

R como un SIG: Extracción de datos climáticos de WorldClim

PAULA N. FERGNANI

*Laboratorio Ecotono. Centro Regional Universitario Bariloche-Universidad Nacional del Comahue, INIBIOMA-CONICET.
San Carlos de Bariloche, Río Negro, Argentina.*

RESUMEN. Caracterizar los sitios de muestreo en base a datos climáticos es una necesidad clave para un número importante de profesionales que trabajan en ecología. De manera más general, también lo es utilizar herramientas SIG (Sistemas de Información Geográfica). Sin embargo, acceder al uso de las bases de datos climáticas disponibles en línea y a las herramientas SIG supone un desafío. En este trabajo se muestra de qué manera usar R como un SIG para obtener valores de variables climáticas para sitios de muestreo georreferenciados. Se utiliza la base de datos WorldClim, que contiene información climática para todo el globo. En primer lugar, se describe en qué consiste la base de datos WorldClim. En base a una búsqueda bibliográfica, se analiza cuánto se la usa en la actualidad y en qué temas. Se concluye que el uso de esta base de datos a través de los años es creciente y en la actualidad es masivo. Estados Unidos es el país en el que se la utiliza con mayor frecuencia. En la Argentina, 420 publicaciones citan a esta base de datos, y los temas que se abordan con mayor frecuencia están vinculados a la ecología. En segundo lugar, se presenta el paso-a-paso para extraer datos de la base de datos de WorldClim en una rutina de R. De manera complementaria, se describen las razones para usar R como un SIG y se incluye una introducción al manejo de objetos espaciales en R. Al completar la rutina se obtienen los valores de 19 variables bioclimáticas relacionadas con la precipitación y temperatura para una serie de sitios de muestreo, y se muestran los resultados en mapas. El desarrollo de la rutina es una introducción al uso de R como SIG. La información presentada tiene aplicación directa para los profesionales relacionados con la ecología. Las explicaciones se complementan con gráficos para que se pueda incluir el material en guías de trabajo práctico en las materias de grado o en cursos de posgrado.

[Palabras clave: ayuda didáctica, R project, datos georreferenciados, análisis espacial, diversidad, macroecología, cambio climático]

ABSTRACT. R as a GIS: WorldClim climate data extraction. Characterizing sampling sites based on climatic data is a key need for a significant number of professionals working in 'ecology. In general terms, so is the use of GIS (Geographic Information Systems) tools. However, accessing the use of available online climatic databases and GIS tools is a challenge. This paper shows how to use R as a GIS to obtain values of climatic variables for geo-referenced sampling sites, using the WorldClim database, which contains climatic information for the entire globe. Firstly, there is a description of the WorldClim database. A bibliographic search shows the frequency of use of WorldClim and the topics of study. It is concluded that the use of this database has been increasing over the years, and currently it is massively used. The United States is the country that uses it most frequently. In Argentina, 420 publications cite this database, and the topics that are most frequently addressed are linked to ecology. Secondly, a step-by-step routine for extracting data from WorldClim data base is presented. In addition, the reasons for using R as a GIS are described and an introduction to the management of spatial objects in R is included. Upon completion of the routine, the values of 19 bioclimatic variables related to precipitation and temperature are obtained for a series of sampling sites and the results are displayed on maps. The routine is also an introduction to the use of R as a GIS. The information presented has direct application for professionals in the field of ecology. The explanations are complemented with graphics so that the material can be included in exercise guides in undergraduate or graduate courses.

[Keywords: teaching aid, R project, georeferenced data, spatial analysis, diversity, macroecology, climatic change]

INTRODUCCIÓN

La mayoría de los fenómenos ecológicos presentan estructura espacial en todas las escalas de estudio. Entonces, no sorprende que los datos con los que trabajamos los ecólogos provengan en su mayoría de sitios de muestreo georreferenciados. En los últimos tiempos, la publicación de grandes bases de datos globales libres que contienen información climática y ambiental ha permitido caracterizar los sitios de muestreo, aun cuando no se cuenten con datos climático-ambientales medidos a campo. La publicación de estas bases ha permitido avanzar en respuestas a preguntas tan generales como ¿Qué determina la biodiversidad y la estructura de los ensambles? hasta otras más aplicadas como ¿Qué tierras seleccionar para las conservación de la biodiversidad? o ¿En qué forma ocurre la dispersión de vectores? Para un número importante de ecólogos es necesario acceder a describir los sitios de muestreo según características climáticas y ambientales disponibles en las bases de datos.

En paralelo a la disponibilidad creciente de bases de datos tuvo lugar un avance en las características técnicas y en la disponibilidad de programas de Sistemas de Información Geográfica (SIG). Inicialmente, las bases de datos se descargaban exclusivamente de los respectivos sitios *web* y se trabajaban en plataformas de SIG como ArcGis (*software* comercial) o QGis (*software* libre). La principal ventaja de estos programas es la forma dinámica en que permiten visualizar datos espaciales y espacio-temporales. Sin embargo, las plataformas SIG se han utilizado, además de para visualizar datos, para gestionar datos climáticos. Esto incluye desde realizar operaciones sobre los datos (e.g., extraer datos climáticos de las grandes bases de datos para los sitios de muestreo) hasta realizar distintos análisis como el cómputo de la riqueza de especies. En la última década, la disponibilidad creciente de paquetes para trabajar con datos georreferenciados en R Project, es decir, para usar R como un SIG, impactó en la disciplina, lo que produjo un salto técnico que optimizó la forma de trabajar con datos espaciales. El proceso se volvió más eficiente y se acortaron enormemente los tiempos de trabajo; la funcionalidad se amplió. No obstante, la curva de aprendizaje de R y la escasa disponibilidad de cursos y de materias para usar R como un SIG ha dificultado su implementación en algunos casos. En este trabajo, presento una rutina *script* para extraer datos climáticos de

la base de datos WorldClim para sitios de muestreo georreferenciados. También presento información complementaria sobre la base de datos WorldClim y el uso de R como un SIG. El objetivo de este trabajo es que la rutina de R se pueda utilizar para extraer datos climáticos, pero también como forma de inicio en el uso de R como un SIG. Su implementación puede ser útil en investigaciones en ecología, para los técnicos de SIG y como material para guías de trabajos prácticos en el marco de materias de grado o de posgrado, para que su uso se pueda multiplicar.

