

StreamCatTools: An R package for working with StreamCat and LakeCat watershed data in R

Marc H. Weber¹, Ryan A. Hill², Travis Hudson¹, Alan Brookes¹, and Selia Markley¹

¹ Office of Water, United States Environmental Protection Agency ² Pacific Ecological Systems Division, United States Environmental Protection Agency

DOI:

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted:

Published:

License

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

Summary

StreamCatTools provides functions for easily working with, visualizing and analyzing the StreamCat(Hill, Weber, Leibowitz, Olsen, & Thornbrugh, 2016) and LakeCat(Hill, Weber, Debbout, Leibowitz, & Olsen, 2018) data and API within **R**. StreamCat and LakeCat provide hundreds of landscape metrics for both the local catchment and full watershed for every stream reach and lake depicted in the medium resolution National Hydrography Dataset Plus Version 2.1 (NHDPlusV21)(McKay et al., 2012)

Statement of Need

Easily accessible, robust, and consistent watershed data is an underpinning of hydrology research, water quality monitoring programs, and predictive modelling applications, to name just a few examples. The StreamCat(Hill et al., 2016) and LakeCat(Hill et al., 2018) datasets fill this need by providing nationally consistent watershed data that has had stringent quality control applied for the conterminous United States (CONUS). The data encompasses hundreds of watershed metrics for every stream reach and lake feature represented in the NHDPlusV21(McKay et al., 2012). StreamCatTools fills this need for easily accessible watershed metrics for CONUS by: (1) providing a simple interface in R to the StreamCat and LakeCat web services, (2) providing convenient functionality to find available StreamCat and LakeCat metric names and information, (3) extracting StreamCat and LakeCat metrics by COMID (a unique identifier in the NHDPlusV2 framework), by state, by county, by NHD Hydroregion, or for all of CONUS. Providing this valuable watershed data via web services in R follows the FAIR principles laid out in (Wilkinson et al., 2016).

Package Overview

How it works

Installing StreamCatTools

```
# library(remotes)
# install_github("USEPA/StreamCatTools", build_vignettes=FALSE)
```

StreamCatTools is loaded into an **R** session:

```
library(StreamCatTools)
```

Examples

Discussion

Let's talk StreamCat!

Acknowledgements

Examples of using StreamCat and LakeCat make extensive use of `nhdplusTools` (Blodgett & Johnson, 2023) and the functions for accessing the API are facilitated through use of `httr2`. Figures were created using `ggplot2` (Wickham, 2016).

We would like to sincerely thank the editor and reviewers for all of their helpful feedback which greatly improved both the software and the manuscript.

The United States Environmental Protection Agency (EPA) GitHub project code is provided on an “as is” basis and the user assumes responsibility for its use. EPA has relinquished control of the information and no longer has responsibility to protect the integrity, confidentiality, or availability of the information. Any reference to specific commercial products, processes, or services by service mark, trademark, manufacturer, or otherwise, does not constitute or imply their endorsement, recommendation or favoring by EPA. The EPA seal and logo shall not be used in any manner to imply endorsement of any commercial product or activity by EPA or the United States Government.

References

- Blodgett, D., & Johnson, M. (2023). *nhdplusTools: Tools for accessing and working with the NHDPlus*. Reston, VA: U.S. Geological Survey; U.S. Geological Survey. Retrieved from <https://doi.org/10.5066/P97AS8JD>
- Hill, R. A., Weber, M. H., Debbout, R. M., Leibowitz, S. G., & Olsen, A. R. (2018). The lake-catchment (LakeCat) dataset: Characterizing landscape features for lake basins within the conterminous USA. *Freshwater Science*, 37(2), 208–221. doi:[10.1086/697966](https://doi.org/10.1086/697966)
- Hill, R. A., Weber, M. H., Leibowitz, S. G., Olsen, A. R., & Thornbrugh, D. J. (2016). The stream-catchment (StreamCat) dataset: A database of watershed metrics for the conterminous united states. *JAWRA Journal of the American Water Resources Association*, 52(1), 120–128. doi:<https://doi.org/10.1111/1752-1688.12372>
- McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., & Reah, A. (2012). *NHDPlus version 2: User guide*. Retrieved from http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. doi:[10.1007/978-0-387-98141-3](https://doi.org/10.1007/978-0-387-98141-3)
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3(1), 1–9. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)