está basada en los estudios de campo ejecutados en el marco del proyecto Diversidad Biológica de la Amazonía Peruana, Perú-Finlandia (BIODAMAZ), en su Fase I. Dicho proyecto de cooperación técnica tiene su origen en un convenio entre los gobiernos del Perú y de Finlandia cuya ejecución en su Fase I tuvo lugar del 1999 al 2002; la Fase II se encuentra en ejecución actualmente (2003-2007). El objetivo general del proyecto en la Fase I fue la conservación y uso sostenible de la diversidad biológica amazónica y su propósito fue brindar instrumentos de gestión y metodologías de investigación para el manejo sostenible de la diversidad biológica de la Amazonía peruana. A través de estas acciones el proyecto últimamente aspiró a asistir al desarrollo socioeconómico sostenible de la Amazonía Peruana. El objetivo incluyó apoyar al país en el desarrollo de herramientas de gestión ambiental, por ejemplo, a través de la elaboración de la estrategia para la conservación y el uso sostenible de la diversidad biológica de la región amazónica; el desarrollo de metodologías de investigación y evaluación de la diversidad biológica de la Amazonía Peruana, con el objetivo de contribuir a mejorar su conocimiento y uso en los procesos de planificación; y el fortalecimiento de capacidades y de colaboración y comunicación interinstitucional e intersectorial en este campo. La contraparte nacional del proyecto BIODAMAZ es el Instituto de Investigaciones de la Amazonía Peruana (IIAP). La contraparte finlandesa en la ejecución del proyecto es un consorcio formado por la empresa finlandesa de consultorías ambientales, Biota BD Oy, y la Universidad de Turku. La Universidad Nacional de la Amazonía Peruana (UNAP) ha colaborado en numerosas formas durante la ejecución del proyecto.

Los estudios de campo del proyecto BIODAMAZ fueron realizados para incrementar el conocimiento sobre la diversidad biológica y los patrones de distribución de la misma, con el objeto de apoyar los procesos de planificación del uso de la tierra de modo que se logre un aprovechamiento sostenible de los recursos naturales, la ocupación ordenada de espacios amazónicos, y la conservación de la diversidad biológica.

Para facilitar y fortalecer el estudio de la Amazonía Peruana, el proyecto BIODAMAZ ha creado un mosaico de las escenas de imágenes de satélite Landsat TM (*Thematic Mapper*), como una meta del fortalecimiento del conocimiento sobre el ambiente y la vegetación de la Amazonía Peruana (ver Bendayán *et al.*, 2002, BIODAMAZ, 2004b). Este mosaico está dispuesto para los usuarios en diversas formas, incluso en el Internet ([http://www.iiap.org.pe/servidor\\_mapas](http://www.iiap.org.pe/servidor_mapas)), formando parte del Sistema de Información de la Diversidad Biológica y Ambiental de la Amazonía Peruana SIAMAZONIA (<http://www.siamazonia.org.pe>).

Esta guía está dividida en cinco partes distintas. La primera parte (Capítulo 2) trata de teorías científicas que han sido desarrolladas para explicar la distribución de especies en la Amazonía y las fuentes de información utilizadas para dar soporte a las suposiciones de tales teorías. La segunda parte (Capítulo 3) introduce el tema de especies indicadoras. La idea de especies indicadoras es buscar grupos de especies cuya composición varía entre localidades, paralelamente con la variación general de la composición florística y/o faunística entre estas localidades. Se presenta seis grupos taxonómicos que fueron utilizados en el proyecto BIODAMAZ como indicadores para validar la idea de especies indicadoras. La tercera parte (Capítulo 4) ofrece una breve introducción a la percepción remota, y su utilidad y uso en inventarios biológicos. La cuarta parte (Capítulo 5) se concentra en los métodos de análisis estadístico para el tratamiento de datos colectados en el campo. La quinta parte (Capítulo 6) describe los métodos aplicados en los estudios del proyecto BIODAMAZ sobre los seis grupos taxonómicos.

Esta guía pretende ser una herramienta para trabajar con metodologías de inventarios biológicos en el campo mediante el uso de especies indicadoras. Las metodologías de inventarios integrados son poco conocidas y escasas dentro de la literatura; por lo tanto, esta guía puede constituir el primer esfuerzo de aporte al conocimiento y de utilidad práctica para hacer evaluaciones de la diversidad biológica a nivel local o regional, basadas en la utilización de grupos indicadores e imágenes de satélite. Hay que notar que los métodos presentados aquí son diseñados para responder a ciertas interrogantes, principalmente para ver cómo cambian (o si cambian) la flora y la fauna de un lugar al otro, y si estos cambios tienen algo que ver con las condiciones medioambientales de los lugares estudiados.



Teorías  
biogeográficas  
y distribución  
de las especies



## 1. INTRODUCCIÓN

Para conocer la distribución de todas las especies, lo ideal sería inventariar a todas estas especies en todos los rincones del planeta. Esto resulta prácticamente imposible de realizar, y por eso tenemos que suponer patrones de distribución con base en muestras, en colecciones o investigaciones puntuales o focalizadas. Las suposiciones tienen que estar basadas en teorías o hipótesis de factores que controlan la distribución. En las siguientes páginas brindamos al lector una introducción a las teorías e hipótesis más frecuentemente citadas en relación con la fauna y flora amazónicas.

## 2. LA TEORÍA DE LOS REFUGIOS PLEISTOCENICOS

Durante la época del Pleistoceno (desde hace 1,5 millones de años hasta unos 10 000 años atrás) se produjeron grandes glaciaciones en el hemisferio norte. Mientras los glaciares fueron aumentando y disminuyendo en este hemisferio, el clima también fue cambiando en otros lugares, incluyendo el cinturón tropical.

Uno de los primeros biólogos que dio importancia biogeográfica a estos cambios fue el ornitólogo británico R. E. Moreau (1963), quien propuso que los bosques de África central se expandieron y contrajeron durante el Pleistoceno. Moreau creyó que la fragmentación reversible de los bosques africanos proporcionó los mecanismos de aislamiento<sup>1</sup> necesarios para explicar la distribución de especies endémicas (geográficamente restringidas) en muchos, ahora, bosques aislados del este africano. Moreau no tuvo un buen sustento al tratar de explicar sus argumentos con evidencias persuasivas de geología y meteorología; por eso fue ignorado por un escéptico público científico.

La propuesta siguiente en aparecer estuvo mejor documentada. Su autor es el geólogo Jürgen Haffer (Haffer 1969); su tesis comienza con la observación de la distribución de lluvias en Sudamérica tropical, que ha dado lugar a hábitats boscosos y no boscosos. Los bosques están principalmente en la Amazonía, mientras que los pastizales, sabanas y otras formaciones con vegetación abierta componen los llanos de Colombia y Venezuela, y el cerrado y campo-cerrado de Brasil. Haffer observó que los límites de las áreas no boscosas coinciden estrechamente con una precipitación anual de 1500 mm. En lugares donde las lluvias anuales están sobre este nivel, la vegetación es predominantemente boscosa, y donde la precipitación es menor, los fuegos de la estación seca convierten el paisaje en pastizales o sabanas.

Haffer propuso luego que durante el Pleistoceno, cuando la glaciación avanzó sobre los continentes del norte, las lluvias en las regiones tropicales disminuyeron. Una baja precipitación se esperaría durante períodos de glaciación debido a una disminución general de la temperatura en la tierra y a la consecuente reducción de la evaporación. Los registros fósiles de polen de los lagos andinos indican que el promedio de temperaturas en Sudamérica estuvo entre 4 y 6°C más bajo que ahora. Con base en la propuesta sequía pleistocénica, Haffer simplemente cortó 25% de la precipitación anual actual en toda la Amazonía y revisó dónde quedarían todavía más que 1500 mm de lluvia, indicando así áreas boscosas durante el Pleistoceno. Esto mostró a los bosques a manera de parches o refugios dentro de una matriz de vegetación abierta.

Los refugios separados del bosque podrían formar una escena perfecta para lo que funcionaría como una “bomba de especiación”, a través de un mecanismo de aislamiento (Prance 1982). Durante los períodos secos, las especies estarían confinadas a refugios ampliamente separados. Cuando las condiciones de humedad retornaron, los parches de bosques se expandieron y eventualmente conectaron dentro de vastos bosques contiguos que normalmente cubrían la mayoría de la cuenca amazónica. Si los períodos secos respondieron a los períodos de avance glacial en el hemisferio norte, habrían durado 50 a 100 mil años o más, posiblemente el tiempo suficiente para permitir la evolución de nuevas especies de poblaciones hermanas aisladas.

Como evidencia de que ocurrió esta evolución, Haffer proporcionó mapas que mostraron la distribución de tucanes, chachalacas, y otras aves, diferentes pero estrechamente relacionadas. La distribución de estas aves coincidió con las regiones de los refugios postulados. Haffer también buscó zonas de hibridización en áreas entre refugios. Haffer pudo documentar la ocurrencia de algunas zonas de hibridización justo donde se podría esperar que existieran.

Posteriormente, varios otros autores han reportado que también en otros grupos de organismos por ejemplo plantas, mariposas y lagartijas (Prance 1982) se puede observar parches de endemismo dentro de la Amazonía. Uno puede interpretar este patrón como un soporte para la teoría de refugios pleistocénicos.

Aunque la Teoría de los Refugios Pleistocénicos recibió inicialmente amplia aceptación (Prance 1982), más recientemente se han levantado muchas críticas contra la teoría. Las críticas vienen de dos frentes. Por un lado, uno puede interpretar los supuestos parches de endemismo igualmente como consecuencia de la desigualdad geográfica en el esfuerzo dedicado a estudios biológicos (Nelson et al. 1990) en vez de realidades medioambientales (Beven et al. 1984, Tuomisto & Ruokolainen 1997). Por otro lado, no hay evidencias geológicas o palinológicas que prueben que el bosque amazónico habría estado dividido en parches separados, aunque aparentemente sí ha retrocedido a veces en sus márgenes (Salo 1987, Colinvaux 1996). Además, se ha cuestionado la idea de la sequía amazónica durante las glaciaciones pleistocénicas (Colinvaux 1996). Para concluir, se puede decir que la Teoría de Refugios Pleistocénicos tiene ahora muchas dudas y problemas. Seguramente el clima ha variado durante la historia, y esas variaciones indudablemente han tenido efectos en la distribución de las especies. Pero será muy difícil demostrar que algunos patrones actuales de distribución de especies tienen su explicación en la historia climática, por la simple razón de que no será fácil reconstruir las condiciones históricas del clima.

<sup>1</sup>Se utiliza el término *vicariance* en inglés.



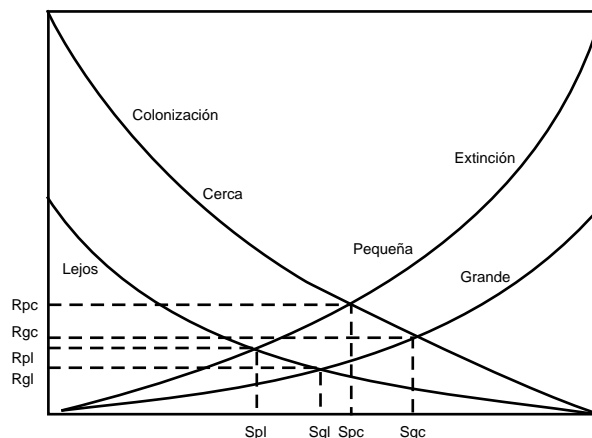
### 3. LA TEORÍA DE LA BIOGEOGRAFÍA DE ISLAS

La Teoría de la Biogeografía de Islas no es una teoría de patrones de distribución, sino de números de especies, específicamente en islas separadas del continente. Sin embargo, proporciona una base conceptual para varias hipótesis que tienen que ver tanto con los patrones de distribución, como con la dinámica poblacional de especies. La teoría (MacArthur & Wilson 1967) tiene como puntos de partida dos suposiciones básicas. Primero, debido a que las especies siempre tienen una capacidad algo restringida de moverse o de distribuir los propagulos, parece obvio que una isla lejana del continente o una isla pequeña reciba menos inmigrantes que una isla cercana o grande. Segundo, cuando la población de una especie es pequeña, tendrá una probabilidad más alta de extinguirse que otra población más numerosa.

La consecuencia lógica de estas suposiciones es que cualquier isla cercana al continente debe de tener más especies que otra isla más lejana del continente, por lo menos si las islas tienen superficies iguales. Esto se debe a que una isla cercana siempre recibe más inmigrantes, y por eso más especies, que una isla lejana. Respecto al tamaño de la isla y el número de especies por isla, existe una correlación positiva: cuanto más grande es la isla, más especies contiene. Esta relación resulta de la probabilidad de extinción. Una isla grande siempre puede soportar poblaciones más grandes y, por eso, menos propensas a extinción.

La Teoría de la Biogeografía de Islas plantea que existe un equilibrio entre la tasa de inmigración de nuevas especies que no están presentes en la isla en un momento dado, y que provienen como colonizadoras del continente, y la tasa de extinción de las especies que habitan la isla (siendo las tasas de inmigración o extinción medidas en número de especies por unidad de tiempo). Es la tasa de recambio en este equilibrio la que determina la diversidad de especies que pueda contener un área dada.

En la historia de una isla originalmente vacía, sin ninguna especie, la tasa de inmigración de nuevas especies es mayor en un principio, y se va reduciendo con el paso del tiempo, en la medida que las primeras especies migrantes van ocupando los hábitats disponibles. Por el contrario, la extinción de las especies es en principio muy baja y con el paso del tiempo va aumentando al existir en un momento dado un número muy grande de especies. Estos dos procesos vienen de direcciones opuestas, se encuentran en un punto, donde se establece el equilibrio, y es allí donde la tasa de inmigración es igual a la tasa de extinción (Figura 1).



**Figura 1.** El modelo del equilibrio de la biogeografía de islas de MacArthur & Wilson (1967) muestra el efecto del tamaño de la isla (diferentes curvas de la tasa de extinción) y aislamiento (diferentes curvas de la tasa de colonización) en el equilibrio del número de especies (S) y la tasa de recambio de especies (R). Las intersecciones de las curvas para las islas de diferente tamaño y distancia (pequeña lejos y grande cerca) pueden ser usadas para predecir el número relativo de especies y las tasas de recambio en el equilibrio.

Si bien la relación área-número de especies explica en gran parte el número de especies presente en una isla, existen, además de los factores mencionados, otros tales como la diversidad de hábitats.

La correlación positiva entre el tamaño de un área y la cantidad de especies que se encuentra allí es algo casi inevitable. Empíricamente esta relación ha sido descrita a través de una fórmula exponencial (Arrhenius 1921):

$$S = CA^z$$

donde S es el número de especies, A es el área, C es una constante que varía ampliamente con los taxa que se estén considerando y la unidad de área, y z es una constante que en la mayoría de los casos estudiados está entre 0,20 y 0,35. El número de especies (S) se incrementa cuando el área aumenta en una relación logarítmica doble.

La Teoría de la Biogeografía de Islas significaba un avance muy significativo para tratar analíticamente, y no sólo según experiencias empíricas, la variación en la cantidad de especies entre diferentes áreas. Posteriormente se ha utilizado la misma base conceptual en el desarrollo de varias otras hipótesis o teorías relacionadas con la biogeografía.

Tal vez las teorías más significativas, o por lo menos más discutidas actualmente, son el concepto de la dinámica de metapoblaciones (Hanski 1999) y la teoría neutral sobre biodiversidad y biogeografía (Hubbell 2001).

Las conclusiones que arroja la Teoría de la Biogeografía de Islas han sido aplicadas también a las prácticas y políticas de conservación. El concepto de isla es aplicable a los fragmentos aislados de hábitats en los cuales se han transformado un gran número de los ecosistemas. Es sobre estos parches donde se selecciona áreas con el fin de preservar las poblaciones y las relaciones representativas de los ecosistemas. En relación con la selección y el manejo de áreas protegidas, hay varias interrogantes a las cuales se busca dar respuestas apoyándose en la Teoría de la Biogeografía de Islas. Ejemplos de tales preguntas son los siguientes. ¿Cuál debe ser el diseño óptimo de estas reservas para que el mayor número de especies sobreviva? Esa muestra del ecosistema que representa la reserva ¿qué porcentaje contiene de las especies que caracterizan la integridad del ecosistema? ¿Cuál es el área óptima para la reserva? ¿Qué forma debe tener? ¿Cuán aislada debe quedar en relación con el parche de hábitat más cercano? ¿Cuáles son las capacidades de dispersión de las especies que habitan la isla? ¿Qué especies desaparecerán primero víctimas del aislamiento?

#### 4. LAS RELACIONES ENTRE LA ABUNDANCIA Y DISTRIBUCION DE ESPECIES

La forma en que las condiciones ambientales y los procesos poblacionales determinan la abundancia y distribución de especies es un problema central de la ecología y la biogeografía. Aunque ha sido ampliamente reconocido que la abundancia y la distribución están íntimamente relacionadas, la naturaleza de esta relación ha sido poco investigada sistemáticamente en las escalas espaciales que van desde poblaciones locales hasta rangos completos de distribución de las especies. A una escala local, las relaciones entre densidades poblacionales y distribuciones espaciales de individuos han sido estudiadas por muchos ecólogos especializados en poblaciones y comunidades. La distribución a gran escala geográfica usualmente ha sido vista como un campo especial de la biogeografía, cuyos practicantes a menudo tienen poca experiencia o interés en la ecología de poblaciones. Así, pocos investigadores han estudiado sistemáticamente la variación en la densidad poblacional a nivel de rango geográfico de distribución de las especies.

Brown (1984) ha desarrollado una teoría para explicar las relaciones entre la abundancia y la distribución de especies que incluye patrones generales dentro y entre las especies. Su teoría está basada en el concepto de nicho.

##### Variación espacial en abundancia dentro de especies

Típicamente, dentro de sus rangos geográficos, los individuos de una especie son relativamente numerosos en algunos hábitats o regiones, mientras que son escasos o ausentes en otros. El límite del rango geográfico se ubica donde la densidad de la población en grandes áreas declina a cero. ¿Existe algún patrón general de variación espacial en la abundancia dentro del área en la cual una especie ocurre normalmente? La respuesta parece ser sí: en general, la densidad es mayor cerca al centro del rango y declina gradualmente hacia los límites.

##### Correlación entre la abundancia y distribución entre las especies

Algunos estudios han demostrado una correlación positiva entre la abundancia local y la distribución geográfica entre especies estrechamente relacionadas y ecológicamente similares. Cuando plantas con la misma forma de vida o animales del mismo gremio son comparados, aquellas especies que tienen las densidades poblacionales locales más altas tienden a habitar una gran proporción de sitios dentro de una región y a tener los rangos geográficos más amplios; por el contrario, las especies que son siempre raras tienen distribuciones espaciales restringidas.

##### LA TEORÍA

Brown propone una sencilla explicación general para estos dos patrones: la disminución relativamente simétrica y monotónica de la abundancia desde el centro de la distribución hacia todos los límites y la correlación positiva entre la densidad de una población local y la extensión de la distribución espacial entre especies similares. Esta teoría está basada en tres principales suposiciones:

- I. Esta suposición se refiere a los requerimientos ecológicos de las especies. La abundancia y distribución de cada especie está determinada por las combinaciones de muchas variables físicas y bióticas que son necesarias para la supervivencia y reproducción de sus individuos. Las variaciones en la densidad de la población de una especie en el espacio se asume que refleja la distribución de la densidad de las combinaciones necesarias de variables ambientales.
- II. Esta suposición se refiere a los patrones de variación espacial en el ambiente, los cuales tienen componentes aleatorios y determinísticos. Algunos grupos de variables (factores) están distribuidos independientemente de cada uno y existe un grado significativo de variación local aparentemente al azar. La variación ambiental también está altamente correlacionada; tal es así que la probabilidad de que ocurran lugares con combinaciones similares de variables ambientales es una función inversa de la distancia entre ellos.
- III. Esta suposición se refiere a la extensión en la cual las especies varían en sus requerimientos. Se asume que las especies ecológicamente similares y estrechamente relacionadas, difieren substancialmente sólo en una dimensión, o a lo más en un pequeño número de dimensiones menores del nicho.

##### Variación espacial en la abundancia de una especie

De las primeras dos suposiciones se sigue que la densidad debe de ser mayor cerca al centro del rango de distribución de una especie y debe de disminuir hacia los límites. Para cada especie debe de haber un sitio favorable donde la densidad de la población debe ser mayor debido a la combinación de variables ambientales correspondientes a los requerimientos de la especie.

Si la variación espacial en el ambiente esta autocorrelacionada, entonces el ambiente sería progresivamente más diferente con el incremento de la distancia desde este sitio, los requerimientos de nicho de la especie se encontrarían menos frecuentemente, y la abundancia disminuiría. Existiría una disminución del número de sitios locales donde los individuos puedan vivir y, de la misma manera, dentro de estos parches la densidad de la población tendería a ser menor debido a que los recursos son escasos y/o las condiciones se acercan a los límites que puedan ser tolerados.

La forma exacta de la variación espacial de abundancia dependería del número y tipo de factores ambientales que abarca el nicho, y de los patrones espaciales de variación de estas variables.

## 5. FUENTES DE INFORMACIÓN SOBRE LA DISTRIBUCIÓN DE ESPECIES

### MUSEOS DE HISTORIA NATURAL

Las colecciones de historia natural en museos contienen datos de primera importancia para la toma de decisiones en la conservación de la biodiversidad. Colectivamente, estos datos, basados en especímenes, documentan las distribuciones conocidas de los taxa en el tiempo y el espacio, y son potencialmente accesibles como las fuentes integradas más confiables de información para la mayoría de especies descritas, como una solución de comparativamente bajo costo a un amplio rango de interrogantes de conservación e investigación (Ponder et al. 2001). Los museos de historia natural guardan colecciones científicas, que usualmente son su más importante patrimonio, que deben incrementar, preservar e investigar. Estas colecciones consisten en ejemplares de plantas y animales preservados, secos o en líquido, con la finalidad de mantenerlos perpetuamente. El valor de estos ejemplares se incrementa además por tener datos precisos del lugar y fecha de colección, sexo, medidas estándar, peso, fotos u otros datos. Actualmente también es posible obtener en el campo muestras de sangre o tejidos congelados para estudios biomoleculares o genéticos.

El cúmulo de estos datos permite resolver problemas científicos que van desde la sistemática básica a la más alta especulación evolutiva. Para esto, toda colección científica procura tener representado el mayor número de especies existentes en el área de interés y, dentro de cada especie, el material suficiente que permita entender su variación morfológica, sexual, y etárea, sus variaciones geográficas, su rol en algún hábitat ecológico, su origen evolutivo y sus relaciones de parentesco con otras especies, entre otros temas de interés.

Países megadiversos como el Perú, y con colecciones científicas aún pequeñas y recientes, con limitado personal científico y con crecientes problemas de conservación de hábitats y especies, deben priorizar el incremento sistemático de estas colecciones, en la mayor diversidad posible de taxa y ambientes (Pacheco 1990).

### BASES DE DATOS

Durante los últimos años, el uso de bases de datos de especímenes depositados en museos y herbarios se está incrementando como una herramienta para resolver problemas de conservación y manejo de la diversidad biológica. Estas bases de datos no solamente proporcionan acceso fácil a una gran cantidad de información de interés para taxónomos y sistemáticos, sino que también pueden ser utilizadas en los trabajos de ecólogos y biogeógrafos para propósitos de aplicación. Información tan variada como estimación de rangos geográficos de especies, cuantificación de variables del nicho fundamental de una especie, y modelos para predecir el número de especies como una función de los parámetros del clima, pueden ser obtenidos desde las bases de datos.

Básicamente, las preguntas que la información de una base de datos debería responder son: (1) ¿Qué especies se encuentran en una determinada localidad? y (2) ¿Cuál es la distribución geográfica de cada especie? (Soberón et al. 2000).

### LISTADOS DE ESPECIES

Las notas de campo y los listados de especies son algunas de las herramientas más tradicionales entre los biólogos. Estas observaciones proporcionan los primeros pasos para trabajos descriptivos que documentan los patrones de distribución y abundancia de las especies en el pasado. Las listas de comprobación son ampliamente utilizadas para catalogar observaciones de plantas y animales en el campo; éstas pueden proveer información confiable de cambios en las poblaciones, fenología, patrones de abundancia geográfica, entre otros, y pueden ser utilizadas para acciones de conservación y manejo (Droege et al. 1998).

### ENCUESTAS

Información referencial acerca de la ocurrencia o distribución de las especies también puede ser obtenida por encuestas o entrevistas a especialistas, aficionados o gente común. Esta información puede ayudarnos a tener alguna idea de las especies que habitan un determinado lugar, cuando no nos es posible trabajar directamente en dicho lugar. Sin embargo, la información proporcionada por gente común debe tomarse con cuidado y ser muy cautos en publicar información sólo referencial. Los especialistas de un determinado grupo taxonómico pueden proporcionar datos confiables, y se les puede incluir en una publicación citándolos como una comunicación personal; asimismo, los aficionados también pueden muchas veces darnos buena información sobre su campo de interés, particularmente los ornitólogos aficionados.

Los especialistas pueden realizar inventarios cuantitativos basados en metodologías científicas, y obtener resultados valiosos y confiables de la distribución y/o abundancia de las especies, mientras que los aficionados, al carecer del conocimiento de estas metodologías, podrían sólo informarnos acerca de la ocurrencia de las especies, por lo que su información resultaría insuficiente si queremos hacer estudios biogeográficos, ecológicos, u otros.



# Especies Indicadoras



## 1. INTRODUCCIÓN

En un estudio que trata sobre algún objeto difícil y complicado de medir, uno puede intentar medir, en lugar del objeto mismo, alguna otra cosa que está correlacionada con el objeto principal del estudio. Por ejemplo, en las áreas urbanas del hemisferio norte se ha utilizado la presencia de los líquenes epifíticos como especies indicadoras de la calidad del aire. Donde se encuentra a los árboles cubiertos por ciertas especies de líquenes, se puede suponer que el aire esté muy contaminado y viceversa. Naturalmente, este método para estimar la calidad del aire es menos preciso que un muestreo y análisis químico del aire mismo, pero como es mucho más rápido y barato inventariar líquenes, y resulta útil estudiarlos porque se puede cubrir áreas mucho más grandes en menos tiempo y con un menor costo.

Cuando hablamos de estudios de la flora o fauna, especialmente en áreas tropicales, encontramos un problema parecido al alto costo económico del muestreo y análisis de la contaminación aérea. Las áreas tropicales son tan ricas en especies de plantas y animales que inventariarlas, hasta dentro de áreas pequeñas, requiere mucho tiempo, sin hablar de regiones algo más grandes, como por ejemplo la Amazonía peruana. Además, el problema no es sólo la alta riqueza de especies, sino también el hecho de que el conocimiento taxonómico de la mayoría de los grupos de organismos es tan limitado que cualquier estudio queda fácilmente atascado por la demora de las identificaciones de las especies.

Es importante precisar para qué aspecto de la fauna y flora buscamos especies indicadoras. Candidatos de aspectos que podrían necesitar algún tipo de indicador son, por ejemplo, la riqueza de especies, la composición de especies, la productividad, y el grado de alteración humana.

La riqueza de especies es un tema que está muy de moda actualmente, y por eso hay mucho interés de encontrar maneras de estudiarla y documentarla sobre áreas grandes. Toda la riqueza de especies de plantas, animales y hongos no se puede documentar fácilmente, y por eso se ha pensado que tal vez fuera suficiente investigar la riqueza de especies sólo en uno o, en todo caso, en pocos grupos de organismos, y utilizar estas mediciones como aproximaciones para la riqueza de especies en otros grupos de organismos. Sin embargo, los resultados hasta ahora indican que hay muy poca esperanza de poder encontrar especies indicadoras de la riqueza de especies (Prendergast *et al.* 1993, Lawton *et al.* 1998, Tuomisto *et al.* 2002).

Productividad es un aspecto ligado a la fertilidad del suelo, y especies indicadoras de este aspecto se conocen tanto fuera (Cajander 1926) como dentro de los trópicos. En la Amazonía peruana los campesinos, por ejemplo, usan la palmera *Phytelephas macrocarpa* como una indicadora de suelos aptos para la agricultura (Kahn & Mejía 1990, Mejía 1992, Coomes 1995). Se conoce también especies indicadoras de la alteración humana (Noss 1990, Landers *et al.* 1998, Alonso 2000). Por ejemplo, la presencia de sachavaca (*Tapirus terrestris*) o especies de paujil (*Crax* spp.) en la zona de estudio indican una presión baja de caza.

En este trabajo nos concentramos a continuación en las especies indicadoras de patrones de distribución de especies. La idea es buscar grupos de especies cuya composición varía entre localidades, paralelamente con la variación general de la composición florística y/o faunística entre estas localidades. En otras palabras, uno puede, por ejemplo, clasificar los lugares de estudio (véase el Capítulo IV) según la composición de especies indicadoras, y si es que las indicadoras realmente sirven, esta clasificación debe dar básicamente el mismo resultado que la clasificación de los lugares según la composición de especies de la flora y/o fauna más general.

Debido a la constante pérdida de hábitats y biodiversidad en el mundo, existe una necesidad urgente de inventariar la composición de especies de plantas y animales para apoyar los planes de conservación. El objetivo principal de la conservación de la naturaleza es mantener poblaciones viables de todas las especies. Cada especie tiene sus propios requerimientos

ecológicos y su propia historia evolutiva y de dispersión, y por eso ninguna especie se encuentra en todas partes. Para poder asegurar la conservación de todas las especies, hay que conservar muestras de todos los tipos de hábitat que albergan especies especializadas. Para encontrar esos hábitats, prácticamente no hay otra alternativa que hacer muchos inventarios tratando de ubicar zonas que tienen una composición especial de especies. El uso de especies indicadoras de patrones de distribución de especies puede dar una solución para esta necesidad (Tuomisto & Ruokolainen 1998).

Naturalmente, toda la idea de especies indicadoras para revelar patrones florísticos o faunísticos generales depende de la organización de las comunidades de especies. En casos extremos se puede distinguir comunidades compuestas de forma determinística por el efecto de factores medioambientales (Brown 1984) o, tal vez, históricos, y allí se podrá encontrar especies indicadoras de este patrón determinístico. Esto es algo obvio cuando hablamos sobre comunidades o agrupamientos de especies claramente distintos. Por ejemplo, si alguien nos cuenta que ha visitado tres zonas del Perú, una con aguaje (*Mauritia flexuosa*), otra con vicuña y la tercera con lobo del mar, ya podemos deducir con buena certeza que la persona ha visitado la selva, sierra y costa, respectivamente, aunque sólo nos informe de la presencia de una especie indicadora por zona.

Cuando tratamos de comunidades no tan marcadamente diferentes, como por ejemplo bosques de tierra firme y bosques inundables en la selva, necesitamos ya información de varias especies presentes por hábitat, para poder decir dónde ha estado uno. Finalmente, cuando se trata de zonas que no tienen diferencias determinísticas tal vez sólo diferencias causadas por acontecimientos básicamente aleatorios en la dispersión de las especies dentro de un ambiente física y químicamente uniforme (Hubbell 2001) no será posible distinguir comunidades distintas, y por eso tampoco hay especies indicadores de patrones de distribución de especies.

