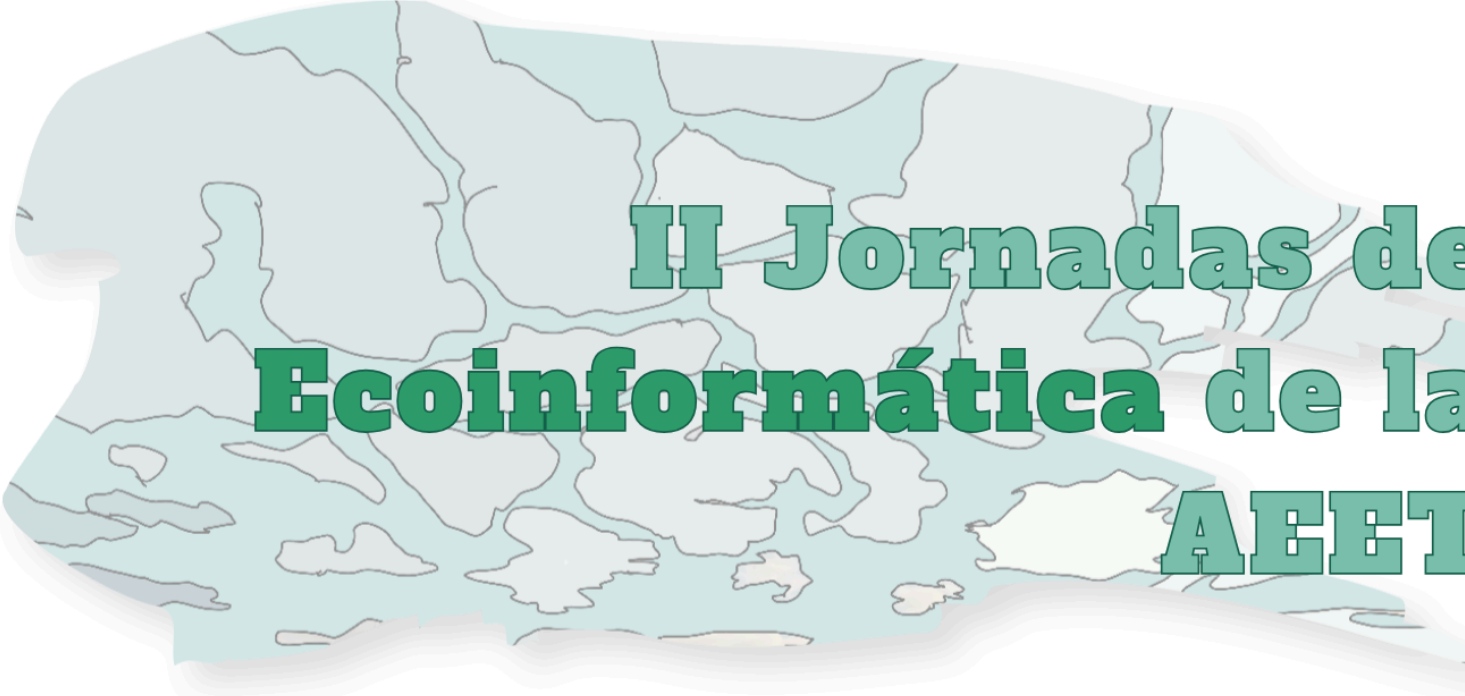




II Jornadas de Ecoinformática de la AEET



Libro de resúmenes

**Sevilla,
1 - 3 de
octubre de 2026**

Colaboran:



Libro de Resúmenes de las II Jornadas de Ecoinformática de la AEET (Sevilla – España, 2026).

COMITÉ ORGANIZADOR

Alberto Cirera

Julia G. de Aledo

Ignacio Ramos-Gutiérrez

Francisco Rodríguez Sánchez

Cristina Ronquillo

COMITÉ CIENTÍFICO

Julen Astigarraga

Verónica Cruz-Alonso

Virginia Domínguez-García

Elena Quintero

PONENCIAS INVITADAS.....	4
IA y análisis de datos en investigación: oportunidades, límites y buenas prácticas (Rosana Ferrero).....	4
De los datos al modelo: cuándo y cómo puedo utilizar la Inteligencia Artificial en mi investigación (Alba Márquez).....	4
Ciencia reproducible: cómo pasar de la teoría a la práctica (Ignasi Bartomeus).....	5
COMUNICACIONES.....	6
All at once: multi-species site-informed Deep Neural Network is a good estimator of species richness, composition and distribution (Admir Cesar de Oliveira Junior).....	6
Claude Code en la investigación ecológica: potencial y límites desde la experiencia de usuario (Alberto Cirera Sancho).....	7
Corrigiendo el efecto de representatividad ambiental mediante el AUC uniforme (Alberto Jiménez-Valverde).....	8
Cuantificación del potencial de secuestro de carbono en bosques monoespecíficos europeos mediante regresiones cuantílicas de límite superior (Cristina Grajera-Antolín).....	9
coverkipR: un paquete de R para la evaluación integrada de sesgos y vacíos de conocimiento en los datos de biodiversidad (Cristina Ronquillo).....	10
CropMED: A harmonized fine-scale crop area dataset across administrations in Europe, the Mediterranean and the Levant (Darío San-Segundo Molina).....	11
Una aproximación ecoinformática al ensamblaje de grandes bases de datos (Francisco Rodríguez-Sánchez).....	11
Hacia la creación de un sistema de detección de decaimiento forestal en Andalucía (José Raúl Raya Castillo).....	13
Utilizando los modelos subestacionales para crear un sistema de alerta de la biodiversidad (Josep M. Serra-Díaz).....	14
LifeWatch Workflow for Automated, Reproducible and Interoperable Environmental Data Integration into a Digital Twin (Julen Torrens-Baile).....	15
Las reconstrucciones filogenéticas revelan patrones evolutivos en la selección de plantas por comunidades indígenas de la Amazonía occidental (Leiba Rodríguez-Mañes).....	16
reserBUGS: A reservoir computing framework for probabilistic forecasting of ecological abundance time series (Miguel A. Mohedano-Muñoz).....	17
NHEX — Niche Hypervolume Explorer: una aplicación interactiva para el análisis y comparación de nichos ecológicos mediante hipervolumenes n-dimensionales (Mónica Gómez Vadillo).....	18
HiBlossoms: una plataforma web para el monitoreo de floraciones de cianobacterias en embalses mediante teledetección con Sentinel-2 (Nicolás Goyena Serveto).....	19
TALLERES.....	20
Taller I: “Pipelines reproducibles del campo al resultado con targets”.....	20
Taller II: “Datos espaciales en Ecología: fundamentos y buenas prácticas”.....	20
Taller III: “Introducción al análisis de redes ecológicas”.....	21
Taller IV: “Ndvi2Gif análisis multi-estacional, fenología, clasificación e hidropериodo con Google Earth Engine”.....	21