¿POR QUÉ USAR R COMO UN SIG?

R es principalmente un lenguaje de programación para el análisis de datos, por lo que todas sus ventajas aplican directamente para analizar datos espaciales y espacio-temporales. En particular, se destaca que 1) R es de código abierto y funciona en cualquier sistema operativo, 2) permite trabajar con grandes bases de datos sin producir fallos en el sistema, 3) los comandos se guardan en una rutina y, por lo tanto, todo lo que se trabaja es reproducible (i.e., se puede rehacer el análisis rápidamente si se necesita, por ejemplo, agregar un sitio de muestreo, incorporar una nueva variable, hacer un cambio que sugiere un revisor, o aplicar el análisis espacial a otro conjunto de datos, etc.), 4) los comandos se pueden automatizar (lo que permite realizar acciones repetidas sobre un número grande de especies o sitios), 5) ¡una multitud de técnicas estadísticas que no se encuentran en plataformas SIG están presentes en paquetes de R! y 6) ¡con R se puede utilizar un único programa para realizar todo el proceso vinculado al análisis de datos espaciales! Por ejemplo, en este artículo se verá cómo desde R se pueden descargar variables de una base de datos climática disponible en internet, visualizar los datos elaborando mapas, extraer las variables climáticas para una serie de coordenadas de muestreo de interés y realizar una estadística descriptiva.

¿QUÉ ES WORLDCLIM Y POR QUÉ UTILIZARLA?

WorldClim es una base de datos global y libre que contiene un conjunto importante de variables climáticas. La primera versión fue publicada en 2005 (Hijmans et al. 2005) y la segunda versión en 2017 (Fick and Hijmans 2017). En enero de 2020 se actualizó la segunda versión a la versión 2.1. Las versiones 1.4 y

2.1 se encuentran disponibles *online* en worldclim.org. Por sus características, WorldClim es una de las bases de datos más utilizadas en ecología.

Información climática

Contiene información promedio para la temperatura mensual mínima, media y máxima, la precipitación, la radiación solar, la velocidad del viento y la presión de vapor de agua (integradas desde el año 1970 hasta el 2000). Además, contiene información de 19 variables bioclimáticas ('bio'), derivadas de la temperatura mensual y la precipitación, representando variables con sentido biológico de tendencia anual (temperatura y precipitación anuales promedio), la estacionalidad en el clima (e.g., el rango anual en la temperatura y la precipitación) y factores ambientales limitantes o extremos (e.g., la temperatura del mes más frío o la precipitación del cuarto del año más seco). Las variables 'bio' se suelen utilizar en estudios de modelado de nicho y de análisis de patrones de diversidad.

Resolución

Las variables climáticas que reúnen información desde el año 1970 hasta el 2000. Se encuentran disponibles en resoluciones desde casi tan solo 1 km² (0.86 km² en el Ecuador, 30 arco-segundos), ¡muy útil para estudios a escala regional/local!, hasta aproximadamente 340 km² (10 arco-minutos).

Información climática del pasado y proyección a futuro

La versión 1 de la base de datos WorldClim (Hijmans et al. 2005) también cuenta con variables climáticas proyectadas para el futuro y variables estimadas para el pasado (paleoclima del holoceno medio y del último máximo glacial). La versión 2.1 (Fick and Hijmans 2017) actualiza las variables climáticas proyectadas para el futuro. Se encuentran disponibles en resoluciones de 10, 5 y 2.5 arco-minutos. Se proyecta que en breve estarán disponibles también a una resolución de 30 arco-segundos.

Gran cantidad de estaciones meteorológicas

Para armar la base WorldClim versión 2 se utilizaron 34542 estaciones meteorológicas para modelar la precipitación, 20268 para la temperatura media, 16883 para la temperatura

mínima, 16988 para la temperatura máxima, 9541 para la presión de vapor, 10149 para la velocidad del viento y 5489 para la radiación solar (Fick and Hijmans 2017). El número de estaciones hace que los datos de partida sean sólidos. Se aplicó un modelo de interpolación para cubrir la superficie completa del globo, incluyendo como covariables la altitud, la distancia a la costa y variables derivadas de información satelital. Recientemente se ha publicado una evaluación del uso de la versión 2 de la base en la Argentina (Bustos et al. 2017) y existen numerosos trabajos sobre la evaluación de la base de datos WorldClim versión 1 en distintas regiones (e.g., Europa) (Marchi et al. 2019). Sin embargo, es importante considerar que las estaciones meteorológicas no se distribuyen de manera homogénea y que la precisión de la base disminuye con la altitud y con la distancia a las estaciones meteorológicas (Fick and Hijmans 2017). Es importante consultar el mapa de las estaciones para conocer el grado de interpolación en la región de estudio (los mapas pueden consultarse en Fick and Hijmans [2017]) y para evaluar la factibilidad de utilizar la base.