Además del criterio básico de que un grupo de especies indicadoras tiene que reflejar los patrones generales de semejanza y diferencia florística y/o faunística entre los lugares de interés, hay otra serie de características que un grupo útil de especies indicadoras debe cumplir. Tuomisto y Ruokolainen (1998) proponen siete de tales características para las plantas indicadoras: (1) deben ser plantas de tamaño y acceso convenientes, que permitan su fácil observación y colección; (2) deben ser fáciles de reconocer y diferenciar de otras plantas en el bosque; (3) deben tener suficientes especies para poder presentar una amplia variación de adaptaciones ecológicas; (4) no deben contener una cantidad excesiva de especies para que ésta no dificulte el trabajo en todas sus fases; (5) deben ser relativamente bien conocidas taxonómicamente para que la identificación no sea complicada; (6) deben ser relativamente comunes en los bosques de diferentes tipos; y (7) no deben ser muy utilizadas por el hombre (esto podría alterar sus patrones de distribución natural). Estos criterios pueden ser aplicados también para otros organismos, y en efecto fueron utilizados, junto con opiniones personales de especialistas de varios grupos, para elegir a seis grupos taxonómicos como indicadores para los estudios de campo que fueran realizados por el proyecto BIODAMAZ.

## 2. SEIS GRUPOS INDICADORES

Bajo la hipótesis de que la distribución y abundancia de plantas está relacionada directamente con factores ambientales y formaciones geológicas, y la distribución y abundancia de animales a las formaciones vegetales y condiciones ambientales, el Proyecto BIODAMAZ seleccionó grupos taxonómicos como potenciales indicadores para validar en el campo esta hipótesis.

Trabajos previos con árboles, melastomatáceas y pteridofitas realizados en los alrededores de Iquitos, en la Amazonía peruana (Ruokolainen *et al.* 1997, Ruokolainen y Tuomisto, 1998) demostraron la existencia de una correlación entre los patrones de semejanzas y diferencias florísticas entre estos tres grupos de plantas. Este patrón común a los tres grupos de plantas se explica principalmente por factores ambientales y por la geología local. Entonces tanto las melastomatáceas como las pteridofitas cumplen la característica básica de un indicador, pero ambos grupos cumplen también las otras características necesarias, propuestas por Tuomisto y Ruokolainen (1998), para ser considerados indicadores.

Las palmeras tienen algunas características importantes para considerarlas como un grupo indicador de diferentes condiciones ambientales: como grupo son fáciles de reconocer, son relativamente bien conocidas taxonómicamente, son abundantes en diferentes tipos de bosque, y existen suficientes especies para reflejar diferencias en condiciones ambientales. Vormisto (2000) afirma que las palmeras reflejan diferencias de condiciones ambientales. Por ejemplo, *Phytelephas macrocarpa* es una palmera conocida como un indicador de suelos agrícolas (Coomes 1995). Existen también otras especies que podrían ser utilizadas como indicadoras de la riqueza o pobreza de los suelos. Las palmeras son muy importantes económicamente, debido a que todas sus partes pueden ser utilizadas: hojas, tronco, meristemo apical, frutos y raíces (Kahn & de Granville 1992). Ésta es una característica que puede disminuir su potencial como indicador, porque los patrones de distribución de especies económicamente importantes no son necesariamente muy naturales. Además, la taxonomía de palmeras está, por lo menos ahora, en una situación algo confusa, porque las autoridades más importantes en taxonomía de las palmeras tienen opiniones claramente divergentes en cuanto a la clasificación de las especies.

Los anfibios anuros aparentemente cumplen los siete criterios antes mencionados, y por eso pueden ser considerados como un grupo indicador potencial para los patrones de distribución de otros grupos de fauna o flora. Aspectos de los anfibios anuros que pueden disminuir su valor como indicadores de patrones de distribución de especies son el uso de algunas especies como animales ornamentales o mascotas (caza selectiva y la consiguiente desaparición de algunas especies de sus áreas naturales de distribución), y la estacionalidad de su actividad. La mayoría de los anfibios anuros están activos principalmente durante períodos específicos del año, y solamente bajo ciertas condiciones climáticas. Por esta razón, puede ser difícil registrar observaciones estrictamente comparables entre sí de la anurofauna en diferentes lugares, porque siempre hay que cuidar de que los inventarios se hayan hecho en condiciones climáticas semejantes. La otra alternativa sería inventariar los anuros durante un año entero en cada lugar, pero haciendo eso se pierde ya la idea original de hacer inventarios rápidos con las especies indicadoras. Muchas especies de anuros tienen un amplio rango de distribución, y potencialmente podrían servir como especies clave para evaluar cambios en el ambiente. Otras especies son especialistas de hábitats o tienen distribuciones restringidas, y podrían servir para evidenciar perturbaciones locales (Heyer *et al.* 1994).

Las aves cumplen relativamente bien con los criterios antes listados, aunque hay también algunos hechos que pueden complicar su uso como indicadores de patrones de semejanzas y diferencias faunísticas y florísticas. Algunas especies son muy perseguidas por el hombre (por su carne, por sus plumas o para mascotas), y por eso ya no reflejan patrones naturales de distribución. También varias especies son migrantes estacionales. Migran de un lugar al otro en diferentes épocas del año, y también la detectabilidad de muchas especies varía según la fase reproductiva. Cuando los machos reclaman sus territorios al inicio del ciclo reproductivo, la especie es mucho más fácil observar que cuando los individuos tratan de esconderse con sus pichones. Una ventaja especial con aves es que es un grupo excepcionalmente bien conocido taxonómicamente, y que también existe mucha gente interesada y hasta entusiasmada por las aves. Por eso es fácil encontrar gente capacitada para trabajar con aves. El interés público también hace que el financiamiento para realizar estudios de aves sea siempre más seguro que para muchos otros grupos de organismos.

El sexto y último grupo de organismos incluido en el programa de trabajos de campo del proyecto BIODAMAZ fue el de las hormigas. Éstas parecen también tener buena potencialidad para ser utilizadas como especies indicadoras de patrones de semejanzas y diferencias faunísticas y florísticas. Se las encuentra en abundancia en casi todas las hábitats terrestres del mundo, y son relativamente fáciles de coleccionar, por lo menos las hormigas terrestres (Majer 1983). Son particularmente apropiadas para programas de inventarios y monitoreo debido a que la mayoría de las especies son sedentarias, y construyen nidos perennes con áreas muy reducidas de forrajeo. Por eso en contraste con otros insectos que se mueven frecuentemente entre hábitats en busca de alimento, pareja o sitios para anidar las hormigas están generalmente presentes en un sitio y pueden ser confiablemente muestreadas y monitoreadas. También son importantes ecológicamente por las funciones que cumplen en diferentes niveles de un ecosistema como depredadores y presa, así como detritívoros, mutualistas y herbívoros (Alonso, 2000).

Estos seis grupos, pteridofitas, melastomataceas, palmeras, anfibios anuros, aves y hormigas fueron seleccionados como grupos para ser investigados durante el proyecto BIODAMAZ. Todos cumplieron la mayoría de requerimientos de un grupo indicador de patrones de distribución de especies. Sin embargo, sólo de los grupos de plantas hubo resultados previos (Ruokolainen et al. 1997, Ruokolainen & Tuomisto 1998, Vormisto et al. 2000) indicando que cumplen también el requerimiento fundamental que los patrones de semejanza y diferencia en composición de especies del grupo observados entre los lugares estudiados, se repiten esencialmente igual en la flora y fauna en general.







Introducción  
a la percepción  
remota



## 1. INTRODUCCIÓN

Los bosques amazónicos cubren un área muy grande, y a menudo el acceso al Bosque es difícil. Dadas estas condiciones, se hace necesario el uso de la percepción remota para estudios a gran escala. La ventaja de la percepción remota está en su potencialidad de cubrir áreas grandes con un esfuerzo relativamente pequeño. El concepto de percepción remota abarca varios métodos diferentes, y cada método tiene sus propias características que lo hacen apropiado para cierto tipo de uso, mientras que para algún otro tipo de uso no es necesariamente tan eficiente. Aquí presentamos sólo generalidades de percepción remota, poniendo énfasis en explicar las imágenes de satélite. Esta presentación está dirigida más que todo a biólogos y forestales que no tienen mucho conocimiento anterior sobre la percepción remota. Para lectores que quieren información más detallada y técnica hay mucha literatura especializada, por ejemplo Harris (1987) y Lillesand & Kiefer (1994).

Los métodos más utilizados en percepción remota son fotografías aéreas, imágenes de radar e imágenes de satélite. Fotografías aéreas son simplemente fotos tomadas de un avión normalmente en fajas, en las cuales las fotos individuales son parcialmente sobrepuestas para posibilitar una vista tridimensional del paisaje. En las fotos se ve el paisaje y sus detalles (árboles, ríos, edificios, cultivos, entre otros) con la nitidez que el equipo óptico y la altura de la toma de las fotos permite. Las fotos son con frecuencia relativamente fáciles de interpretar, porque captan características del paisaje básicamente de la misma forma que el ojo humano.

Las imágenes de radar se pueden sacar de un avión o de un satélite. Desde el avión o satélite se emite radiación de onda larga. Esta radiación es reflejada de la superficie de la tierra y capturada por un sensor instalado en el mismo avión o satélite. La radiación reflejada se registra en el sensor en una superficie dividida en celdas uno puede imaginarlas como celdas en una hoja de cálculo. Las celdas pueden corresponder a diferentes áreas de la superficie, y el área de la celda determina la resolución de la imagen. Una celda es comúnmente llamada un píxel. En cada celda o píxel hay un número que indica la intensidad de la radiación reflejada. Cuando uno da a cada número una tonalidad correspondiente desde negro (cero radiación) hasta blanco (recibida toda la radiación emitida), con diferentes tonalidades de gris con valores intermedios, se puede construir una imagen blanco y negro de la superficie de la tierra. La radiación utilizada penetra muy bien el agua atmosférica, y por eso las imágenes de radar no tienen problemas con nubosidad. Por otra parte, la radiación también penetra la vegetación, y por eso lo que se ve en las imágenes de radar es mayormente la variación topográfica del terreno. Por este motivo, normalmente las imágenes de radar tienen un valor limitado para estudiar la vegetación.

Claro que se puede emitir varios tipos de radiación, y en esta manera obtener también información de la vegetación. Una aplicación interesante de este sentido es el lidar (altímetro de laser), que emite luz laser. Cuando se emite luz laser con una apertura relativamente grande (unos 20 m de diámetro), se puede obtener de la calidad del reflejo una estimación muy buena de la altura del dosel, y también de la biomasa del bosque hasta en bosques pluviales tropicales (Drake et al. 2002).

Las imágenes de satélite son tomadas por instrumentos pasivos, como fotos aéreas, pero son digitales y no necesariamente registran el reflejo de la luz visible. El tamaño del píxel varía de un instrumento al otro, y muchos satélites registran en el mismo píxel simultáneamente varios rangos de longitud de onda o varias bandas de luz. La mayor ventaja para un biólogo de las imágenes de satélite está precisamente en la capacidad de registrar valores de varias bandas del mismo punto. La resolución es muchas veces de menor importancia, especialmente si uno quiere obtener una idea de la variación sobre áreas grandes. El sistema satelital más útil para la mayoría de biólogos, o por lo menos para los que quieren trabajar con la variación de los bosques amazónicos, es el sistema Landsat.

## 2. IMÁGENES LANDSAT

Las imágenes Landsat tienen dos resoluciones, el Landsat MSS (*Multispectral Scanner*, ya no activo) tiene un píxel de 80 x 80 m, y el Landsat TM (*Thematic Mapper*) de 30 x 30 m. Esto significa que la imagen no puede mostrar ningún detalle menor en el paisaje. El satélite Landsat circula alrededor de la tierra entre los polos registrando todo el tiempo una faja (path) de 180 km de ancho. Vuelve a registrar el mismo punto más o menos dos veces por mes. Las imágenes Landsat que uno puede comprar son normalmente cortadas de las fajas para tener una extensión de 180 x 180 km.

Para cada píxel, el sensor registra la intensidad de radiación recibida con un número entre el rango de 0 a 255. Se registra la radiación en cuatro (MSS) o siete (TM) diferentes rangos de longitud de onda, y estos rangos son numerados empezando de 1, que representa la longitud más corta. Cada rango forma así una capa, o banda, con valores de radiación registrados en los píxeles. En el MSS, dos bandas (1 y 2) registran luz visible (verde y rojo) y dos bandas (3 y 4) luz infrarroja. En el TM, tres bandas (1, 2 y 3) registran luz visible (azul, verde y rojo), y tres bandas (4, 5 y 7) son para luz infrarroja. La última banda (número 6) registra radiación termal con una longitud de onda mayor.

Se puede mirar la imagen de cualquiera de las bandas separadamente, en blanco y negro, asignando el tono para cada píxel según su valor digital. Pero también se puede hacer imágenes de color utilizando tres bandas diferentes simultáneamente. Por ejemplo, podemos usar las tres bandas infrarrojas de Landsat TM. Para cada banda asignamos subjetivamente uno de los colores rojo, verde o azul, por decir, banda 4 corresponde a rojo, 5 a verde y 7 a azul. Después podemos imaginar que colocamos tres lámparas en cada píxel de la imagen, una lámpara para cada banda. Ahora tenemos una superficie de píxeles, dentro de los cuales hay tres lámparas con tres colores. Para ver la imagen se prende las lámparas, cada una separadamente, con una tensión que corresponde al valor de radiación registrada en el píxel de la banda correspondiente.

Por ejemplo, si el sensor ha registrado en un píxel los valores 80, 40 y 20 en las bandas 4, 5 y 7, respectivamente, en este píxel la lámpara roja (asignada para la banda 4) tendrá una luminosidad doble en comparación con la lámpara verde (que representa la banda 5) y una luminosidad cuádruple en comparación con la lámpara azul (que representa la banda 7). El color del píxel será la mezcla de los colores que las lámparas estén iluminando y, consecuentemente, este píxel tendrá un color dominado por el rojo. Otro píxel, en cual el sensor haya registrado un valor muy alto, en la banda 7, tendrá un color más azulado. Así, siguiendo esta lógica, toda la superficie consistirá en píxeles que presentan colores diferentes según las radiaciones registradas en las bandas.

Vale recordar que los colores en una imagen de satélite no tienen ningún sentido absoluto. El usuario puede obtener diferentes colores para los mismos objetos en la imagen según los colores que se asigna a las bandas. Los patrones no cambian, pero sus coloraciones sí, y es una decisión subjetiva elegir la combinación de colores que parece más apropiada.

Típicamente, los valores de radiación en las bandas se concentran mucho alrededor de un rango relativamente restringido, especialmente cuando las características reflectantes de la superficie no son muy variables. Por ejemplo, en imágenes presentando los bosques amazónicos, todas las bandas tienen muy poca variación en sus valores de radiación. Esto resulta en una imagen que muestra a los bosques con una coloración bastante homogénea. Si uno quiere hacer más visible la variación existente, uno tiene que hacer un realce a los valores de radiación.

En un realce se redistribuye los valores originales de radiación más homogéneamente a lo largo de los valores posibles entre 0 y 255. O en otras palabras, el orden de los valores se mantiene, pero el distanciamiento se cambia. Hay varias maneras de hacer un realce, pero el realce más fuerte es la ecualización del histograma. En éste, la idea es simplemente redistribuir los valores originales de tal manera que el histograma que emerge de los valores tiene la forma que se aproxima lo más posible a una cuadrícula. Esto significa que los valores comunes resultarán más distanciados, mientras los valores raros resultarán menos distanciados que en el histograma original.

Para cubrir áreas grandes con imágenes de satélite, es necesario juntar varias imágenes a forma de un mosaico. Este proceso es en principio muy sencillo, pero lo que muchas veces resulta difícil es ajustar los valores de radiación en las imágenes, de tal manera que la misma superficie saldrá con los mismos valores de radiación en todas las imágenes. Es que las imágenes contiguas de diferentes fajas fueron tomadas en momentos diferentes, y por eso la atmósfera probablemente tenía diferentes características ópticas. Además, también las imágenes de la misma faja reflejan a menudo condiciones atmosféricas diferentes, porque con frecuencia las nubes tapan la vista y obligan a usar imágenes vecinas de distintas fechas.

Por ejemplo, uno puede tener dos imágenes con pantanos exactamente iguales de la palmera *Mauritia flexuosa*. Como son bosques iguales, uno debe de poder registrar los píxeles de estos bosques los aguajales con igual radiación en las dos imágenes. Sin embargo, puede ser que al momento de registrar la otra imagen, la atmósfera tuvo, por ejemplo, ceniza muy fina proveniente de las quemadas de chacras, y por eso los valores de radiación son diferentes en las dos imágenes para superficies iguales. Para armonizar los valores de radiación entre dos imágenes, uno se vale de las partes sobrepuestas de dos imágenes vecinas para calcular una función de corrección entre los valores de las dos imágenes.

Muchas veces uno trata de utilizar las imágenes de satélite para clasificar el paisaje en diferentes unidades, como por ejemplo áreas cultivadas, urbanas y boscosas, si es que se tiene por lo menos un conocimiento parcial de qué significan los diferentes colores que se ven en la imagen. También se puede simplemente ubicar diferentes parches según sus características de radiación, sin saber qué son estos parches en términos reales. Una clasificación de la imagen se puede hacer visualmente, simplemente delimitando áreas que subjetivamente parecen distintas. También se puede utilizar diferentes métodos numéricos para clasificar la imagen.

Las áreas cubiertas por bosques naturales presentan, lamentablemente, un problema especial para clasificaciones numéricas. Esto se debe a que el dosel de un bosque natural tiene una estructura bien variable. Los árboles son de diferentes tamaños y especies, y hay claros pequeños y grandes. Todo esto hace que los píxeles vecinos de un bosque pueden tener valores muy diferentes de radiación. Normalmente este tipo de variación muy fina no es interesante para un investigador que quiere obtener una idea de la variación del terreno sobre superficies muy grandes. El ojo humano es muy hábil para descartar esta variación menos interesante y concentrarse en los rasgos más generales y grandes. Sin embargo, las clasificaciones numéricas muchas veces se tropiezan con esta variación (Tuomisto *et al.* 1994). Además, en una clasificación numérica uno tiene que definir cuántas clases quiere obtener, o cuánta variación interna aceptará dentro de cada clase. De esta manera se introduce inevitablemente un cierto grado de subjetividad en la clasificación.

Debido a los problemas que hay en la clasificación de las imágenes de satélite, resulta muchas veces más útil usar una imagen no clasificada y solamente procesada con un realce. De una imagen así uno puede quedar relativamente seguro de que los patrones visibles realmente reflejan características existentes en el paisaje. Solamente hay que tener cuidado con el oscurecimiento gradual y artificial de oeste hacia este que se observa en las imágenes Landsat sobre los bosques amazónicos, y que ha sido descubierto sólo recientemente (T. Toivonen y R. Kalliola, comunicación personal), pero su origen no está todavía aclarado. Según una investigación reciente que abarcó decenas de imágenes de la Amazonía occidental (Toivonen *et al.*, manuscrito), parece que el origen del fenómeno es complejo. Varía según el tipo de cubierta de la tierra y puede guardar mucha de la información general de la imagen. Muy probablemente es debido a la reflectancia bidireccional de la atmósfera y de la tierra, la que generalmente no está considerada en el sensor Landsat TM. Además, las sombras pueden incrementar el fenómeno. A pesar de este problema, las imágenes de satélite son una herramienta importante para biólogos que trabajan en los bosques amazónicos, porque ofrecen actualmente la fuente más segura de información sobre la variación geográfica de las características del paisaje (Tuomisto 1998).





Los principios  
de muestreo,  
análisis de datos e  
interpretación de  
resultados en estudios  
de biogeografía y de  
ecología de  
comunidades





## 1. INTRODUCCIÓN A LOS ANÁLISIS DE PATRONES DE COMUNIDADES DE ESPECIES

Biogeógrafos y ecólogos que estudian comunidades naturales de especies tienen datos y observaciones que son, de muchas maneras, muy diferentes de los datos de biólogos que trabajan con experimentos en laboratorios o en condiciones controladas. El interés en la biogeografía y ecología de las comunidades está en el conjunto de las especies, no en una especie particular. Muchas veces uno quiere, por ejemplo, estudiar cambios en la composición de especies en relación al tiempo, la geografía o las condiciones medioambientales. Las observaciones a ser estudiadas forman una variable multidimensional. Uno estudia el conjunto de especies como un carácter, y aspira a entender su reacción a diferentes factores simultáneamente. Las observaciones se hacen en el campo, normalmente sin posibilidades de manipular experimentalmente las variables de interés. Consecuentemente, muchas variables incontroladas afectan a la variación de la composición de las especies, y además muchas veces uno no puede ni siquiera medir o estimar todas estas variables. Ese tipo de datos raras veces puede ser analizado con los métodos tradicionales de la estadística. La intención de esta presentación es explicar por qué la estadística tradicional puede ser inadecuada. Además, tratamos de presentar los métodos más comunes para analizar la variación en la composición de las especies.

Esta presentación no intenta ser una guía exhaustiva de los métodos para estudiar comunidades naturales. La intención es que el lector pueda entender la esencia de cada método, para poder estimar su aplicabilidad en analizar sus propios datos, y para poder interpretar los resultados obtenidos con el método seleccionado. Tratamientos más profundos de métodos multivariados pueden ser encontrados en varios libros especializados, en los cuales uno puede hallar los fundamentos matemáticos para los métodos presentados aquí. A aquéllos que quieran profundizar sus conocimientos podemos recomendarles, en especial, los siguientes libros:

Jongman R.H.G., ter Braak, C.J.F. & van Tongeren, O.F.R. 1995. Data analysis in community and landscape ecology. Cambridge University Press, Cambridge, Gran Bretaña.

Legendre P. & Legendre L. 1998. Numerical ecology. Second English Edition. Elsevier, Ámsterdam, Países Bajos.

La fuente más importante consultada para esta presentación ha sido Legendre & Legendre (1998), la cual es una obra muy extensa de ecología numérica. Queremos también expresar nuestra gratitud a Pierre Legendre, quien amablemente nos prestó algunas figuras en formato digital para que sean más fáciles de usar en este documento. También agradecemos a Daniel Borcard por los materiales que nos proporcionó para esta guía

## 2. ¿QUÉ SON DATOS DE BIOGEOGRAFÍA Y ECOLOGÍA DE COMUNIDADES?

### Preguntas básicas en los estudios de patrones de comunidades

Un biogeógrafo o ecólogo que estudie los patrones de semejanza y diferencia entre comunidades naturales, típicamente quiere saber cómo cambian la flora o la fauna de un lugar y momento a otro. Este cambio en sí puede ser lo suficientemente interesante para justificar el estudio, pero muchas veces se trata de relacionar este cambio con alguna característica medida o estimada del medio ambiente.

Ejemplos de preguntas de interés son las siguientes: ¿Cómo cambia la avifauna desde el sur de la Amazonía peruana hacia el norte? ¿Existen diferentes comunidades de hormigas en purmas de edades diferentes? ¿Ha cambiado la fauna de peces en el río Nanay durante el siglo XX? ¿Afecta la fertilidad del suelo a la composición de especies de plantas en los bosques de tierra firme de Loreto?

Por razones de simplicidad vamos a hablar en esta presentación mayormente de parcelas en las cuales se ha estudiado la flora. Pero nuestra unidad de muestreo (parcela) o nuestro grupo de organismos (plantas) puede variar: uno puede cambiar a transectos para el conteo de aves, a trampas de colecta de insectos o cualquier otra combinación, sin afectar la generalidad de los métodos. Otra convención que usamos a continuación es que hablamos de muestreos que abarcan diferentes puntos geográficos. Pero igualmente uno puede pensar en el muestreo realizado en un punto geográfico distribuido en diferentes puntos en el tiempo. Los métodos para analizar una serie temporal de muestras son iguales que en el análisis de una serie espacial de muestras.

### ¿Qué tipo de datos se obtiene?

Los datos de comunidades naturales se componen normalmente de tablas o matrices, en las cuales se presentan las parcelas como filas y las características que describen las parcelas como columnas o viceversa. Comúnmente, una matriz contiene información de las especies observadas en las unidades de muestreo (Tabla 1), y otra matriz presenta los caracteres medioambientales (Tabla 2) que se utiliza para explicar la variación en la primera matriz. La tercera matriz puede indicar, por ejemplo, la ubicación de las unidades de muestreo con coordenadas geográficas (Tabla 3).

**Tabla 1.** Un ejemplo no real de datos de campo de un estudio de ecología de comunidades, en el cual se ha estudiado la abundancia de cinco especies de plantas en seis parcelas.

Parcela / Especie	Huasai	Ungurahui	Pona	Irapay	Chambira
Parcela 1	70	0	0	30	0
Parcela 2	5	40	0	55	0
Parcela 3	0	0	40	50	0
Parcela 4	0	50	40	10	0
Parcela 5	0	80	10	0	10
Parcela 6	0	80	0	0	20

Elaborada por el equipo del proyecto

**Tabla 2.** Un ejemplo no real de datos de campo de un estudio de ecología de comunidades, que describen algunas características edáficas de seis parcelas.

Parcela/ Características edáficas	pH	Suelo	Nitrógeno (mg / kg)	Humedad (1 seco - 5 húmedo)
Parcela 1	3,5	arena	0,7	5
Parcela 2	4,0	arena	1,2	4
Parcela 3	5,3	limo	2,7	1
Parcela 4	4,2	limo	5,6	2
Parcela 5	5,8	arcilla	8,2	3
Parcela 6	6,2	arcilla	15,9	3

Elaborada por el equipo del proyecto

**Tabla 3.** Un ejemplo no real de datos de campo de un estudio de ecología de comunidades con las coordenadas geográficas de las parcelas de estudio.

Parcela	Latitud	Longitud
Parcela 1	3°21' S	76°38' O
Parcela 2	3°26' S	76°44' O
Parcela 3	3°44' S	76°27' O
Parcela 4	3°47' S	76°23' O
Parcela 5	3°35' S	76°11' O
Parcela 6	3°36' S	76°42' O

Elaborada por el equipo del proyecto

De cada especie se registra su abundancia en cada una de las unidades de muestreo. La medida de abundancia puede variar según las características de la comunidad y el aspecto que se estudia. La manera más simple es anotar la presencia (valor 1) y ausencia (valor 0). Una variable con sólo dos valores posibles se llama variable binaria. Variables cuantitativas pueden tener muchos valores. Las variables cuantitativas son, por ejemplo, número de individuos, biomasa y cobertura. Uno puede medir estas variables con una medida continua u ordinal. Por ejemplo, uno puede estimar la cobertura de plantas en porcentajes o en clases de cobertura como:

0=	0%
1=	1% -12,5%;
2=	12,5% -25%;
3=	25% -50%;
4=	50% -100%;

Igualmente, uno puede medir los factores medioambientales con variables binarias, continuas u ordinales.

Todas las variables tienen que estar expresadas numéricamente antes de hacer los análisis. Hay que tomar en cuenta que una variable categórica siempre tiene que expresarse como una serie de variables binarias. Por ejemplo, una variable categórica del sustrato “suelo mineral/ hojarasca/ palo podrido/ roca dura” no puede transformarse a una variable con valores ordinales desde 1 hasta 4. Esto es imposible, ya que estas características no tienen ningún orden lógico de tamaño, humedad u otro aspecto ecológicamente relevante. Cada tipo de sustrato necesita ser considerado como una variable binaria, que obtenga el valor 1 cuando la parcela tenga este sustrato y 0 cuando no lo tenga.

### ¿Cómo realizar el muestreo?

Como en cualquier ejercicio científico, incluyendo el muestreo, también en biogeografía y ecología de comunidades es importante planificar y ejecutar el muestreo puntualmente. Los métodos tienen que ser definidos y descritos tan detalladamente que cualquier persona pueda repetir el estudio para ver si llega a los mismos resultados. Consecuentemente, es importante que se realice el muestreo de una manera objetiva. Así, no es buena idea tratar de observar una cantidad “suficiente” de individuos de plantas en sus lugares “típicos” de crecimiento, porque esto muy probablemente dirige el muestreo para reforzar las ideas personales que el investigador tuvo ya antes del muestreo. Cuando se planifica el muestreo, uno tiene que responder a las preguntas siguientes:

1. ¿Cuál será el tamaño de la parcela (o unidad de muestreo)?
2. ¿Cuántas parcelas se levantarán?
3. ¿De qué manera estarán espaciadas las parcelas una de otra?

El tamaño de las parcelas depende de las preguntas estudiadas y, por ejemplo, del tamaño físico y la densidad poblacional de las especies en mención. Cuanto más grandes sean las parcelas, más representativa será la muestra de las especies locales. Por otro lado, dentro de una parcela grande siempre hay más variación ambiental, y por eso una parcela pequeña sirve mejor para estudiar la relación entre la distribución de especies y los factores ambientales.

La primera y la segunda pregunta están relacionadas de tal manera que parcelas grandes requieren siempre más tiempo y trabajo para ser levantadas, y por eso uno no puede establecer varias parcelas grandes. Claramente, uno está siempre forzado a hacer compromisos en el diseño del muestreo. Muchas veces una revisión de estudios anteriores ayuda a definir la manera adecuada del muestreo de los organismos.

Para hacer un muestreo objetivo, el investigador tiene que definir primero su área general de estudio el área o las áreas de la(s) cual(es) uno quiere obtener una buena estimación de su flora o fauna. Esta definición se hace según la pregunta principal del estudio. Por ejemplo, uno puede tener el objetivo de comparar áreas inundables con las de tierra firme. Otro ejemplo podría ser el objetivo de estudiar manchas de diferentes colores visibles en las imágenes de satélite. En ambos ejemplos, uno primero delinearía en un mapa las áreas donde se encuentran los tipos de ambiente que son de interés en el estudio. Después se hará muestreos en cada uno de los tipos de ambiente (o áreas generales de interés), para poder comparar las áreas según la información obtenida en el muestreo. El muestreo dentro del área de interés puede ser aleatorio o sistemático.

La literatura más antigua recomienda normalmente un muestreo aleatorio, porque se ha considerado que ésta es la manera más apropiada para analizar estadísticamente los resultados. Por otra parte, la realización de un muestreo aleatorio es relativamente laboriosa, porque uno tiene que ubicar cada parcela exactamente según las coordenadas obtenidas en un proceso aleatorio (por ejemplo, de una tabla de números aleatorios o de un generador de números aleatorios).

En un muestreo aleatorio, el número de parcelas tiene que ser relativamente grande, porque con una cantidad pequeña sucede que algunos sectores del área de interés quedan sin parcela. También es posible que uno tenga que rechazar varias parcelas, porque atraviesan carreteras o ríos en un estudio que pretende muestrear vegetación natural terrestre. Estos problemas pueden ser evitados si existe un mapa de la zona que permita definir sectores que definitivamente no son de interés.

La ventaja de un muestreo sistemático es que es mucho más fácil realizarlo en términos prácticos. Cada parcela está situada sistemáticamente en relación con otras, y por eso es mucho más fácil ubicar los lugares de muestreo. Uno tiene que definir a priori si las parcelas van a estar pegadas una con otra, o si van a estar separadas de alguna manera sistemática. También hay que decidir si se realiza el muestreo de una forma unidimensional (línea o serie de tiempo), bidimensional (una red o parrilla de parcelas), o tridimensional (por ejemplo en un estudio de plancton).

En años recientes han sido desarrollados varios métodos de análisis que funcionan mejor con datos colectados sistemáticamente, en comparación a datos muestreados al azar. Por eso ya no es tan obvio qué estrategia de muestreo, sistemático o aleatorio, es mejor. En consecuencia, es recomendable conocer los métodos disponibles antes del muestreo, para saber qué método responde mejor a nuestras preguntas y cuáles son los requerimientos que cada método tiene en relación con los datos a analizar.