PONENCIAS INVITADAS

IA y análisis de datos en investigación: oportunidades, límites y buenas prácticas

Ponente: Rosana Ferrero (Máxima Formación SL.)

La irrupción de la IA generativa está transformando el análisis de datos, la construcción de modelos y la toma de decisiones en investigación científica. Esta ponencia propone una reflexión crítica sobre sus oportunidades y límites, con especial atención a la reproducibilidad, la calidad del dato, el rigor estadístico y el papel del pensamiento crítico en los flujos de trabajo actuales. A través de ejemplos prácticos aplicados a contextos como la Ecología, se discutirán buenas prácticas para integrar la IA y la ciencia de datos de forma rigurosa, transparente y útil en la investigación aplicada.

De los datos al modelo: cuándo y cómo puedo utilizar la Inteligencia Artificial en mi investigación.

Ponente: Alba Márquez (Universidad de Cádiz)

La Inteligencia Artificial está transformando la forma en que se recopilan, procesan e interpretan los datos ecológicos, impulsando nuevas formas de analizar imágenes, sonidos, datos espaciales y otras fuentes de información. Sin embargo, la aparición de nuevos algoritmos, modelos y herramientas también está generando una pregunta cada vez más frecuente; ¿cuándo merece la pena utilizar Inteligencia Artificial y cómo puede integrarse dentro del método científico?

Esta charla propone un marco práctico para responder a esta pregunta. A partir de ejemplos de distintos ámbitos de la ecología, se presentarán conceptos fundamentales como la diferencia entre datos estructurados y no estructurados, la evolución desde el *Machine Learning* hasta los actuales modelos fundacionales y algunas estrategias como el *few-shot learning* y enfoques *human-in-the-loop*. Se discutirá en qué situaciones estas técnicas pueden complementar o aportar ventajas frente a los enfoques estadísticos tradicionales. El objetivo será proporcionar criterios para seleccionar la herramienta adecuada en función de la pregunta ecológica, el tipo de datos disponibles y los objetivos del estudio.

Más allá de la descripción tecnológica, la charla abordará el papel de la investigación en un escenario donde la automatización es cada vez mayor, entendiendo la IA como una herramienta para ampliar la capacidad de análisis y no para sustituir el criterio científico ni el conocimiento ecológico. La sesión pretende ofrecer herramientas para que cada participante pueda identificar cómo integrar estas metodologías en sus propios proyectos de investigación, así como las oportunidades, limitaciones y buenas prácticas necesarias para desarrollar flujos de trabajo reproducibles y científicamente robustos.

Ciencia reproducible: cómo pasar de la teoría a la práctica

Ponente: Ignasi Bartomeus (Estación Biológica de Doñana-CSIC)

A pesar del consenso entre la nueva generación de investigadores e investigadoras sobre la importancia de hacer una ciencia reproducible, sigue habiendo barreras tanto sociales como técnicas para su implementación. En este taller daremos una serie de herramientas fáciles de usar para que hacer ciencia reproducible no tenga un coste técnico significativo, y discutiremos cómo lidiar con barreras de índole más filosófico.

All at once: multi-species site-informed Deep Neural Network is a good estimator of species richness, composition and distribution.

Admir Cesar de Oliveira Junior^{1*}

¹*Instituto Botánico de Barcelona/CSIC, Barcelona, España*

*Autor de correspondencia: admircjunior@gmail.com

Species Distribution Models, Deep Neural Networks, Community Ecology, Macroecology, Latent Factors.

Species Distribution Models (SDMs) are key tools in Ecology, holding significant relevance for both theoretical and applied research. However, because they are typically built within a single-species framework, they often lack the community-level information required to provide comprehensive insights for conservation planning. Since community-level factors, such as species composition and richness, are known to be important drivers of species distributions, several approaches have been proposed to address this limitation. One approach involves stacking multiple single-species SDM predictions to construct richness maps and extract community matrices (Stacked-SDMs). Another involves modeling a multi-species matrix as a response to the same set of predictors, capturing the entire assemblage's relationship with the environment simultaneously (Joint-SDMs). While popular, these approaches present significant limitations, whether in their conceptual frameworks (Stacked-SDMs) or computational complexity (Joint-SDMs). Using empirical community matrices from forest surveys in the Gran Chaco region, I addressed this gap through a comparative exploration across species and algorithms. I used two distinct multi-species Deep Neural Networks, including one site-informed network designed to model site wise latent factors (iDNN), alongside four single-species SDM algorithms to generate spatial predictions for (1) species richness, (2) composition, and (3) individual species distributions. The models were evaluated and compared based on their predictive performance of the three factors above, using Mean Absolute Error, Sorensen's similarity index, and True Skill Statistics metrics. The iDNN consistently outperformed competing methods, demonstrating marked superiority in richness prediction and greater reliability in modeling composition and species distribution. These findings highlight the robust potential of site-informed neural networks for community-level distribution modeling and underscore the value of further developing this approach to inform conservation strategies.

