

wcswatin: An R Package for Processing Multi-Source Weather and Climate Data for SWAT Hydrological Modeling

Réginal Exavier^{1¶} and Peter Zeilhofer²

¹ University of São Paulo (USP), Postgraduate Program in Physical Geography, São Paulo - SP, Brazil.

² Federal University of Mato Grosso (UFMT), Postgraduate Program in Geography, Cuiabá - MT, Brazil.

¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970

Published: unpublished

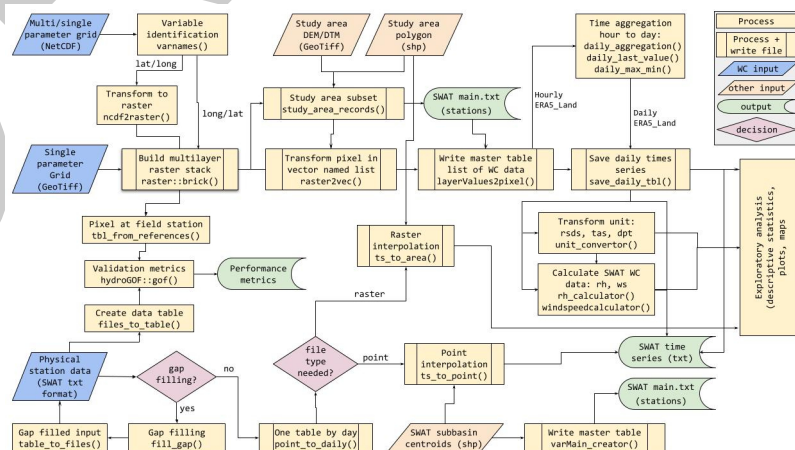
License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

Weather and climate variables constitute essential boundary conditions for hydrological modeling frameworks; however, the acquisition, harmonization, preprocessing, and quality control of heterogeneous datasets originating from multiple sources remain computationally intensive and susceptible to systematic and random errors. *wcswatin* is an R package developed to standardize and automate the preprocessing of multi-source weather and climate datasets, with a specific emphasis on generating rigorously formatted meteorological inputs for the Soil and Water Assessment Tool (SWAT), a globally established, process-based ecohydrological model.

The package provides integrated tools for extracting, cleaning, validating, and formatting Global Weather and Climate Grids (GWCG), as well as for processing measured station data through spatial interpolation. By supporting common data formats including NetCDF, GeoTIFF, and CSV, *wcswatin* enables consistent and reproducible preparation of weather and climate inputs for SWAT modeling. Figure 1 illustrates the main processing pipelines implemented in the package.



lowers methodological and technical barriers for hydrologists and environmental scientists,
particularly in applications requiring the integration of multiple climate data sources.

Statement of need

The Soil and Water Assessment Tool (SWAT) (Arnold et al., 2012) is widely used for watershed-scale hydrological modeling worldwide (Tan et al., 2020) and requires accurate weather and climate inputs, including precipitation, temperature, solar radiation, wind speed, and relative humidity. While Global Weather and Climate Grids (GWCG) derived from reanalysis products, satellite missions, and other modeled datasets provide increasingly high-resolution alternatives to sparse station networks, their preprocessing for SWAT remains technically demanding.

These datasets typically require multiple processing steps prior to model application, including temporal aggregation, unit conversion, spatial subsetting, validation against in situ observations, and conversion into model-specific input formats. Performing these steps manually or using multiple disconnected tools is time-consuming and increases the risk of errors, particularly when comparing alternative climate products or working with large spatio-temporal datasets.

`wcswatin` addresses these challenges by providing a generalized and integrated workflow for preprocessing multisource weather and climate data for SWAT modeling. The package supports the processing of gridded datasets in common formats, offers specialized functionality for compiling complete SWAT-compatible weather datasets from ERA5-Land, and includes tools for gap-filling, spatial interpolation, and validation using measured station data. By consolidating these tasks within a single R-based environment, `wcswatin` facilitates reproducible and efficient preparation of weather and climate inputs for hydrological and ecohydrological studies

Functionality

The `wcswatin` package supports the extraction, preprocessing, exploratory analysis, spatial interpolation, and validation of weather and climate data from multiple sources, with a focus on generating SWAT-compatible inputs.

For gridded datasets, the package provides functions to process Global Weather and Climate Grids supplied in NetCDF or GeoTIFF formats. This includes automated handling of temporal resolution, aggregation of hourly data to daily values, unit conversion, and derivation of additional variables required by SWAT, such as wind speed and relative humidity. These capabilities have been applied to multi-parameter reanalysis products including ERA5-Land.

In addition to gridded data, `wcswatin` supports the use of measured in situ station data. Station observations can be imported in standard SWAT format, screened for missing values, gap-filled using regression-based approaches, and spatially interpolated using trend surface analysis. The resulting interpolated datasets can be generated as point-based or gridded outputs and subsequently converted into SWAT-compatible weather inputs. Figure 2 demonstrates an example of the spatial interpolation capabilities, showing a third-order trend surface interpolation for daily precipitation data.