### 3. AUTOCORRELACIÓN ESPACIAL Y SUS CONSECUENCIAS

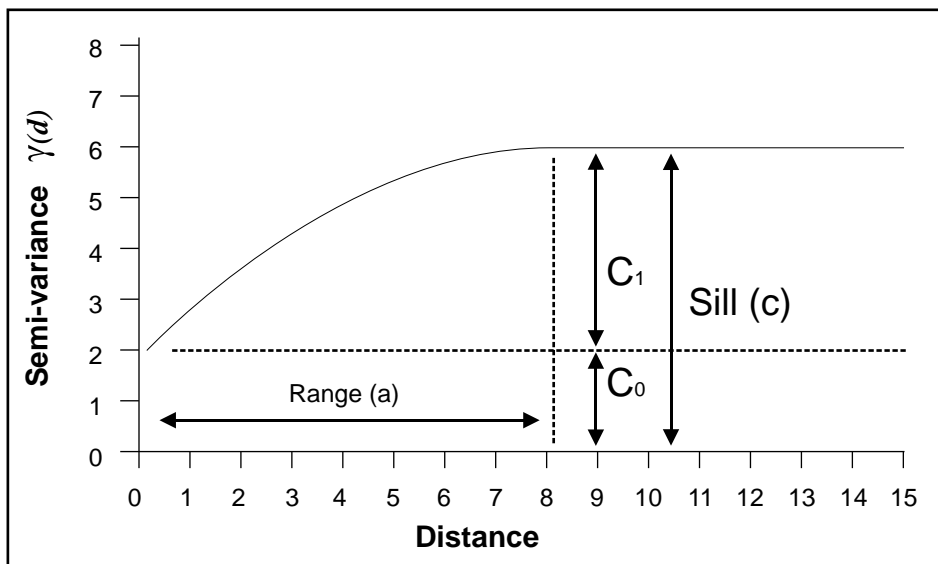
En la naturaleza se puede observar frecuentemente que cuanto más cerca están dos parcelas entre sí, es más probable que éstas sean semejantes. Así que conociendo la composición de especies o cualidades medioambientales de un lugar, podemos predecir o estimar más o menos estas características en otro lugar cercano. Este fenómeno se llama autocorrelación espacial.

La autocorrelación espacial nace de la distribución no aleatoria en la naturaleza de los factores ambientales y organismos. La mayor parte de las semillas de plantas caen muy cerca a la planta madre, y por eso dos parcelas geográficamente cercanas comparten normalmente más especies que dos parcelas lejanas. El flujo de agua de lluvia homogeniza diferencias entre puntos vecinos de suelo. Los procesos geológicos afectan normalmente áreas grandes, formando a la vez parches con características internamente homogéneas, como terrazas fluviales y manchas de arena blanca. Los vientos hacen que la temperatura, humedad y otras características climatológicas entre áreas cercanas sean parecidas.

La ocurrencia de la autocorrelación espacial en los datos puede ser verificada con un semivariograma (llamado como variograma también; Figura 2). En el semivariograma, el eje x indica clases de distancia geográfica entre las parcelas. El eje y es semivarianza, que es calculada para cada clase de distancia d:

$$\gamma(d) = \frac{1}{2W} \sum_{h=1}^n \sum_{i=1}^n w_{hi} (y_h - y_i)^2 \quad \text{en cual } h \neq i.$$

En la fórmula,  $y_h$  e  $y_i$  son valores de la variable  $y$  en parcelas  $h$  e  $i$ ,  $W$  es la cantidad de pares de parcelas con distancia  $d$  entre sí, y el factor  $w_{hi}$  es 1 cuando  $h$  e  $i$  tienen una distancia  $d$  entre sí, y 0 en todos los otros casos. En otras palabras, la semivarianza mide cuánto varía el valor de la variable medida en parcelas con una cierta distancia geográfica entre sí. En un caso como el presentado en la Figura 2, la autocorrelación es observable hasta la clase de distancia 8; después de eso, la semivarianza ya no aumenta, pero se estabiliza en un nivel que es el mismo que la varianza de la variable  $y$  (*Sill*). Se conoce como *Nugget effect* ( $C_0$  en la Figura 2) a la varianza en distancias que es menor a la resolución de los datos. A ese efecto están incluidos, por ejemplo, errores de medición.  $C_1$  corresponde a la variación que no está estructurada espacialmente.



**Figura 2.** Autocorrelación espacial expresada en un semivariograma. Después de la clase 8 de distancia ya no se puede observar autocorrelación. **Fuente:** Legendre & Legendre 1998, p. 729.

En caso de que se observe autocorrelación en los datos, uno tiene que tomar esto en cuenta para los análisis. La autocorrelación en sí crea una organización espacial en los datos que dificulta el análisis de las relaciones entre las especies y factores medioambientales.

Cuando tanto la variable para explicar (las especies) como la variable explicativa (el medio ambiente) están autocorrelacionadas espacialmente, cualquier dependencia entre las variables puede ser creada por la autocorrelación o por una interacción verdadera entre las variables. Esto significa que se acepta fácilmente una correlación estadísticamente significativa entre especies y medio ambiente. La autocorrelación también hará más difícil interpretar diferencias (como una diferencia en el promedio de alguna característica medioambiental) entre zonas de muestreo. Si las unidades de muestreo son independientes (sin autocorrelación), cada unidad nueva llevará una cantidad igual de nueva información para usarla como evidencia positiva o negativa de la diferencia entre las zonas. Consecuentemente, uno podría usar en el análisis  $n-1$  grados de libertad, siendo  $n$  el número de unidades de muestreo. Pero en una situación con autocorrelación entre las unidades de muestreo, las unidades no son independientes, y por eso la información nueva de cada nueva unidad de muestreo tampoco vale un grado de libertad. En lugar de tener  $n-1$  grados de libertad, uno tendrá al final tal vez  $0,53(n-1)$  grados de libertad. Entonces uno interpretaría fácilmente una diferencia como estadísticamente significativa.

Con datos autocorrelacionados se habla a menudo de pseudo replicación. Esta palabra quiere decir que el investigador fácilmente piensa que sus datos tienen más replicaciones (unidades de muestreo u observaciones) de las que existen en realidad. Las replicaciones autocorrelacionadas no tienen la misma fuerza de evidencia que las replicaciones independientes, y por eso verdaderas.

En términos prácticos, es imposible estimar si una unidad de muestreo lleva consigo medio grado de libertad, o tal vez alguna otra fracción de un grado entero. Por eso, si uno tiene datos autocorrelacionados, tiene que ser muy cuidadoso con los análisis estadísticos que estiman la significancia estadística de algún parámetro. Normalmente no será posible estimar la significancia estadística con la estadística tradicional, porque no se conoce el número de grados de libertad. Una manera de afrontar este problema es usar aleatorizaciones para estimar la significancia estadística; un ejemplo de esto es la prueba de Mantel que se explica más adelante.

#### 4. MEDIR DIFERENCIAS Y SEMEJANZAS

Muchos análisis en biogeografía y ecología de comunidades están basados en medidas de semejanzas y diferencias entre las unidades de muestreo con la aplicación de algún índice. También es común interpretar los valores del índice en relación con algunas características medioambientales que se ha medido o estimado en las unidades de muestreo. Entonces, para entender este tipo de análisis, vale la pena saber cómo se miden los índices de las semejanzas y diferencias, y qué suposiciones tienen cada uno de ellos.

Para medir diferencias en la composición de especies han sido desarrollados varios índices, que pueden indicar similitud o distancia entre las unidades de muestreo. Similitud y distancia son solamente las dos caras de una sola cosa. Si dos unidades de muestreo tienen casi la misma composición de especies, la distancia entre ellas es muy pequeña pero la similitud muy grande. De igual manera, dos unidades de muestreo con especies casi completamente diferentes tienen una distancia muy grande entre ellas, pero una similitud muy baja. Cuando uno entiende esta correspondencia, es relativamente intrascendente si uno habla de índices de distancia o de similitud (Figura 3).

Normalmente, un índice de similitud varía entre 0 y 1, 1 cuando la similitud es completa o total, y 0 cuando la diferencia es total.

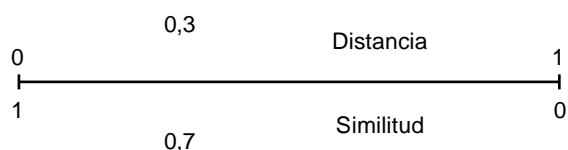


Figura 3. La distancia y similitud son efectivamente la misma cosa vista desde dos direcciones opuestas.

Cualquier índice de similitud ( $S$ , *similitud*) es convertible a un índice de distancia ( $D$ , *distancia*) a través de una de estas fórmulas:

$$D = 1 - S \quad D = \sqrt{1 - S} \quad D = \sqrt{1 - S^2}$$

Los índices de distancia a veces pueden crecer sin límite, pero también estos pueden convertirse después de una normalización a similitudes entre 0 y 1. La normalización se hace de las siguientes maneras:

$$D_{norm} = \frac{D}{D_{max}} \quad D_{norm} = \frac{D - D_{min}}{D_{max} - D_{min}}$$

Las distancias normalizadas se pueden convertir a similitudes con las fórmulas dadas arriba resolviendo el S.

Existen docenas de índices de distancia y similitud. Aparentemente, una razón que ha contribuido a esta gama es que no existe ninguna regla universal que nos guíe para definir qué índice es el mejor. Uno tiene que definir la aplicabilidad del índice según las preguntas, y también según los datos que uno tiene. A continuación vamos a ofrecer una introducción a algunos índices comúnmente usados. Se puede encontrar un tratamiento más profundo en Legendre & Legendre (1998).

### Medidas simétricas y asimétricas de distancia

Para un biogeógrafo o ecólogo, la característica más importante de un índice de distancia es probablemente su simetría o asimetría. Un índice simétrico usa todos los valores de la tabla de observaciones de una manera idéntica, mientras que un índice asimétrico trata el valor 0 de una manera diferente a todos los demás valores. Los índices simétricos son normalmente aplicables para datos medioambientales en los cuales el cero es solamente un valor entre los otros (por ejemplo, la temperatura 0°C).

Para datos de especies, sin embargo, es mucho mejor utilizar índices asimétricos, porque un valor 0 de observación no es cualquier porción de información de abundancia, sino que significa que no se ha observado la especie. Se puede justificar el tratamiento particular con las respuestas que las especies tienen a gradientes de factores medioambientales. Cada especie tiene una tolerancia respecto a los factores medioambientales. En ambientes fuera del rango de tolerancia, la especie ya no puede sobrevivir. En caso de que se haya observado la especie en dos lugares, podemos deducir que los lugares son de cierta manera semejantes. En ambos lugares todos los factores medioambientales tienen que estar dentro de los límites de tolerancia de la especie. Pero en el caso de que la especie no esté en ninguno de los dos lugares, tenemos varias opciones para interpretar la situación. Ambos lugares pueden ser, por ejemplo, demasiado fríos o demasiado calientes, o un lugar es demasiado caliente mientras el otro es demasiado frío, o entonces un lugar es demasiado frío y el otro demasiado húmedo. En otras palabras, la ausencia de observación de una especie en dos lugares no nos puede decir mucho sobre la similitud de los lugares. Consecuentemente, lo mejor es no tomar en cuenta las dobles ausencias de especies cuando uno calcula las similitudes entre lugares.

Más adelante explicamos más sobre análisis basados en índices simétricos y distancia euclidiana.

### Índices binarios de similitud

La medición de similitud puede basarse en datos binarios o cuantitativos. Los índices binarios son relativamente simples para calcular y entender, y por eso los primeros índices binarios fueron desarrollados hace unos cien años. En principio, en todos estos índices se calcula cuántas de las variables consiguen valores iguales y cuántas valores diferentes (Tabla 4), pero existen unas diferencias sutiles entre los índices. Todos los índices de similitud tratados aquí varían entre 0 (lugares completamente diferentes) y 1 (lugares idénticos).

El índice más sencillo calcula cuál es la proporción de variables con valores iguales de todas las variables en los dos lugares de comparación. Este índice se llama en inglés *simple matching coefficient*, y se calcula como:

$$S_1 = \frac{a + d}{a + b + c + d} \quad , \text{ para los datos en la Tabla 4} \quad S_1 = \frac{2 + 2}{2 + 1 + 0 + 2} = 0,8$$

Este índice es simétrico, porque el d está incluido en su fórmula. El índice está principalmente sugerido para datos medioambientales binarios.

**Tabla 4.** La significancia de los símbolos usados en las fórmulas de índices binarios de similitud. La cantidad total de variables medidas (por ejemplo especies) es  $p=a+b+c+d$ ;  $a$  es la cantidad de variables que obtuvieron en ambas parcelas el valor 1,  $b$  es la cantidad de variables que obtuvieron el valor 0 en la primera parcela y el valor 1 en la segunda parcela, etc. Abajo se presenta un ejemplo de datos binarios utilizados para calcular los valores ejemplares de cada índice.

		Parcela 1	
		1	0
Parcela 2	1	a	b
	0	c	d

	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5
Parcela 1	1	0	0	1	0
Parcela 2	1	1	0	1	0
	a	b	d	a	d

Cuando se excluye  $d$  de la fórmula, se consigue el índice de similitud de Jaccard, que es muy utilizado en diferentes estudios:

$$S_7 = \frac{a}{a + b + c} \quad , \text{ para los datos en la Tabla 4} \quad S_7 = \frac{2}{2 + 1 + 0} = 0,67$$

El índice de Jaccard es asimétrico, y por eso es bien aplicable a datos de especies. A través de una multiplicación por 100 uno consigue el porcentaje de especies compartidas de todas las especies observadas en las dos parcelas.

Otro índice binario que es apto para datos de especies es el índice de Sørensen. Este índice es bien parecido al Jaccard, pero da mayor peso a especies compartidas entre parcelas:

$$S_8 = \frac{2a}{2a + b + c} \quad , \text{ para los datos de la Tabla 4} \quad S_8 = \frac{2 \times 2}{2 \times 2 + 1 + 0} = 0,8$$

#### Índices cuantitativos de similitud e índices de distancia

Del índice de Sørensen existe también una variante, que toma en cuenta información cuantitativa. Este índice ha sido modificado varias veces, pero probablemente el nombre más comúnmente usado es el índice de Steinhaus, y se calcula como:

$$S_{17} = \frac{2W}{A + B} \quad , \text{ para los datos en la Tabla 5} \quad S_{17} = \frac{2(1 + 5 + 8)}{14 + 28} = 0,67$$

En la fórmula,  $A$  es la suma de todas las abundancias en una parcela y  $B$  la suma de todas las abundancias en la otra parcela.  $W$  es la suma de abundancias de todas las especies compartidas, pero calculado de tal manera que se ha incluido sólo la abundancia de la especie menos abundante. En la Tabla 5, la similitud entre las parcelas 1 y 3 (también 2 y 3) es 0, porque estas parcelas no comparten ni una especie.



**Tabla 5.** Ejemplo de datos cuantitativos usado para calcular valores de índices cuantitativos de similitud.

	Especie 1	Especie 2	Especie 3	Especie 4	Especie 5	Especie 6
Parcela 1	1	0	5	0	8	0
Parcela 2	0	0	10	0	16	0
Parcela 3	2	1	0	2	0	1

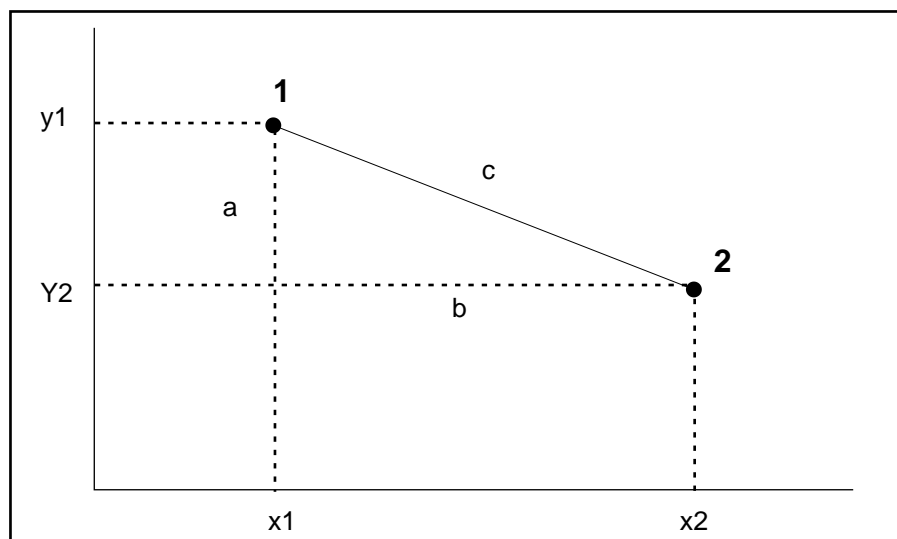
La versión de distancias del índice de Steinhaus ha sido inventada varias veces, y consecuentemente existen varios nombres que refieren al mismo índice. Los nombres más comunes de este índice de distancia son índice de Odum e índice de Bray-Curtis. El índice se puede calcular entre parcelas  $x_1$  e  $x_2$  de la manera siguiente:

$$D_{14}(x_1, x_2) = 1 - \frac{2W}{A + B} = \frac{\sum_{j=1}^p |y_{1j} - y_{2j}|}{\sum_{j=1}^p (y_{1j} + y_{2j})}$$

En la fórmula,  $p$  es el número total de especies,  $y_{1j}$  es la abundancia de especie  $j$  en la primera parcela, e  $y_{2j}$  es la abundancia de especie  $j$  en la segunda parcela.

Entre parcelas 1 y 2 de la tabla 5  $D_{14}(x_1, x_2) = \frac{|1 - 0| + |5 - 10| + |8 - 16|}{1 + 2 + 5 + 10 + 8 + 16} = 0,33$

El índice cuantitativo de distancia más sencillo es la distancia euclidiana. Esta es la distancia que uno puede medir, por ejemplo, con una regla entre dos puntos dibujados en un papel. En el caso que se haya dibujado los puntos dentro de un sistema de coordenadas  $x$  e  $y$ , uno puede calcular la distancia entre los puntos aplicando la fórmula de Pitágoras (Figura 4).



**Figura 4.** Con base en las coordenadas se puede calcular  $a=|y_1-y_2|$  y  $b=|x_1-x_2|$ . La distancia euclidiana  $c$  entre los puntos 1 y 2 es según la fórmula de Pitágoras:  $c = \sqrt{a^2 + b^2}$

La distancia euclidiana entre dos lugares se puede calcular también cuando existen más de dos caracteres que describen los lugares:

$$D_1(x_1, x_2) = \sqrt{\sum_{j=1}^p (y_{1j} - y_{2j})^2}$$

Entre parcelas 1 y 2 de la tabla 5  $D_1(x_1, x_2) = \sqrt{(1-2)^2 + (5-10)^2 + (8-16)^2} = 9.5$

La distancia euclidiana es muy útil cuando uno quiere definir la distancia entre lugares con base en valores de factores medioambientales. El mínimo de la distancia euclidiana es 0, y naturalmente no tiene un valor máximo.

Hay que tener presente que la distancia euclidiana da un peso grande a las variables que varían mucho, porque se eleva al cuadrado las diferencias de las variables antes de la sumatoria. Se puede evitar esta distorsión con transformaciones que son explicadas en la sección siguiente.

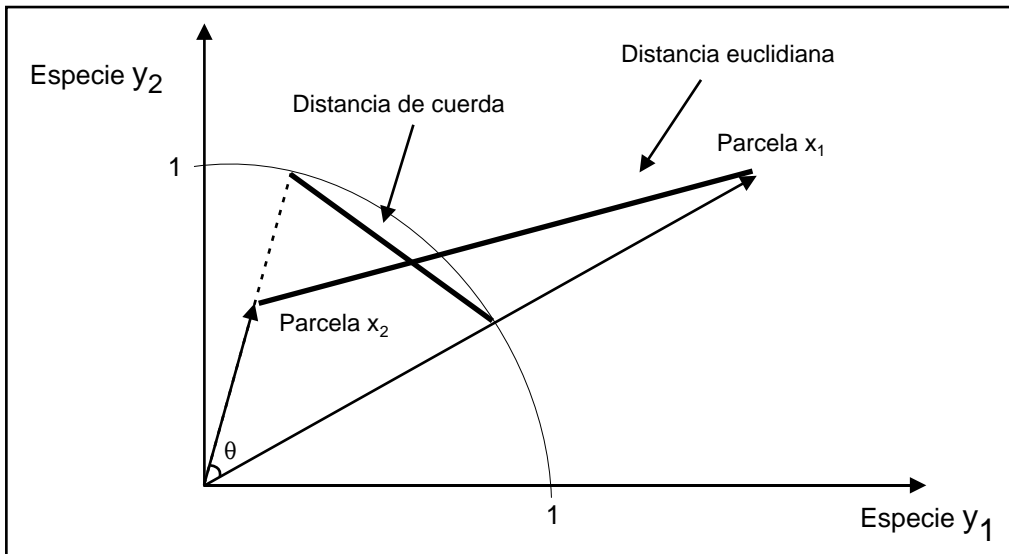
La distancia euclidiana es un índice simétrico de distancia, y por eso no puede ser recomendado para datos de especies. Cuando prácticamente todas las especies están presentes en todas las parcelas (pocos ceros en los datos), se puede considerar el uso de la distancia euclidiana. Esta condición se cumple a veces cuando se estudia gradientes muy cortos de factores medioambientales. En estas condiciones, las diferencias entre los lugares se expresan más con cambios en las abundancias de las especies, y no tanto con cambios en la composición de especies.

El problema de los dobles ceros incluye la siguiente paradoja: cuando uno calcula la distancia euclidiana entre las parcelas 1 y 3 de la Tabla 5, el resultado es 9,8. Éste es sólo un poco mayor que la distancia entre las parcelas 1 y 2, aunque todas las especies en las parcelas 1 y 2 están compartidas entre ellas, y no haya ninguna especie compartida entre las parcelas 1 y 3. La distancia euclidiana entre las parcelas 2 y 3 es 19,1, que es mucho mayor que la distancia entre las parcelas 1 y 3. Tampoco este resultado es ecológicamente coherente, porque si la parcela 3 no comparte especies ni con la parcela 1 ni con la 2, la expectativa real será que ambas parcelas tengan la misma distancia a la parcela 3. Este resultado incoherente tiene su origen en la característica de la fórmula de distancia euclidiana, que interpreta que una distancia de 9 a 10 individuos es igual a la distancia de 0 a 1 individuo.

El problema de los dobles ceros en la distancia euclidiana está solucionado en la distancia de cuerda (*chord distance*):

$$D_3(x_1, x_2) = \sqrt{2 \left( 1 - \frac{\sum_{j=1}^p y_{1j} y_{2j}}{\sqrt{\sum_{j=1}^p y_{1j}^2 \sum_{j=1}^p y_{2j}^2}} \right)} = \sqrt{2(1 - \cos \theta)}$$

La explicación intuitiva de la distancia de cuerda está presentada en la Figura 5. La diferencia significativa con la distancia euclidiana es que no se toma en cuenta las abundancias absolutas de especies, sino abundancias relativas. Según la distancia de cuerda, las parcelas 1 y 2 de la Tabla 5 son idénticas ( $D_3=0$ ), porque las parcelas tienen las mismas especies con las mismas abundancias relativas. La distancia desde la parcela 3, tanto a la parcela 1 como a la 2, es la mayor posible  $\sqrt{2} = 1,41$



**Figura 5.** Para medir la distancia, uno puede pensar que ambas parcelas son vectores con direcciones y puntas definidas según las abundancias de las especies observadas en las parcelas. La distancia euclidiana es la distancia entre las puntas de los vectores. Para medir la distancia de cuerda (*chord distance*), primero hay que dibujar un círculo con el centro en el origen y con un radio de 1. La distancia de cuerda es la distancia entre los puntos donde los vectores cruzan el perímetro del círculo. *Adaptado de:* Legendre & Legendre 1998, página 279.

El último índice de distancia que se presenta aquí es la distancia  $\chi^2$  ( $\chi$  es la letra chi del alfabeto griego). La distancia  $\chi^2$  se calcula así:

$$D_{16}(x_1, x_2) = \sqrt{y_{++}} \sqrt{\sum_{j=1}^p \frac{1}{y_{+j}} \left( \frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2}$$

El  $y_{1j}$  es la abundancia de la especie  $j$  en la parcela  $x_1$ ,  $y_{2j}$  es la abundancia de la especie  $j$  en la parcela  $x_2$ ,  $y_{1+}$  es la abundancia total de todas las especies en la parcela 1,  $y_{2+}$  es la abundancia total de todas las especies en todas las parcelas. Desde la Tabla 5 conseguimos la distancia  $\chi^2$  entre las parcelas 1 y 2:

$$D_3(\text{parcela1}, \text{parcela2}) = \sqrt{42} \sqrt{\frac{1}{3} \left( \frac{1}{14} - \frac{2}{28} \right)^2 + \frac{1}{15} \left( \frac{5}{14} - \frac{10}{28} \right)^2 + \frac{1}{24} \left( \frac{8}{14} - \frac{16}{28} \right)^2} = 0$$

La distancia entre las parcelas 1 y 3 (igual que entre 2 y 3) es 3,5.

La distancia  $\chi^2$  es un índice asimétrico de distancia, y por eso útil para datos de especies. Este índice está basado, como la distancia de cuerda, en abundancias relativas. La distancia  $\chi^2$  da mucho peso a especies raras en los datos. Esto es aparente en su fórmula, porque la diferencia en abundancias de cada especie está multiplicada por la inversa de la abundancia total de la especie. Cuanto más abundante es la especie en general en los datos, más grande tiene que ser la diferencia en abundancias relativas de esta especie, para que el efecto de la especie al valor del índice se mantenga igual.

La distancia  $\chi^2$  es más aplicable a datos en los cuales se ha muestreado prácticamente todas las especies presentes en cada lugar. Entonces las especies raras pueden ser las más especializadas, y dan más información del medio ambiente de diferencias pequeñas que no son relevantes para especies más abundantes. Pero en caso de que se haya muestreado sólo una parte de las especies presentes lo que muchas veces sucede en ambientes ricos con especies las ocurrencias de especies raras reflejan más suerte de observación que otra cosa. En estos casos, el dar más peso a especies raras aumenta el efecto del error de observación/medición en los resultados.

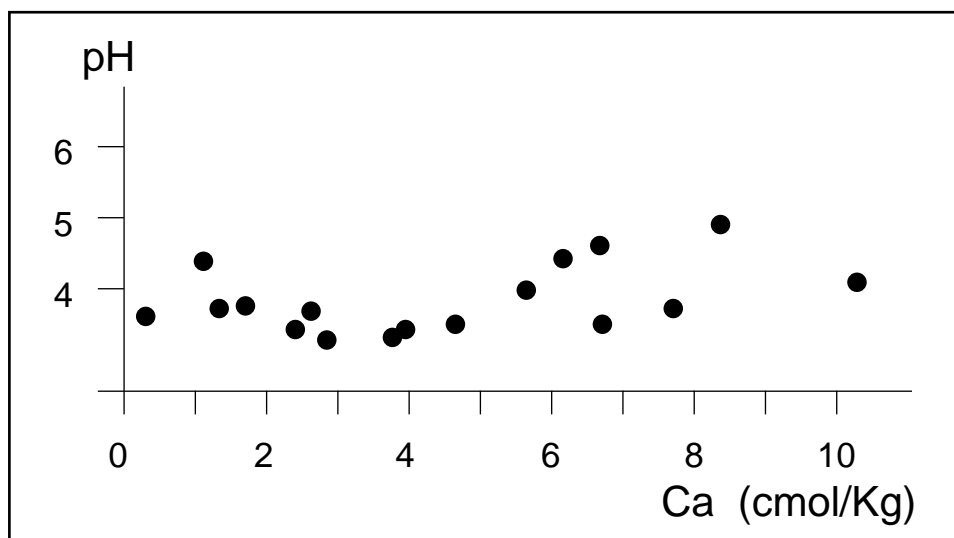
La distancia  $\chi^2$  es un buen método para conocer, porque es el índice de distancia que es implícitamente utilizado en el análisis de correspondencia. Explicaremos con más detalle este análisis cuando presentemos los métodos de ordenación.

### Transformaciones de los datos

Muchas veces se mide las variables medioambientales con diferentes unidades y escalas de medición. Por ejemplo, la acidez del suelo está expresada con las unidades de pH, el contenido de nutrientes en miligramos o moles por kilo, y la proporción de arena en porcentajes de todas las clases de tamaño de granos. Todas estas variables tienen rangos bien diferentes, y por eso uno no puede usarlas libremente para calcular distancias euclidianas entre los lugares de muestreo. Si se utilizan los valores originales de las variables, la distancia quedaría mayormente determinada por la variable con el rango más grande de valores (Figura 6), y esta variable podría ser cualquiera según las unidades usadas. Por ejemplo, indicando el contenido de nutrientes en g/kg en lugar de mg/kg, disminuirían todos los valores numéricos a su milésima parte, e igualmente disminuiría el efecto de la variable en la distancia euclidiana.

Para evitar este problema uno puede utilizar alguna transformación de los datos originales. Podemos restringir el rango de variación de la variable entre 0 y 1 con las mismas fórmulas que ya presentamos arriba:

$$y' = \frac{y}{y_{max}} \quad \text{O} \quad y' = \frac{y - y_{min}}{y_{max} - y_{min}}$$



**Figura 6.** El pH y el contenido de calcio en muestras de suelo colectadas en bosques de tierra firme. El cálculo de distancias euclidianas entre las muestras con base en esta información resultaría determinado principalmente por el calcio, porque su rango numérico de variación es más grande.

Lo más común es usar una estandarización (en inglés *standardisation*) para igualar las variables. Después de la estandarización, el promedio de la variable es 0 y varianza 1. Los valores estandarizados de una variable se calculan de la manera siguiente:

$$z_i = \frac{y_i - \bar{y}}{s_y} \quad \text{donde } s_y \text{ es la variación estándar de la variable y o} \quad s_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}}$$

Aunque los métodos presentados aquí no requieren que las variables tengan una distribución normal, es normalmente ventajoso si las variables presentan una distribución aproximadamente simétrica. Es muy común que la distribución de la variable sea alargada hacia derecha. Esto sucede, por ejemplo, con los datos de abundancias de las especies y de contenido de nutrientes, porque tienen el mínimo absoluto de cero, pero carecen de un valor máximo predefinido. Los valores excepcionales pueden afectar significativamente al resultado del análisis, porque esos valores determinan mucho el valor del promedio o correlación lineal, por ejemplo, y pueden dominar casi completamente la variación en distancias euclidianas. Este problema puede ser resuelto a través del cálculo de un logaritmo de cada valor de la variable. La fórmula básica es:

$$y' = \ln(a + by)$$

El logaritmo del cero no está definido, y por eso hay que añadir una constante  $a$  a cada valor. El tamaño de esta constante se determina según la exactitud de las mediciones en los datos. Por ejemplo, si se trata de números de individuos la constante  $a$  es normalmente 1. Pero si se ha medido biomasa con una exactitud de 0,01 unidades,  $a$  puede ser 0,01. El factor  $b$  es normalmente 1 cuando se trata de números de individuos, pero cuando los datos tienen valores menores de 1,  $b$  puede ser definido como 100 para eliminar logaritmos negativos. En un caso como éste, el valor de  $a$  se determinará según la exactitud de los valores después de la multiplicación con  $b$ .

La transformación logarítmica de las variables no debe hacerse automáticamente, pues primero hay que observar la distribución de las variables para ver si la transformación es necesaria.