Claude Code en la investigación ecológica: potencial y límites desde la experiencia de usuario.

Alberto Cirera Sancho¹*

¹ Instituto Pirenaico de Ecología IPE-CSIC

*Autor de correspondencia: acirera@ipe.csic.es

Claude Code, Inteligencia artificial, Vibe coding, Ecoinformática, Ciencia abierta y reproducible.

Los modelos de lenguaje de inteligencia artificial han experimentado un crecimiento acelerado en los últimos años y su uso se está generalizando en ámbitos muy diversos, incluida la investigación científica. Dentro de este panorama han surgido herramientas especializadas en programación y construcción de código, como Codex de OpenAI o Claude Code de Anthropic (que, en la actualidad, ya no sirven exclusivamente para generar código).

Esta comunicación aborda, desde la experiencia de usuario y con una mirada crítica, qué puede y qué no puede aportar Claude Code al proceso de investigación en ecología. La premisa es que conocer con precisión el funcionamiento de estas herramientas permitirá discriminar en qué casos su uso puede resultar beneficioso y en cuáles puede ser perjudicial, tanto para la calidad de la investigación como para el propio proceso de aprendizaje de quien la utiliza.

A través de ejemplos y casos prácticos se revisarán brevemente los elementos clave de la herramienta: contexto, modos de trabajo, gestión de permisos; modelos disponibles, esfuerzo y consumo de tokens; y skills, rutinas y agentes. Desde una perspectiva de buenas prácticas y de ciencia abierta y reproducible, se comentarán los límites que conviene tener presentes.

La sesión se plantea como una introducción a un aprendizaje compartido, y pretende generar un espacio de debate posterior orientado a poner en común experiencias y dudas de los participantes.

Corrigiendo el efecto de representatividad ambiental mediante el AUC uniforme.

Alberto Jiménez-Valverde^{*}, Pablo Peña-Aguilera¹

¹ *Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN), CSIC, Madrid, España*

*Autor de correspondencia: alberto.jimenez@mncn.csic.es

Modelos de distribución de especies, uAUC, validación

El *uAUC* es una propuesta metodológica diseñada para corregir el denominado efecto de representatividad ambiental, un problema que afecta a la evaluación de los modelos de distribución de especies. Este efecto surge porque el poder de discriminación de un modelo depende no solo de su capacidad real para distinguir presencias y ausencias, sino también de la distribución de los valores de idoneidad ambiental presentes en el conjunto de datos de validación. Como consecuencia, un mismo modelo puede obtener valores de discriminación diferentes en distintos territorios o conjuntos de datos, incluso cuando la relación entre la especie y el ambiente permanece inalterada. Para resolver este problema, el *uAUC* impone una distribución uniforme de los valores de idoneidad mediante la ponderación de los casos. De este modo, la métrica deja de depender de la frecuencia relativa de ambientes favorables o desfavorables y pasa a reflejar de forma más fiel la capacidad intrínseca de discriminación del modelo. La formulación original utilizó remuestreo estratificado con ponderación inversa, mientras que una propuesta posterior basada en la incorporación directa de pesos reduce el sesgo del estadístico y abre la posibilidad de incorporar diversas fuentes de incertidumbre asociadas a los datos en el proceso de validación. Así, el *uAUC* permite comparar de manera coherente modelos evaluados en diferentes regiones o períodos, revelando situaciones en las que el *AUC* convencional sobrestima o subestima el rendimiento real. Además, la metodología puede extenderse a escenarios con datos de presencia-*background* y a otras métricas de discriminación. En conjunto, el *uAUC* proporciona una base más sólida para interpretar diferencias biológicas reales y evaluar la transferibilidad de los modelos.

Cuantificación del potencial de secuestro de carbono en bosques monoespecíficos europeos mediante regresiones cuantílicas de límite superior.

Cristina Grajera-Antolín^{1*}, Julen Astigarraga², Thomas A.M. Pugh², Marina Rodes Blanco¹, Verónica Cruz-Alonso¹, Miriam Bravo-Hernández¹, Paloma Ruiz-Benito¹, Miguel A. Zavala³, Sara Villén-Pérez⁴

¹ Universidad de Alcalá, Madrid, Spain

² Universidad de Lund, Lund, Sweden

³ Universidad de Alcalá, Departamento de Ciencias de la Vida, Forest Ecology and Restoration Group (FORECO), Avenida de Madrid, s/n, Alcalá de Henares, 28802, Madrid, Spain

⁴ Universidad de Alcalá, GloCEE – Global Change Ecology and Evolution Research Group, Departamento de Ciencias de la Vida, Alcalá de Henares 28805, Madrid, Spain

*Autor de correspondencia: cristina.grajera@uah.es

Almacenamiento de carbono, biomasa, perturbaciones, regresiones cuantílicas, tasas de crecimiento.

La evaluación del potencial de secuestro de carbono en los bosques europeos es fundamental para el diseño de estrategias efectivas de mitigación del cambio climático. En este trabajo estimamos el almacenamiento potencial de carbono en bosques monoespecíficos a escala europea. Para ello, utilizamos datos de 60.681 parcelas procedentes de inventarios forestales nacionales de 10 países para aplicar regresiones cuantílicas de límite superior a la biomasa de las especies arbóreas más ampliamente distribuidas del continente. Como resultado de estos modelos, obtuvimos la biomasa potencial considerando diversas variables bióticas y abióticas que actúan como las principales limitantes de los bosques. Posteriormente, se transformó en carbono potencial mediante factores de conversión específicos. Además, incorporamos las tasas de crecimiento de las especies y el régimen de perturbaciones regionales para estimar la probabilidad de alcanzar dicho máximo de almacenamiento a lo largo del tiempo. De este modo, cuantificamos el potencial de secuestro de carbono como la diferencia entre el carbono observado y el carbono máximo alcanzable en función de dicha probabilidad. Esta aproximación metodológica, que integra el potencial de almacenamiento de carbono estático con la dinámica temporal de crecimiento y perturbaciones de los bosques, ofrece una metodología novedosa y realista para orientar el diseño de estrategias de mitigación al cambio climático.

coverkipR: un paquete de R para la evaluación integrada de sesgos y vacíos de conocimiento en los datos de biodiversidad.