Descarga libre y gratuita

Los archivos correspondientes a las variables de la base de datos se encuentran disponibles en forma gratuita (worldclim.org). Descargar las variables 'bio' a una resolución de 10 arco-minutos requiere descargar alrededor de 47 MB; en cambio, a 30 arco-segundos requiere descargar alrededor de 10 GB. La descarga se puede realizar de manera sencilla desde R project (esto se encuentra en la rutina presentada en este trabajo).

Formato

Cada variable de la base se encuentra disponible en formato raster, representado por celdas cuadradas que cubren la superficie continua del globo. Este formato de objeto espacial puede leerse en R Project o en plataformas SIG como QGIS.

Sistema de coordenadas

Cada celda de cada variable de WorldClim se encuentra asociada a una latitud y una longitud. El sistema es el WGS84, que es el sistema de coordenadas más usado en la actualidad y es el más utilizado por el GPS (Olaya 2014). Gracias a esto, la superposición de los puntos de muestreo con la base climática se puede realizar de forma directa (i.e., sin

pasos intermedios de cambio de sistemas de coordenadas o proyecciones).

Algunas consideraciones sobre el uso de WorldClim

Además de que se debe tener en consideración el número (y la distribución espacial) de las estaciones meteorológicas que se encuentran en el área de estudio, hay un aspecto más a considerar. Se ha identificado que las variables que combinan información de precipitación y temperatura en la base de datos de WorldClim muestran anomalías en forma de discontinuidades entre celdas vecinas (Escobar et al. 2014), por lo que algunos estudios excluyen su uso (e.g., Escobar et al. 2014; Núñez-Penichet et al. 2021).

¿CUÁNTO SE UTILIZA LA BASE DE DATOS WORLDCLIM EN EL MUNDO Y EN LA ARGENTINA?

Para responder esta pregunta se realizó una búsqueda bibliográfica de los trabajos que citan la base de datos WorldClim con

fecha 08/10/2019, en el conjunto de todas las disciplinas. Se utilizó Scopus como base de datos bibliográfica y analicé los resultados en R 3.6.1 (R Core Team 2019) con el paquete bibliometrix (Aria and Cuccurullo 2017).

El uso de la base de datos de WorldClim en el conjunto de las distintas disciplinas es masivo. Las versiones 1 y 2 fueron citadas 11694 veces, de las cuales 420 citas corresponden a trabajos con al menos un autor argentino (Figura 1a,b y Figura 2). La tendencia de citas en el tiempo es creciente para la versión 1 de la base (Figura 1a,b). Se observa una tendencia similar en artículos extranjeros, como en artículos con al menos un autor argentino (Figura 1b). Como es de esperar, en los trabajos internacionales de los últimos 2 años, las citas a la versión 1 de la base se comenzaron a reemplazar por las citas a la versión 2. Sin embargo, hasta el momento no hay en Scopus citas de trabajos con autores argentinos de la versión 2 de la base. Esto puede evidenciar una dificultad para acceder a la misma, ya que los autores argentinos continúan citando únicamente la versión 1. Aunque no es posible determinar

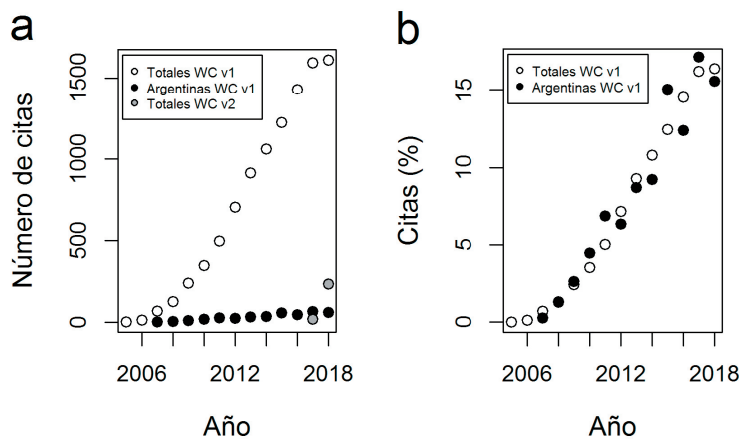


Figura 1. Cantidad de citas a la base de datos WorldClim en función del año, estimadas como (a) número de citas en valores absolutos a las bases de WorldClim versión 1 (WC v1) y versión 2 (WC v2), y (b) porcentaje de citas a la base WorldClim versión 1 (WC v1). Totales=citas en Scopus. Argentinas=citas realizadas por trabajos que incluyen al menos un autor argentino. Se observa una tendencia creciente en el número de citas a la base de WorldClim versión 1 en el conjunto de los trabajos y en los trabajos con al menos un autor argentino (Figura 1a). Al expresar la cantidad de citas a WC v1 en porcentaje se observa que la tendencia tiene una forma similar en ambos conjuntos de datos (Figura b). En 2017, los trabajos externos a la Argentina comienzan a citar a WC v2 (Figura 1a). En 2018 se frena el crecimiento en el número de citas a la v1 (Figura 1a). Sin embargo, los trabajos con al menos un autor argentino no incluyen citas a WC v2 (Figura 1a).

Figure 1. Number of citations to the WorldClim database according to the year, estimated as (a) number of citations in absolute values 'to the databases of WorldClim version 1 (WC v1) and version 2 (WC v2), and (b) percentage of citations to the WorldClim version 1 database (WC v1). Totales=cites in Scopus. Argentinas=citations made in articles that include at least one Argentine author. The number of citations to the WorldClim version 1 database increases with years according to both the whole set of articles and according to articles with at least one Argentine author (Figure 1a). When expressing the number of citations to WC v1 in percentage, it is observed that the trend has similar shape in both data sets (Figure b). In 2017, articles from outside Argentina begin to cite WC v2 (Figure 1a). In 2018, the growth in the number of cites to v1 is slowed (Figure 1a). However, works with at least one Argentine author do not include citations to WC v2 (Figure 1a).

a qué se debe esto, algunas razones podrían ser el desconocimiento de la actualización, las limitaciones de la red para realizar la descarga, el uso de funciones desactualizadas en R Project que trabajan con la primera versión de la base (e.g., la función `getData()` del paquete *raster* [Hijmans 2020]).