Con algunas variables medioambientales la transformación logarítmica puede ser justificable por razones ecológicas. Esta clase de variables son, por ejemplo, aquéllas que indican la riqueza del suelo, como nitrógeno o potasio. Es lógico pensar que un cambio con cierta cantidad de unidades de medición en la riqueza del suelo es más significativo cuando el nivel de riqueza es bajo, y menos cuando la riqueza es alta. Se espera que las plantas reaccionen más a cambios relativos que absolutos en el suelo. Con una transformación logarítmica uno logra eso; por ejemplo, que un aumento doble, de 1 a 2 tenga, el mismo efecto en la distancia euclidiana que el aumento doble desde 100 a 200. La misma lógica está detrás de los valores de pH. El pH es el logaritmo del contenido de iones  $H^+$ . Consecuentemente, uno no debe de tomar el logaritmo de pH por segunda vez

**Las matrices de similitud y distancia**

La mayoría de los métodos presentados aquí están basados en matrices de similitud o distancia. Esta matriz consta de valores de distancia (o similitud) entre cada par de parcelas. En la Tabla 6 se presenta un ejemplo de una matriz de distancia, y en la Tabla 7 una matriz de similitud. En la diagonal de la matriz se encuentran los valores conseguidos cuando se compara la parcela consigo misma. Una matriz de distancia tiene sólo ceros en la diagonal, mientras que una matriz de similitud tiene unos a lo largo de la diagonal.

**Tabla 6.** Una matriz de distancia que presenta distancias euclidianas entre las parcelas según los datos de la Tabla 1. La matriz de distancia está presentada aquí en su totalidad, pero vale observar que la matriz es idéntica en ambos lados de la diagonal.

	Parcela 1	Parcela 2	Parcela 3	Parcela 4	Parcela 5	Parcela 6
Parcela 1	0	80	83	97	111	112
Parcela 2	80	0	57	61	70	71
Parcela 3	83	57	0	64	99	104
Parcela 4	97	61	64	0	45	55
Parcela 5	111	70	99	45	0	14
Parcela 6	112	71	104	55	14	0

**Tabla 7.** Una matriz de similitud con base en los datos de la Tabla 1, utilizando el índice de Jaccard. Para matrices simétricas de distancia o similitud se presenta muchas veces sólo el triángulo inferior o superior, como ocurre aquí.

	Parcela 1	Parcela 2	Parcela 3	Parcela 4	Parcela 5	Parcela 6
Parcela 1	1	0.67	0.33	0.25	0.00	0.00
Parcela 2		1	0.25	0.50	0.20	0.25
Parcela 3			1	0.67	0.25	0.00
Parcela 4				1	0.50	0.25
Parcela 5					1	0.67
Parcela 6						1

Se puede cambiar una matriz de distancia a otra de similitud y viceversa, a través de las fórmulas dadas arriba. En términos prácticos, no importa mucho qué tipo de matriz se presenta. Tal vez la única situación en la cual es importante es cuando se hace el análisis, porque muchos programas de computadoras requieren uno u otro tipo de matriz.

La mayoría de las matrices de distancia en estudios biogeográficos y de ecología de comunidades son simétricas. Esto quiere decir que la distancia desde la parcela 1 a la 2 es igual a la distancia desde la parcela 2 a la 1. En algunas situaciones la matriz de distancia puede ser asimétrica. Por ejemplo, cuando uno estudia la ocurrencia de especies en sistemas fluviales, puede estar justificado utilizar una matriz de distancia en la cual se toma en cuenta la dirección de la corriente. Si la parcela 1 está ubicada aguas arriba de la parcela 2, la distancia desde 1 a 2 puede ser mucho más corta que la distancia desde 2 a 1, porque es más difícil moverse contra la corriente.

## 5. IDENTIFICANDO LOS GRADIENTES

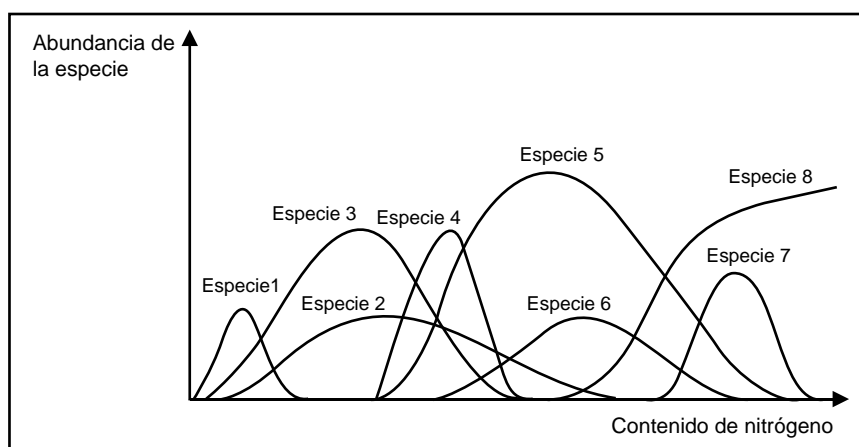
### Análisis directo de gradiente

En un análisis directo uno estudia la respuesta de la especie a la variación, medida o estimada, de los factores medioambientales. La ocurrencia de la especie a lo largo del gradiente puede ser examinada a través de una presentación gráfica que demuestra el cambio en abundancia de la especie cuando el factor medioambiental varía (Figura 7).

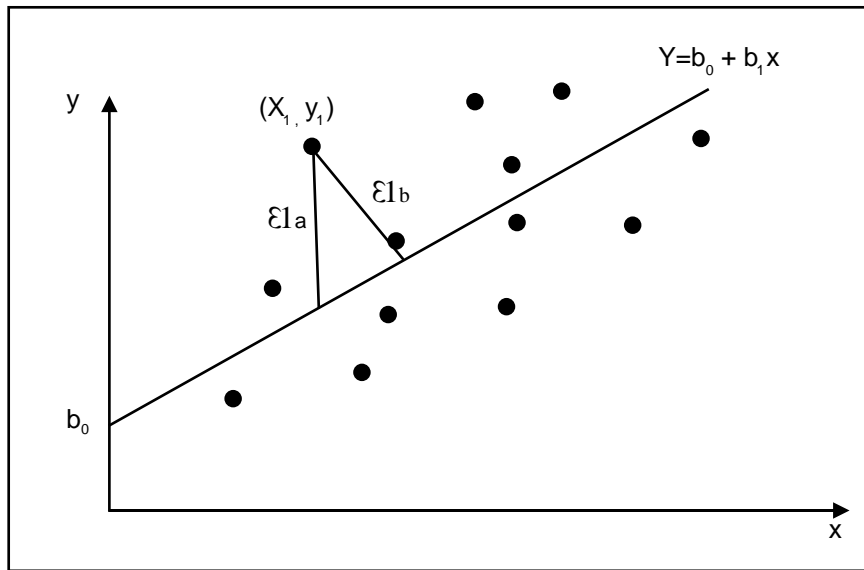
Se puede estudiar numéricamente la respuesta de la especie a través de un análisis de regresión. En este caso se presenta la abundancia de la especie como una función de la variable medioambiental. La fórmula básica de una regresión lineal es:

$$y_i = b_0 + b_1 x_i + \hat{\epsilon}_i$$

La constante  $b_0$  indica el valor donde la línea de regresión cruza el eje  $y$ ,  $b_1$  es la constante que indica el ángulo de la línea, y  $\hat{\epsilon}_i$  es el factor de error o el residual (Figura 8). Se determina la fórmula de la regresión de tal manera que el total de los errores sea el mínimo posible. En otras palabras, se minimiza la suma de residuales cuadrados.



**Figura 7.** En una situación ideal, las respuestas de diferentes especies al contenido de nitrógeno en el suelo podrían manifestarse así. Cada especie tiene un rango de tolerancia, dentro del cual su abundancia primero aumenta y después disminuye, cuando el contenido de nitrógeno aún sigue aumentando. El óptimo nivel de nitrógeno asociado con la mayor abundancia de cada especie está en lugares diferentes a lo largo del gradiente, y los valores máximos en abundancia son también diferentes. Algunas especies tienen zonas más amplias de distribución y otras más restringidas a lo largo del gradiente.



**Figura 8.** El principio de regresión lineal: se busca una línea que tenga una suma mínima posible de los errores  $\hat{\epsilon}_{1b}$  al cuadrado. Al mismo tiempo, se maximiza la cantidad de variación hasta que la línea de regresión pueda explicar las posiciones de los puntos. Cuando se busca la línea de regresión de la suma normal de errores al cuadrado (en inglés *ordinary least-squares regression*), se mide el error  $\hat{\epsilon}_{1b}$  en la dirección del eje  $y$ . Uno puede medir el error también en la dirección perpendicular a la línea de regresión ( $\hat{\epsilon}_{1a}$ , inglés *major axis*)

Normalmente, la respuesta de las especies a los gradientes medioambientales es lineal sólo en el caso de que el gradiente estudiado sea muy corto. En un gradiente largo, la respuesta es normalmente unimodal, es decir, tiene sólo un pico como en la Figura 7. Por lo tanto, uno puede también buscar una fórmula no-lineal para la regresión. Sobre estos y otros aspectos de la regresión se puede encontrar más información en los libros Legendre & Legendre 1998, y Jongman *et al.* 1995.

Es fácil generalizar la regresión para situaciones que tengan más de un factor explicativo:

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + \epsilon_i$$

Esta fórmula se utiliza en los métodos de ordenación que son explicados más adelante.

Con base en las observaciones, es posible calcular un valor indicativo de cada especie en relación con cada una de las variables medioambientales. El valor indicativo se consigue tras calcular el promedio de la variable medioambiental de las parcelas donde se ha encontrado la especie, dando diferentes pesos a las parcelas según la abundancia de la especie:

$$u^* = \frac{y_1x_1 + y_2x_2 + \dots + y_nx_n}{y_1 + y_2 + \dots + y_n}$$

En la fórmula,  $y_i$  es la abundancia de la especie  $y$ ,  $x_i$  es el valor de la variable medioambiental en la parcela  $i$ .

Si se ha calculado el valor indicativo de cada especie para, por ejemplo, el contenido de nitrógeno en el suelo, podemos predecir con base en esta información el contenido de nitrógeno en parcelas donde no se ha registrado el nitrógeno pero sí la composición de especies. Este método se llama calibración (en inglés *calibration*; Jongman *et al.* 1995). Así, podemos probar cuán confiables son los valores indicativos a través de una comparación entre los valores estimados y los reales medidos de nitrógeno. Estimar el contenido de nitrógeno a través de valores indicativos de la especie puede ser útil cuando no es posible, por alguna razón, medirlo directamente. El valor estimado se consigue con la fórmula siguiente:

$$\hat{x}_0 = \frac{y_1u_1 + y_2u_2 + \dots + y_mu_m}{y_1 + y_2 + \dots + y_m}$$

En la fórmula,  $y_i$  es la abundancia de la especie  $i$  y  $u_i$  su valor indicativo para la variable medioambiental, cuyo valor estimado para la parcela es:  $\hat{x}_0$

Es importante tener en cuenta que este tipo de cálculos da resultados útiles sólo cuando el muestreo ha sido representativo en relación con la variación de la variable medioambiental. Por ejemplo, si se ha muestreado el gradiente de nitrógeno con 30 parcelas, de las cuales 25 están establecidas en lugares con contenido de nitrógeno de 2 a 5, y en cinco lugares con nitrógeno cuyo contenido sobrepasa el valor de 10, no se puede conseguir resultados muy confiables del análisis del valor indicativo de las especies, porque una gran parte del rango de los valores de nitrógeno (de 5 a 10) ha quedado sin muestreo.

### Análisis indirecto de gradiente u ordenación

Uno puede estudiar la composición de especies en las unidades de muestreo y después tratar de arreglarlas en una tabla o diagrama de tal manera que sea fácil detectar los cambios graduales en la composición de especies. Este análisis se llama análisis indirecto de gradiente u ordenación.

En términos generales, el objetivo de la ordenación es resumir la información multidimensional de especies o variables medioambientales a una presentación con un número reducido de dimensiones para poder visualizar, de una manera entendible, las relaciones entre las unidades de muestreo con, por ejemplo, diagramas que representan el espacio de especies o variables medioambientales.

Con la ordenación es posible estudiar los cambios y gradientes en la composición de especies, aunque se desconozca los factores medioambientales relevantes. Pero si se ha hecho mediciones de variables medioambientales, se puede estudiar si los gradientes en la composición de especies corresponden a los gradientes en variables medioambientales.

### Análisis de componentes principales (principal component analysis, PCA)

En caso de que se haya observado dos especies en el estudio, podemos presentar las similitudes entre las parcelas fácilmente en un diagrama, en el cual la abundancia de la primera especie está representada por el eje *x* y la abundancia de la segunda especie por el eje *y*. Por ejemplo, la Figura 9A presenta un diagrama de las parcelas de la Tabla 8 según las abundancias de las especies A y B. En la figura uno se ve rápidamente que en el espacio de especies definido por las especies A y B, la trampa 5 produjo un resultado muy diferente a la trampa 3, mientras que la trampa 4 sí se parecía mucho a la trampa 3.

**Tabla 8.** Las abundancias de seis especies de mariposas nocturnas capturadas por cinco trampas durante una noche.

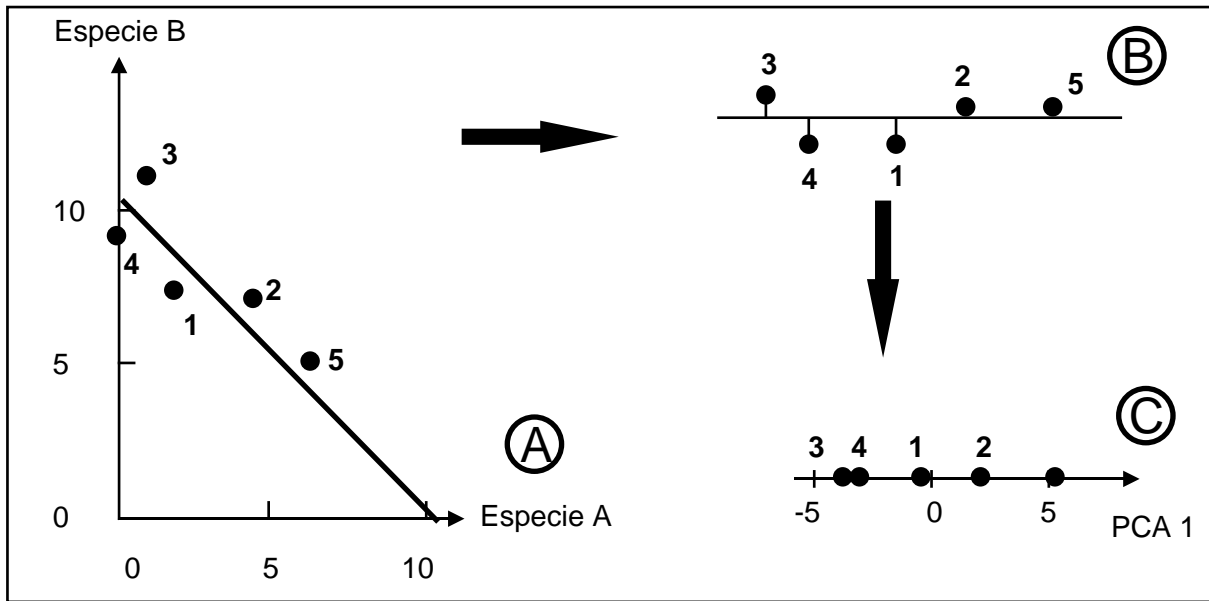
Trampa	Especie A	Especie B	Especie C	Especie D	Especie E	Especie F
Parcela 1	2	7	8	0	0	3
Parcela 2	5	6	9	1	3	2
Parcela 3	1	11	0	0	1	6
Parcela 4	0	9	2	0	0	8
Parcela 5	7	4	5	1	4	0

Si hacemos el mismo tipo de diagrama usando las abundancias de tres especies, conseguimos una presentación tridimensional. Cada especie está representada por un eje propio, y todos los ejes están en 90° con relación a otros ejes. Para representar toda la tabla de especies, uno necesitaría seis dimensiones o seis ejes, uno para cada una de las especies. Matemáticamente es completamente posible definir un espacio de seis dimensiones (seis ejes teniendo 90° en relación con todos los otros ejes), pero en nuestro mundo tridimensional no es muy fácil hacer una representación visual de este espacio con seis dimensiones. Además, normalmente los datos biogeográficos y de ecología de comunidades tienen mucho más especies que seis fácilmente docenas o cientos, hasta miles.

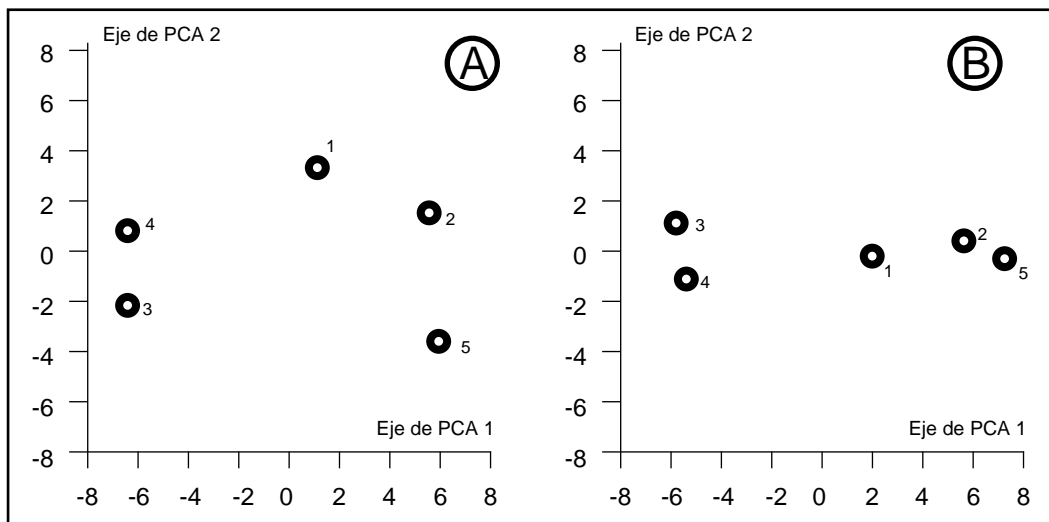
Como no es fácil visualizar toda la variación en la composición de especies, uno puede tratar de visualizar simplemente la mayor parte posible de la variación. En otras palabras, uno puede buscar, en el espacio multidimensional de las especies, unas pocas nuevas dimensiones o ejes que puedan conservar al máximo la información original sobre la variación en la composición de especies. En el análisis de componentes principales (en inglés: *principal component analysis*, PCA), se coloca una línea de regresión, con el método de *major axis regression*, a la nube de las unidades de muestreo (espacio definido por las abundancias de las especies, o simplemente espacio de especies) (ver Figura 8). Las proyecciones de las unidades de muestreo en esta línea forman una ordenación unidimensional. Cuando se coloca el punto cero de la línea de regresión en el promedio de las proyecciones de las unidades de muestreo, se consigue el primer componente principal (Figura 9B-C).

El segundo componente principal es otra línea de regresión, que se forma de tal manera que la línea abarque la mayor cantidad posible de la variación no explicada por el primer componente principal. Al mismo tiempo, el segundo componente principal o eje tiene que tener 90° en relación con el primer componente. En el caso de la Figura 9, los dos componentes principales ya son suficientes para explicar toda la variación en el espacio de especies, porque este espacio no tuvo más que dos dimensiones. Cuando existen más dimensiones, el tercer componente principal, tanto como los componentes siguientes, se forma de la misma manera que el segundo eje: cada nuevo eje tiene que tener 90° en relación a todos los ejes existentes. Vale notar que el ángulo de 90° entre dos ejes significa que son independientes uno del otro. La Figura 10 presenta la ordenación de toda la información de abundancias de especies de la Tabla 8, o en otras palabras, el patrón de semejanzas y diferencias en la composición de especies entre las trampas.





**Figura 9.** A) Las posiciones de las cinco trampas de la Tabla 8 en un espacio definido por las abundancias de las especies A y B. Se ha colocado una línea de regresión en la nube de puntos que representan las trampas. B) Se ha cambiado la línea de regresión de la figura A a una posición horizontal proyectando los puntos a la línea. C) A través de la proyección se consigue una ordenación unidimensional de las trampas a lo largo del eje del primer componente principal.

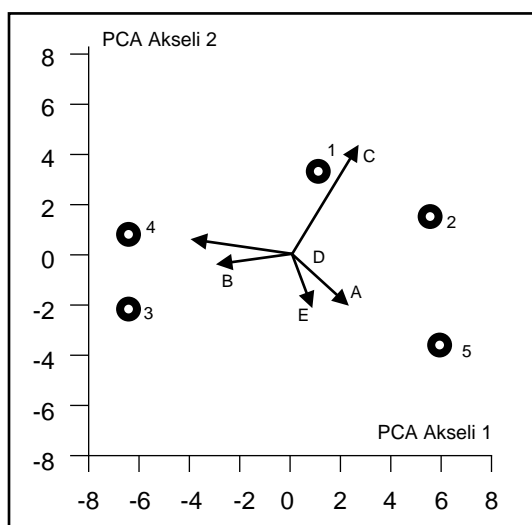


**Figura 10.** Una representación gráfica de resultados de un análisis de componentes principales de los datos de la Tabla 8. A) El primer y el segundo componente principal. B) El primer y el tercer componente principal. El primer componente explica un 79% de la variación total en la composición de especies, el segundo un 17% y el tercero un 2,2%. Casi toda la variación ha podido ser capturada en las dos primeras dimensiones.

La cantidad de variación capturada por el componente es llamada *eigenvalue*. El *eigenvalue* de un componente en relación con la suma de todos *eigenvalues* describe la proporción de la variación total explicada por este componente principal. El primer componente siempre explica más que el segundo, el segundo más que el tercero, y así sucesivamente. Esto se puede notar en la Figura 10: las distancias son más largas entre las trampas a lo largo de la dirección del eje 1 que a lo largo de los ejes 2 ó 3. Con buena suerte, uno puede explicar mucho de la variación total con sólo dos o tres primeras dimensiones. Si los últimos componentes principales no explican nada de la variación, se conseguirá el valor cero.

Uno puede también pensar el análisis de los componentes principales de esa manera, que primero se mueva el origen del sistema de coordenadas de los ejes originales de las abundancias de las especies al centro de este espacio. Después se da vueltas al sistema de coordenadas hasta que las direcciones nuevas de los ejes correspondan a las direcciones determinadas por las líneas de regresión de cada uno de los componentes principales. El espacio de especies se mantiene intacto, lo que significa que la ordenación conserva distancias euclidianas entre las unidades de muestreo. Esto significa que el PCA es apto sólo para datos aplicables al cálculo de distancias euclidianas (ver arriba).

Es posible dibujar las variables originales (especies o variables medioambientales) en el mismo diagrama de ordenación con las unidades de muestreo. Así, se consigue un *biplot* que indica tanto las distancias entre las unidades de muestreo como las variables que más han contribuido a la formación de cada uno de los componentes principales. Como las variables corresponden a los ejes en los datos originales, se marca normalmente las variables con vectores en un *biplot* (Figura 11).



**Figura 11.** La ordenación de la Figura 10 con vectores que representan cada una de las variables originales (las especies). La longitud de la proyección del vector a lo largo del eje de un componente principal cuenta el rol que la variable ha tenido en la construcción del componente. Por ejemplo, el vector de la especie D es tan corto que casi no puede verse, y por eso esta especie no ha contribuido prácticamente con nada a los dos primeros componentes presentados en la figura. La proyección de una unidad de muestreo a un vector de una determinada especie da una aproximación (no necesariamente muy exacta) de la abundancia de la especie en esta unidad de muestreo. Por ejemplo, la especie C debe de ser más abundante en las trampas 1 y 2, mientras su abundancia mínima debe encontrarse en las trampas 3 y 4. Akseli = eje.

El análisis de componentes principales (como otros métodos de ordenación) está basado en el álgebra de matrices. Cuando uno hace el PCA, tiene que decidir si se calcula los componentes principales desde una matriz de covarianza o desde una matriz de correlación. La decisión se puede hacer pensando primero cuál sería la manera más apropiada de calcular distancias euclidianas entre las unidades de muestreo. Si se puede calcular la distancia euclidiana usando directamente los valores originales de las variables, se calcula el PCA según la matriz de covarianza. Pero si tuviésemos que estandarizar las variables antes del cálculo de la distancia euclidiana, hay que escoger la matriz de correlación. En caso de que se estandarice las variables antes del análisis, el resultado es igual al obtenido con la matriz de correlación.

Como el PCA conserva distancias euclidianas entre las unidades de muestreo, es un método apto para datos sobre variables medioambientales, pero para datos de especies este método es aplicable sólo con la salvedad de que todas las especies ocurran en prácticamente todas las unidades de muestreo (ver arriba la explicación sobre la distancia euclidiana). Más adelante analizaremos métodos de ordenación que son más aptos para datos de especies.

### **Análisis de correspondencia (correspondence analysis, CA)**

El análisis de correspondencia es similar al PCA. La mayor diferencia es que en el CA no se define el espacio de especies en términos de distancias euclidianas, pero de una manera que conserva la distancia  $\pm 2$  entre las unidades de muestreo. Cuando el PCA busca relaciones lineares en los datos, el CA es especialmente apto para situaciones en las cuales la variación en la abundancia tiene una forma unimodal (Figura 7). Consecuentemente, el PCA es aplicable sólo para datos que cubren gradientes cortas, mientras el CA sirve mejor para estudiar gradientes largos. Pero es bueno recordar lo que se notó arriba sobre el efecto de especies raras en los valores de distancia  $\pm 2$ , porque esto afecta también los resultados de la ordenación.

Para el análisis de correspondencia se utiliza también el nombre de "*reciprocal averaging*", porque por primera vez se presentó el método de tal forma que se aplicaron los promedios usados en el análisis directo de gradiente (véase arriba las fórmulas para calcular los valores indicativos).

El punto de partida es una tabla de observaciones, como la Tabla 8. Primero se asigna a cada trampa un valor arbitrario de sitio (en inglés *site scores*). No importa que valores sean, sólo que cada trampa tenga un valor único; por ejemplo, uno puede usar el número de la trampa. Para cada especie se calcula un valor (en inglés *species scores*) que es el promedio de los valores de sitio de todos los individuos de la especie. Después se calcula para cada trampa un nuevo valor de sitio, que es el promedio de valores de especie de cada individuo de la trampa. Se estandariza los valores de sitio y después se vuelve a calcular los valores de especie. Así iterando la misma tarea muchas veces, los valores van a estabilizarse al final.

Los valores de sitio ya estabilizados forman el *eigenvector* del primer eje de ordenación, o en otras palabras, las coordenadas para cada trampa en la primera dimensión del espacio de especies. Los valores de especie dan, respectivamente, las coordenadas para cada especie en la primera dimensión del espacio de trampas. La segunda dimensión se puede resolver con el mismo método iterativo, solo que se incluye una fase en la cual se garantiza que el segundo eje va a tener 90° con el primero.

En efecto, también el PCA puede hacerse a través de iteraciones. La diferencia es solamente que los *eigenvectors* son calculados usando sumas y no promedios de valor de especies, o sitio de los individuos por trampa o por especie.

En el CA uno puede producir, tanto como en el PCA, *biplots* presentando en el mismo tiempo las unidades de muestreo y las especies. Para el *biplot*, uno puede escalarizar los ejes para las unidades de muestreo y las especies de una manera algo diferente, afectando así un poco a la interpretación del *biplot*. Más comúnmente se utiliza una escalarización que conserva las distancias  $X^2$  entre las unidades de muestreo. En este caso, el mayor peso en la interpretación está en las interrelaciones de las unidades de muestreo. Las especies van a aparecer cerca a esas unidades de muestreo, donde cada especie particular es más abundante. El programa Canoco, que es el programa más utilizado para hacer el CA, utiliza en esta escalarización la expresión: "*sample scores are weighted averages of species scores*".

Otra escalarización útil conserva las distancias  $X^2$  entre las especies, y consecuentemente el interés entonces se concentra en las interrelaciones de las especies. Las unidades de muestreo se encuentran en las cercanías de las especies que son más abundantes en la unidad de muestreo. Canoco expresa esta escalarización así: "*species scores are weighted averages of sample scores*".

Un tratamiento más profundo de PCA y de CA se encuentra en los libros Legendre & Legendre 1998, y Jongman *et al.* 1995.

### **Análisis de coordenadas principales (*principal coordinates analysis, PCoA*)**

Es posible hacer una ordenación desde la tabla original de observaciones, pero también se puede hacer a partir de las distancias calculadas de los valores de la tabla original. Un método que hace esto es el Análisis de Coordenadas Principales (PCoA). La primera fase en este análisis es escoger un índice de distancia que sea adecuado para los datos y la pregunta. Después se calcula las distancias formando una matriz de distancias. Con base en la matriz de distancias, se reconstruye las posiciones de las unidades de muestreo en un espacio definido por el índice aplicado de distancia (o similitud).

Podemos imaginar la tabla de distancias viales entre las ciudades peruanas. Desde esta tabla es posible reconstruir un mapa que presenta las ubicaciones de las ciudades en un plano. Esto se puede hacer forzando y probando varias veces hasta que una nube de puntos satisfactoria haya sido encontrada, de lo contrario puede hacerse el mapa de una manera matemática. En el PCoA se opta por la matemática. Los ejes de la ordenación se determinan de igual modo que en el PCA, es decir, colocando líneas de regresión en la nube de puntos de las unidades de muestreo. Después podemos dibujar en un papel la ordenación de las unidades de muestreo según los ejes preferidos de las coordenadas principales.

En un mapa vial, todas las ciudades se encuentran en el mismo plano, y se puede obviar la altura sobre el nivel del mar, porque es más corta que las distancias horizontales. En un estudio sobre la variación en la composición de las especies normalmente hay muchas más dimensiones. Uno puede siempre colocar dos unidades de muestreo en diferentes puntos a lo largo de una línea, de tal manera que la distancia geográfica en la línea corresponda completamente a la distancia en términos de composición de especies. Igualmente, siempre es posible construir una presentación bidimensional completa de las distancias en relación a las especies entre tres lugares. En general, se puede decir que siempre se puede representar completamente las distancias entre  $n$  unidades de muestreo en un espacio que tiene  $n-1$  dimensiones. En otras palabras, con el PCoA se consigue  $n-1$  *eigenvectors*. En caso de que el número de variables ( $p$ , por ejemplo especies) para el cálculo de distancias entre las unidades de muestreo haya incluido menos variables que  $n-1$ , se consigue como máximo  $p$  *eigenvectors* con *eigenvalue* mayor a cero.

La gran ventaja de PCoA es que la ordenación no es dependiente de ningún índice específico de distancia. Podemos seleccionar libremente el índice más apropiado para el estudio. La ordenación puede basarse, por ejemplo, en valores de distancia según el índice de Jaccard o Steinhaus, que normalmente describen muy bien la variación en la composición de especies. Si se hace el PCoA usando distancias euclidianas, el resultado será una ordenación igual a la conseguida con el PCA. Un PCoA de una matriz de distancias  $X^2$  resulta en una ordenación igual al CA.

El PCoA está basado en una matriz de distancia, y por eso la información de los valores de variables no se conserva en el análisis. Así, no es posible producir un *biplot* que presentaría tanto las unidades de muestreo como las especies. Para estudiar la relación entre especies, es necesario hacer una ordenación en la cual los objetos ordenados sean las especies, y las distancias entre las especies sean calculadas con base en sus abundancias en diferentes parcelas.

#### **Escalarización multidimensional (multidimensional scaling, MDS; non-metric multidimensional scaling, NMDS)**

Otro método para hacer una ordenación según la información presentada en una matriz de distancia es la escalarización multidimensional (MDS). A diferencia de los demás métodos de ordenación, este método no está basado en un cálculo de los *eigenvectors*, pero sigue el método de prueba y error. Una consecuencia de esto es que la MDS se origina en los mismos datos de ordenaciones, que siempre son algo diferentes unos de otros.

