title: 'Netcdfplot: Streamlined Analysis and Visualization of NetCDF Data in R for Earth System Sciences' tags: - R - Grads - meteorology - oceanography - hydrology authors: - given-names: Sandra Isay surname: Saad orcid: 0000-0002-4380-3175 affiliation: 1 affiliations: - name: Department of Atmospheric and Climate Sciences (DCAC), Federal University of Rio Grande do Norte (UFRN), Natal, Rio Grande do Norte, Brazil index: 1 date: 24 February 2026 bibliography: paper.bib

Summary:

The Network Common Data Form (NetCDF) is a widely used data format in atmospheric and climate sciences for storing multidimensional geospatial datasets such as model outputs and observational data (UCAR 2026). Although several R packages provide tools to read NetCDF files, working with these data often requires multiple processing steps to extract spatial and temporal information prior to visualization or analysis.

The `netcdfplot` package provides a set of functions to facilitate the loading, processing, visualization, and analysis of georeferenced NetCDF data in R. It aims to simplify common tasks such as extracting spatial dimensions, handling temporal variables, visualizing gridded fields, and analyzing time series at specific geographic locations.

Statement of need:

Many Earth system sciences, such as Meteorology and Oceanography, require the analysis of four-dimensional data defined by latitude, longitude, vertical level, and time. This type of analysis demands both technical expertise and familiarity with commonly used state-of-the-art tools. In operational weather forecasting, for example, atmospheric cross-sections at specific pressure levels are routinely analyzed to identify relevant systems such as jet streams in the upper troposphere or frontal systems near the surface.

A key development that enabled advances in Earth system sciences was the creation of a standardized data format suitable for storing multidimensional geophysical variables: the Network Common Data Form (NetCDF), developed and maintained by the University Corporation for Atmospheric Research (UCAR 2026). NetCDF is a binary and self-describing format that includes information about dimensions and variables and the data itself, allowing efficient storage and interoperability across multiple platforms (UCAR 2026) (Rew and Davis 1990).

Several software tools have been developed specifically for the analysis and visualization of NetCDF data. For instance, the Grid Analysis and Display System (GrADS) allows users to load files and generate spatial maps or time

series with relatively few commands (Silva, Silva, and Fiorino 2021). However, despite its ease of use for basic visualization tasks, it is limited when compared to modern programming environments that support more advanced data processing workflows, such as Java, Python, or R.

The R environment is a free and open-source programming language widely used for statistical computing and graphical analysis in scientific research (R Core Team 2026). It provides packages such as ncdf4 for reading NetCDF files; however, performing common operations on multidimensional geophysical data can still require relatively complex workflows, particularly for users with limited experience in programming or geospatial data analysis.

The `netcdfplot` package was developed to simplify the visualization and analysis of NetCDF data within the R environment by reducing the amount of code required to perform common tasks such as spatial plotting and time series extraction from multidimensional datasets.

State of the field:

Several tools are available for handling NetCDF-formatted geospatial data, including command-line utilities and specialized visualization software such as the Grid Analysis and Display System (`GrADS`) (Silva, Silva, and Fiorino 2021), as well as R packages such as `ncdf4` (Pierce 2019) and `metR` (Campitelli 2025). While these tools provide robust functionality for reading or visualizing gridded datasets, performing common analysis tasks in R often requires combining multiple packages and executing several preprocessing steps.

For instance, plotting a spatial field (e.g., surface temperature) at a given time step typically involves loading the variable along with the corresponding latitude and longitude vectors using the `ncdf4` package (Pierce 2019), transforming the data into a tabular format, and generating a plot using `ggplot2` functions such as `geom_raster` (Wickham 2016). Similarly, visualizing wind vector fields may require combining `ggplot2` with additional packages such as `metR`, which can be challenging for users with limited experience in geospatial data processing.

The `netcdfplot` package integrates these steps into a unified workflow by providing high-level functions that directly handle spatial dimensions, temporal indexing, and visualization of gridded variables. It also includes utilities to convert geographic coordinates (latitude/longitude) into spatial indices, facilitating the extraction of time series at user-defined locations, as well as functions to compute spatial statistics (e.g., mean, minimum, and maximum values) within areas defined by shapefiles.

By reducing the number of preprocessing steps and minimizing the need to combine multiple packages, `netcdfplot` simplifies the analysis of NetCDF datasets within the R environment, promoting reproducible and efficient workflows in atmospheric and geoscientific applications.

Software design:

The package is implemented in R and relies on the `ncdf4` package for reading NetCDF files (Pierce 2019), while graphical outputs are generated through integrations with `ggplot2` (Wickham 2016) and `metR` (Campitelli 2025). Package documentation was generated using the `roxygen2` framework (Wickham et al. 2025). The software design prioritizes usability by providing high-level functions that abstract common preprocessing and visualization workflows required for multidimensional geophysical data analysis.

References

Campitelli, Elio. 2025. *MetR: Tools for Easier Analysis of Meteorological Fields.* https://CRAN.R-project.org/package=metR.

Pierce, David W. 2019. *Ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files.* https://CRAN.R-project.org/package=ncdf4.

R Core Team. 2026. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rew, R., and G. Davis. 1990. "NetCDF: An interface for scientific data access." *IEEE Computer Graphics and Applications* 10. https://doi.org/10.1109/38.56302.

Silva, A. da, D. Silva, and M. Fiorino. 2021. "Open Grid Analysis and Display System." *Source Forge.* https://sourceforge.net/. https://sourceforge.net/projects/opengrads/.

UCAR. 2026. "NetCDF Software Libraries." *Source Forge.* UCAR. https://www.unidata.ucar.edu/software/netcdf/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Peter Danenberg, Gábor Csárdi, and Manuel Eugster. 2025. *Roxygen2: In-Line Documentation for R.* https://roxygen2.r-lib.org/.