Estados Unidos es el país que más cita la base de datos WorldClim en sus versiones 1 y 2 (Figura 2), algo esperable por ser el país con mayor producción científica. La contribución de citas de la Argentina es relevante, ubicándose en el catorceavo lugar en el ranking de uso por país (Figura 2). La base de datos WorldClim versión 1 es el segundo artículo más citado en los artículos argentinos que usan el enfoque de la macroecología (Fergnani and Filloy 2021).

¿PARA QUÉ UTILIZAN LA BASE DE DATOS WORLDCLIM LOS AUTORES ARGENTINOS Y EN QUÉ ZONAS SE REALIZAN LOS ESTUDIOS?

En base a la búsqueda realizada en Scopus, se realizó un análisis de las palabras claves de los

manuscritos que utilizan WorldClim y están publicados por al menos un autor argentino. Los resultados sugieren que la base de datos se usa principalmente en el área de ecología. Los temas más frecuentes son el ‘cambio climático’, el estudio de la ‘biodiversidad’ y los ‘ecosistemas’ (Figura 3).

Las palabras claves sugieren que una gran parte de los estudios se realizaron en la Argentina (i.e., primera palabra clave) (Figura 3). Dentro de la Argentina, la región más mencionada es la Patagonia (Figura 3). Otras zonas de estudio frecuentemente mencionadas son Chile, Brasil y América del Sur (Figura 3).

OBJETOS ESPACIALES EN R

Para facilitar la comprensión del código de la rutina se hará una síntesis sobre algunos aspectos de los dos grandes modelos de datos espaciales y su implementación en R. El modelo vectorial y el modelo raster.

El modelo vectorial refiere a estructuras de puntos, líneas o polígonos (Figura 4). En R, este formato de objeto espacial se trabaja

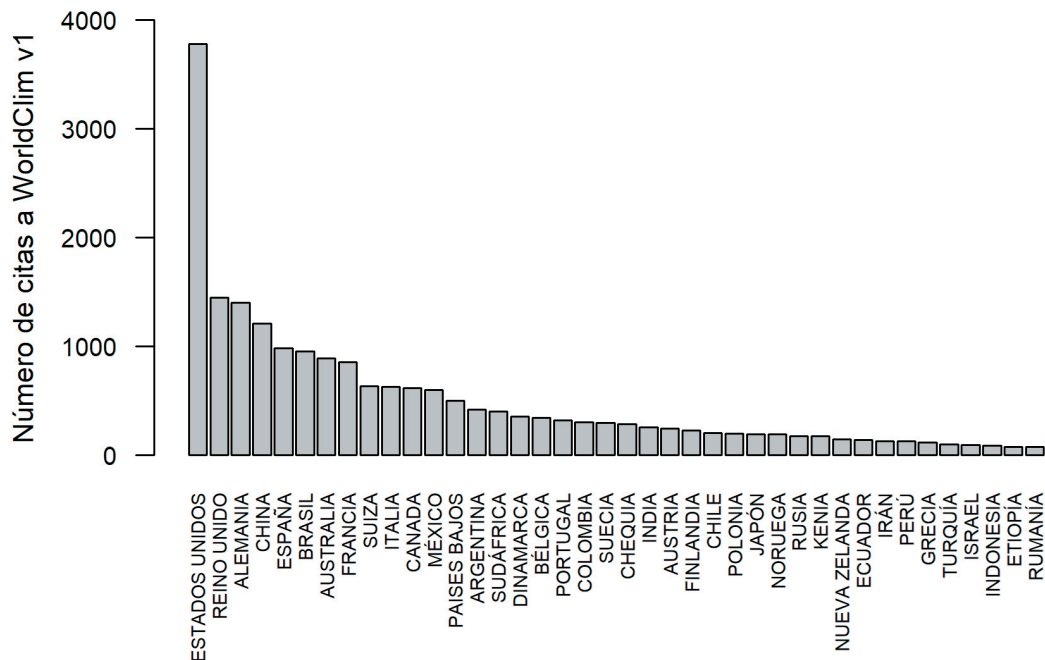


Figura 2. Número de citas a WorldClim versión 1 (v1) por país. Se muestran los valores para los 40 países que más citan a la base de datos. Estados Unidos es el país que produce más citas, Argentina se encuentra en el lugar catorceavo.

Figure 2. Number of citations to WorldClim version 1 (v1) per country. Values are shown for the 40 countries that most cite the database. The United States is the country that produces the most citations, Argentina is in the fourteenth place.