Cristina Ronquillo*

¹ *Dpto. Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, España*

*Autor de correspondencia: cristinaronquillo@mncn.csic.es

Compleitud de inventario, cobertura ambiental, paquete R, sesgos de muestreo, vacíos de conocimiento.

La creciente compilación de datos de biodiversidad procedentes de repositorios ha aumentado sustancialmente en las últimas décadas, sin embargo, estos datos son inherentemente incompletos y presentan sesgos, lo que compromete la fiabilidad e interpretación de los patrones ecológicos. *coverkipR* es un paquete de R diseñado para macroecólogos, biogeógrafos y gestores de datos que ofrece un marco reproducible con el que evaluar vacíos y cobertura de conocimiento y detectar zonas prioritarias para nuevos muestreos.

El paquete integra siete funciones que procesan registros de ocurrencia de especies y capas ambientales para generar mapas en tres dimensiones complementarias de diagnóstico de la información: (i) mapas del grado de conocimiento (evaluado mediante completitud de inventario y distancia geográfica a celdas bien muestreadas); (ii) mapas de cobertura temporal y ambiental de los datos mediante la reconstrucción de espacios multidimensionales; y (iii) mapas de accesibilidad (proximidad a infraestructuras). El conjunto de mapas obtenidos clasifica cada celda con un valor de prioridad de modo que aquellas áreas con bajos niveles de conocimiento y baja accesibilidad o cobertura pueden requerir expediciones específicas, mientras que regiones con alta accesibilidad y bajo conocimiento representan oportunidades de movilización de datos más eficaces, optimizando los esfuerzos de muestreo.

coverkipR amplía el conjunto de herramientas metodológicas disponibles, reforzando la gestión informada y una planificación estratégica del muestreo y conservación de la biodiversidad.

CropMED: A harmonized fine-scale crop area dataset across administrations in Europe, the Mediterranean and the Levant.

Darío San-Segundo Molina^{*}, Sara Villén-Pérez¹, Ignacio Morales-Castilla¹

¹ *Departamento de Ciencias de la Vida, GloCEE –Global Change Ecology & Evolution Research Group, Universidad de Alcalá, Alcalá de Henares, España.*

*Autor de correspondencia: dario.san@uah.es

Hectares, georeferenced, NUTS3, crop richness, agriculture.

The Mediterranean Basin and Europe have been cultivated for millennia. Official statistical agencies routinely collect and report crop areas for hundreds of crops at fine administrative resolution. Yet, despite decades of European statistical harmonization and the routine collection of these data by national administrations, no harmonized cross-country dataset currently aggregates them while preserving their original subnational resolution. *CropMED* is a comprehensive database compiling sown, harvested, or utilised hectares of 141 unique crops across 1,709 administrative units of 47 countries in Europe, the Mediterranean region, and the Levant. The dataset preserves the original reporting resolution of each national source, facilitating their use in policy and land management applications. Over 92% of units correspond to Eurostat NUTS3 or equivalent divisions (e.g., governorates, wilayas, prefectures, provinces, or counties). The most widely cultivated crops include cereals (e.g., wheat, barley, maize, oats), potatoes, oilseeds (e.g., sunflower, rapeseed), fruit orchards (e.g., olive groves, vineyards), and legumes (e.g., beans, peas). Compared with existing georeferenced datasets, *CropMED* provides harmonized data directly reported by the responsible administrations, minimizing the uncertainty introduced by spatial interpolation and serving as a baseline for future gridded datasets and continuous updates as source data are newly submitted by their respective institutions. Beyond describing crop distributions, the database enables new ecological applications, including modelling realized crop niches and incorporating host plant availability into forecasts of biological invasions, particularly for pest risk assessments. To facilitate reuse by the research community, *CropMED* is openly available through Zenodo (<https://zenodo.org/records/20717354>) and the project website (<https://dario-ssm.github.io/CropMED/>).

Una aproximación ecoinformática al ensamblaje de grandes bases de datos.

Francisco Rodríguez-Sánchez^{1*}, Elena Quintero^{1,2}, Ignacio Ramos-Gutiérrez^{1,3}

¹ *MAYBE_lab, Departamento de Biología Vegetal y Ecología, Universidad de Sevilla*

² *Evolving Networks Lab, Swedish University of Agricultural Sciences*

³ *Departamento de Biología, Universidad Autónoma de Madrid*

*Autor de correspondencia: frodriguez@us.es

Ecoinformatics, database, reproducibility, GitHub, open science.

En esta charla se abordarán los desafíos existentes cuando se trata de ensamblar bases de datos a partir de grandes volúmenes de información de calidad variable, y compartiremos los errores, aciertos y lecciones aprendidas durante el ensamblaje de FRUSEED, una base de datos global, dinámica, abierta y reproducible sobre rasgos de frutos y semillas y modos de dispersión en plantas que cuenta con varios millones de registros. Se ilustrarán algunas de las técnicas y herramientas disponibles para abordar la tarea con eficiencia y rigor, incluyendo git, GitHub, Rmarkdown, Quarto, control de calidad automático y actualización continua con GitHub Actions, y consulta rápida y eficiente con DuckDB. También se discutirán algunos de los desafíos pendientes, incluyendo la validación taxonómica y el reconocimiento justo de la contribución de los recolectores originales de los datos. Por último, se discutirá el posible papel de la inteligencia artificial en el ensamblaje de bases de datos masivas a partir de fuentes de datos no estructuradas.