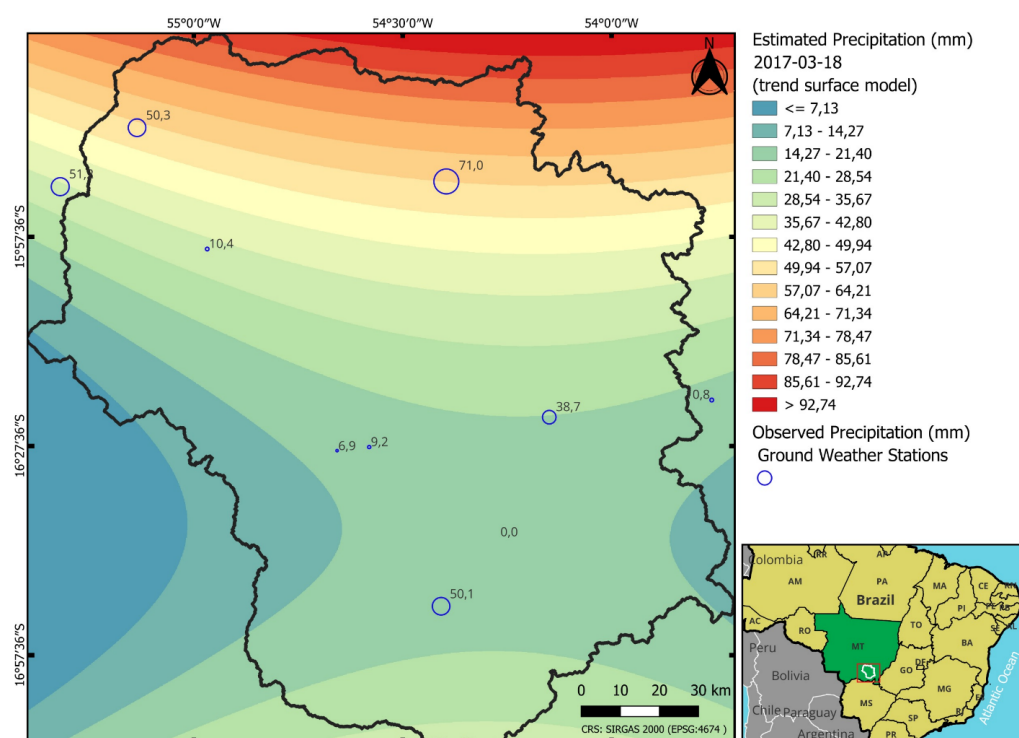


Figure 2: Example of third-order trend surface interpolation for precipitation data (March 18, 2017) using `wcswatin`'s spatial interpolation function.

The package also includes tools for validating gridded climate products against station measurements through statistical metrics and visual inspection, enabling users to compare alternative weather and climate datasets prior to hydrological modeling. Performance optimizations, including parallel processing and reduced memory overhead, allow `wcswatin` to handle large spatio-temporal datasets efficiently on standard computing systems.

Related work

Several existing tools partially support the preparation of weather and climate data for hydrological modeling. The `NASAaccess` R package (Ibrahim, 2023) provides functionality for downloading and reformatting selected NASA datasets for use in models such as SWAT, but is limited to a small number of variables and data sources. The `pRecipe` package (Godoy & Markonis, 2023) focuses on global precipitation datasets and visualization but does not offer comprehensive workflows for generating complete SWAT weather inputs.

Web-based tools like the NCEP Climate Forecast System Reanalysis portal (<https://global-weather.tamu.edu/>) are restricted to specific time periods (1979-2014) and have demonstrated weak performance in discharge simulations (Senent-Aparicio et al., 2021). In contrast, `wcswatin` emphasizes integrated preprocessing, validation, and formatting of multisource weather and climate data, supporting both gridded and station-based inputs within a unified and extensible workflow tailored to SWAT modeling.

Acknowledgements

The authors would like to acknowledge the Critical Ecosystem Partnership Fund (CEPF) for partially funding the software development under grant CEPF-104254.

References

- 82
- 83 Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R.,
84 Santhi, C., Harmel, R. D., Griensven, A. van, Van Liew, M. W., Kannan, N., & Jha, M. K.
85 (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4),
86 1491–1508. <https://doi.org/10.13031/2013.42256>
- 87 Godoy, M. R. V., & Markonis, Y. (2023). pRecipe: A global precipitation climatology toolbox
88 and database. *Environmental Modelling & Software*, 165, 105711. [https://doi.org/10.](https://doi.org/10.1016/j.envsoft.2023.105711)
89 [1016/j.envsoft.2023.105711](https://doi.org/10.1016/j.envsoft.2023.105711)
- 90 Ibrahim, M. (2023). *NASAaccess: Downloading and reformatting tool for NASA earth*
91 *observation data products*. National Aeronautics; Space Administration, Goddard Space
92 Flight Center. <https://github.com/nasa/NASAaccess>
- 93 Senent-Aparicio, J., Blanco-Gómez, P., López-Ballesteros, A., Jimeno-Sáez, P., & Pérez-
94 Sánchez, J. (2021). Evaluating the potential of GloFAS-ERA5 river discharge reanalysis
95 data for calibrating the SWAT model in the grande san miguel river basin (el salvador).
96 *Remote Sensing*, 66(16: 3299). <https://doi.org/https://doi.org/10.3390/rs13163299>
- 97 Tan, M. L., Gassman, P. W., Yang, X., & Haywood, J. (2020). A review of SWAT applications,
98 performance and future needs for simulation of hydro-climatic extremes. *Advances in Water*
99 *Resources*, 143, 103662. <https://doi.org/10.1016/j.advwatres.2020.103662>

DRAFT