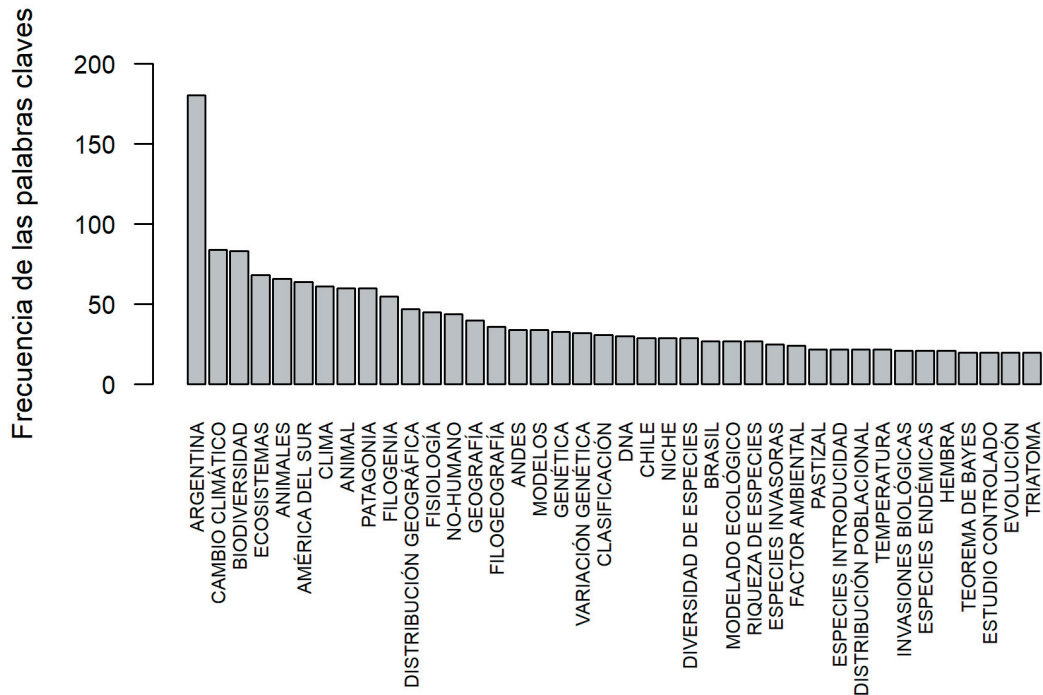


Figura 3. Frecuencia de las 40 palabras claves más utilizadas en los 420 artículos con al menos un autor argentino que citan a WorldClim versión 1. Las palabras están traducidas al español. Se deduce que la mayoría de los trabajos se realizaron en América del Sur, a distintas escalas, siendo la Argentina el lugar más frecuente. Los temas de estudio más frecuentes son el cambio climático y la biodiversidad.

Figure 3. Frequency of the 40 most used keywords in the 420 articles with at least one Argentine author who cite WorldClim version 1. The words are translated into Spanish. Most of the work was carried out in South America, at different scales, with Argentina being the most frequent place. The most frequent topics of study are climate change and biodiversity.

generalmente en dos librerías alternativas, `sp` o `sf`. La librería `sp` se encuentra cubierta en el libro de Bivand et al. (2008) y la librería `sf` en el libro en castellano de Mas (2018). La presente rutina se desarrolla con el paquete `sp`. El formato vectorial contiene como mínimo la información de dónde se encuentran las estructuras (i.e., puntos, líneas o polígonos), y también puede tener asociado un marco de datos (*data frame*). Por ejemplo, para un objeto espacial de puntos que representan sitios de muestreo se puede tener un marco de datos asociado con valores de una variable respuesta o una variable predictora que se midió en el sitio. En la rutina se muestra cómo acceder a la información relevante de un objeto espacial de puntos. Escribiendo el nombre del objeto y '@data' se muestran los valores del marco de datos. El nombre del objeto y '@coords' se muestra la lista de coordenadas. El nombre del objeto y '@proj4string' se muestra el sistema de coordenadas o proyección geográfica del objeto. Aplicando la función `summary()` al nombre del objeto se muestra la información sumaria del mismo. Esta información es un

breve resumen de algunas de las características más importantes del objeto.

El modelo raster representa una variable en una superficie continua de celdas cuadradas (Figura 4). Cada raster se caracteriza por la resolución de la celda (i.e., todas las celdas tienen la misma resolución) y la extensión que cubre todo el conjunto de las mismas (i.e., los límites del área geográfica que abarca). En R, los rasters se crean y trabajan utilizando el paquete `raster`. Escribiendo el nombre del objeto raster, se obtiene una información sumaria del mismo. El modelo vectorial y el modelo raster se grafican en mapas usando el comando `plot()`.

RUTINA DE R PARA EXTRAER DATOS CLIMÁTICOS DE WORLDCLIM

La rutina es autoexplicativa y se encuentra adjunta como Material Suplementario. En síntesis, a partir de un archivo que contiene coordenadas de muestreo, se obtiene la información de WorldClim para las variables