Hacia la creación de un sistema de detección de decaimiento forestal en Andalucía.

José Raúl Raya Castillo*, Cristina Acosta Muñoz¹

¹ Universidad de Córdoba

*Autor de correspondencia: a02racaj@uco.es

Teledetección, cambio climático, decaimiento forestal, pinares de repoblación, estado fitosanitario.

El cambio climático está alterando de forma significativa el funcionamiento de los ecosistemas forestales, especialmente en regiones mediterráneas caracterizadas por una elevada aridez. El incremento de la concentración atmosférica de CO₂, junto con el aumento de las temperaturas y la intensificación de las sequías, está modificando los patrones de crecimiento, productividad y distribución de las masas forestales, favoreciendo episodios de decaimiento y mortalidad arbórea. Estos procesos tienen importantes repercusiones ecológicas, económicas y sociales, lo que hace imprescindible desarrollar herramientas que permitan evaluar de forma objetiva el estado fitosanitario de los bosques y anticipar su respuesta frente al cambio global.

Tradicionalmente, la evaluación de la salud forestal se ha basado en inventarios de campo y en indicadores de supervivencia y mortalidad, procedimientos que requieren un elevado esfuerzo técnico y económico. En este contexto, la teledetección constituye una alternativa especialmente útil al proporcionar información continua sobre grandes extensiones de territorio. Sin embargo, pese a los importantes avances tecnológicos alcanzados en las últimas décadas, todavía no existe una metodología universal, sencilla y robusta que permita monitorizar de manera sistemática el estado fitosanitario de las masas forestales.

Este trabajo surge con el objetivo de contribuir al desarrollo de una metodología de seguimiento aplicada a los pinares de repoblación de Andalucía, un tipo de bosque que en los últimos años ha mostrado episodios de mortalidad creciente. La propuesta pretende facilitar la detección temprana de procesos de decaimiento mediante el uso de productos derivados de la teledetección, proporcionando una herramienta que contribuya a mejorar la gestión y conservación de estos ecosistemas en un escenario de cambio climático.

Utilizando los modelos subestacionales para crear un sistema de alerta de la biodiversidad.

Josep M. Serra-Díaz^{1*}, Lauren C. Andrews², Michiel D. Pillet³, Adam M. Wilson⁴, Andreas Schwarz Meyer⁵ Christopher H. Trisos^{5,6}, Brian Maitner³, Mark C. Urban⁷, Cory Merow⁷

¹ *Institut Botànic de Barcelona (CSIC-CMCNB), 08038 Barcelona, España*

² *Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA*

³ *Department of Integrative Biology, University of South Florida, St. Petersburg, FL 33701, USA.*

⁴ *Departments of Geography, Environment & Sustainability, University at Buffalo, Buffalo, NY, USA*

⁵ *Climate Risk Lab, African Climate and Development Initiative, University of Cape Town, Cape Town, South Africa*

⁶ *African Synthesis Centre for Climate Change, Environment and Development (ASCEND), University of Cape Town, Cape Town, South Africa*

⁷ *Department of Ecology and Evolutionary Biology and Center of Biological Risk, University of Connecticut, Storrs, CT, USA*

*Autor de correspondencia: pep.serra.diaz@csic.es

Cambio climático, sistemas de alerta, modelos climáticos subestacionales, ola de calor.

El cambio climático intensifica la frecuencia y magnitud del calor extremo, pero la mayoría de las predicciones sobre biodiversidad se centran en horizontes de décadas o describen los impactos de forma retrospectiva, ofreciendo poca orientación para la acción conservacionista inmediata. Presentamos un sistema de alerta temprana que anticipa la exposición de la biodiversidad al calor sin precedentes combinando las previsiones climáticas subestacionales-estacionales de la NASA (GEOS-S2S) con los límites térmicos históricos propios de cada especie (percentil 99 desde 1980). El sistema pronostica, con entre 1 y 9 meses de antelación, la exposición de 30.585 vertebrados terrestres (anfibios, reptiles, aves y mamíferos). En una primera aplicación retrospectiva (mayo de 2024) predijimos que más de 3.500 especies, incluidas más de 1.250 de interés para la conservación, quedarían expuestas a temperaturas extremas; México, el África subsahariana y el Himalaya concentraron el mayor riesgo, y evidencias posteriores —como mortandades de monos aulladores en México— confirmaron impactos poblacionales. Aplicando el mismo marco al fuerte episodio de El Niño previsto para 2026-2027, pronosticamos que 4.159 especies (17% de los vertebrados terrestres) afrontarán calor sin precedentes en más del 30% de su distribución en menos de nueve meses, con la exposición concentrada en el Hemisferio Sur y el verano austral; anfibios y reptiles son los grupos más afectados, y un criterio mensual duplica las especies expuestas. Estos resultados muestran cómo los modelos subestacionales permiten pasar de documentar pérdidas de biodiversidad a anticiparlas, orientando el seguimiento y la mitigación allí donde más urgen.

LifeWatch Workflow for Automated, Reproducible and Interoperable Environmental Data Integration into a Digital Twin.

Julen Torrens-Baile¹, Nicolás Goyena Serveto¹ Esther Lasheras¹, Carolina Santamaría¹, David Elustondo¹

¹ Instituto BIOMA, Universidad de Navarra

*Autor de correspondencia: Sheila Izquieta-Rojano (sizquieta@unav.es)¹

LifeWatch; Digital Twin, Open Science, Environmental Monitoring, Data Interoperability, ICP Forests, ICP Integrated Monitoring.

