

{agromet}: un paquete en R para el análisis de datos meteorológicos

Natalia Gattinoni¹, Paola Corrales², Elio Campitelli², Yanina Bellini³ y Gabriel Rodríguez¹

¹ INTA, Instituto de Clima y Agua, gattinoni.natalia@inta.gob.ar, rodriguez.gabriel@inta.gob.ar, ² CIMA – FCEN-UBA, paola.corrales@cima.fcen.uba.ar, elio.campitelli@cima.fcen.uba.ar, ³ INTA- EEA Anguil yabellini@gmail.com

Keywords: índices agrometeorológicos, representación geoespacial, seguimiento agrometeorológico

Las condiciones climáticas ejercen un efecto favorable o perjudicial sobre la producción agropecuaria. Es así que es primordial hacer un seguimiento y análisis de las distintas variables agrometeorológicas que permitan describir dichas condiciones y su efecto a largo del periodo de la producción agrícola y pecuaria.

Con este enfoque se ha desarrollado un paquete en R llamado {agromet} diseñado para ser utilizado por usuarios pertenecientes al Instituto Nacional de Tecnología Agropecuaria (INTA) pero con suficiente generalidad como para ser útil a una amplia comunidad. El presente trabajo tiene como objetivo mostrar las principales características del paquete y posibles implementaciones.

El paquete {agromet} incluye una serie de funciones que pueden ser utilizadas de manera habitual para el cálculo de índices y estadísticos agrometeorológicos. Los datos meteorológicos de entrada se trabajan bajo la filosofía de datos tidy o datos ordenados, donde cada columna será una variable meteorológica y cada fila será una observación (el dato de una determinada fecha) [cita 1]. De esta forma, las funciones del paquete son genéricas, pudiendo ser aplicadas a cualquier set de datos tabulares sin importar su origen, orden o nombre de las columnas. De todas formas, atento a los requerimientos internos del INTA, el paquete también incorpora herramientas para leer datos en un formato ascii utilizado por la institución. Dentro del paquete se considera como ejemplo de datos de entrada un archivo ascii perteneciente al Observatorio Agrometeorológico de INTA Castelar-Bs.As. Argentina.

A continuación se detallan algunas de las funciones destacadas del paquete:

umbrales() permite contar la cantidad de observaciones que cumplen una determinada condición. Por ejemplo, cantidad de días con temperaturas mínimas inferiores a 0°C durante el periodo crítico del cultivo de trigo o la cantidad de días con temperaturas máximas superiores a los 38°C para el cultivo de soja. *dias_promedios()* devuelve el primer y último día del año promedio de ocurrencia de un evento. Esta función puede utilizarse para el análisis de datos históricos de heladas, pudiendo determinar la fecha media de primera y última helada. *olas()* identifica periodos de persistencia de un evento definido a partir de una condición lógica, por ejemplo día consecutivos con temperaturas mínimas inferiores a 0°C para análisis de heladas o con temperatura máxima superior a un umbral para el análisis de olas de calor. *ith()* esta función permite calcular

el Índice de Temperatura y Humedad (ITH) ampliamente utilizado para el seguimiento del estrés calórico o confort del ganado lechero. El cálculo se realiza a nivel diario y considerando la ecuación Earl C. Thom (1959) [cita 2]. *decil()* y *anomalía_porcentual()* permiten el cálculo de estos estadísticos para distintos intervalos de tiempo y pueden resultar útiles para el seguimiento de las precipitaciones mensuales permitiendo identificar meses en los cuales las mismas fueron superiores o inferiores a los valores medios históricos. Otras funciones como *spi()* y *spei()* permiten calcular los índices estandarizados de precipitación (SPI) mundialmente utilizados para el seguimiento de los periodos con excesos y déficit hídricos, en este caso se adaptaron las funciones de otros paquetes para ser compatibles con el manejo de datos tidy.

Además de las funciones para calcular índices, el paquete cuenta con funciones para trabajar con series de datos como por ejemplo *completar_serie()* esta función permite completar la serie de datos temporales definiendo alguna resolución disponible. Esta función es útil, por ejemplo, para completar los días ausentes o el registro ausente en una serie de datos diarios, completando dicho registro con NA (Not Available) en todas las variables. Así también el paquete provee de funciones relacionadas con el mapeo geoespacial de datos meteorológicos a nivel nacional. Para la representación geoespacial se encuentran disponibles distintas cartografías tanto a nivel nacional, provincial, como de países limítrofes (<https://www.naturalearthdata.com/>) y departamental (<https://www.ign.gob.ar>), cuya función es *mapa_argentina*. Para el gráfico de contornos se disponen de distintas escalas ya organizadas según intervalos de las variables y escala de colores (por ejemplo, *escala_temp_min()*). En tal caso que el usuario trabaje desde SURFER® (Golden Software, Inc.) puede importar su escala definida previamente a partir de la función *leer_surfer()* [cita 3]. El paquete incluye una función de graficado de datos georeferenciados *mapear()* con el estilo y logo propios de INTA. Finalmente, el paquete provee de la posibilidad de disponer de un reporte final en RMarkdown para los profesionales de la institución, el cual incluye un espacio para mapa y tabla de valores destacados. Este tipo de informe estandariza las salidas y cálculos de información producida por INTA relacionadas con la agrometeorología.

Si bien algunas de las funcionalidades incluidas en el paquete ya existen en otros paquetes, su elaboración permitió unificarlas para que funcionen con los datos en una estructura genérica y generar funciones específicas como la creación de mapas y reportes con el estilo institucional del INTA. Así también es un ejemplo más de como utilizar sistemas de desarrollo colaborativo utilizando herramientas de control de versiones como GitHub. El repositorio se encuentra disponible en <https://github.com/AgRoMeteorologiaINTA/agromet>. {agromet} es de código abierto, está disponible gratuitamente, y en continuo desarrollo permitiendo mejorar las funciones que ya contiene como poder incluir nuevas según las necesidades de los usuarios. La licencia de uso es Licencia pública general de GNU y el proyecto cuenta con un código de conducta y una guía para quienes quieran contribuir.

Bibliografía:

1. Wickham, H. Tidy Data. Journal of Statistical Software, 59(10), 1 - 23. doi:<http://dx.doi.org/10.18637/jss.v059.i10> (2014).
2. SURFER, versión 9.0, Golden Software. Sitio web: <https://www.goldensoftware.com/>.
3. Earl C. Thom. The Discomfort Index. Weatherwise 12, 57-59, (1959).