Al inicio de la MDS tenemos que decidir cuantas dimensiones queremos utilizar en la ordenación. Muchas veces uno quiere dos dimensiones, para poder simplemente presentar un diagrama de las unidades de muestreo, pero a veces algún otro número de dimensiones puede ser más apropiado para la pregunta que uno está formulando. La idea del MDS es colocar las unidades de muestreo en un espacio con un número definido de dimensiones, de tal manera que las distancias entre las unidades de muestreo en la ordenación correspondan al máximo a las distancias originalmente calculadas entre las unidades de muestreo.

La diferencia entre las distancias medidas en la ordenación y las distancias originales se calcula con una función de estrés. El objetivo es conseguir una ordenación con un estrés mínimo posible. Una vez que se ha calculado el estrés en una ordenación, se cambia ligeramente las posiciones de las unidades de muestreo en la ordenación, y después se calcula nuevamente el estrés. Este procedimiento se continúa hasta que ya no sea posible disminuir el estrés. En esta fase, la última ordenación va a ser considerada como la solución final, en la cual se ha conservado en un número reducido de dimensiones la máxima cantidad posible de la información original del espacio multidimensional que presentan las unidades de muestreo.

Normalmente, en la MDS no se trata de conservar los valores originales de distancia, sino el orden relativo de las distancias entre las unidades de muestreo. En este caso se habla comúnmente de una escalarización multidimensional no-métrica, NMDS, porque la escalarización no está siguiendo mediciones métricas. Por eso no se puede interpretar en la ordenación la distancia entre dos unidades de muestreo, pero sí se puede interpretar la relación de esta distancia con las otras en la ordenación si es más corta o menos corta que alguna otra distancia. Una ventaja en este método no-métrico es que no interviene el efecto de arco. Este problema es tratado a continuación.

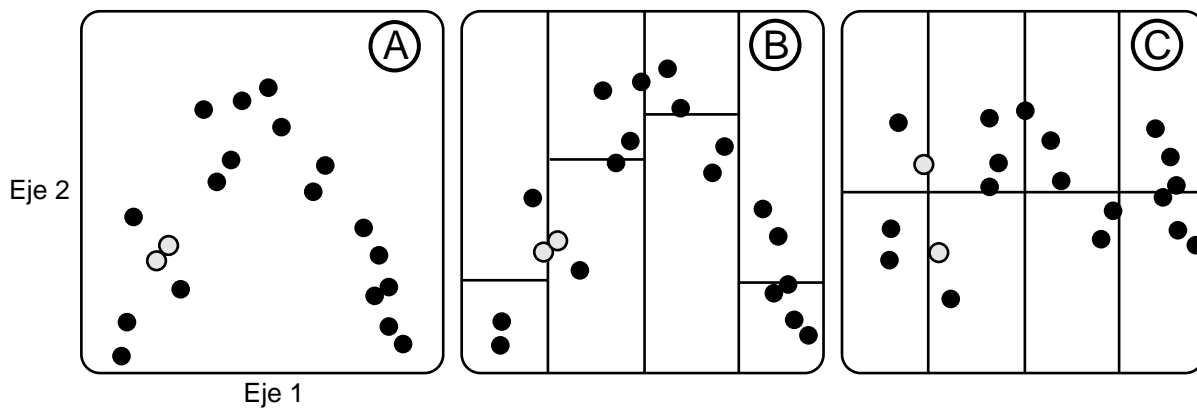
La desventaja de la MDS es que se obtienen resultados diferentes de los mismos datos según la primera aproximación de la ordenación. Para atenuar este problema podemos comenzar de una ordenación pensada como buena aproximación de la ordenación final. Este puede ser, por ejemplo, el resultado de un análisis de coordenadas principales o de la misma MDS, pero con un número mayor de dimensiones. También es posible repetir el análisis varias veces, siempre comenzando de diferentes nubes iniciales, y al final tomando la solución que tuvo el menor estrés. Esto último es normalmente posible sólo con datos relativamente pequeños, porque la MDS requiere mucho poder de cálculo de la computadora.

#### **Efecto del arco y un intento llamado DCA para superarlo**

Cuando se habla de que los ejes de la ordenación son perpendiculares uno al otro, se refiere a la independencia lineal entre éstos. Pero una dependencia no-lineal es completamente posible y, en efecto, este tipo de dependencia se observa casi siempre entre los ejes, especialmente cuando en los datos existe un solo gradiente fuerte. En el resultado gráfico de la ordenación se puede observar esta dependencia como un arco formado por las unidades de muestreo (en inglés: *arch effect* o *horseshoe effect*; Figura 12A). La NMDS es un método excepcional de ordenación porque no sufre del efecto de arco.

El efecto de arco hace la interpretación de los resultados de ordenación algo complicada. Primero, las distancias rectas entre las unidades de muestreo en la ordenación no reflejan necesariamente muy bien las diferencias en la composición de especies entre las unidades de muestreo. Los puntos del arco están a veces hasta más cerca uno del otro que del centro del arco, aunque en realidad las unidades de muestreo en los extremos del arco son las unidades más diferentes en términos de la composición de especies. En otras palabras, la segunda dimensión de la ordenación no es completamente independiente de la primera, pero está más o menos repitiendo el mensaje de ésta.

Consecuentemente, uno no debe de interpretar el gradiente de cambio en la composición de especies de una manera lineal, pero sí de algún modo curvilíneo.



**Figura 12.** A) Ordenación según el CA de 20 parcelas de estudio de vegetación. En realidad, sólo existe un gradiente de cambio florístico que pasa aproximadamente en la dirección del eje 1. Aunque el eje 2 es linealmente independiente del eje 1, existe claramente una correlación con valores cuadrados del eje 1. Por eso, el gradiente de cambio florístico está curvado en forma de un arco. B-C) Presentación esquemática del procedimiento del DCA con segmentación. B) El primer eje de la ordenación se divide en segmentos, y para cada segmento se calcula el promedio de los valores en el eje 2. C) Se define que el promedio de cada segmento sea cero. El arco desaparece, pero las posiciones relativas de las parcelas pueden cambiar mucho, como se ve en el caso de las dos

La curvatura del gradiente no necesariamente representa un problema si la intención de la ordenación es solamente la presentación visual de las relaciones entre las unidades de muestreo. En este caso, uno puede tomar la curvatura en cuenta cuando hace la interpretación. Aunque, naturalmente, la aparición del arco hace que otros gradientes potencialmente interesantes aparezcan en dimensiones más altas en número, lo que puede dificultar su observación.

Muchas veces se trata de interpretar los ejes de ordenación de tal manera que se calcula correlaciones entre las coordenadas de las parcelas en un eje y variables medioambientales medidas en las parcelas. La idea es que si las coordenadas a lo largo de un eje que describe un cambio en la composición de especies crecen cuando, por ejemplo la fertilidad del suelo crece, uno puede considerar la fertilidad del suelo como un explicador de los cambios en la composición de las especies. Naturalmente, esta lógica funciona solamente cuando la relación es lineal entre el eje y el cambio en la composición de especies. Cuando aparece el efecto de arco, la relación no es lineal.

El DCA (en inglés *detrended correspondence analysis*) es un método desarrollado para alinear el arco a una dimensión. El punto de partida es un análisis normal de correspondencia. El arco se elimina o con una segmentación del primer eje o con una aplicación a los valores del segundo eje de una función polinomial de los valores del primer eje.

En la segmentación se divide el primer eje en varios segmentos de longitudes iguales. Después se cambia la posición vertical de cada segmento de tal manera que el promedio de los valores en el eje 2 de las unidades de muestreo dentro del segmento resulten 0 (Figura 12 B-C). Uno puede decir que el arco se segmenta, y luego se coloca cada segmento al mismo nivel horizontal. Podemos eliminar el arco en los otros ejes con este mismo método. El método es relativamente violento, y por eso la información ofrecida a lo largo del eje 2 y los demás puede fácilmente distorsionarse hasta que pueda resultar prácticamente inútil o hasta engañosa. Los resultados pueden ser cambiados también según el número de segmentos usados. Normalmente, la interpretación resulta más fácil cuando el arco está presente en su forma original para poder reconocerlo.

El estiramiento del arco con una función polinomial es en principio más elegante, porque así se pueden evitar los problemas causados por los movimientos abruptamente verticales y diferentes entre segmentos. Pero el investigador tiene que definir de antemano qué grado de función polinomial utilizará. No existe ninguna regla universal para ayudar en esta decisión, y por eso entra un componente grande de subjetividad en el análisis. Siempre existe el riesgo de que cuando se hace desaparecer el arco, también se pierda información ecológicamente relevante. Es factible también que se creen estructuras artificiales en los datos sin ninguna conexión con la realidad.

La conclusión de todo esto es que no se puede recomendar el uso de DCA para analizar datos de observaciones de especies.

### Explicando con variables medioambientales la variación en la composición de especies visible en las ordenaciones

Muchas veces la intención del estudio es entender si las variables medidas del medio ambiente en general, o algunas variables en particular, pueden explicar la variación en la composición de especies. Podemos encontrar respuestas a estas preguntas utilizando ordenaciones de tres maneras diferentes.

### Visualización de la variación medioambiental

La manera más simple de estudiar las relaciones entre algunas pocas variables medioambientales y la variación en la composición de especies expresada en una ordenación, es utilizar diferentes símbolos o colores para marcar las unidades de muestreo que tengan diferentes características medioambientales. También se puede separar con líneas esas unidades de muestreo que pertenecen a las mismas categorías. Además, se puede escribir el valor de la variable medioambiental al lado del punto que representa la unidad de muestreo, aunque esto resulta difícil de fijar si es que hay muchas unidades de muestreo.

Veamos que sucede con los datos de 15 parcelas de tierra firme, en las cuales se ha inventariado los helechos terrestres. Se midió las similitudes florísticas entre las parcelas usando el índice de Jaccard, y después se hizo una ordenación NMDS para visualizar el patrón de semejanzas y diferencias florísticas entre las parcelas. Las coordenadas de las parcelas en las dos primeras dimensiones se presentan en la Tabla 9. En la misma tabla se presenta también la edad del bosque y la cobertura herbácea. La ordenación se presenta en la Figura 13, utilizando diferentes símbolos y alineamientos para indicar la variación de las variables medioambientales. Según la figura, parecería que la edad del bosque afecta más a la composición florística de helechos que a la cobertura herbácea.

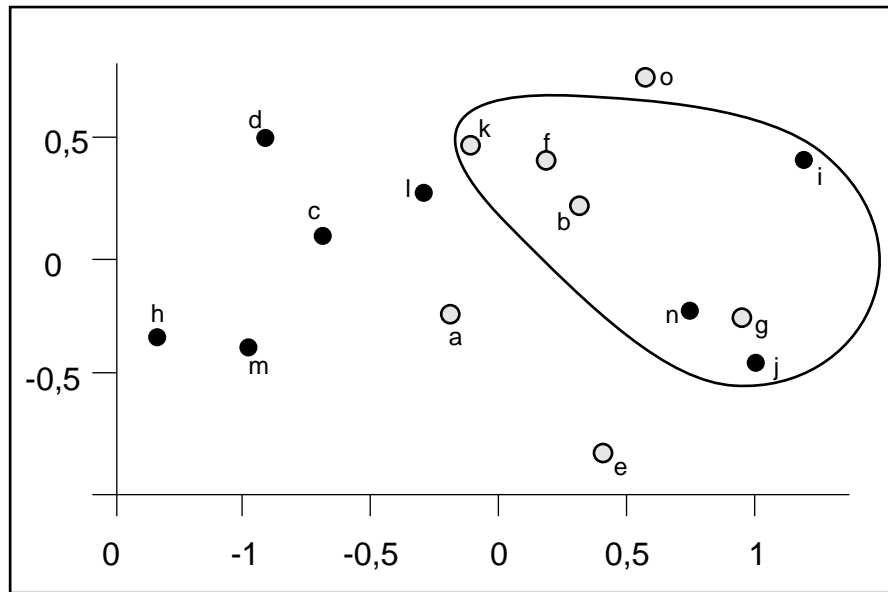
No es posible presentar simultáneamente la variación de varias variables medioambientales en una figura, por lo que muchas veces es más claro estructurar un diagrama o una ordenación propia para las variables medioambientales.

La ordenación determina automáticamente en su proceso el origen de los ejes de ordenación y las direcciones de los ejes. Sin embargo, la dirección de crecimiento o disminución de los valores de los ejes no tiene ninguna importancia definida. Por eso, uno siempre puede dar vueltas a la ordenación o multiplicar todos los valores de coordenadas por -1 sin afectar el contenido informativo de la ordenación. Esta clase de manipulaciones puede a veces ser más fácil de hacer entre dos o más ordenaciones, o entre una ordenación y un diagrama simple dibujado, por ejemplo, con base en variables medioambientales.

Claro que las comparaciones visuales que hemos descrito hasta aquí no contienen ninguna prueba estadística. Pero muchas veces este tipo de visualizaciones simples ayudan a entender e interpretar los resultados, porque describen los patrones en los datos de una manera sencilla y fácil de comprender.

**Tabla 9.** Las coordenadas en dos dimensiones de NMDS de 15 parcelas de tierra firme en las cuales se ha inventariado las especies de helechos. Las similitudes florísticas entre las parcelas han sido calculadas con el índice de Jaccard. Además, se presenta la edad del bosque y la cobertura de hierbas.

Parcela	Coordenadas en la dimensión 1 de la NMDS	Coordenadas en la dimensión 2 de la NMDS	Edad del bosque (años)	Cobertura herbácea (%)
a	-0.2	-0.2	46	54
b	0.45	0.21	75	35
c	-0.7	0.07	32	88
d	-0.86	0.47	14	130
e	0.42	0.82	23	66
f	0.34	0.4	72	78
g	0.97	-0.24	68	76
h	-1.32	-0.36	21	113
i	1.25	0.44	105	111
j	1.01	-0.42	58	97
k	-0.12	0.48	56	47
l	-0.25	0.33	48	134
m	-0.99	-0.39	39	85
n	0.74	-0.22	69	101
o	0.63	0.74	40	69



**Figura 13.** Una ordenación NMDS con dos dimensiones de 15 parcelas de tierra firme. La ordenación está basada en valores del índice de Jaccard calculados con base en la composición de especies de helechos. Los círculos abiertos son parcelas que tienen una cobertura mayor de 80% de hierbas. Dentro de la línea se encuentran parcelas que tienen más de 50 años de edad (Tabla 9).

**Correlación de variables medioambientales con los ejes de ordenación**

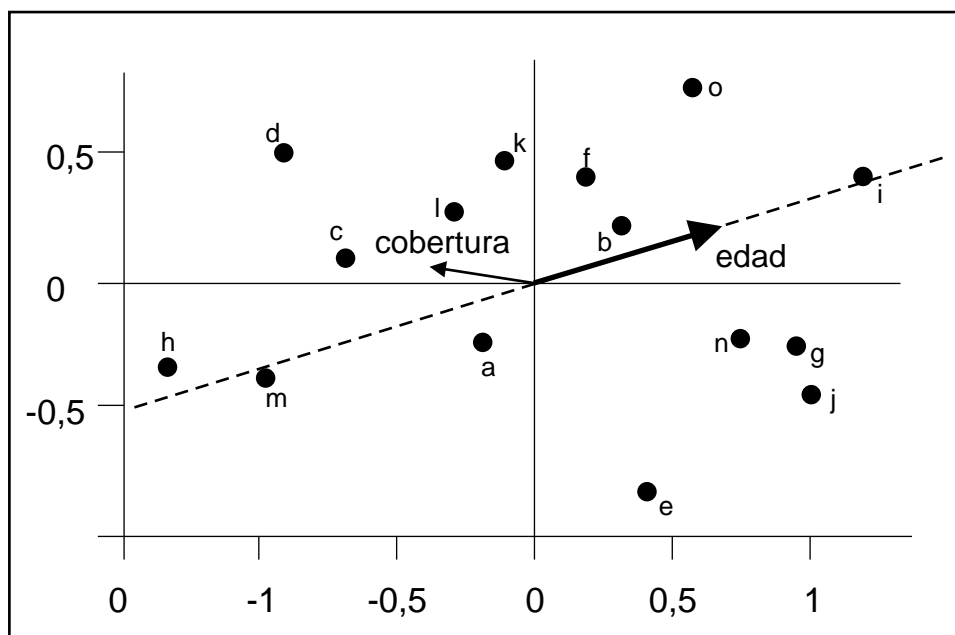
La segunda manera de investigar la relación entre la composición de especies y las variables medioambientales es calculando la correlación entre los ejes de ordenación y las variables medioambientales, y visualizar esta correlación en la figura de ordenación con vectores.

Las correlaciones entre las coordenadas en los ejes de NMDS y la edad del bosque y cobertura herbácea (Tabla 9) son presentadas en la Tabla 10. Podemos visualizar estas correlaciones en la figura de ordenación de tal manera que se dibuja para cada variable medioambiental un vector que salga del origen de la ordenación y termine en un punto determinado por los coeficientes de correlación (Figura 14). La longitud del vector indica cuán predecible es el crecimiento de la variable. La flecha o vector de la edad del bosque es claramente más larga que la de la cobertura herbácea, lo que indica que la edad explica mejor los cambios en la composición de las especies de helechos. La dirección de la flecha indica la dirección del crecimiento de la variable en el espacio de la ordenación. En nuestro ejemplo, la edad del bosque crece desde la esquina inferior izquierda del cuadrante superior derecho hacia la esquina superior derecha del mismo cuadrante. La línea cortada está marcada sólo para explicar este principio en la interpretación. Cuando se reportan los resultados de los análisis, normalmente sólo se dibuja la flecha para mantener la figura más limpia y más entendible.

Según la Figura 14 las parcelas i y o están en bosques relativamente viejos, mientras h y m están en bosques jóvenes. Relacionando esta observación con los datos originales de la Tabla 9, se puede notar que la figura es sólo parcialmente correcta en este sentido. Ésta es una situación típica, porque raras veces los modelos describen exactamente la verdad.

**Tabla 10.** Las correlaciones entre los ejes de NMDS y la edad del bosque y la cobertura herbácea. Los datos vienen de la Tabla 9.

Eje de NMDS	Edad del Bosque	Cobertura herbácea
1.	0,74	-0,23
2.	0,26	0,07



**Figura 14.** Una ordenación bidimensional de la NMDS de 15 parcelas de tierra firme. La ordenación está basada en una matriz de valores del índice de Jaccard calculados con base en la composición de especies de helechos. Las flechas indican la intensidad de correlación entre los ejes de NMDS y la edad del bosque y la cobertura herbácea.

#### **Análisis canónico: análisis y redundancia (RDA, redundancy analysis) y análisis canónico de correspondencia (CCA, canonical correspondence analysis)**

En un análisis canónico se analiza simultáneamente dos (o más) tablas de observaciones que describen las mismas unidades de muestreo. Son varias tanto las variables para explicar o variables dependientes (normalmente especies), como las variables explicativas (normalmente variables medioambientales o geográficas).

Un análisis canónico es una combinación de un análisis directo e indirecto de gradiente: se realiza una ordenación y una regresión simultáneamente. La ordenación está basada en variables dependientes (especies), y los ejes que salen de esta ordenación son restringidos con una regresión de tal manera que tienen que construirse linealmente a partir de las variables explicativas. El resultado es una ordenación en la cual los ejes describen la máxima variación en la composición de especies que es explicable por las variables medioambientales o geográficas.

Existen varias formas para hacer un análisis canónico, según el método de ordenación que uno elige para el punto de partida. La forma basada en el análisis de componentes principales es llamada **análisis de redundancia (RDA)**, y la forma basada en el análisis de correspondencia es llamada **análisis canónico de correspondencia (CCA)**. Estos dos métodos son los más usados por ecólogos y biogeógrafos. Cuando uno decide utilizar estos métodos, tiene que considerar exactamente los mismos requisitos y restricciones como en los casos del PCA y CA. El RDA supone que las especies cambian su abundancia de una manera lineal a lo largo de los gradientes y conservan la distancia euclidiana entre las unidades de muestreo. El CCA supone que la respuesta de las especies a los gradientes es unimodal y conserva una distancia  $\chi^2$  entre las unidades de muestreo.

Uno consigue tantos ejes canónicos como variables explicativas existen. La ordenación en sí produce tantos ejes como la ordenación normal, pero el número de ejes canónicos es sólo el número de variables explicativas. En caso de que exista un número igual de variables

Explicativas al número de ejes en la ordenación, las variables explicativas no restringen, de ninguna manera, la ordenación, y el resultado es el mismo que en una ordenación normal. Consecuentemente, si uno ha medido varias variables explicativas, puede ser ventajoso dejar algunas variables fuera del análisis.

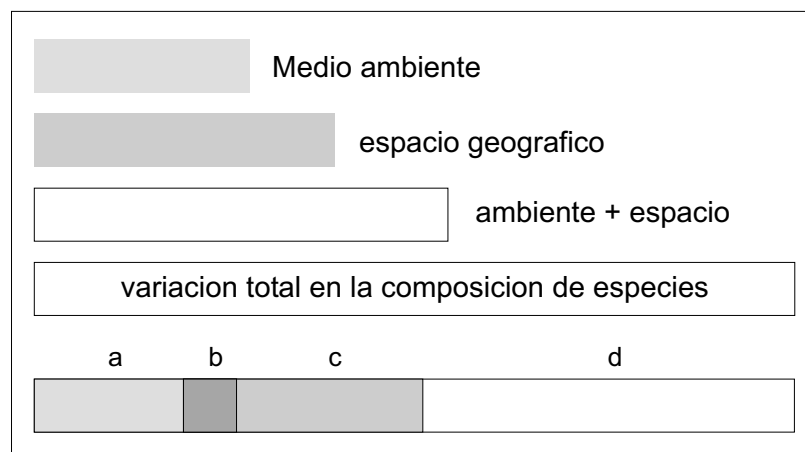
Uno puede presentar el resultado de una ordenación canónica en un diagrama como una ordenación normal. El rasgo adicional es que uno puede presentar adicionalmente en el diagrama los vectores que describen las variables medioambientales. Se puede así presentar un *tripplot*, es decir, un diagrama donde aparecen simultáneamente las posiciones de las especies y unidades de muestreo en la ordenación, y los vectores de las variables medioambientales. Además, uno puede hacer *biplots* con varias combinaciones (especies+medio ambiente, unidades de muestreo+medio ambiente, especies+unidades de muestreo).



Podemos pensar así que el análisis canónico produce dos ordenaciones: una basada sólo en la tabla de observaciones de especies, y otra basada en las tablas de especies y variables medioambientales. En consecuencia se puede probar cuánto de la variación en la composición de especies es explicable por variables medioambientales, y además cuáles son las variables que mejor explican esta variación. A través de un cálculo de la suma de los *egenvalues* de los ejes canónicos, y la comparación de esta suma con la suma total de todos los ejes, se consigue una estimación de la cantidad explicable de la variación en la composición de especies.

La posibilidad de estimar la cantidad explicable de la variación en la composición de especies abre una ruta interesante para examinar qué cantidad de esta variación puede ser explicada por factores diferentes (Figura 15). Por ejemplo, se puede calcular cuánta variación en la composición de especies es explicable solamente por variables medioambientales, cuánta solamente por variables geográficas (coordenadas geográficas de las unidades de muestreo), y cuánta por estos dos grupos de variables juntos. De estos cálculos se puede construir una presentación simple que ayude a entender qué tipo de variables son más importantes para entender la variación en la composición de especies (en inglés *variation partitioning*). De la misma manera, podemos estudiar qué tipo de variables medioambientales son las más relevantes, por ejemplo si las variables edáficas son más importantes que las variables climáticas.

En la interpretación de los resultados es necesario que el análisis sea sobre una presentación de la existencia de correlaciones, no de casualidades. Por ejemplo, es correcto decir que una variable "ene" explica un porcentaje x de la variación en la composición de especies, pero no es justificable deducir de este resultado que la variable necesariamente ha causado la variación observada. La correlación puede fácilmente ser causada por un tercer factor, que afecta a ambas variables.



**Figura 15.** Presentación de la variación en la composición de especies en un análisis canónico de ordenación: (a) variación explicada sólo por variables medioambientales, (b) variación explicada por variables medioambientales y geográficas juntas, (c) variación explicada sólo por coordenadas geográficas, y (d) variación inexplicada.

## 6. CLASIFICACION

### Métodos diferentes de clasificación

El acto de clasificación es casi automático para la mente humana. En una clasificación tratamos de agrupar cosas o fenómenos de tal manera que en una clase existan objetos que sean más o menos semejantes entre sí, cuando se observan según ciertos criterios.

La lógica y la manera de proceder pueden variar regularmente en el proceso de clasificación. Uno puede, por ejemplo, dividir las clasificaciones en jerárquicas y no jerárquicas. En una clasificación jerárquica cada clase es una parte de otra clase más grande, hasta que todas las clases al final forman una "superclase" a la cual pertenecen todos los objetos. Una clasificación no jerárquica divide los objetos en clases que son exclusivas, donde cada objeto puede pertenecer a un solo grupo.

La clasificación puede proceder de manera divisiva o aglomerativa. El punto de partida en una clasificación divisiva es tener todos los objetos en una sola clase, que se va a dividir poco a poco en clases más pequeñas. Una clasificación aglomerativa resulta de una situación en la que cada objeto pertenece a su propia clase, y después se agrupan estas clases en grupos más y más grandes. Las clasificaciones aglomerativas están basadas en matrices de distancia. Las clasificaciones divisivas pueden formar los grupos según la variación en una sola variable (métodos monotéticos) o en varias variables simultáneamente (métodos politéticos).

Debemos tener presente que una clasificación trata de encontrar discontinuidades en los datos, y usar éstas para definir los grupos más homogéneos posibles de objetos. Si los datos representan un gradiente continuo que ha sido bien muestreado, la definición de los límites de los grupos es necesariamente arbitraria y la clasificación resulta ser muy difícil de interpretar. En estos casos, normalmente es más fácil interpretar los patrones a través de un análisis de gradiente. La clasificación y la ordenación revelan rastros algo diferentes de los datos, y por eso pueden ser métodos complementarios y útiles para utilizarlos juntos.

Aquí presentamos mayormente métodos jerárquicos de clasificación, de los cuales uno es divisivo y los otros aglomerativos. Al final presentamos una clasificación aglomerativa que puede ser restringida a objetos vecinos en espacio o tiempo (en inglés *constrained clustering*).

### TWINSpan

TWINSpan (Two Way INDicator SPecies ANALysis) es uno de los métodos más usados de clasificación en estudios ecológicos y biogeográficos. Es bueno conocerlo, aunque no recomendamos su uso por razones que explicamos más abajo. La popularidad de TWINSpan proviene de las décadas 1970 y 1980, cuando este método era uno de los pocos disponibles como un programa de computadora.

TWINSpan es un método divisivo que utiliza información de varias variables (especies) a la vez. Está basado en el análisis de correspondencia. La clasificación se inicia dividiendo el primer eje del CA en dos partes desde su punto cero. Después se hace un nuevo CA separadamente para las dos mitades, las cuales se dividen en dos otra vez según el primer eje del CA. Este proceso se continúa hasta que las clases son ya tan pequeñas que no vale la pena dividir las.

TWINSpan tiene dos problemas clave. Primero, debido a que utiliza sólo la información del primer eje, toda la variación en los otros ejes queda sin importancia en la clasificación. Esto no es tan grave si de verdad toda la variación relevante ya está representada por el primer eje, pero normalmente uno no sabe si la situación realmente es tan sencilla o no antes de investigar bien los datos.

El otro problema es que el método no toma en cuenta las discontinuidades en los datos a lo largo del primer eje del CA. Como el criterio de división está estrictamente definido por el centro del primer eje, significa que un grupo de unidades de muestreo, aunque sea un grupo muy claramente separado de los otros, será dividido en dos si se encuentra justo en ambos lados del cero del primer eje. Legendre & Legendre (1998) citan todavía algunas otras debilidades del TWINSpan, pero las dos mencionadas ya son suficientes para evitar el uso del método.

### Clasificación del vecino más cercano (*nearest neighbour clustering, o single linkage clustering*)

Ahora existen varios métodos de clasificación que están basados en las distancias calculadas entre los objetos. Los métodos defieren uno del otro según la forma de calcular la distancia entre dos grupos, cuando por lo menos uno de los grupos tiene más que un objeto. Una manera consiste en ver a qué distancia están dos vecinos cuando se mide la distancia de una forma predefinida. La clasificación del vecino más cercano es el método más fácil de entender los grupos.

Vamos a considerar datos que constan de ocho parcelas de tierra firme, en las cuales se ha calculado la cantidad de individuos de irapay y ungurahui. Se hace la clasificación usando distancias euclidianas entre las parcelas. En la Figura 16 se muestra las parcelas en un sistema de coordenadas definido por las abundancias de estas dos especies. El diagrama puede presentar los datos de manera completa, porque solo existen dos dimensiones en los datos.

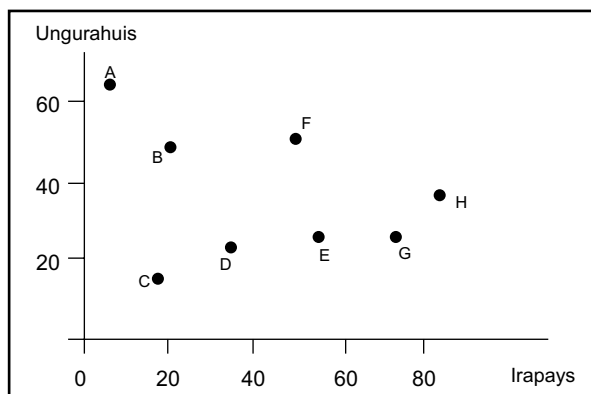
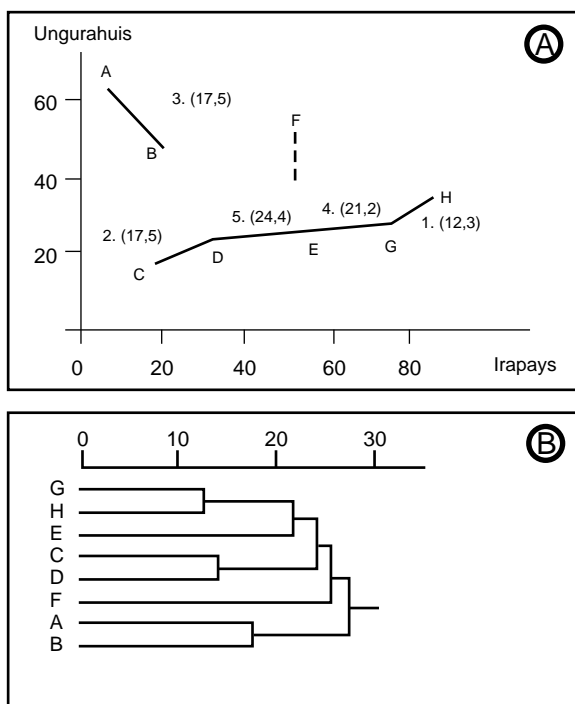


Figura 16. La ordenación de ocho parcelas (A-H) según el número de individuos de ungurahui e irapay.

La primera fase en la clasificación del vecino más cercano es arreglar las distancias medidas entre las parcelas en un orden según sus valores. Luego se toma la distancia más pequeña del listado y se juntan las dos parcelas. A continuación se toma la segunda distancia más corta y se juntan las parcelas respectivas a un grupo. En caso de que una de estas parcelas ya perteneció al primer grupo, van todas las parcelas de ese grupo también al grupo nuevo. Así se continúa hasta que todas las parcelas pertenezcan a un solo grupo (Figura 17A). Se puede decir que cada parcela va al grupo donde se encuentra su vecino más cercano. Cuando se piensa cuales son los dos grupos más cercanos, la distancia entre grupos siempre es la misma que la distancia más pequeña entre cualquiera de los dos miembros de los dos grupos.