This work presents a LifeWatch-supported workflow for automated preprocessing, validation, storage, and continuous updating of environmental monitoring data, using water chemistry data from the European Forest Monitoring Networks, specifically ICP Forests and ICP Integrated Monitoring, as a case study. LifeWatch provides a digital research infrastructure that supports open, reproducible, and interoperable science by enabling researchers to develop, execute, share, and reuse computational workflows within a transparent framework.

In this context, the workflow addresses the need to transform heterogeneous monitoring data into reliable, standardized, and continuously available information. ICP monitoring programmes generate long-term datasets that are valuable for assessing ecosystem responses to atmospheric pollution and climate change, but their use requires robust procedures for data harmonization, quality control, and integration.

The workflow covers the complete data lifecycle. Field and laboratory data are collected and recorded in standardized templates, which are automatically checked in a first LifeWatch component. This component contains executable code that verifies data structure, format, and completeness. In LifeWatch, workflows are built by connecting computational components, each based on code developed in Python, R, or other languages, with defined inputs and outputs. After this initial check, the data pass through subsequent components for preprocessing, where raw measurements are cleaned, transformed, and harmonized, and then for chemical validation, where standardized quality-control rules assess consistency and reliability. Validated outputs are compiled into final datasets and stored in a structured database integrating historical and newly generated data.

This database will continuously feed a Digital Twin, keeping it dynamically updated as new validated data become available. Thus, the Digital Twin becomes a living representation of monitored ecosystems, supporting visualization, research, and long-term environmental monitoring while ensuring automation, traceability, reproducibility, interoperability, and continuous data availability.

Las reconstrucciones filogenéticas revelan patrones evolutivos en la selección de plantas por comunidades indígenas de la Amazonía occidental.

Leiba Rodríguez-Mañes^{1,2*}, Julia G. de Aledo³, Laura Matas-Granados³, Rafael Molina-Venegas³, Manuel J. Macía^{4,5}

¹ Universidad Internacional Menéndez Pelayo, C. de Isaac Peral, 23, Moncloa - Aravaca, 28040 Madrid, España.

² Real Jardín Botánico, CSIC, Plaza de Murillo, 2, Madrid, España.

³ Estación Biológica de Doñana (EBD), CSIC, C/Américo Vespucio 26, 41092, Sevilla, España.

⁴ Departamento de Biología, Unidad de Botánica, Universidad Autónoma de Madrid, Calle Darwin 2, 28049 Madrid, España.

⁵ Centro de Investigación en Biodiversidad y Cambio Global (CIBC - UAM), Universidad Autónoma de Madrid, Calle Darwin 2, 28049 Madrid, España.

*Autor de correspondencia: leibarod@gmail.com

Conocimiento tradicional, convergencia etnobotánica, incertidumbre filogenética, rasgos funcionales, señal filogenética, versatilidad de uso.

Las filogenias constituyen una herramienta clave para comprender los patrones de selección y uso de las plantas, pero su aplicación en regiones con una gran diversidad como la cuenca de la Amazonía está limitada por la escasa disponibilidad de información molecular. En este trabajo desarrollamos un flujo reproducible para expandir filogenias incompletas de plantas utilizando el árbol de referencia más actualizado hasta la fecha, que incluye información molecular de 128.270 especies, y estandarizando los nombres según Plants of the World Online. El estudio incluyó 1.787 especies de plantas leñosas tropicales registradas en Ecuador, Perú y Bolivia. Los objetivos fueron desarrollar una expansión filogenética de estas especies, analizar si la versatilidad de uso presenta una señal evolutiva y evaluar si las distintas categorías de uso muestran agrupamiento filogenético. Para ello, se generaron 100 filogenias mediante *'randtip'*, permitiendo incorporar la incertidumbre filogenética. Sobre estas filogenias, para analizar la señal evolutiva utilizamos la λ de Pagel y para el agrupamiento filogenético las métricas complementarias MPD (*Mean Pairwise Distance*) y MNTD (*Mean Nearest Taxon Distance*). La versatilidad de uso mostró una señal filogenética débil a nivel de especie, pero no de familia. Además, las especies utilizadas para utensilios, herramientas, construcción, combustible y alimentación presentaron agrupamiento filogenético, mientras que las empleadas con fines medicinales y culturales se distribuyeron aleatoriamente. La comparación entre las 100 reconstrucciones permitió evaluar la sensibilidad de estas métricas teniendo en cuenta la incertidumbre filogenética. Este trabajo demuestra que la integración de reconstrucciones filogenéticas y el desarrollo de un flujo reproducible mejoran la interpretación de los patrones etnobotánicos y proporcionan una metodología aplicable a otros sistemas tropicales con información molecular limitada.

reserBUGS: A reservoir computing framework for probabilistic forecasting of ecological abundance time series.

Miguel A. Mohedano-Muñoz^{1,2}, Javier Galeano^{1,3}, Juan M. Pastor^{1,3}, Julia G. de Aledo⁴, Ignasi Bartomeus⁴, Alfonso Allen-Perkins^{1,3*}

¹Grupo de Sistemas Complejos (GSC), Universidad Politécnica de Madrid, Madrid, Spain

²Departamento de Informática y Estadística, Universidad Rey Juan Carlos, Madrid, Spain

³Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain

⁴Estación Biológica de Doñana (EBD-CSIC), Avda. Americo Vesputio 26, Seville E-41092, Spain

*Autor de correspondencia: alfonso.allen.perkins@upm.es

Reservoir computing, Ecological forecasting, Probabilistic forecasting, Population dynamics, Time-series analysis

Predicting how animal populations will change in the future is essential for biodiversity conservation, ecological management, and understanding the impacts of environmental change. Yet population dynamics arise from complex interactions among climate, habitat, species interactions, and stochastic processes, making reliable forecasts particularly challenging, especially when only short ecological time series are available.