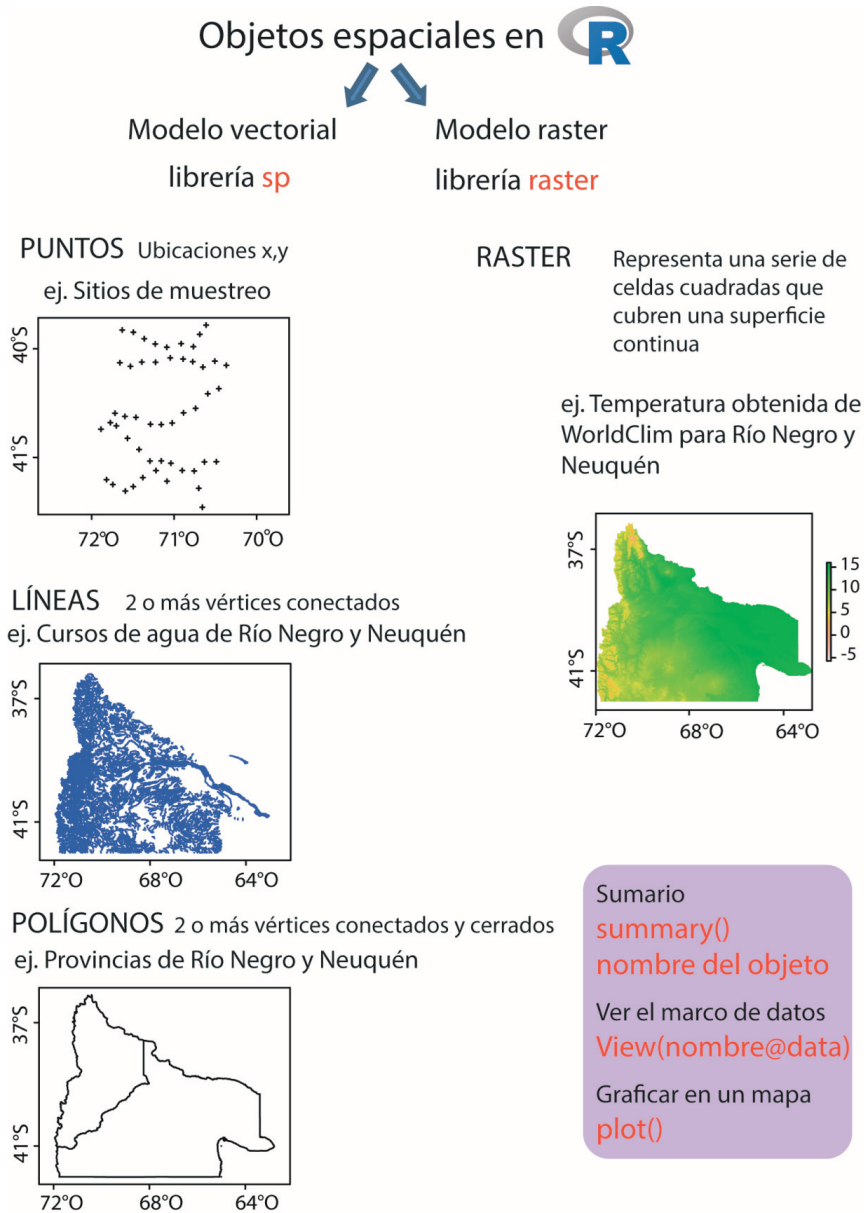


Figura 4. Objetos espaciales en R project. Modelos vectorial y raster según las dos librerías utilizadas en la rutina presente (sp y raster). Los ejemplos provienen de la rutina, con la excepción del mapa de cursos de agua.

Figure 4. Spatial objects in R project. Vector and raster models according to the two libraries used in the present routine (sp and raster). The examples come from the routine, with the exception of the watercourse map.

'bio' y un mapa de las coordenadas de muestreo (Figura 5). La rutina requiere una conexión a internet.

Realizar la rutina involucra adquirir las siguientes capacidades: 1) elaborar objetos espaciales en R a partir de listas de coordenadas, 2) graficar mapas en R, 3) realizar operaciones básicas con objetos espaciales, 4) crear y exportar objetos espaciales (mapas) para ser leídos en otras plataformas SIG (e.g.,

QGIS) y 5) extraer datos de bases de datos (WorldClim).

Instalación de R

Para correr la rutina es necesario instalar R Project (r-project.org). Aunque no es imprescindible, se recomienda instalar en segundo lugar la interfaz RStudio (rstudio.com), que permite acceder de forma simple a toda la potencia de R. Entre otras

