

Taller avanzado de R para el análisis de datos en ciencias ambientales

Posgrados en Ciencias Ambientales – Facultad de Ciencias – UDELAR
Docente a cargo: Dr. Marcos Texeira

Carga horaria: 36 horas de docencia directa

Créditos: 4

Unidad Análisis e integración

Días y horarios:

Seis (6) sesiones distribuidas en el mes de julio del 2026. Horario de 9:00 a 12:00 y de 14:00 a 17:00.

Programa

INTRODUCCIÓN:

R es un poderoso lenguaje de programación y un entorno dedicado principal pero no únicamente al análisis de datos. Este taller pretende reforzar los contenidos introducidos en el taller inicial de R y presentar funcionalidades avanzadas. Esta dirigido a usuarios con conocimientos básicos, provenientes de las ciencias ambientales y disciplinas relacionadas.

OBJETIVOS:

El objetivo del taller es presentar funcionalidades avanzadas y modernas de R / Rstudio para el (pre)procesamiento, análisis de datos, comunicación y presentación de resultados en Cs. Ambientales.

Los datos en Cs ambientales siempre se estructuran espacial y temporalmente, y suelen provenir de distintas fuentes (muestreos a campo y experimentos, data loggers, sensores remotos, modelos y bibliografía). En el taller presentamos herramientas y métodos para integrar, manejar y analizar en R este tipo de información. Se presentarán las buenas prácticas de programación y comunicación de resultados, lo que permitirá a los estudiantes adquirir las habilidades para producir resultados repetibles y transparentes.

CONTENIDOS:

Día 1: Repaso fundamentos de R

Objetivo: refrescar los fundamentos de R

- Estructuras de datos y acceso
- estructuras de programación: funciones
- estructuras de programación: control de flujo y ciclado
- lectura y escritura de datos (I/O)
- paquetes

Día 2: El ecosistema de trabajo moderno en R

Objetivo: introducir a las mejores practicas de uso actuales

- Introducción a **tidyverse**.
- Organización de proyectos con here, renv, config.
- Conexiones a bases de datos y archivos grandes (parquet, feather).
- Importación robusta de datos ambientales (NetCDF, HDF, shapefiles, GeoTIFF), mediante **terra**, **sf**, **arrow** y **ncdf4**.

►► Actividad práctica:

Construir un proyecto reproducible que integre datos espaciales + tabulares (ej. uso de suelo, sensores, clima).

Día 3: Datos con estructura espacial – sensores remotos

Objetivo: introducir a los flujos de trabajo modernos de análisis geoespacial.

- Manipulación avanzada de geometrías con **sf** y raster/vector con **terra**.
- Reproyección, limpieza, overlays sofisticados.
- Cálculo de métricas ambientales: NDVI, pendientes, fragmentación, buffers ecológicos, modelos DEM.
- Automatización de flujos geoespaciales (**purrr** + functions).
- Integrar R con Google Earth Engine (**r-spatial** + **rgee**)

►► **Actividad práctica:**

Generar indicadores ambientales de una región (ej. degradación, conectividad, hotspots, índice de oferta de servicios ecosistémicos).

Día 4: Datos con estructura temporal – series temporales

Objetivo: introducir a los flujos de trabajo modernos de análisis básico de series temporales.

- Manipulación de fechas-horas con **lubridate**
- Estructuras de series temporales: **tsibble** y **tidyverts**.
- Visualización de series múltiples: índices de vegetación, temperatura, precipitación
- Análisis de tendencias y estacionalidad: Descomposición de series
- Detección de anomalías en datos de monitoreo continuo
- Autocorrelación y patrones temporales

►► **Actividad práctica:**

Caracterizar la dinámica de las series temporales presentadas (NDVI, precipitación, IOSE) para una determinada región

Día 5: Modelos estadísticos

Objetivo: Conectar los datos con modelos estadísticos que capturen procesos ecológicos y ambientales.

- Modelos lineales y lineales generalizados (LM/GLM/GLMM) con **lme4** y **glmmTMB**.
- Modelos aditivos (GAM) para fenómenos no lineales: **mgcv**.
- Integrar la estructura espacial (día 3) y/o temporal (día 4) en los modelos.
- Validación cruzada ecológica, selección de modelos, interpretación adecuada.

►► **Actividad práctica:**

Modelar los controles ambientales de alguna de las variables/indicadores ambientales descritos en esta sesión y en la previas.

Día 6: Automatización, comunicación y reproducibilidad

Objetivo: presentar las bases para trabajos transparentes, replicables y comunicables.

- Automatización de análisis con **targets** o **drake**.
- Reportes dinámicos con **Quarto** (HTML, PDF, dashboards).
- Creación de funciones y mini-paquetes propios (**devtools**, **usethis**).
- Buenas prácticas de versionado, documentación y control de calidad.

►► **Actividad práctica:**

Empaquetar el proyecto de la semana (lo trabajado en las sesiones!): pipeline - modelo - reporte dinámico y visualizaciones.

APROBACIÓN DEL CURSO

El taller será aprobado mediante la entrega de un proyecto reproducible completo (pipeline + análisis espacial + modelado + reporte). Este proyecto es el que los estudiantes irán desarrollando/trabajando en el transcurso del taller. Tendrán 15 días tras el fin del taller para la entrega del mismo.

BIBLIOGRAFÍA:

- Chang, W. 2013. R graphics cookbook. O'Reilly.
- Gillespie, C. y Lovelace, R. 2017. Efficient R programming. O'Reilly.

- Inchausti, P. 2023. Statistical modeling with R: A dual frequentist and bayesian approach for life scientists. Oxford University Press.
- McElreath, R. 2020. Statistical Rethinking: A bayesian course with examples in R and Stan. Chapman & Hall.
- Wegmann, M., Leutner, B. y Dech, S. 2016, Remote sensing and GIS for ecologists using open source software.
- Wickham, H y Grolemond, G. 2017. R for data science. O'Reilly.