Uno puede presentar los pasos tomados en la clasificación con un dendrograma (Figura 17B). La situación inicial está al lado izquierdo, donde cada parcela forma su propio grupo, y la distancia dentro de cada grupo es cero. Moviéndose hacia la derecha, la distancia dentro de cada grupo crece. Podemos imaginar el dibujo de un círculo creciente alrededor de cada parcela en la Figura 17A. Cuando el radio de los círculos crece lo suficiente, dos parcelas quedan al final dentro de un solo círculo, y en este momento estas parcelas se unen en la distancia que es igual al radio del momento. Para indicar la unión, se juntan los vecinos con una línea vertical en el dendrograma. Después se continúa la ampliación de los círculos hasta que se encuentran los siguientes vecinos más cercanos, y así sucesivamente.



**Figura 17.** La clasificación de las parcelas de la Figura 16 con el método del vecino más cercano. **A)** Primero se juntan las parcelas más cercanas G y H (al costado de la línea que une las parcelas se ha escrito el orden de la distancia, 1.) en la distancia de 12,3 (entre paréntesis), en segundo lugar se juntan C y D, y en tercer lugar A y B. La cuarta unión lleva la parcela E al grupo donde ya están G y H, y la quinta junta los grupos donde están D y E. En el próximo pase (marcado con línea cortada) F se une al grupo que contiene E, porque la distancia F-E es la más corta entre grupos. **B)** Un dendrograma que cuenta la historia de las uniones. La distancia de cada línea horizontal desde la margen izquierda indica la distancia de unión entre parcelas o grupos de parcelas.

A veces se consiguen jerarquías con niveles claramente diferentes, y en estos casos es relativamente fácil determinar cuántas clases se puede distinguir en los datos. Pero muchas veces no es nada sencillo decidir el número de clases. Es muy subjetivo si el investigador interpreta, por ejemplo, la Figura 17 de tal manera que existan dos clases (A+B como una clase y el resto como otra), o si se piensa que existan tal vez cuatro clases (A+B; F; C+D; G+H+E). Cuando uno está identificando las clases, se puede, por ejemplo, decidir que sólo esas clases que se forman antes de una cierta distancia van a ser consideradas. Así, una distancia límite de 23 resultaría en la identificación de cuatro clases en el caso de la Figura 17. Aunque el número de clases es un asunto de opinión, los métodos de clasificación en sí son objetivos y repetitivos.

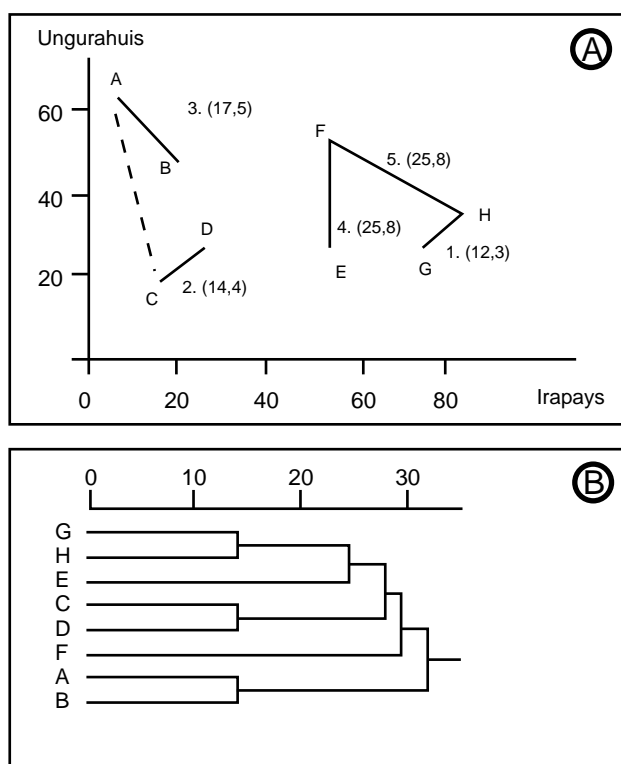
El dendrograma en sí no indica nada acerca de la significación o relevancia ecológica de los grupos. Esto es algo que uno tiene que estimar siempre aparte, con base en las observaciones medioambientales de las unidades de muestreo.

La clasificación del vecino más cercano llega fácilmente a una situación en la cual no aparecen estructuras jerárquicas claras, pero las parcelas se juntan una tras otra a las clases existentes, formando cadenas en las cuales pueden existir parcelas muy diferentes. Por eso este método no es muy recomendable para, por ejemplo, una clasificación de vegetación. Por otra parte, la encadenación da otra información de los datos que puede ser útil en estudios de gradientes.

**Clasificación del vecino más distante (*furthest neighbour clustering, o complete linkage clustering*)**

La clasificación del vecino más distante está basada en el mismo principio de la clasificación del vecino más cercano, poniendo primero las distancias entre parcelas en un orden decreciente de distancia, empezando a juntar las parcelas desde la distancia más corta. La diferencia está en que en la comparación de grupos con más de una parcela, se determina la distancia entre grupos según la distancia más grande (no más corta) que se encuentra entre dos parcelas de diferentes grupos (Figura 18). Si se piensa esto en términos de los círculos crecientes, dos grupos se unen sólo cuando todas sus parcelas están dentro de un solo círculo.

En esta clasificación los grupos tienden a formarse muy compactos, redondos si uno mira la situación en un diagrama como la Figura 18A. La clasificación del vecino más lejano es un buen método si uno quiere definir en sus datos las clases que están claramente separadas (Figura 18B).



**Figura 18.** La clasificación de las parcelas de la Figura 16 con el método del vecino más lejano. La figura está preparada de igual manera que la Figura 17. **A)** Las tres primeras uniones son iguales a la clasificación anterior, pero la cuarta unión lleva la parcela E junto con F, porque la distancia entre éstas es más corta que la distancia entre E y H. Después se une el grupo A+B con el grupo C+D, porque la distancia A-C (47,4) es más corta que A-H y C-H. **B)** El dendrograma que describe la historia de las uniones.

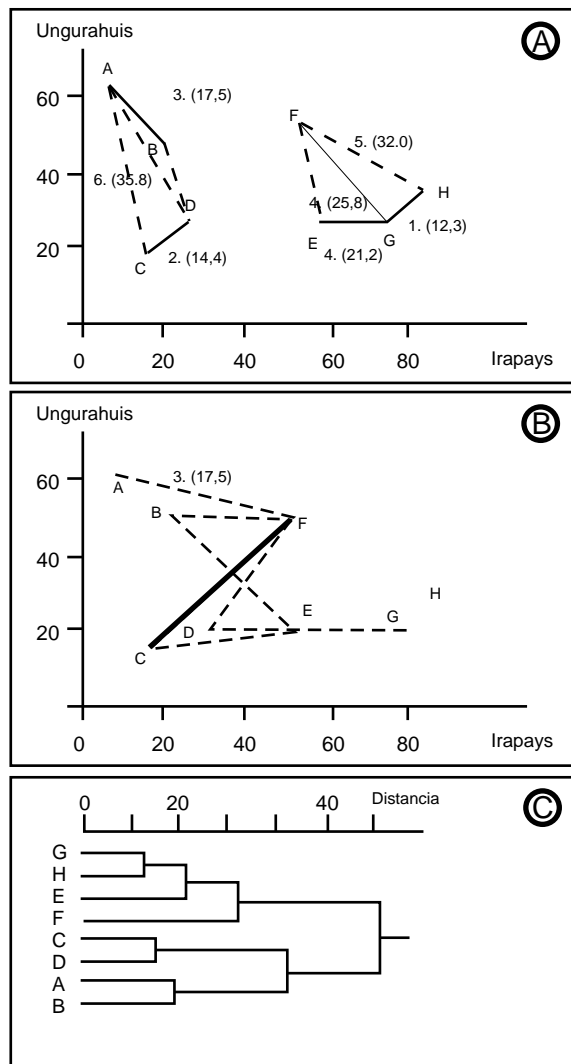
**Clasificación del vecino relativo (*intermediate linkage clustering, o proportional link linkage*)**

Las dos clasificaciones anteriores representan, de cierta manera, lógicas extremas. Una junta dos grupos enseguida cuando existe sólo una parcela de ambos grupos cercana entre sí, mientras la otra clasificación acepta la unión de dos grupos sólo si todas las parcelas de los grupos están dentro de una distancia corta.

La clasificación del vecino relativo es un método entre los dos extremos. En esta clasificación el usuario puede decidir qué porcentaje de los enlaces entre dos grupos (distancias entre parcelas en dos grupos diferentes) hay que tomar en cuenta cuando se define la distancia entre estos dos grupos. El grado de enlace (en inglés *connectedness*) puede variar entre 0 y 100%.

Por ejemplo, si se comparan dos grupos con tres parcelas cada uno, existen nueve enlaces entre estos grupos; de cada parcela del primer grupo se puede calcular la distancia (enlace) a tres parcelas en el otro grupo. Estas distancias se colocan en orden según su longitud. Si se ha definido que el grado de enlace es 50%, se divide las distancias ordenadas por la mitad, y se toma la máxima distancia de la mitad que tiene distancias menores. Si hay un número impar de distancias, la distancia entre los grupos será la distancia mediana. Uno puede pensarlo también de este modo: los grupos se unen en la distancia que produce un círculo suficientemente grande como para abarcar el 50% de los enlaces entre los dos grupos.

El grado de enlace en la clasificación del vecino más cercano es 100%, y en la clasificación del vecino más lejano el grado de enlace es 0%. En la Figura 19 se presenta una clasificación que tiene el grado de enlace de 50%.



**Figura 19.** Una clasificación de las parcelas de la Figura 16 con el método del vecino relativo, cuando el grado de enlace es 50%. La figura está construida con el mismo principio que las Figuras 17 y 18. **A)** Las cuatro uniones primeras se hacen de la misma manera que la clasificación del vecino más cercano. En la quinta unión, la parcela solitaria F se une a la clase E+G+H en la distancia 32,0, que es la distancia mediana de todos los enlaces (otros enlaces están marcados con una línea cortada). En la sexta unión la clase A+B se une con la clase C+D. La distancia de esta unión es 35,8, porque éste es el segundo más corto de los cuatro enlaces entre estas clases. **B)** Las dos últimas clases (en los dos círculos) se unen con el enlace dibujado con línea gruesa. Los enlaces más cortos aparecen con líneas cortadas, y los enlaces más largos no están dibujados. **C)** El dendrograma que presenta la historia de las uniones.

**Clasificación por el promedio aritmético (arithmetic average clustering; UPGMA y WPGMA)**

En todas las clasificaciones de vecino, la distancia entre los grupos está determinada por una de las distancias existentes entre las parcelas. En la clasificación por el promedio aritmético la situación es diferente, pero siempre después de la formación de un grupo nuevo se calcula una distancia promedio desde este grupo a todos los otros grupos (o parcelas solitarias) con base en las distancias de los enlaces que este grupo tiene con los otros. La próxima unión se hace en la distancia promedio más corta entre los grupos.

Uno puede calcular el promedio de las distancias de los enlaces de dos maneras: bien sea considerando que cada enlace tiene el mismo peso, o bien que cada grupo tiene el mismo peso independiente del número de parcelas que contiene. En el primer caso el promedio es simplemente la suma de las distancias de enlaces dividida por el número de enlaces (Tabla 11, Figura 20). La abreviatura de este método es UPGMA (en inglés *Unweighted Pair-Group Method using arithmetic Averages* o *unweighted arithmetic average clustering*). En el otro caso, se calcula el promedio de distancias de los enlaces de tal manera que el peso de cada enlace disminuye en relación con el número de agrupamientos en los cuales ya se ha participado (Tabla 12). Este método usa la abreviatura WPGMA (en inglés *Weighted Pair-Group Method using Arithmetic averages* o *weighted arithmetic average clustering*).

**Tabla 11.** Las distancias calculadas en la clasificación UPGMA de las parcelas A-E de la Figura 16. En la tabla superior están las distancias euclidianas originales. La distancia mínima, que es la distancia que resulta de la unión, está escrita en negrita en cada tabla. Debido a las uniones, cada tabla nueva es siempre más pequeña que la anterior. Las distancias nuevas después de la unión, más las formas de calcularlas, están representadas en cada tabla nueva. El dendrograma que sale de estas distancias está representado en la Figura 20.

	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
A	17,5	47,4	43,0	57,0
B		35,8	27,9	39,2
C			14,4	37,0
D				24,4

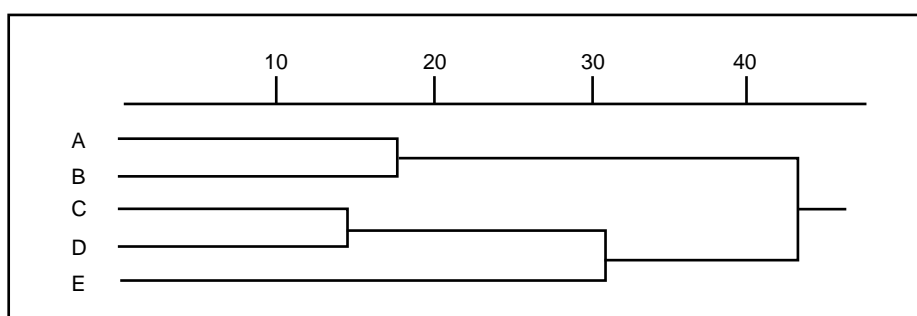
	<b>B</b>	<b>C+D</b>	<b>E</b>
A	17,5	42,2=(47,4+43)/2	57,0
B		31,85=(35,8+27,9)/2	39,2
C+D			30,7=(37+24,4)/2

	<b>C+D</b>	<b>E</b>
A+B	38,53= (47,4+43+35,8+27,9)/4	48,1=(57+39,2)/2
C+D		30,7=(37+24,4)/2

	<b>C+D+E</b>
A+B	41,72 =(47,4+43+35,8+ 27,9+57+39,2)/6



**Figura 20.** El dendrograma de la clasificación UPGMA de la Tabla 11.

UPGMA es uno de los métodos de clasificación más usados en estudios biogeográficos y ecológicos, porque puede mantener relativamente bien la estructura de la matriz original de distancias entre las unidades de muestreo. La precondition para usar este método es que el muestreo haya sido sistemático o aleatorio. En caso de que los datos estén representados de una manera sesgada, las diferentes comunidades biológicas o áreas geográficas con más unidades de muestreo determinarían demasiado el resultado de la clasificación. En estos casos es mejor usar el método de WPGMA.

**Tabla 12.** Las distancias calculadas en la clasificación WPGMA de las parcelas A-E de la Figura 16. En la tabla superior están las distancias euclidianas originales. La distancia mínima que resulta de la unión está escrita en negrita en cada tabla. Debido a las uniones, cada tabla nueva es siempre más pequeña que la anterior. Las distancias nuevas después de la unión, más las formas de calcularlas, están representadas en cada tabla nueva. Aquí, el promedio de distancias ha sido calculado con base en los promedios de la fase anterior de clasificación, y no con las distancias originales de la tabla anterior. La diferencia con UPGMA viene aparentemente sólo en las dos últimas tablas.

	B	C	D	E
A	17,5	47,4	43,0	57,0
B		35,8	27,9	39,2
C			14,4	37,0
D				24,4

	B	C+D	E
A	<b>17,5</b>	$42,2=(47,4+43)/2$	57,0
B		$31,85=(35,8+27,9)/2$	39,2
C+D			$30,7=(37+24,4)/2$

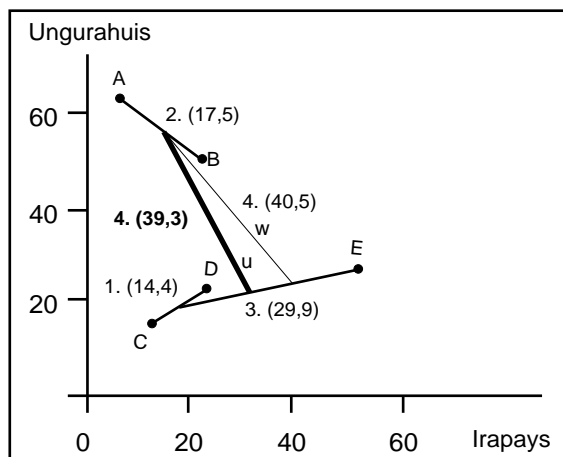
	C+D	E
A+B	$38,53=(45,2+31,85)/2$	$48,1=(57+39,2)/2$
C+D		<b>30,7</b> $=(37+24,4)/2$

	C+D+E
A+B	<b>43,31</b> $=(38,53+48,1)/2$

**Clasificación por el centroide (centroid clustering; UPGMC y WPGMC)**

La clasificación por el centroide es muy semejante a la clasificación por el promedio aritmético, pero en lugar de calcular distancias promedio entre los grupos, se calcula la posición central dentro del grupo. Éste es el punto de gravedad del grupo (centroide), y la distancia entre grupos es la misma que la distancia entre los centroides. Como en el caso con la clasificación por el promedio aritmético, también en la clasificación por el centroide hay dos formas diferentes. En la UPGMC (en inglés *unweighted centroid clustering*), cada unidad de muestreo tiene el mismo peso en el proceso de la determinación del centroide. En otras palabras, cuando se unen dos grupos, el centroide del nuevo grupo está siempre más cerca al subgrupo que tuvo más unidades de muestreo (que tuvo más peso). En la WPGMC (en inglés *weighted centroid clustering*) las unidades de muestreo en el grupo más grande tienen peso individual más pequeño, de tal manera que ambos subgrupos tienen exactamente el mismo peso. Entonces, el centroide del nuevo grupo está exactamente en el centro de la distancia entre los dos centroides de los subgrupos (Figura 21).



**Figura 21.** La clasificación por el centroide de las 5 parcelas de la Tabla 12. El orden de la unión está indicado al costado de cada línea que representa una unión. La distancia de la unión está entre paréntesis. Las líneas de unión están dibujadas entre los centroides, porque la clasificación se basa en distancias entre éstos. Después de la tercera unión se consigue la clase C+D+E, la cual tiene su centroide en posiciones diferentes en las dos formas de clasificación por el centroide. Por eso, en la cuarta unión se obtiene distancias diferentes en las dos formas de clasificación (negrita se refiere a UPGMC). Los centroides después de la cuarta unión están marcados con cuadrángulos parados (u=UPGMC, w=WPGMC).

Los centroides se mueven durante el análisis, y por eso es posible que después de una unión de dos grupos, su distancia a un tercer grupo sea más corta que la que cualquiera de los dos primeros grupos tuvo al tercero antes de la unión. Esto causa problemas al dibujar dendrogramas (en inglés *reversal*).

De igual modo que la clasificación por el promedio, es recomendable utilizar la forma de pesos iguales de las parcelas de la clasificación por el centroide (UPGMC) sólo cuando se ha colectado los datos de una manera sistemática o aleatoria. En los otros casos es mejor usar WPGMC, si es que uno quiere generalizar los resultados.

#### **Clasificación restringida a los vecinos (*constrained clustering, chronological clustering*)**

Si se desea, se puede restringir la clasificación de tal manera que sólo las unidades de muestreo que son vecinas en tiempo o en espacio puedan ser juntadas al mismo grupo. Esta opción es especialmente interesante si queremos buscar solamente los puntos de discontinuidad en los datos. Por ejemplo, en un estudio sobre la sucesión vegetal puede ser interesante saber si la vegetación se desarrolla de una manera gradual, o si hay saltos más rápidos interrumpidos por relativa estabilidad. De igual manera, puede ser interesante saber si la composición de especies cambia a lo largo de una línea de estudio en ciertos puntos. Las clases distinguidas en una clasificación restringida forman unidades más compactas que las clases de una clasificación sin restricción, y por eso puede ser también que los grupos de una clasificación restringida sean más fáciles de interpretar.

La clasificación restringida a los vecinos puede basarse en cualquier método aglomerativo (y jerárquico) de clasificación. Únicamente se introduce la restricción de que sólo las unidades vecinas de muestreo pueden ser agrupadas en la misma clase. Dos unidades semejantes de muestreo que sean situadas más lejos entre sí pueden entrar en la misma clase sólo si todas las unidades intermedias también son lo suficientemente semejantes como para entrar a la misma clase.

El usuario puede libremente definir cuáles son las unidades vecinas de muestreo. Normalmente, en las líneas o series de tiempo los vecinos son solamente las unidades próximas (uno antes y uno después). Cuando el muestreo ha sido realizado de una manera bidimensional, se puede definir la vecindad de más de una forma. A veces puede ocurrir, especialmente en estudios acuáticos, que las unidades de muestreo se encuentren en un espacio tridimensional, y en este caso el vecino puede estar tanto al costado como arriba o abajo.

Normalmente las clasificaciones no contienen ninguna prueba estadística. Sin embargo, en la clasificación restringida es fácil incluir una prueba estadística. La hipótesis nula es que todas las unidades de muestreo pertenecen a la misma población de unidades, y que crear clases separadas es solo un acto artificial creado por el algoritmo de clasificación. Con una prueba de aleatorizaciones se puede estimar si se puede descartar la hipótesis nula, si las diferencias en la composición de especies entre grupos son más grandes que dentro de los grupos (el principio de una prueba de aleatorizaciones se explica abajo junto a la prueba de Mantel).

En el análisis uno tiene que definir un nivel de significancia ( $\alpha$ ), que mide el riesgo de equivocarse si uno rechaza la hipótesis nula. Si se ha puesto el  $\alpha$  a un valor muy pequeño (por ejemplo 0,001), sólo se acepta los cambios más fuertes en la composición de especies para definir límites entre clases. Pero si el  $\alpha$  es grande (por ejemplo 0,2), unas diferencias relativamente pequeñas en la composición de especies causarían la separación de diferentes clases.

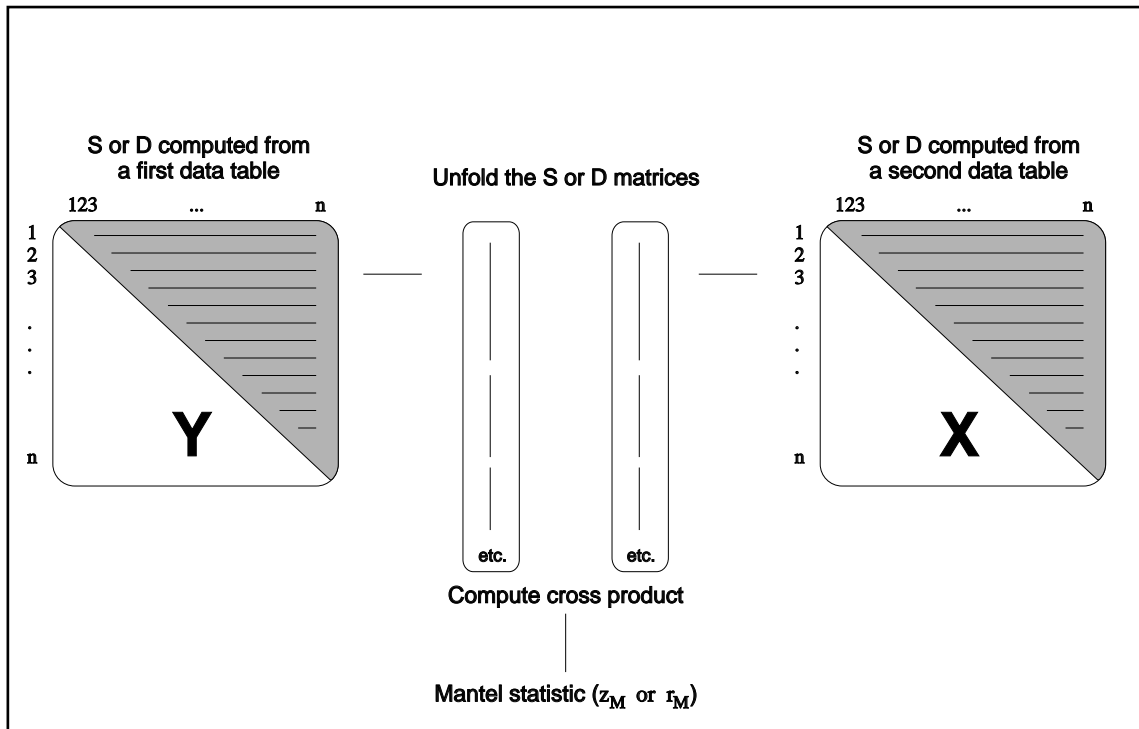
La clasificación restringida está siempre basada en una clasificación jerárquica, pero a pesar de eso el resultado es un listado no jerárquico de clases. Hay que recordar que el interés verdadero está en resolver los puntos de discontinuidad en los datos. Claro que después de haber identificado estos puntos, uno puede examinar si una clase está dispersa entre varias unidades distinguidas por la clasificación restringida. Por ejemplo, es posible que una línea de investigación florística cruce el mismo tipo de flora en varias áreas separadas.

## **7. LA PRUEBA DE MANTEL: CORRELACIÓN ENTRE MATRICES DE DISTANCIA**

La prueba de Mantel es útil para conocer, por ejemplo, si las diferencias florísticas entre las parcelas de estudio están relacionadas con diferencias medioambientales. En caso de que exista un coeficiente de correlación de Mantel ( $r_M$ ) positivo, la interpretación significa que cuanto más parecida es la flora entre dos parcelas, más parecidas son también las condiciones medioambientales. Si el coeficiente es negativo, la interpretación en principio es que cuanto más parecida es la flora entre dos parcelas, más diferente es el medio ambiente. Sin embargo, esta última interpretación no tiene ningún sentido ecológico, y por eso en términos prácticos la interpretación es que no hay correlación entre las semejanzas florísticas y medioambientales.

Para calcular la prueba de Mantel se necesita dos matrices de distancia o similitud, X e Y (Figura 22). Es recomendable usar o matrices de distancia o de similitud y no mezclarlos en el mismo análisis. Cuando se calcula la correlación entre una matriz de distancia y otra de similitud, el signo del coeficiente de correlación es opuesto a lo que se espera intuitivamente, como presentamos en el párrafo anterior (no afecta al valor numérico del coeficiente).





**Figura 22.** El principio de la prueba de Mantel. Primero se transforman ambas matrices a vectores, de tal manera que las líneas de las matrices se transponen una tras otra a dos columnas. Después se calcula la correlación entre las columnas (normalmente la correlación de Pearson). **Fuente:** Legendre & Legendre 1998, p. 553.

En ambas matrices X e Y tiene que existir el mismo número de unidades de muestreo en el mismo orden. Los índices de distancia pueden ser diferentes en las dos matrices. Por ejemplo, la matriz Y puede usar el índice de Jaccard o Steinhaus, convertido a distancia, y la matriz B puede presentar distancias euclidianas con base en variables medioambientales.

Las variables en las matrices X e Y tienen que ser independientes una de otra. Por ejemplo, si queremos estudiar si dos grupos de plantas dan información parecida del patrón florístico entre parcelas de estudio, tenemos que tener presente que los dos grupos estén separados, sin especies que aparezcan en ambos grupos. Por ejemplo, una comparación entre palmeras y gramíneas es posible, pero otra entre palmeras y monocotiledóneas no lo es. Todas las especies de palmeras pertenecen también a las monocotiledóneas.

Originalmente el término estadístico de Mantel fue determinado simplemente como la suma de los productos entre cada valor de distancia, medido entre las mismas unidades de muestreo:

$$z = \sum_{i=1}^{n-1} \sum_{j=i+1}^n x_{ij} y_{ij} \quad \text{donde } i \text{ y } j \text{ son los números de líneas y columnas en la matriz de distancia.}$$

Este término estadístico puede obtener cualquier valor positivo, y por eso es difícil interpretarlo. Más comúnmente se utiliza la forma estandarizada de la correlación de Mantel, que varía entre -1 y 1:

$$r_M = \frac{1}{d-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{x_{ij} - \bar{x}}{s_x} \frac{y_{ij} - \bar{y}}{s_y}$$

en la fórmula  $d = \frac{n(n-1)}{2}$  es la cantidad de distancias en el triángulo superior de la matriz de distancia sin la diagonal.

En efecto, la fórmula de la correlación estandarizada de Mantel es la misma que la de correlación de Pearson, y por eso es fácil ver que la correlación de Mantel mide una correlación lineal entre las matrices. Uno puede también calcular otro tipo de correlaciones entre las distancias en las matrices para estudiar correlaciones de otras formas. Por ejemplo, con la correlación de Spearman se puede estudiar si existe una relación monótona entre las matrices.

Aunque la prueba de Mantel está basada en coeficientes de correlación, hay que tener presente que la correlación está calculada entre las distancias y no entre variables originales observadas en el campo. Claro que tampoco será posible calcular una correlación de Pearson o Spearman de tal manera que exista más que dos variables. En caso de que uno obtenga sólo dos variables de los datos de campo, y calcule las correlaciones de Mantel y de Pearson con base en estas variables, se consiguen valores diferentes. Comúnmente la correlación de Mantel (correlación entre distancias) tiene un valor más bajo.

Una parte esencial de la prueba de Mantel es el cálculo de una significancia estadística para la correlación observada. La significancia tiene que ser estimada utilizando aleatorizaciones, porque las distancias dentro de una matriz no son completamente independientes entre sí, y por eso las distancias no cumplen el requerimiento de independencia de observaciones del cálculo de significancia estadística para la correlación de Pearson o Spearman. La dependencia de las distancias en la matriz es evidente, porque si se sabe que, por ejemplo, la parcela 1 tiene casi la misma flora que las parcelas 2 y 3, entonces existe una buena razón para esperar que las parcelas 2 y 3 también tengan floras iguales.

El principio de la prueba de aleatorizaciones (prueba Monte Carlo o en inglés *permutation test*) es que si la hipótesis nula es verdadera, la correlación observada entre las matrices X e Y no puede ser más grande que una correlación típica entre las matrices, en el caso de que se hubiera mezclado (aleatorizado) el orden de parcelas en una de las matrices. Podemos probar esta hipótesis aleatorizando el orden de una de las matrices varias veces (por ejemplo 999 veces), y calculando el valor de la correlación entre X e Y después de cada aleatorización. Así, es posible comparar el valor de la correlación original con los valores obtenidos a través de las aleatorizaciones. El riesgo de rechazar la hipótesis nula, aunque en realidad sea verdadero, se calcula de la manera siguiente:

**El número de correlaciones iguales o mayores al original**  
**Número total de aleatorizaciones+1**

La correlación original (observada) se incluye en el cálculo. En el dividendo está el "número de correlaciones iguales", y el divisor está representado por el término "+1". Por ejemplo, si la correlación original era más grande que todas las correlaciones obtenidas a través de la aleatorización de 999 veces, es la correlación original estadísticamente significativa con el riesgo de  $p < 0,001$ .

Es posible hacer la prueba de Mantel con tres matrices, y en este caso se habla de una prueba parcial de Mantel (en inglés *partial Mantel test*). Con una prueba parcial de Mantel se puede examinar si existe una correlación entre las matrices X e Y después de haber tomado en cuenta la correlación que estas dos matrices tienen con una tercera matriz Z. Con esta prueba uno puede investigar, por ejemplo, si las distancias geográficas entre las parcelas pueden explicar toda de la correlación que se ha observado entre las distancias florísticas y medioambientales. La matriz Z es en este caso la matriz de distancias geográficas. En caso de que existe una correlación estadísticamente significativa entre la flora y el medio ambiente todavía después de haber tomado en cuenta el efecto de las distancias geográficas, uno puede decir que la correlación entre flora y medio ambiente es por lo menos parcialmente independiente del efecto de distancias geográficas entre las parcelas.