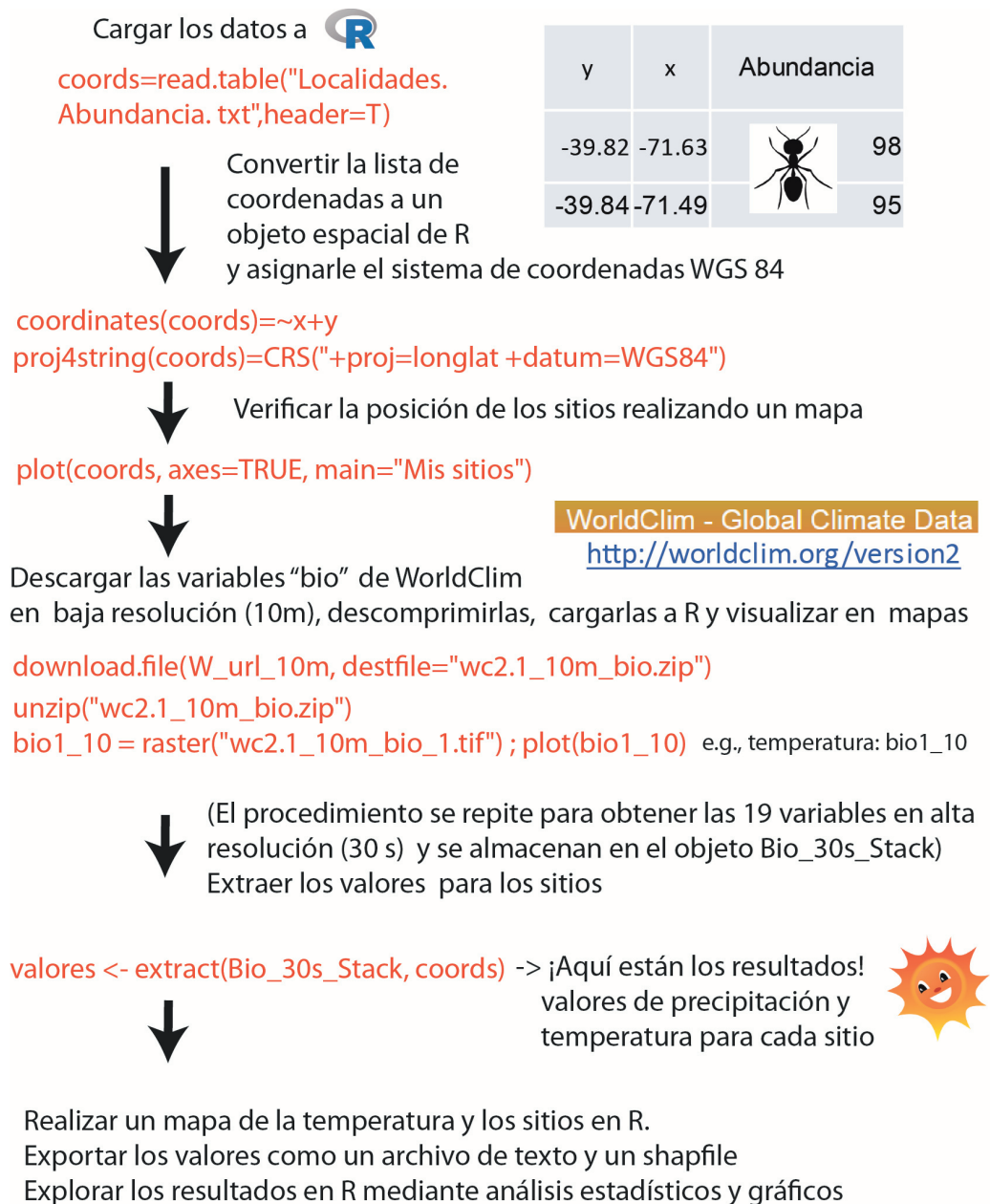


Figura 5. Resumen de los principales pasos de la rutina para obtener datos de WorldClim para una lista de coordenadas de muestreo en R project. El primer paso consiste en instalar y cargar los paquetes de R necesarios para la rutina y establecer el directorio de trabajo. Este paso no se incluye en la figura.

Figure 5. Summary of the main steps of the routine to obtain WorldClim data for a list of sampling coordinates in R project. The first step (not included in the figure) is to install and load the R packages needed for the routine and establish the working directory.

cosas, RStudio asiste con la sintaxis y organiza las ventanas de ayuda y de gráficos; además, gestiona muy bien las descargas de internet. Se recomienda abrir la rutina directamente desde RStudio. Cada línea de la rutina se envía a la consola de R clickeando en el botón 'Run'.

Partes principales de la rutina

La rutina se encuentra dividida en siete partes breves. Cada una se encuentra explicada en detalle en la rutina. Aquí se menciona cada sección y se recapitulan solo algunos aspectos principales.

Parte 1. Instalar y cargar los paquetes de R necesarios para la rutina. Establecer el directorio de trabajo. El directorio de trabajo debe establecerse en la carpeta donde se encuentra el archivo de coordenadas; en el ejemplo, 'Localidades.Abundancia.txt'. Desde este directorio se cargará el archivo, y es donde quedarán guardados los archivos que se exporten y los que se descarguen de internet desde R. Es la única parte de la rutina en la que se debe modificar una línea de código. El resto de la rutina no requiere modificaciones y está escrita de manera general para que se la pueda utilizar con otros datos de partida.

Parte 2. Cargar las coordenadas de los sitios de muestreo y crear un objeto espacial de puntos. En la rutina se utiliza como ejemplo una lista de coordenadas de 50 sitios de muestreo estudiados en Fernani et al. (2010). Estos sitios, ubicados en el noroeste de la Patagonia argentina, se georreferenciaron en el campo por medio de un GPS en el sistema de coordenadas WGS84. Como ejemplo de variable respuesta se consigna a la abundancia de hormigas de cada sitio. El archivo 'Localidades.Abundancia.txt' contiene para cada sitio las variables 'y' (latitud), 'x' (longitud) y 'Abundancia' (de hormigas). Es un archivo de texto '.txt' separado por tabulaciones. Contiene puntos en lugar de comas porque R utiliza puntos para los decimales. En la rutina, la función para cargar estos datos es 'read.table()'. Luego, con la función 'coordinates()' se crea un objeto espacial de puntos en base a las coordenadas y se le asigna el sistema de coordenadas WGS84 con la función 'proj4string()' (ambas funciones del paquete 'sp'). Es importante notar que la georreferenciación se realizó en el campo con un GPS utilizando el mismo sistema de coordenadas que la base de WorldClim, por lo que ambos mapas se podrán superponer en forma directa.

Parte 3. Graficar las coordenadas en un mapa para verificar que están situadas correctamente. Antes de realizar cualquier operación con objetos espaciales es importante verificar que los sitios de muestreo se encuentren ubicados de manera correcta en el espacio. Para ello, en la rutina se superpone un mapa de la Argentina con el objeto espacial de puntos creado a partir de las coordenadas. La función que se utiliza para graficar cada mapa es 'plot()'. Incorporando el argumento 'add=TRUE', un mapa se superpone sobre el otro.

Parte 4. Explorar las variables 'bio' de la base de datos climática. Obtener los datos de WorldClim mundiales y visualizarlos usando una escala de celda grande (10 minutos, ~340 km²). Antes de extraer los valores puntuales para los sitios a partir de cualquier base de datos climática es recomendable conocer la base de datos. Para ello, en la rutina se descargan las variables de interés 'bio' de la base climática de WorldClim versión 2.1 para el globo a una carpeta física en un archivo zip, se descomprimen, se cargan las variables a R y se visualiza la variación espacial en las variables. La resolución seleccionada fue de 10 minutos de grados (~340 km²) — es una primera aproximación — para que la descarga se pueda realizar de forma rápida. En la rutina se indica cómo modificar la resolución. La función 'download.file()' se utiliza para descargar la base de datos. Se indica en los argumentos de la función la dirección web del archivo (URL) y el nombre con el que se descargará el archivo. En la página web de WorldClim (worldclim.org) se puede acceder a los links de descarga de todas las variables disponibles en la base de datos de WorldClim. La función 'download.file()' se puede utilizar para descargar cualquier base de datos o archivo de internet cambiando el URL.