We present reserBUGS, an open-source computational framework for probabilistic forecasting of insect population abundance. The tool is based on reservoir computing, a machine-learning approach that uses a large network of randomly connected artificial neurons to capture complex temporal dependencies without the computational cost of training deep neural networks. Unlike conventional forecasting methods, reserBUGS produces full predictive distributions, allowing users not only to estimate future abundance but also to quantify forecast uncertainty and identify situations in which predictions are likely to be unreliable.

We evaluated reserBUGS using thousands of insect population time series representing hundreds of species from around the world. Reservoir-based models consistently outperformed widely used statistical and machine-learning approaches in both point accuracy and probabilistic forecast quality. Incorporating environmental variables, including temperature and precipitation, further improved predictive performance, particularly at forecast horizons of three to four years.

By combining accurate forecasts with explicit estimates of uncertainty and forecast reliability, reserBUGS provides ecologists and conservation practitioners with a practical tool for anticipating population change and supporting evidence-based management decisions under uncertainty.

NHEX – Niche Hypervolume Explorer: una aplicación interactiva para el análisis y comparación de nichos ecológicos mediante hipervolúmenes n-dimensionales.

Mónica Gómez Vadillo^{1,2*}, Emilio García Roselló³, Joaquín Hortal¹

¹ *Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, España*

² *Programa de Doctorado en Ecología, Universidad Autónoma de Madrid, Madrid, España*

³ *Departamento de Informática, Universidad de Vigo, Vigo, España*

*Autor de correspondencia: monica.gomez@mncn.csic.es

Nicho ecológico; hipervolúmenes n-dimensionales; solapamiento de nicho; visualización interactiva; software científico

Los hipervolúmenes n-dimensionales permiten representar los nichos ecológicos como regiones de un espacio ambiental multivariado y cuantificar su amplitud, posición y solapamiento. Sin embargo, su aplicación suele requerir flujos de trabajo complejos y conocimientos de programación, lo que limita su accesibilidad y reproducibilidad. Presentamos una aplicación interactiva desarrollada en Python con Dash para facilitar la construcción, exploración y comparación de nichos multivariados mediante hipervolúmenes.

La aplicación utiliza como entrada registros de especies previamente asignados a una cuadrícula espacial y una tabla de variables ambientales asociadas a cada celda. El usuario puede explorar la distribución espacial de los registros, seleccionar las especies y variables de interés y evaluar estas últimas mediante matrices de correlación, histogramas y diagramas de caja. A partir de esta selección, la herramienta estima hipervolúmenes mediante modelos gaussianos de densidad de kernel o máquinas de vectores de soporte de una clase, implementados en el paquete R hypervolume. Los resultados incluyen el volumen del nicho de cada especie y métricas pareadas de similitud y diferenciación: índices de Jaccard y Sørensen, fracciones exclusivas, volúmenes de intersección y unión y distancia entre hipervolúmenes.

En esta comunicación se presentará el flujo de trabajo implementado, su interfaz y una demostración con datos ecológicos. También se expondrá su desarrollo futuro, que incluye análisis de componentes principales, incorporación directa de datos ambientales y de ocurrencia, cálculo de coocurrencia y evaluación de su relación con el solapamiento de nicho. Finalmente, los módulos desarrollados se integrarán en la aplicación de análisis de datos masivos de biodiversidad ModestR mediante la conexión ya establecida entre Python y C#.

HiBlooms: una plataforma web para el monitoreo de floraciones de cianobacterias en embalses mediante teledetección con Sentinel-2.

Nicolás Goyena Serveto^{1*}, David Elustondo¹, Yasser Morera Gómez¹, Esther Lasheras Adot¹, Jesús Miguel Santamaría¹, Carolina Santamaría Elola¹, Aimee Valle Pombrol¹, Sheila Izquieta Rojano¹

¹ Instituto BIOMA, Universidad de Navarra (UNAV), Pamplona, España.

*Autor de correspondencia: ngoyena@unav.es

Cianobacterias, teledetección, Sentinel-2, calidad del agua, embalses.

Las floraciones de cianobacterias son uno de los principales problemas de calidad del agua en embalses de uso público, con consecuencias directas sobre la salud humana y el funcionamiento de los ecosistemas acuáticos. En España, donde más del 40% de los embalses se consideran susceptibles a episodios de proliferación, la situación se agrava en un contexto de cambio climático que prevé un aumento en la frecuencia e intensidad de estos eventos.

HIBLOOMS es un proyecto de investigación financiado por el Ministerio de Ciencia e Innovación cuyo objetivo es reconstruir la dinámica histórica de las cianobacterias en embalses españoles e implementar un sistema de monitoreo integrado que combina muestreo en campo, análisis biogeoquímico e isotópico, y teledetección satelital.

En el marco de este proyecto se ha desarrollado una plataforma web de código abierto que permite monitorear floraciones a partir de imágenes Sentinel-2 procesadas mediante Google Earth Engine. La herramienta incluye el cálculo automático de índices espectrales sensibles a la biomasa algal, la visualización interactiva de series temporales y la aplicación de modelos de machine learning calibrados con datos de campo para estimar concentraciones de clorofila y ficocianina.

Además, la plataforma ha sido integrada como componente de workflow reutilizable en LifeWatch ERIC, la infraestructura europea de e-ciencia para la biodiversidad, facilitando que cualquier investigador pueda reproducir y adaptar los análisis a sus propios casos de estudio bajo principios FAIR.

Como caso de estudio se presenta el embalse de El Val (Zaragoza), donde las imágenes Sentinel-2 se han combinado con datos físico-químicos de la red SAICA de la Confederación Hidrográfica del Ebro. Los resultados indican que los índices calibrados permiten discriminar episodios de floración con buena precisión, abriendo la puerta a sistemas de alerta temprana basados en satélite aplicables a múltiples embalses de forma simultánea.