Metodologías  
de muestreo  
para los grupos  
indicadores



## 1. INTRODUCCION

Existen diferentes metodologías para el estudio de la flora y la fauna. Cada una de éstas tiene un fundamento y puede ser aplicada según el tipo de estudio que se quiera hacer (objetivos), y según las condiciones para hacerlo (accesibilidad, logística). Así por ejemplo, el establecimiento de parcelas, parcelas con sub parcelas de áreas variables, trochas en espiral, trochas radiales, puede ser empleado para inventarios de flora, donde se registran individuos a partir de cierta medida en el diámetro a la altura del pecho (DAP), incluyendo o no individuos arbustivos o epífitos.

Los inventarios de aves pueden ser realizados también de diferentes maneras: a través de trochas elegidas al azar o sistemáticamente y de una longitud determinada, por observaciones directas con caminatas al azar, mediante el establecimiento de un sistema de trochas y aplicando en él esfuerzos de muestreo constantes y estandarizados, utilizando trampas para la captura, entre otros métodos. Los anfibios anuros pueden ser inventariados mediante colectas diurnas y/o nocturnas en parcelas, transectos o trochas de cierta longitud y anchura, con búsqueda intensiva durante un intervalo de tiempo, y mediante grabaciones de cantos u otros métodos. Los insectos, debido a la gran cantidad de especies, géneros o familias que existen, pueden ser estudiados mediante el uso de trampas que capturan gran cantidad de individuos o utilizando trampas específicas para cierto grupo. Los mamíferos pueden ser inventariados con metodologías de captura, censos y registro de huellas, bien sea en transectos, parcelas, en un sistema de trochas, o al azar.

La presente guía está enfocada en métodos de inventario en seis grupos taxonómicos: palmeras, pteridofitas, melastomátáceas, aves, anfibios anuros e insectos de hojarasca (hormigas), y está basada en los estudios de campo del Proyecto Diversidad Biológica de la Amazonía Peruana (BIODAMAZ), Perú-Finlandia. Para estos estudios se seleccionó una metodología para el inventario de plantas y las más adecuadas, según los propósitos del proyecto, para los inventarios de aves, anfibios anuros e insectos de hojarasca (hormigas).

## 2. CRITERIOS PARA UBICAR LOS LUGARES DE MUESTREO

El uso de imágenes de satélite es un factor muy importante y determinante para la selección de lugares de muestreo. El proyecto BIODAMAZ confeccionó un mosaico de imágenes de satélite que cubre la selva baja amazónica peruana. El mosaico está disponible en el Internet a través del Sistema de Información de la Diversidad Biológica y Ambiental de la Amazonía Peruana SIAMAZONIA ([www.siamazonia.org.pe](http://www.siamazonia.org.pe); ver también BIODAMAZ, 2004b). Dicho mosaico permite formular una hipótesis sobre los patrones de distribución de especies de plantas y animales. Como los patrones de coloración en las imágenes de satélite reflejan algún tipo de diferencias en la superficie, se puede esperar que las manchas del mismo color (mismas características de reflectancia de la superficie) representan hábitats similares, y las manchas de otro color corresponden a hábitats diferentes. Consecuentemente, manchas diferentes deben de tener una flora y fauna algo diferente, mientras que las manchas del mismo color deben ser iguales tanto florística como faunísticamente. Esta regla se espera que funcione relativamente bien sobre distancias cortas (ca. 100 km), mientras que sobre distancias más largas puede ser que factores evolutivos e históricos hagan que la flora y fauna puedan ser diferentes, aunque físicamente el ambiente tal vez sea igual.



**Figura 23.** Las imágenes de satélite son una herramienta importante en la planificación de los inventarios florísticos y faunísticos y en la ubicación de sitios de muestreo.

El proyecto BIODAMAZ hizo una primera estimación sobre la diversidad de vegetación en la selva baja peruana, identificando y caracterizando 24 unidades vegetales en la región amazónica, que corresponden a los patrones y características que se puede visualizar en el mosaico de las imágenes de satélite (BIODAMAZ, 2004a). Además, estas unidades fueron presentadas como un mapa de diversidad de vegetación que está también disponible en SIAMAZONIA.

Sin embargo, además de los parches visibles en las imágenes de satélite, se puede considerar también el conocimiento geológico cuando se planifica la ubicación de los lugares de muestreo. La razón para considerar la geología es que el origen geológico del suelo determina mucho las características pedológicas relevantes para el crecimiento de las plantas (Linna *et al.* 1998). Como el suelo afecta mucho la composición de especies de plantas, por lo menos localmente (Ruokolainen *et al.* 1997, Ruokolainen & Tuomisto 1998), se puede esperar que este efecto esté presente también en una escala más gruesa. En la escala gruesa, o macro, uno puede considerar a la Amazonía Peruana como un paisaje con un número relativamente pequeño de gradientes geológicos.

La Amazonía Peruana está compuesta por algunas áreas que están hundiéndose, divididas por otras áreas de levantamiento (Räsänen *et al.* 1990, 1992, Räsänen 1993). En las áreas de hundimiento se encuentran sedimentos recientes en la superficie, porque son áreas donde la sedimentación fluvial está activa ahora. Desde estas áreas hacia las áreas de levantamiento se encuentran en la superficie sedimentos más antiguos, cuanto más lejos se avanza desde los centros de hundimiento.

Un ejemplo de esto se puede ver entre las ciudades de Nauta e Iquitos (Räsänen *et al.* 1998). Nauta está al borde de un centro grande de hundimiento, la cuenca de Ucamara (ríos Marañón y Ucayali), donde se encuentra la Reserva Nacional Pacaya Samiria. Iquitos, por su parte, está en una zona de levantamiento, o por lo menos en una zona que no se ha hundido desde hace mucho tiempo (arco de Iquitos). La consecuencia de esto es que entre Iquitos y Nauta existe un gradiente de edad de los sedimentos superficiales, de tal manera que la zona de Nauta tiene sedimentos más jóvenes y la zona de Iquitos sedimentos más viejos. Cerca de Nauta se ven sedimentos del Mioceno tardío, de la época final del lago Pebas (Hoorn 1993), cuando el lago desapareció por lo menos de esa zona, y el sistema sedimentológico se convirtió de lacustre a fluvial. Hacia Iquitos se encuentra sucesivamente sedimentos más antiguos, correspondiendo a las épocas del lago Pebas, con sus conexiones con el agua salobre del mar. En términos físicos y químicos del suelo, este gradiente se nota de tal manera que cerca a Nauta el suelo es mucho más pobre en nutrientes (no hay efecto de los sales ni de los fósiles del mar), y contiene más limo y arena (un sistema fluvial tiene más fuerza que un sistema lacustre).

Esta clase de gradientes geológicos, que supuestamente se expresan también en la calidad del suelo, existe en toda la Amazonía Peruana. Claro que el patrón no es en realidad tan sencillo como el ejemplo de arriba puede hacer creer. Aunque existen gradientes obvios sobre distancias largas, localmente se encuentra siempre sedimentos más recientes en la superficie que parcialmente esconden el gradiente general. También es posible que algún río haya erosionado la superficie en otra zona mucho más que en los alrededores de esta zona, y por eso, otra vez, el gradiente va a tener alteraciones locales. Pero a pesar de estas alteraciones, uno puede esperar que existan gradientes ambientales en la Amazonía, creados por los gradientes geológicos. Por esta razón, es aconsejable planificar el muestreo de tal manera que siga la dirección del gradiente. Si se observa un patrón predecible según el gradiente, se podrá hacer con bastante certeza extrapolaciones de los patrones de distribución de especies en áreas donde no se ha estudiado la flora o la fauna pero se conoce el patrón geológico.

Una vez aclarados los criterios científicos para ubicar las zonas de muestreo, queda todavía la tarea de buscar y determinar las áreas donde realmente se va a hacer el muestreo. En términos de la metodología científica, lo mejor sería que primero se ubicase todas las áreas que cumplen los criterios científicos de un área de estudio, y después se escogiese aleatoriamente los sitios de muestreo en el terreno. Pero en términos prácticos, por lo menos en la Amazonía, hay que considerar también la factibilidad logística. Por eso, por ejemplo en el proyecto BIODAMAZ, se seleccionó las áreas de muestreo tomando en cuenta las áreas de gradientes geológicos y la facilidad de acceso a ellas.

En el proyecto BIODAMAZ se seleccionó las áreas de muestreo con el criterio de que representasen gradientes geológicos conocidos en condiciones climáticas algo diferentes y que tengan acceso fácil. Dentro de estas áreas se ubicó los puntos de muestreo en las imágenes de satélite, tratando de abarcar tanta variación en las imágenes como fuera posible. Al llegar en el terreno a una de las manchas elegidas para ser muestreada, se cuidó de no realizar los inventarios muy cerca de chacras (una distancia mínima de unos 200 m) ni de ríos, porque se trató de mantener siempre los inventarios dentro del bosque de tierra firme.

### 3. METODOLOGÍA PARA EL INVENTARIO DE PLANTAS

#### Establecimiento del transecto de muestreo

Una vez seleccionado el punto de inicio y la dirección del transecto (un lugar no alterado y una dirección que permite mantener el muestreo dentro de una mancha visible en las imágenes de satélite), se abre una trocha central de 500 m de largo. La trocha debe tener el menor grado posible de intervención, lo suficiente como para poder seguirla y medir su topografía con un clinómetro (Figura 24).



**Figura 24.** Midiendo la topografía de un transecto con un clinómetro. *Foto: Sanna-Kaisa Juvonen.*

El transecto (Figura 25) se establece de la siguiente manera:

1. Se coloca una estaca de ca. 1,5 m en el inicio del transecto.
2. Se fija el cabo del metro en el suelo, en el punto de inicio, con la ayuda de un palito (el punto 0 del metro debe coincidir con la estaca).
3. Uno de los trocheros avanza, controlando la dirección con la brújula, abriendo la trocha y llevando el metro. En la instrucción del uso de la brújula hay que tener presente que no se debe portar machetes u otros objetos de metal en el momento de tomar la dirección.
4. Otro trochero sigue al primero colocando una estaca pequeña (60-70 cm) cada 5 m y una estaca grande (1-1,5 m) cada 25 m. Debe limpiarse la corteza del extremo superior de las estacas y colocar una cinta marcadora de color rojo, anaranjado o azul, por ejemplo, para facilitar la visualización. Se escribe en la estaca grande el metraje recorrido. En una zona accidentada, la cinta métrica tiene que seguir el terreno y no cruzar libremente las quebradas.
5. Cuando el primer trochero (que porta el metro) ha avanzado 50 m, colocará allí una estaca de 1-1,5 m escribiendo el metraje en la estaca. En caso que el otro trochero todavía no haya avanzado 50 m, el primero le ayudará a poner las estacas pequeñas.
6. Cuando todas las estacas pequeñas están colocadas cada 5 m y las estacas grandes con sus metrajes escritos, se inicia la próxima medida siguiendo los pasos 1 a 5 hasta marcar 500 m de trocha.



**Figura 25.** Los investigadores estudian plantas a lo largo de un transecto establecido. *Foto: Juan José Rodríguez.*



### El registro y el muestreo de plantas

El registro y el muestreo de plantas (Figura 26) se hace dentro de un transecto de 500 x 5 m, dividido en sub unidades de 25 x 5 m. Se usa la trocha y las estacas marcadas para ubicar la línea central del transecto. Se registran todos los individuos de las especies de plantas de interés ubicadas dentro una faja de 2,5 m contados desde la línea central a ambos lados. Para determinar los 2,5 m se puede utilizar un palo de esa longitud o usar un metro (hasta que el ojo se acostumbre a medir).

Para cada zona de muestreo (una zona de muestreo contiene aproximadamente diez transectos que se establecen en una salida de campo), se colecta por lo menos dos muestras representativas de cada especie, con la excepción de palmeras grandes muy conocidas (*Euterpe precatória*, *Iriartea deltoidea*, *Mauritia flexuosa*, *Oenocarpus bataua* y *Socratea exorrhiza*). Una muestra de la especie se colecta siempre cuando está encontrada por primera vez. Con la muestra debe registrarse características de la planta que no van a aparecer de la muestra misma: altura del individuo, lugar de crecimiento (colina, valle, claro, entre otros), hábito (arbusto, liana, epífita, palmera acaule, entre otros); en palmeras y helechos grandes, dimensiones de la hoja, y en palmeras, tipo de venación, textura y posible coloración.

En el Anexo se presenta un ejemplo de una ficha de campo para el registro de plantas dentro de un transecto de 20 sub unidades.

Una muestra consiste de un triplicado, o de cinco copias si presentan flores y/o frutos. Cuando la muestra tiene flores o frutos debe registrarse la coloración. El material colectado debe ser suficiente para considerar las copias.



Figura 26. Inventario de plantas. Fotos: Sanna-Kaisa Juvonen.

Para la colecta de helechos debe considerarse sólo los individuos que tengan por lo menos una hoja mayor o igual a 10 cm de longitud, e individuos epífitos a dos metros o menos de altura en los troncos. Individuo aquí significa un tallo o eje separado. El registro sólo incluye la cantidad de individuos por especie por sub unidad de transecto. La altura y otras características del individuo se registran únicamente con relación a las colectas.

En el campo se coloca la muestra en una bolsa plástica (es recomendable tener la bolsa dentro de un costal) para que la muestra no se seque durante el día. Con un marcador se escribe el número de colecta en la hoja de la muestra o en una ficha separada. Para muestras muy pequeñas es recomendable usar bolsitas separadas, para que la muestra no quede aplastada por muestras más grandes.

### El prensado de muestras

En el campamento se prensa las muestras entre hojas de periódico, y se escribe con un lápiz el número de colecta en la hoja de periódico. Al prensar la muestra es importante que ésta esté incluida en su totalidad dentro del periódico. En una buena muestra se puede ver ambas caras de las hojas (éstas no deben estar en exceso una encima de otra) (Figura 27).



Figura 27. Prensado de muestras de plantas. Fotos: Sanna-Kaisa Juvonen.

Para muestras de helechos es importante que éstas contengan una buena parte del rizoma, además es importante anotar si el rizoma es erecto u horizontal. Cuando el rizoma es muy grande, puede empaquetárselo separadamente dentro de un periódico, escribiendo el número de colecta en el exterior y dentro del paquete grande para evitar confusión si es que los números se borran durante la manipulación. De igual manera, es conveniente hacer paquetes para frutos grandes o cualquier otra parte gruesa sin hojas u otras estructuras delicadas de la colecta. Para escribir los números de colecta en las hojas de periódico, es necesario utilizar un lápiz si es que se va a alcoholizar las muestras. Un lapicero o marcador puede ser usado si se seca las muestras en el secador inmediatamente después de la colecta, sin alcoholización.

En helechos y palmeras se encuentra a menudo hojas grandes que no caben dentro de una sola hoja de periódico; este material puede ser cortado en dos o tres partes y colocado en periódicos separados y numerados, de tal manera que se identifique la pertenencia de las partes a la hoja. En el caso de que la hoja sea demasiado grande para ser incluida totalmente en tres partes dentro de los periódicos, debe hacerse muestras sólo de la base de la hoja, de la mitad y del ápice de ésta. En los apuntes se registra la longitud y la amplitud máxima de la hoja. Las muestras preparadas se colocan una sobre otra para obtener una pila de muestras. En la pila es preferible tener las hojas del periódico siempre abriendo hacia la misma dirección. Para evitar el desequilibrio en la pila, se colocan las muestras dentro de los periódicos con orientaciones variadas. Cuando la pila tiene unos 20 cm de grosor se la ata con hilo grueso para formar un paquete compacto. Entre el hilo y las hojas con muestras hay que poner algunos papeles extras, protegiendo así los periódicos con muestras. La cara del paquete hacia la cual abren los periódicos se deja libre, sin papel protector. El paquete listo se coloca en una bolsa plástica y se vierte 0,7-1 litro de alcohol ca. 50% vol. (a través de la cara libre). Lo importante es que todas las muestras en el paquete queden mojadas por el alcohol, para evitar la aparición de hongos. No se necesita alcohol si las muestras pueden ser secadas en el lapso de 24 horas después de la colecta.

Una vez alcoholizado el paquete, deber ser cerrado herméticamente. Es recomendable guardar las bolsas plásticas en un costal para prevenir huecos en la bolsa. Es necesario revisar que los paquetes dentro de las bolsas estén mojados hasta el momento del secado. Así guardadas, las muestras se conservan bien hasta meses.

### **El secado de las muestras**

Las muestras tienen que ser prensadas y secadas. Es muy importante un buen prensado (sin arrugas) y que las muestras estén secas; así se evitará la presencia de hongos, y el material podrá conservarse en principio hasta la eternidad.

Para secar las muestras son necesarios una prensa y un secador. Existen diversos secadores, pero el principio es calentar aire que pueda circular entre las muestras prensadas para extraer la humedad.

La prensa se construye de tal manera que cada muestra quede entre un par de papeles secantes y láminas corrugadas de aluminio o cartón. El papel secante es necesario para evitar la impresión de las láminas o cartones corrugados en las muestras. Los papeles secantes pueden ser cartones especiales, o también se puede usar 8 a 10 hojas de papel periódico. Si se utiliza papel periódico, es recomendable mantenerlos siempre en una dirección al abrirlos, y con muestras en la dirección opuesta. Esto disminuye la posibilidad de confusión.

Cuando las muestras hacen prensas grandes, es recomendable utilizar láminas corrugadas de aluminio, ya que los cartones corrugados se aplastan fácilmente. Así mismo, para muestras que tengan tanto hojas como pedazos gruesos de tallo, rizoma, fruto, entre otros, debe colocarse algún bulto (almohada) adicional como hojas arrugadas de periódico sobre dichas muestras y cerca a la parte abultada, para que sequen bien prensadas y no onduladas.

La prensa se amarra fuertemente con sogas o cintas resistentes, para que las muestras queden bien aplastadas y sin ondulaciones. No se debe practicar demasiada fuerza, porque los cartones pueden aplastarse e impedir la circulación del aire por los canales. Antes de amarrar la prensa, debe ajustarse sus bordes para que quede bien enmarcada, sin algunas láminas o cartones afuera. Al hacer el ajuste, hay que tener cuidado de no usar fuerza en exceso, para que las láminas no se distorsionen. Una vez preparada la prensa, se coloca en el secador, cuidando de que la soga o cinta no cuelgue hacia la fuente de calor, creando así un peligro de incendio.

En un secador que funcione óptimamente, el secado dura unas 20 horas o hasta menos, pero el tiempo depende mucho de la calidad del secador, los papeles secantes y las muestras. Los paquetes con rizomas u otras partes gruesas se secan en el secador al doble o triple del tiempo normal. Se puede acelerar el secado volteando la prensa y colocando la cara opuesta hacia la fuente de calor. Cuando se piensa que la mayoría de las muestras ya están secas, se abre la prensa. Debe revisarse que ninguna de las muestras esté húmeda. Una muestra seca es dura y poco flexible, y no se percibe nada de humedad en la hoja del papel periódico. Una muestra húmeda es todavía flexible (tener cuidado con tallos, inflorescencias, frutos y otras partes gruesas); muchas veces se siente humedad en la hoja del periódico, y a menudo se ven manchas más oscuras de colores en las hojas que corresponden a las áreas húmedas. Una muestra de la cual uno no está seguro si está seca, siempre tiene que volver al secador.

Al final de la presente publicación en la sección de la bibliografía se presenta un listado de bibliografía recomendada para los estudios florísticos.

### Materiales específicos

- Cinta métrica de 50 m
- Brújula
- GPS
- Machetes para abrir la trocha y para coleccionar palmeras grandes
- Bolsas de plástico de aproximadamente 40x70 cm para las colectas en el bosque
- Bolsas plásticas de más o menos 1 x 0,5 m para guardar los paquetes alcoholizados
- Bolsas plásticas de 20x30 cm para las muestras pequeñas
- Tijera podadora
- Libreta de campo o un tablero con papel, preferentemente cuadriculado
- Lápices, Lapiceros, Marcadores indelebles
- Pabilo para amarrar paquetes
- Alcohol
- Balde con tapa para mezclar el alcohol y guardar el exceso
- Papel periódico
- Costales plásticos
- Cinta marcadora
- Impermeable o paraguas (un paraguas permite trabajar aún con un poco de lluvia).

## 4. METODOLOGÍA PARA EL INVENTARIO DE LA FAUNA

### Generalidades

Los inventarios de fauna deben ser realizados dentro del hábitat seleccionado para los estudios de plantas. Así, para el inventario de aves se establecerá un sistema de trochas en parrilla; para el inventario de anfibios anuros podrá utilizarse uno de esos transectos, o establecer un transecto cercano, al igual que para el inventario de hormigas. De esta manera se podrá correlacionar las variables tomadas en consideración (suelo, topografía, área basal, pH, entre otros) con todos los grupos taxonómicos.

### Hormigas



El inventario de hormigas puede ser realizado utilizando una de las trochas de 500 metros previamente establecida por el grupo de ornitólogos, o uno de los transectos botánicos, o establecer un nuevo transecto muy cerca de los anteriores, pero con una longitud de 625 metros. Es decir, si se utiliza el transecto de aves o el botánico, debe aumentarse 125 metros.

### Trampas Winkler

Estas trampas han sido diseñadas para capturar insectos que viven en la hojarasca, principalmente hormigas terrestres. En un transecto de 625 metros, se toma muestras de hojarasca cada 25 metros (25 puntos de muestreo) y a 3,5 metros a la derecha o izquierda del transecto (los muestreos deben realizarse a un mismo lado del transecto, a menos que el punto a muestrear tenga mucha humedad). Cada muestra será colocada en una trampa Winkler.

### Fase de campo 1

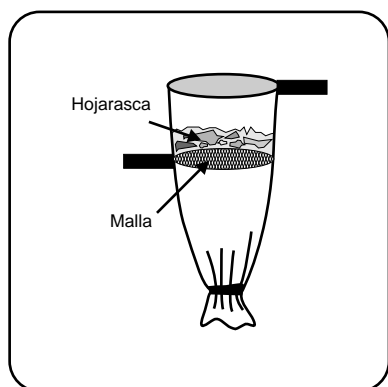
Una vez ubicados los 25 puntos de muestreo, se delimita un área de 1 m<sup>2</sup> de hojarasca, y luego con un machete se remueve dicha área (cortar las ramas y el resto del material en pedazos pequeños) con la finalidad de perturbar los nidos de hormigas (Fisher 1999) y lograr coleccionar todos los insectos, inclusive los que se encuentran dentro de ramas, palitos y materia en descomposición. Utilizando guantes de cuero (badana) se reúne todo el material orgánico y se coloca en un cernidor de hojarasca, que es sacudido durante 3 minutos (Figura 28 y Figura 29). El material obtenido se coloca en una bolsa de tela, la cual es trasladada hasta el laboratorio o campamento.

**Fase de campo 2**

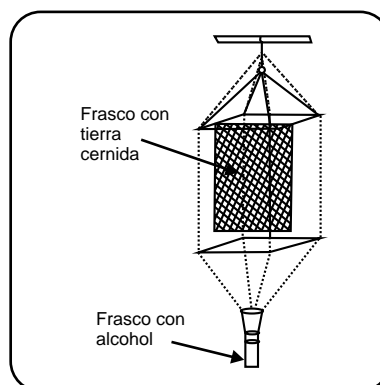
La muestra es colocada en el laboratorio en las trampas Winkler. Estas trampas se pueden describir como sacos largos que tienen en la base un pequeño agujero, en el cual se coloca una bolsa Whirl-Pack o un frasco de plástico, ambos con alcohol al 70%; dentro de estos sacos cuelgan unas bolsas rectangulares confeccionadas con malla, en las que se coloca las muestras obtenidas en la fase anterior (una muestra por trampa) (Figura 30 y Figura 31). Este material permanece colgado durante 48 horas. En este tiempo los insectos, y en especial las hormigas, debido a la desecación de la materia orgánica, se mueven a través de ésta hacia la parte exterior de la malla, quedándose atrapados en el saco y teniendo como única salida y fuente de humedad las bolsas Whirl-Pack o los frascos con alcohol, donde caen y se preservan para su posterior separación e identificación. La identificación debe hacerse con un microscopio estereo de al menos 100X de aumento.



**Figura 28.** Sacudiendo el cernidor de hojarasca. Fotos: Sanna-Kaisa Juvonen.



**Figura 29.** La hojarasca con tierra es colocada dentro del cernidor, tras una serie de sacudidas, se obtiene sólo tierra con insectos.



**Figura 30.** La tierra cernida es puesta en las trampas Winkler. Las hormigas caerán al recipiente con alcohol.



**Figura 31.** Las muestras colectadas en el campo se colocan en las trampas Winkler. Las hormigas están separadas y preservadas en frascos con alcohol. Fotos: Sanna-Kaisa Juvonen.

Al final de la presente publicación en la sección de la bibliografía se presenta un listado de bibliografía recomendada para los estudios de hormigas.

### Equipo de campo y laboratorio

- 25 trampas Winkler;
- 1 ó 2 cernidores;
- Bolsas Whirl Pack o frascos plásticos;
- Guantes de cuero;
- Bolsas de tela 40 x 30 cm (25 o más);
- Alcohol 70%;
- Tubitos de vidrio para insectos;
- Papel Canson;
- Estilógrafo;
- Estereo microscopio >80X;
- Pinzas finas;
- Placas petri;
- Ligas;
- Cinta marcadora;
- Machete;
- Cuerda de 4m.

#### Anfibios anuros

Se evaluará un transecto o faja de 500 x 5 metros dentro del hábitat seleccionado para los inventarios florísticos. Si se utiliza uno de los transectos establecidos para el inventario de flora o aves, es necesario dejar pasar 2-3 días para iniciar el inventario, ya que el transecto fue algo perturbado por las actividades anteriores. De lo contrario, se podrá realizar un nuevo transecto cercano al ya evaluado.

La búsqueda de anfibios anuros se realiza en horas diurnas y nocturnas, mediante registro visual y auditivo. El inventario diurno consiste en la remoción de la capa de hojarasca y restos de madera en descomposición (Figura 32). El tiempo empleado en el inventario durante el día varía dependiendo de la longitud del transecto, pudiendo ser de cuatro a seis horas para 100-175 m de transecto. En la noche se emplea aproximadamente tres horas avanzando de 125 a 200 m; se utiliza una linterna frontal, realizando el muestreo hasta una altura de tres metros. Las distancias recorridas dependerán de la cantidad de restos de materia orgánica en descomposición (hojas, raíces, ramas secas, entre otros), de la vegetación del sotobosque (tupida o clara), y de la topografía (ondulada o plana). En el transecto se identifica todos los individuos presentes, y cuando la identificación taxonómica en el campo no es posible, se colecta los ejemplares para su identificación en el laboratorio (Figura 32). Los especímenes colectados serán fijados en formalina al 10% y posteriormente transferidos a alcohol al 70 %.



Figura 32.

a) El inventario diurno de anfibios anuros consiste en la remoción de la capa de hojarasca y restos de madera en descomposición.

Foto: Sanna-Kaisa Juvonen.

b) Se identifica las ejemplares de anfibios colectados en el laboratorio.

Foto: Sanna-Kaisa Juvonen.

c) *Dendrobates* sp. Las especies de esta familia son de hábitos diurnos y generalmente de colores vistosos; se encuentran sobre la hojarasca, y en huecos. Los dendrobátidos han sido considerados como un grupo indicador para los fines del proyecto BIODAMAZ.

Foto: Carlos Rivera.

Al final de la presente publicación, en la sección de la bibliografía, se presenta un listado de bibliografía recomendada para los estudios de anfibios anuros.

### Equipo de campo y laboratorio

- Linterna frontal;
- Bolsas plásticas 10 x 15 cm;
- Bolsas de tela;
- Cinta marcadora;
- Libreta de campo;
- Anestésico;
- Agujas;
- Jeringas;
- Formol 10%;
- Alcohol 70%;
- Tapers.

#### Aves

El inventario de aves se llevará a cabo en un sistema de trochas establecidas a manera de una parrilla, con cuatro transectos paralelos de 500 metros y separados entre sí por 400 metros. Estos transectos estarán unidos por sus extremos por trochas de 1200 metros (trochas de acceso) (Figura 33). De esta manera se evaluará un total de 2000 metros de transectos por día.

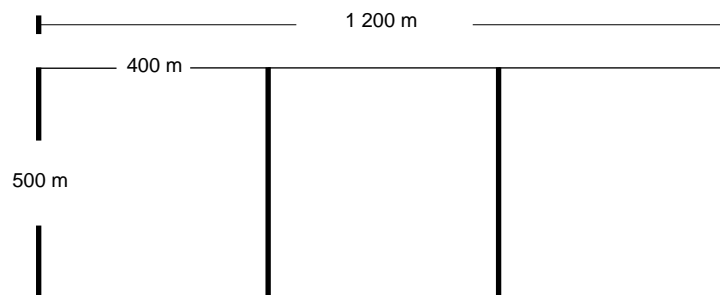


Figura 33. El registro de aves se realizará en las cuatro trochas de 500 metros.

Si el terreno es de difícil accesibilidad para establecer la parrilla, el inventario podrá hacerse en un solo transecto de 2000 metros.

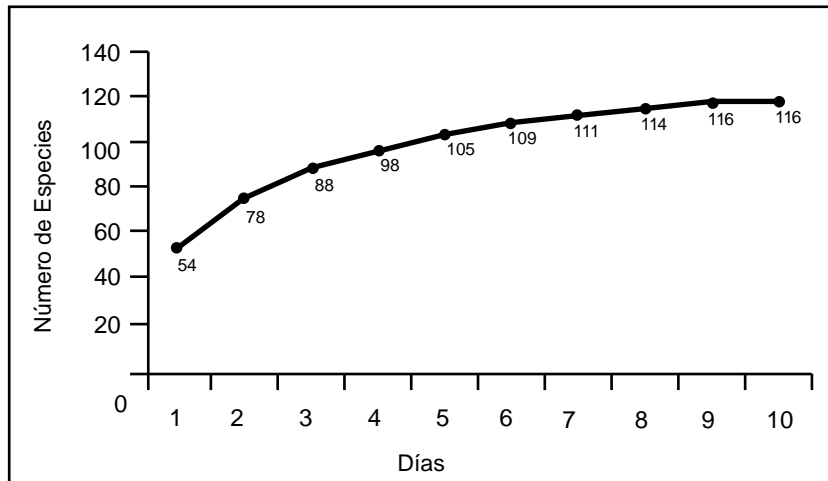
En los inventarios se registrará las especies y el número de individuos presentes. Las identificaciones de especies se realizarán principalmente por identificación de cantos o vocalizaciones y por observaciones directas (Figura 34). El uso de una grabadora con micrófono direccional será muy útil para grabar cantos y reproducirlos con la finalidad de atraer a algunas especies cuya identificación resulte dudosa. Así, el investigador tendrá la oportunidad de observar directamente al espécimen.



Figura 34. Observación directa empleada en los inventarios de aves. Fotos: Sanna-Kaisa Juvonen.

El inventario deberá hacerse desde horas muy tempranas de la mañana (de 04:30 ó 05:00 a 10:30 ó 11:00) y algunas horas en la tarde (15:00 a 17:00). Para el inventario de especies nocturnas se recorrerá los transectos por la noche.

El tiempo de evaluación se calculará con base en una curva acumulativa de especies y podría variar de acuerdo al área de estudio (Figura 35).



**Figura 35.** Ejemplo de una curva acumulativa de especies. Se observa que al cabo de 10 días de evaluación, los registros de aves ya no aumentaron. Así, el tiempo de evaluación puede ser de diez días.