La información de las variables climáticas cargada en R como objetos espaciales en modelo raster se grafican en mapas usando la función 'plot()'. Los nombres de las 19 variables obtenidas se encuentran en la rutina. Cabe mencionar que existe una función implementada en el paquete raster 'getData()' que descarga la base de datos de WorldClim, pero descarga versión 1.4 obsoleta, por lo que no está incorporada en la rutina.

Parte 5. Obtener y extraer los datos de la base de datos de WorldClim para las coordenadas de muestreo. En primer lugar se obtienen los datos de WorldClim en alta resolución (30 segundos de grado, ~1 km²). Para ello se descargan las variables utilizando la misma función que en la parte 4 del *script*. Sin embargo, es importante mencionar que esta descarga puede demorar horas cuando la conexión a internet es muy lenta. La recomendación para lugares con conexión a internet deficiente es dejar la descarga activa en la noche y utilizar para la descarga la interfaz de RStudio. En segundo lugar, se extraen los valores puntuales de las variables 'bio' para las coordenadas de

muestreo utilizando la función `extract()` del paquete `raster`. Los resultados se exportan de dos maneras: como un archivo de texto `'txt'` con la función `write.table()` y como un archivo vectorial en formato `shapefile`, que puede ser leído en R o en una plataforma SIG (funciones `SpatialPointsDataFrame()` del paquete `sp` y `writeOGR()` del paquete `rgdal`).

Parte 6. Graficar un mapa de los sitios de muestreo utilizando como fondo los valores obtenidos de la base de datos de WorldClim.

Parte 7. Estadística elemental. Se muestra el uso de algunos comandos para explorar los datos de variables climáticas en los sitios de muestreo y la relación entre variables. Éstos incluyen análisis de correlación `cor()`,

regresión lineal simple `lm()`, histogramas `hist()`, gráfico de cajas `boxplot()` y gráfico de dispersión `plot()`.

AGRADECIMIENTOS. Esta ayuda didáctica es en parte el resultado de un taller de R como un SIG que dicté en R-Ladies Bariloche. Quiero agradecer también a Adriana Ruggiero quién me motivó a avanzar en el uso de plataformas SIG y R. A Damián Rajoy por su ayuda con la conexión para el acceso a las bases de datos. A Gabriela Liz Fioriti por la revisión de los textos en inglés. A Darío Fernani por la revisión de la rutina. Los comentarios de dos revisores/as anónimos/as contribuyeron a mejorar el manuscrito. Financiamiento: ANPCyT-FONCYT, PICT 2016 3489 y UNCOMA 04/B228.

REFERENCIAS

- Aria, M., and C. Cuccurullo. 2017. Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* 11:959-975. <https://doi.org/10.1016/j.joi.2017.08.007>.
- Bivand, R. S., E. J. Pebesma, V. Gómez-Rubio, and E. J. Pebesma. 2008. *Applied spatial data analysis with R*. Springer.
- Bustos, M. L., F. Ferrelli, and M. C. Piccolo. 2017. Estudio comparativo de tres modelos climáticos en Argentina. *Universidade Federal do Rio de Janeiro. Anuário do Instituto de Geociências* 40:34-43. https://doi.org/10.11137/2017_1_34_43.
- Escobar, L. E., A. Lira-Noriega, G. Medina-Vogel, and A. T. Peterson. 2014. Potential for spread of the white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: use of Maxent and NicheA to assure strict model transference. *Geospatial Health* 9(1):221-229. <https://doi.org/10.4081/gh.2014.19>.
- Fernani, P. N., and J. Filloy. 2021. Desde la areografía de Rapoport a los 30 años de la macroecología: Síntesis de la contribución argentina. *Ecología Austral* 31:101-113. <https://doi.org/10.25260/EA.21.31.1.0.1134>.
- Fernani, P. N., P. Sackmann, and A. Ruggiero. 2010. Richness environment relationships in epigaeic ants across the Subantarctic Patagonian transition zone. *Insect Conservation and Diversity* 3:278-290. <https://doi.org/10.1111/j.1752-4598.2010.00105.x>.
- Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1'km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37:4302-4315. <https://doi.org/10.1002/joc.5086>.
- Hijmans, R. J. 2020. Raster: Geographic Data Analysis and Modeling. R package version 3.4-5. <https://CRAN.R-project.org/package=raster>.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965-1978. <https://doi.org/10.1002/joc.1276>.
- Marchi, M., I. Sinjur, M. Bozzano, and M. Westergren. 2019. Evaluating WorldClim Version 1 (1961-1990) as the Baseline for Sustainable Use of Forest and Environmental Resources in a Changing Climate. *Sustainability* 11:3043. <https://doi.org/10.3390/su11113043>.
- Mas, F. J. 2018. *Análisis espacial con R: Usa R como un Sistema de Información Geográfica*. European Scientific Institute, Macedonia.
- Núñez-Penichet, C., L. Osorio-Olvera, V. H. González, M. E. Cobos, L. Jiménez, D. A. DeRaad, A. Alkische, R. G. Contreras-Díaz, A. Nava-Bolaños, and K. Utsumi. 2021. Geographic potential of the world's largest hornet, *Vespa mandarinia* Smith (Hymenoptera: Vespidae), worldwide and particularly in North America. *PeerJ* 9:e10690. <https://doi.org/10.7717/peerj.10690>.
- Olaya, V. 2014. *Sistemas de Información Geográfica*. Un libro libre de Victor Olaya. URL: volaya.github.io/libro-sig.
- R Core Team. 2019. *R: A language and environment for statistical computing* R Foundation for Statistical Computing, Vienna, Austria. URL: R-project.org.