TALLERES

Taller I: “Pipelines reproducibles del campo al resultado con targets”.

Guillermo Fandos^{1*}

¹ *Departamento de biodiversidad, ecología y evolución, Universidad Complutense de Madrid*

*Autor de correspondencia: gfandos@ucm.es

El monitoreo pasivo mediante fototrampeo y grabadoras acústicas como AudioMoth ha disparado el volumen y la complejidad de los datos en ecología. Con ello aparecen problemas serios de reproducibilidad, porque los flujos de trabajo tienen muchas etapas (ingesta de archivos, detección y clasificación, agregación y análisis) que son caras de recalcular y difíciles de auditar, sobre todo en proyectos de varios años o con infraestructura compartida. El paquete targets de R aborda esto con un enfoque declarativo tipo make: define el pipeline como un grafo de dependencias, ejecuta solo lo que ha cambiado y mantiene los resultados sincronizados con el código y los datos. Este taller práctico introduce targets usando como hilo conductor dos casos reales, uno de fototrampeo y otro de ecoacústica, aunque los principios sirven para cualquier flujo de trabajo. Los participantes construirán un pipeline completo, desde los archivos crudos hasta las tablas de actividad y los modelos. Por el camino aprenderán a estructurar objetivos y seguir dependencias, a usar branching dinámico para procesar cientos de despliegues o grabaciones sin duplicar código, y a manejar archivos grandes con formatos de tipo fichero. Al terminar, cada asistente se llevará una plantilla funcional y criterios claros para reproducir y escalar sus propios análisis.

Taller II: “Datos espaciales en Ecología: fundamentos y buenas prácticas”.

Marina Rodes Blanco^{1,2*}, Sofía Miguel Romero^{1*}

¹ *Universidad de Alcalá. Department of Geology, Geography and Environment. Environmental Remote Sensing Research Group (GITA), Alcalá de Henares 28805, Spain*

² *Universidad de Alcalá. Department of Life Sciences, Forest Ecology and Restoration Group (FORECO), Universidad de Alcalá, Alcalá de Henares 28805, Spain*

*Autor de correspondencia: marina.rodes@uah.es; sofia.miguelr@uah.es

Datos espaciales, ecología, teledetección, incertidumbre, análisis espacial.

El uso de datos espaciales procedentes de mediciones de campo, estaciones meteorológicas o sensores satelitales, es cada vez más frecuente en ecología gracias a su creciente disponibilidad y a la posibilidad de combinarse entre sí. Dado que los resultados derivados de estas aplicaciones pueden tener consecuencias en la gestión y en la toma de decisiones, es esencial hacer un buen uso de los datos espaciales, conocer sus limitaciones, y las implicaciones de las decisiones metodológicas tomadas en las distintas fases del proceso de analítico.

El objetivo de este taller es presentar los principales aspectos y limitaciones a considerar cuando trabajamos con datos espaciales, así como buenas prácticas para minimizar y comunicar adecuadamente la incertidumbre asociada al uso de información georreferenciada en estudios ecológicos.

La estructura del taller será:

- Aspectos fundamentales de los datos espaciales: sistemas de coordenadas, tipos de datos espaciales y extracción de información.
- Buenas prácticas para el uso e interpretación de información espacial disponible.
- Consideraciones y recomendaciones en el uso de productos derivados de teledetección.

Taller III: “Introducción al análisis de redes ecológicas”.

Virginia Domínguez-García*

¹ *Estación Biológica de Doñana (EBD-CSIC)*

*Autor de correspondencia: virginia.dominguez@ebd.csic.es

En este taller veremos cómo pueden usarse las redes complejas para entender las comunidades ecológicas. Además de una breve introducción teórica sobre la estructura de las redes (qué representan, cómo cuantificar su estructura) se harán ejercicios prácticos en Python y R de análisis básicos de la estructura, y si da tiempo, también visualización.

Taller IV: “Ndv2Gif análisis multi-estacional, fenología, clasificación e hidroperiodo con Google Earth Engine”.

Diego García Díaz ¹*

¹ *Laboratorio de SIG y Teledetección (LAST-EBD), Estación Biológica de Doñana (CSIC), Sevilla, España*

*Autor de correspondencia: diegogarcia@ebd.csic.es

Google Earth Engine, Python, series temporales, fenología, teledetección.

Ndv2Gif es un paquete Python de código abierto, construido sobre Google Earth Engine (GEE) y geemap, que simplifica el análisis multi-temporal de datos de teledetección. Su idea central es organizar la información en composites estadísticos multi-estacionales mediante una estructura jerárquica basada en bandas, lo que permite comparar ventanas temporales fijas (estaciones, meses, quincenas o periodos personalizados) sin escribir código complejo de filtrado, enmascaramiento de nubes y agregación. El paquete da acceso unificado a siete plataformas satelitales (Sentinel-1/2/3, Landsat 4-9, MODIS) y a datos climáticos (ERA5-Land, CHIRPS), con más de 50 índices espectrales y radar predefinidos.

El taller es eminentemente práctico. Tras una breve introducción a la arquitectura de la librería, cada participante trabajará sobre su propio portátil siguiendo unos notebooks guiados que cubrirán: (1) la generación de composites estacionales para una región de interés propia; (2) la extracción de métricas fenológicas (inicio, pico y fin de estación) y la detección de tendencias mediante series temporales;

(3) una clasificación de coberturas del suelo sencilla a partir de los composites; y
(4) un ejemplo de cálculo de hidroperiodo para el seguimiento de láminas de agua. El objetivo es que, al terminar, las personas asistentes sean capaces de replicar estos flujos en sus propias zonas de estudio cambiando una sola palabra (índice, estadístico o sensor) y se lleven plantillas reutilizables. El taller está pensado para ser accesible a cualquier nivel: basta con nociones básicas de Python.