Se recomienda el uso de una grabadora Sony TCM-5000EV y un micrófono Sennheiser ME66 con un Adaptador de Alimentación Sennheiser K6. Para la observación de aves es necesario contar con buenos binoculares con un alcance de 10 x 40 o aproximados. Entre las mejores marcas están Zeiss, Leica y Swarovsky. En vista de que estas marcas son caras y difíciles de encontrar en el mercado nacional, podrían utilizarse otras marcas, como Minolta, Nikon y Olympus.

Es muy importante que estos equipos permanezcan secos en el campo, ya que la humedad puede dañarlos considerablemente. El uso de bolsas o sacos plásticos e impermeables conteniendo absorbentes de humedad (sílica gel, por ejemplo) ayudarán a conservar los equipos en buen estado.

Una libreta de campo para anotaciones diversas y una planilla de registros complementan los materiales básicos para la evaluación.

Al final de la presente publicación en la sección de la bibliografía se presenta un listado de bibliografía recomendada para los estudios de anfibios anuros.



## CONCLUSIONES

Los bosques amazónicos cubren un área grande, el acceso es típicamente difícil y la biota es extremadamente rica en especies que muy a menudo son taxonómicamente poco conocidas. Esta combinación de características presenta un reto formidable a los biólogos, forestales, planificadores del uso de tierra y otros profesionales que necesitan saber los patrones geográficos de distribución y abundancia de las especies amazónicas. En estas condiciones es inevitable que los estudios de la biota tengan que concentrarse en alguna pequeña parte de la totalidad. Además, la percepción remota ofrece una opción potencialmente muy fuerte para poder extrapolar los resultados obtenidos de los puntos estudiados a regiones más amplias.

Esperamos que este documento aporte al reconocimiento de las limitaciones prácticas que existen en la documentación de las distribuciones y abundancias de las especies amazónicas. Sin embargo, al mismo tiempo esperamos que este documento pueda también ayudar a investigar críticamente como sería posible utilizar el esfuerzo de trabajo disponible de la manera más eficaz posible. Hay que limitarse en ciertos grupos de la biota, pero antes de que uno pueda llamar algún grupo como un grupo indicador, hay que probarlo y hay que definir exactamente cuál es la característica por la cual el grupo sirve como un indicador. Los métodos aquí presentados ayudan especialmente a definir si un grupo de especies puede servir como un indicador de los patrones de distribución de otro grupo de especies. Éste es un concepto más difícil de captar y analizar que, por ejemplo, un indicador de riqueza local de especies. Sin embargo, para la mayoría de las aplicaciones prácticas, tanto en el uso sostenible como en la conservación de las especies amazónicas, será mucho más útil poder pronosticar si una localidad inexplorada es muy semejante o diferente en su composición de especies a lugares explorados y bien conocidos, que poder predecir cuántas especies tiene. Para la conservación de la biota amazónica, un lugar con un número grande de especies puede ser insignificante si es que las especies del lugar ocurren en todas partes. Por otra parte, un lugar con unas pocas especies puede ser insustituible para la conservación si es que sus especies no ocurren en otras partes.

De los grupos de plantas tratados aquí ya se han hecho estudios que demuestran que pueden servir como indicadores de la distribución de especies de árboles. De los grupos de animales de este documento no se ha publicado todavía estudios que podrían ayudar a estimar su utilidad como grupos de indicadores de patrones de distribución de especies de otros grupos de organismos. Los resultados preliminares sugieren que las distribuciones de especies en estos grupos de animales posiblemente no tienen mucho que ver con las especies de plantas, pero claramente hacen falta de estudios más amplios. Esperamos que este documento ayude en realizar tales estudios.

No hemos entrado en detalles muy finos ni a los métodos numéricos más avanzados. La intención de este trabajo es poder servir como una base relativamente fácil para ayudar a entender los conceptos y términos básicos en la percepción remota y los análisis numéricos de comunidades de especies. En los métodos de campo para estudiar ciertos grupos de plantas y animales, penetramos a un nivel más profundo, pero por otra parte, dejamos muchas técnicas tanto de éstos, como de otros grupos fuera de este libro. La selección de los grupos tratados en este libro fue basada en una serie de criterios que son explicados en su propio capítulo. Según estos criterios, uno podría haber llegado a una selección algo diferente, pero probablemente no completamente diferente. Más bien, uno podría haber incluido más grupos de organismos; además, está claro que cambios en los criterios pueden resultar en la selección de otros grupos. Esta guía ha podido cumplir con uno de sus objetivos, si alguien quiere demostrar que hay grupos que claramente cumplen mejor los criterios presentados, y/o que los criterios en sí son sesgados de alguna manera.







## BIBLIOGRAFIA GENERAL

- ALONSO, L.E. 2000. Ants as indicators of diversity. *En*: Agosti, D., Majer, J. D., Alonso, L. E. & Schultz, T. R. (eds.). *Ants. Standard Methods for Measuring and Monitoring Biodiversity*. Biological Diversity Handbook Series. Smithsonian Institution Press, Washington, EEUU. pp. 80-88.
- ARRHENIUS, O. 1921. Species and area. *Journal of Ecology* 9: 95-99.
- BENDAYÁN, L., SANJURJO, J., KALLIOLA, R. & RODRÍGUEZ, F. 2003. Experiencia de la tecnología de percepción remota para la elaboración del mosaico de imágenes de satélite Landsat TM en la selva baja de la Amazonía peruana. *Folia Amazónica* 14 (1): 73-86
- BESUCHET, C., BURCKHARDT, D.H. & LONL, I. 1987. The "Winkler/Moczarski" selector for fungus and litter Coleoptera. *The Coleopterists Bulletin* 41:393-394.
- BEVEN, S., CONNOR, E.F. & K. BEVEN. 1984. Avian biogeography in the Amazon basin and the biological model of diversification. *Journal of Biogeography* 11: 383-399.
- BIODAMAZ. 2004a. Diversidad de vegetación de la Amazonía peruana expresada en un mosaico de imágenes de satélite. Documento Técnico 12. Serie BIODAMAZ-IIAP. Iquitos, Perú.
- BIODAMAZ. 2004b. Manual para elaboración de mosaicos de imágenes de satélite Landsat TM para la selva baja peruana. Documento Técnico 03. Serie BIODAMAZ-IIAP. Iquitos, Perú.
- BROWN, J.H. 1984. On the relationship between abundance and distribution of species. *American Naturalist* 124: 255-279.
- CAJANDER, A.K. 1926. The theory of forest types. *Acta Forestalia Fennica* 29 (3): 1-108.
- COLINVAUX, P.A. 1996. Quaternary environmental history and forest diversity in the Neotropics. *En*: Jackson, J.B.C., Budd, A.F. & Coates, A.G. (eds.). *Evolution and environment in tropical America*. University of Chicago Press, Chicago, EEUU. pp. 359-405.
- COOMES, O.T. 1995. A century of rain forest use in Western Amazonia. Lessons for extraction-based conservation of tropical forest resources. *Forest & Conservation History* 39: 108-120.
- DRAKE, J.B., DUBAYAH, R.O., CLARK, D.B., KNOX, R.G., BLAIR, J.B., HOFTON, M.A., CHAZDON, R.L., WEISHAMPEL, J.F. & PRINCE, S.D. 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment* 79: 305-319.
- DROEGE, S., CYR, A. & LARIVÉE, J. 1998. Checklists: an under-used tool for the inventory and monitoring of plants and animals. *Conservation Biology* 12 (5): 1134-1138.
- FISHER, B.L. 1999. Improving inventory efficiency: A case study of leaf-litter ant diversity in Madagascar. *Ecological Applications* 9 (2): 714-731.
- HAFFER, J. 1969. Speciation in Amazonian forest birds. *Science* 165: 131-137.
- HANSKI, I. 1999. *Metapopulation Ecology*. Oxford University Press, Oxford, Gran Bretaña.
- HARRIS, R. 1987. *Satellite remote sensing: An introduction*. Routledge & Kegan Paul. Londres, Gran Bretaña.
- HEYER, W.R., DONNELLY, M.A., MCDIARMIND, R.W., HAYEK, L.C. & FOSTER, M. S. (eds.). *Measuring and monitoring biological diversity. Standard methods for amphibians*. Biological Diversity Handbook Series. Smithsonian Institution Press, Washington, EEUU. 363 pp.
- HOORN, C. 1993. Geología del nororiente de la Amazonía peruana: La Formación Pebas. *En*: Kalliola, R., Puhakka, M. & Danjoy, W. (eds.). *Amazonía peruana vegetación húmeda tropical en el llano subandino*. Proyecto Amazonía de la Universidad de Turku & Oficina Nacional de Evaluación de Recursos Naturales. Jyväskylä, Finlandia. pp. 69-85.
- HUBBELL, S. P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ, EEUU.
- JONGMAN, R.H.G., TER BRAAK, C.J.F. & VAN TONGEREN, O.F.R. 1995. *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, Gran Bretaña.
- KAHN, F. & DE GRANVILLE, J.J. 1992. *Palms in forest ecosystems of Amazonia*. Springer Verlag, Berlín, Alemania.

- KAHN, F. & MEJÍA, K. 1990. Palm communities in wetland forest ecosystems of Peruvian Amazonia. *Forest Ecology and Management* 33/34: 169-179.
- LANDERS, P.B., VERNER, J. & THOMAS, J.W. 1998. Ecological uses of vertebrate indicator species: A critique. *Conservation Biology* 2: 316-328.
- LAWTON, J.H., BIGNELL, D.E., BOLTON, B., BLOEMERS, G.F., EGGLETON, P., HAMMOND, P.M., HODDA, M., HOLT, R.D., LARSEN, T.B., MAWDSLEY, N. A., STORK, N.E., SRIVASTAVA, D.S. & WATT, A.D. 1998. Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature* 391: 72-76.
- LEGENDRE P. & LEGENDRE L. 1998. Numerical ecology. Second English Edition. Elsevier, Amsterdam, Países Bajos.
- LILLESAND, T.M. & KIEFER, R.W. 1994. Remote Sensing and Image Interpretation. Tercera edición. John Wiley & Sons, Nueva York, EEUU.
- LINNA, A., IRION, G., KAUFFMAN, S., WESSELINGH, F. & KALLIOLA, R. 1998. Heterogeneidad edáfica de la zona de Iquitos: origen y comprensión de sus propiedades. En: Kalliola, R. & Flores, Paitán, S. (eds.). Geoecología y desarrollo amazónico: estudio integrado en la zona de Iquitos, Perú. *Annales Universitatis Turkuensis Ser A II* 114: 461-480.
- MACARTHUR, R.R. & WILSON, E.O. 1967. The theory of island biogeography. Princeton University Press. Princeton, New Jersey, EEUU.
- MAJER, J.D. 1983. Ants: bioindicators of minisite rehabilitation, land-use and land conservation. *Environmental Management* 7 (4): 375-383.
- MEJÍA, K. 1992. Las palmeras en los mercados de Iquitos. *Bulletin de l'Institut Français d'Études Andines* 21: 565-597.
- MOREAU, R.E. 1963. Vicissitudes of the African biomes in the Pleistocene. *Proc. Zool. Soc. London* 125: 253-295.
- MUNN, R.E. 1988. The design of integrated monitoring systems to provide early indications of environmental ecological changes. *Environmental Monitoring and Assessment* 11: 203-217.
- NELSON, B.W., FERREIRA, C.A., DA SILVA, M.F. & KAWASAKI, M.L. 1990. Refugia, endemism centers and collecting density in Brazilian Amazonia. *Nature* 345: 714-716.
- NOSS, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 335-364.
- OLIVER, I. & BEATTIE, A.J. 1996. Invertebrate morphospecies as surrogates for species: A case study. *Conservation Biology* 10: 99-109.
- PACHECO, V. 1990. ¿Por qué mantener colecciones científicas de historia natural? *Bol. Mus. Hist. Nat. Universidad Nacional Mayor de San Marcos (UNMSM) Nueva Serie* 2: 1-2.
- PEARSON, D.L. & CASSOLA, F. Cassola. 1992. World-wide species richness patterns of tiger beetles (Coleoptera: Cicindelidae): Indicator taxon for biodiversity and conservation studies. *Conservation Biology* 6: 376-391.
- PONDER, W.F., CARTER, G.A., FLEMONS, P. & CHAPMAN, R.R. 2001. Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology* (15) 3: 648-657.
- PRANCE, G.T. (ed.). 1982. Biological diversification in the tropics. Columbia University Press, Nueva York, EEUU. 714 pp.
- PRENDERGAST, J.R., QUINN, R.M., LAWTON, J.H., EVERSHAM, B.C. & GIBBONS, D.W. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* 365: 335-337.
- RUOKOLAINEN, K. & TUOMISTO, H. 1998. Vegetación natural de la zona de Iquitos. En Kalliola, R. & Flores Paitán, S. (eds.). Geoecología y desarrollo amazónico: estudio integrado en la zona de Iquitos, Perú. *Annales Universitatis Turkuensis Ser A II* 114: 253-365.
- RUOKOLAINEN, K., LINNA, A. & TUOMISTO, H. 1997. Use of Melastomataceae and pteridophytes for revealing phytogeographical patterns in Amazonian rain forests. *Journal of Tropical Ecology* 13: 243-256.
- RÄSÄNEN, M. 1993. La geohistoria y geología de la Amazonía peruana. En: Kalliola, R., Puhakka, M. & Danjoy, W. (eds.). Amazonía peruana vegetación húmeda tropical en el llano subandino. Proyecto Amazonía de la Universidad de Turku & Oficina Nacional de Evaluación de Recursos Naturales. Jyväskylä, Finlandia. pp. 43-67.
- RÄSÄNEN, M., NELLER, R., SALO, J. & JUNGNER, H. 1992. Recent and ancient fluvial depositional systems in the Amazonian foreland basin, Peru. *Geological Magazine* 129: 293-306.

RÄSÄNEN, M., SALO, J., JUNGNER, H. & ROMERO PITTMAN, L. 1990. Evolution of the Western Amazon lowland relief: impact of Andean foreland dynamics. *Terra Nova* 2: 320-332.

RÄSÄNEN, M., LINNA, A., IRION, G., REBATA HERNANI, L., VARGAS HUAMAN, R. & WESSELINGH, F. Geología y geoformas de la zona de Iquitos. En: Kalliola, R. & Flores Paitán, S. (eds.). *Geoecología y desarrollo amazónico: estudio integrado en la zona de Iquitos, Perú*. *Annales Universitatis Turkuensis Ser A II* 114: 59-137.

SALO, J. 1987. Pleistocene forest refuges in the Amazon: evaluation of the biostratigraphical, lithostratigraphical and geomorphological data. *Annales Zoologici Fennici* 24: 203-211.

SIAMAZONIA Sistema de Información de la Diversidad Biológica y Ambiental de la Amazonía Peruana. [www.siamazonia.org.pe](http://www.siamazonia.org.pe).

SOBERÓN, J.M., LLORENTE, J.B. & OÑATE, L. 2000. The use of specimen-label database for conservation purposes: an example using Mexican Papilionid and Pierid butterflies. *Biodiversity and Conservation* 9: 1441-1466.

TOIVONEN, T., KALLIOLA, R., RUOKOLAINEN, K. & MALIK, R. Across-path DN gradient challenge the Landsat TM interpretations in the Amazonian lowlands. Manuscrito.

TUOMISTO, H. 1998. What satellite imagery and large-scale field studies can tell about biodiversity patterns in Amazonian forests? *Annals of Missouri Botanical Garden* 85: 48-62.

TUOMISTO, H. & RUOKOLAINEN, K. 1997. The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. *Biodiversity & Conservation* 6: 347-357.

TUOMISTO, H. & RUOKOLAINEN, K. 1998. Uso de especies indicadoras para determinar características del bosque y de la tierra. En: Kalliola, R. & Flores Paitán, S. (eds.). *Geoecología y desarrollo amazónico: estudio integrado en la zona de Iquitos, Perú*. *Annales Universitatis Turkuensis Ser A II* 114: 481-491.

TUOMISTO, H., LINNA, A. & KALLIOLA, R. 1994. Use of digitally processed satellite images in studies of tropical rain forest vegetation. *International Journal of Remote Sensing* 15 (8): 1595-1610.

TUOMISTO, H., RUOKOLAINEN, K., POULSEN, A.D., MORAN, R.C., QUINTANA, C., CAÑAS, G. & CELI, J. 2002. Distribution and diversity of pteridophytes and Melastomataceae along edaphic gradients in Yasuni National Park, Ecuadorian Amazonia. *Biotropica* 34: 516-533.

VORMISTO, J. 2000. Palms in the rainforests of Peruvian Amazonia: uses and distribution. Ser. All TOM 138. *Annales Universitatis Turkuensis*.

VORMISTO, J., PHILLIPS, O., RUOKOLAINEN, K., TUOMISTO, H. & VÁSQUEZ, R. 2000. A comparison of small-scale distribution patterns of four plant groups in an Amazonian rainforest. *Ecography* 23: 349-359.

#### **BIBLIOGRAFÍA RECOMENDADA PARA ESTUDIOS DE PLANTAS**

ANDERSON, A. 1988. Use and management of native forests dominated by açai palm (*Euterpe oleracea* Mart.) in the Amazon estuary. *Adv. Econ. Bot.* 6: 144-154.

AYORA, N. & ORELLANA, R. 1993. Physicochemical soil factors influencing the distribution of two coastal palms in Yucatan, Mexico. *Principes* 37: 82-91.

BRADFORD, D. & SMITH, C. 1977. Seed predation and seed number in Scheelea palm fruits. *Ecology* 58: 667-673.

CHAZDON, R. 1985. Leaf display, canopy structure, and light interception of two understory palm species. *Amer. J. Bot.* 72: 1493-1502.

CHAZDON, R. 1986. Physiological and morphological basis of shade tolerance in rain forest understory palms. *Principes* 30:92-99.

EGUIARTE, L., BÚRQUEZ, A., RODRÍGUEZ, J., MARTÍNEZ-RAMOS, M., SARUKHÁN, J. & PIÑERO, D. 1993. Direct and indirect estimates of neighborhood and effective population size in a tropical palm, *Astrocaryum mexicanum*. *Evolution* 47: 75-87.

GENTRY, A. 1986. Notes on Peruvian palms. *Ann. Missouri Bot. Gard.* 73: 158-165.

HENDERSON, A. 1995. *The Palms of the Amazon*. Oxford University Press, Nueva York, EEUU.

KAHN, F. & MEJÍA, K. 1987. Notes on the biology, ecology, and use of a small Amazonian palm, *Lepidocaryum tessmannii*. *Principes* 31: 14-19.

KILTIE, R. 1981. Distribution of palms fruits on rain forest floor: Why white-lipped peccaries forage near objects. *Biotropica* 13: 141-145.

## BIBLIOGRAFÍA RECOMENDADA PARA ESTUDIOS DE HORMIGAS

AGOSTI, D., MAJER, J., ALONSO, L. & SCHULTZ, T.R. (eds.). 2000. Sampling Ground-dwelling Ants: Case Studies from the worlds' Rain Forest. Curtin University of Technology, Perth, Western Australia.

ANDERSEN, A.N. 1990. The use of ants communities to evaluate changes in Australian terrestrial ecosystems: a review and a recipe. *Proc. Ecol. Aust.* 16: 347-357.

BENSON, W.W. & HARADA, A.Y. 1988. Local diversity of tropical and temperate ant faunas (Hymenoptera: Formicidae). *Acta Amazónica* 18(3/4): 275-289.

BESTELMEYER, B.T. & WIENES, J.A. 1997. The effects of land use on the structure of ground-foraging ant communities in the Argentine chaco. *Ecol. Appl.* 6: 1225-1240.

BOLTON, B. 1994. Identification Guide to the Ant Genera of the World. Harvard University Press, Cambridge, Massachusetts, EEUU.

BURBRIDGE, A.H., LEICESTER, K., MCDAVITT, S. & MAJER, J.D. 1992. Ants as indicators of disturbance at Yanchep National Park, Western Australia. *Journal of the Royal Society of Western Australia* 75: 89-95.

COVER, S.P., TOBIN, J.E. & WILSON, E.O. 1990. The ant community of a tropical lowland rainforest site in Peruvian Amazonia. *En: Veeresh, G.K., Malik, B. & Viraktamath, C.A. (eds.). Social Insects and the Environment. Proceedings of the 11<sup>th</sup> International Congress of IUSI, 1990. Oxford and IBH Publishing, New Delhi, India. pp.699-700.*

CRANSTON, P.S. & TRUEMAN, J.W.H. 1997. "Indicator" taxa in invertebrate biodiversity assessment. *Memoirs of the Museum of Victoria* 56(2): 267-274.

FOWLER, H.G. 1995. Biodiversity estimates: Ant communities and the rare species (Hymenoptera: Formicidae) in a fauna of a sub-tropical island. *Revista de Matemática e Estatística, São Paulo* 13: 29-38.

KASPARI, M. 1996. Litter ants patchiness at the 1 m<sup>2</sup> scale: Disturbance dynamics in three Neotropical forests. *Oecologia* 107: 265-273.

KIM, K.C. 1993. Biodiversity, conservation and inventory: why insects matter. *Biodiversity and Conservation* 2: 191-214.

KING, J.R., ANDERSEN, A.N. & CUTTER, A.D. 1998. Ants as bioindicators of habitat disturbance: Validation of the functional group model for Australia's humid tropics. *Biodiversity and Conservation* 7: 1627-1638.

KREMEN, C., COLWELL, R.K., ERWIN, T.L., MURPHY, D.D., NOSS, R.F. & SANJAYAN, M.A. 1993. Terrestrial arthropod assemblages: Their use in conservation planning. *Cons. Biol.* 7 (4): 796-808.

MAJER, J.D. & DELABIE, J.H.C. 1994. Comparison of the ant communities of annually inundated and terra firme forests at Trombetas in the Brazilian Amazon. *Insec. Soc.* 41:343-359.

ROTH, D.S., PERFECTO, I. & RATHCKE, B. 1994. The effects of management systems on ground-foraging ant diversity in Costa Rica. *Ecological Applications* 4: 423-436.

## BIBLIOGRAFÍA RECOMENDADA PARA ESTUDIOS DE ANFIBIOS ANUROS

BURNHAM, K.P. 1984. The need for distance data in transect counts. *Journal of Wildlife Management* 48: 1248-1254.

BURY, R.B. & RAPHAEL, M.G. 1983. Inventory methods for amphibians and reptiles. *En: Bell, J.F. & Atterbury, T. (eds.). Renewable Resources Inventories for Monitoring Changes and Trends. Oregon State University, Corvallis, Oregon, EEUU. pp. 416-419.*

CAMPBELL, H.W. & CHRISTMAN, S.P. 1982. Field techniques for herpetofaunal community analysis. *En: Scott, Jr., N.J. (ed.). Herpetological Communities. United States Department of the Interior, Fish and Wildlife Service, Wildlife Research Report 13. pp. 193-200.*

CORN, P.S. & BURY, R.B. 1990. Sampling methods for terrestrial amphibians and reptiles. United States Department of Agriculture, Forest Service, General Technical Report PNW-GTR-256.

DUELLMAN, W.E. 1990. Herpetofaunas in Neotropical rainforest: Comparative compositions, history and resource use. *En: Gentry, A.H. (ed.). Four Neotropical Rainforests. Yale University Press, New Haven, Connecticut, EEUU. pp. 455-505.*

FROST, D.R. (ed.). 1985. Amphibian species of the World: A taxonomic and Geographical Reference. Allen Press and Association of Systematics Collections, Lawrence, Kansas, EEUU.

HEDGES, B.S. & THOMAS, R. 1991. The importance of systematics research in the conservation of amphibian and reptile populations. *En: Moreno, J.A. (ed.). Status y Distribución de los Reptiles y Anfibios de la Región de Puerto Rico. Departamento de Recursos Naturales de Puerto Rico. Publicación Científica Miscelánea 1.*

HEYER, W.R. & BERVEN, K.A. 1973. Species diversity of herpetofaunal samples from similar microhabitats at two tropical sites. *Ecology*54: 427-432.

LIEBERMAN, S.S. 1986. Ecology of the leaf litter herpetofauna of a Neotropical rain forest: La Selva, Costa Rica. *Acta Zoológica Mexicana* 15: 1-72.

MYERS, C.W. & DALY, J.W. 1983. Dart-poison frogs. *Scientific American*248 (2):120-133.

PECHMANN, J.H.K., SCOTT, D.E., SEMLITSCH, R.D., CALDWELL, J.P., VITT, L.J. & GIBBONS, J.W. 1991. Declining amphibian populations: The problem of separating human impacts from natural fluctuations. *Science*253: 892-895.

PISANI, G.R. 1973. A guide to preservation techniques for amphibians and reptiles. Society for the Study of Amphibians and Reptiles. Herpetological Circular 1.

SCOTT, Jr., N.J. 1976. The abundance and diversity of the herpetofaunas of tropical forest litter. *Biotropica*8:41-58.

STEWART, M.M. & POUGH, F. H. 1983. Population density of tropical forest frogs: Relation to retreat sites. *Science*74:1-2.

WELLS, K.D. 1977. The social behavior of anuran amphibians. *Animal Behaviour*25:666-693.

#### **BIBLIOGRAFÍA RECOMENDADA PARA ESTUDIOS DE AVES**

DETMERS, R., BUEHLER, D., BARTLETT, J. & KLAUS, N. 1999. Influence of point count length and repeated visits on habitat model performance. *The Journal of Wildlife Management* (63): 815-823.

MEYER DE SCHAUENSEE, R. 1970. A guide to the birds of South America. Livingston Publishing Company, Wynnewood, PA, EEUU. 470 pp.

RALPH, J., GEUPEL, G., PYLE, P., MARTIN, T., DESANTE, D. & MILA, B. 1996. Manual de métodos de campo para monitoreo de aves terrestres. United States Department of Agriculture, California, EEUU. 45 pp.

SHIELDS, W. 1994. The effect of time of day on avian census results. *The Auk* (94): 380-383.

STOTZ, D., FITZPATRICK, J., PARKER, T. & MOSKOVITS, D. 1996. Neotropical Birds. Ecology and Conservation. The University Press Ltd., Londres, Gran Bretaña. 450 pp.

WAIDE, R. & NARINS, P. 1988. Tropical forest bird counts and the effect of sound attenuation. *The Auk* (105): 296-302.



**Anexo.** Ejemplo de una ficha de campo para el registro de plantas dentro de un transecto de 20 sub unidades. En las casillas de sub unidades se anota el número de individuos presentes. La Ficha puede ser modificada según los criterios del investigador.

# Ficha de Campo Botánica

Zona de estudio: \_\_\_\_\_ Transecto: \_\_\_\_\_

Coordenadas: \_\_\_\_\_

Fecha: \_\_\_\_\_

Nº de colecta y especie	Nº duplicados	Hábito y/o descripción	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Sub 7	Sub 8	Sub 9	Sub 10	Sub 11	Sub 12	Sub 13	Sub 14	Sub 15	Sub 16	Sub 17	Sub 18	Sub 19	Sub 20







*Instituto de Investigaciones  
de la Amazonía Peruana*

BIODAMAZ  
Perú - Finlandia

## EQUIPO TÉCNICO DEL PROYECTO

### DIRECCIÓN DEL PROYECTO:

Hernán Tello Fernández  
Sanna-Kaisa Juvonen  
Jukka Salo

Director Nacional  
Coordinadora del Proyecto  
Coordinador Científico (Univ. de Turku)

### COMPONENTE 1: Estrategia, Planes de Acción y Sistema de Información

Luis Campos Baca  
Martín Cárdenas Vásquez  
Ada Castillo Ordinola  
Rosana Gonzáles Arzubialdes  
Pedro Gratelly Silva  
Luis Gutiérrez Morales  
Antonietta Gutiérrez-Rosati  
Yolanda Guzmán Guzmán  
Sanna-Kaisa Juvonen  
Risto Kalliola  
Carlos Linares Bensimon  
José Maco García  
Jean Mattos Reaño  
José Mena Álvarez  
Víctor Miyakawa Solís  
Víctor Montreuil Frías

Jukka Salo  
Hernán Tello Fernández  
Tuuli Toivonen  
Jimmy Vargas Moreno  
Lissie Wahl

### COMPONENTE 2: Análisis Regional de la Diversidad Biológica

Fernando Rodríguez Achung  
León Bendayán Acosta  
Glenda Cárdenas Ramírez  
Luis Cuadros Chávez  
Alicia De la Cruz Abarca  
Filomeno Encarnación Cajañaupa  
Lizardo Fachín Malaverri  
Darwin Gómez Ventocilla  
Sanna-Kaisa Juvonen  
Risto Kalliola  
Nelly Llerena Martínez  
José Luis Hurtado  
Sandra Ríos Torres  
Carlos Rivera Gonzáles  
Juan Rodríguez Gamarra  
Kalle Ruokolainen  
Edwin Salazar Zapata  
José Sanjurjo Vílchez

Pekka Soini  
Salvador Tello Martín  
Víctor Vargas Paredes

### **COMPONENTE 3: Conservación *In Situ* Ex Situ**

Kember Mejía Carhuanca  
Nélida Barbagelata Ramírez  
Martín Cárdenas Vásquez  
Ada Castillo Ordinola  
Tania de la Rosa  
Roosevelt García Villacorta  
Alicia Julián Benites  
Sanna-Kaisa Juvonen  
Risto Kalliola  
Markku Kanninen  
Ari Linna  
José Maco García  
Matti Räsänen  
Francisco Reátegui Reátegui  
Kalle Ruokolainen  
Ilari Sääksjärvi  
Jukka Salo  
Matti Salo

Hernán Tello Fernández  
Heiter Valderrama Freyre  
Nélida Valencia Coral  
Rodolfo Vásquez Martínez  
Julio Villacorta Ramírez  
Mari Walls

### **APOYO A LA GESTIÓN DEL PROYECTO**

Edwin Arrarte Flores  
Víctor Chung Bartra  
Jessica Díaz Alvarado  
Valentín García Ríos  
Miguel Pinedo Arévalo  
Carlos Suárez  
Diana Tang Tuesta



## SIGLAS

ANP:	Área Natural Protegida
BIODAMAZ:	Proyecto Diversidad Biológica de la Amazonía Peruana. Convenio Perú-Finlandia
CDB:	Convenio sobre la Diversidad Biológica
ERABI:	Ecorregión del Río Amazonas y Bosques Inundables
ERDBA:	Estrategia Regional de la Diversidad Biológica Amazónica
GPS:	Sistema de Posicionamiento Mundial
IIAP:	Instituto de Investigaciones de la Amazonía Peruana
SIAMAZONÍA:	Sistema de Información de la Diversidad Biológica y Ambiental de la Amazonía Peruana
SIG:	Sistema de Información Geográfica
UA:	Unidades Ambientales
UEE:	Unidades Ecológicas Económicas
UNMSM:	Universidad Nacional Mayor de San Marcos
UNAP:	Universidad Nacional de la Amazonía Peruana
WRS:	( <i>Worldwide Referente System</i> ) Sistema de notación global para datos Landsat
ZEE:	Zonificación Ecológico-Económica

**Por:**

Kalle Ruokolainen, Juan José Rodríguez, Hanna Tuomisto, Pekka Soini,  
Glenda Cárdenas Ramírez, Alicia De La Cruz Abarca, Nelly Llerena  
Martínez, Carlos Rivera Gonzáles, Edwin Salazar Zapata, Víctor Hugo  
Vargas Paredes y Sanna-Kaisa Juvonen

**En colaboración con:**

Grupo de Investigación Amazonía de la Universidad de Turku de  
Finlandia (UTU-ART)

**Edición:**

Editora responsable: Sanna-Kaisa Juvonen  
Edición: José Álvarez Alonso, Sanna-Kaisa Juvonen

**Diagramación:**

Dominius Publicidad