

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

**Marc H. Weber<sup>1</sup>, Ryan A. Hill<sup>2</sup>, Travis Hudson<sup>1</sup>, Selia Markley<sup>1</sup>, and Alan Brookes<sup>1</sup>**

<sup>1</sup> Office of Water, United States Environmental Protection Agency <sup>2</sup> Office of Research and Development, 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](#)).

## Contents

## Summary

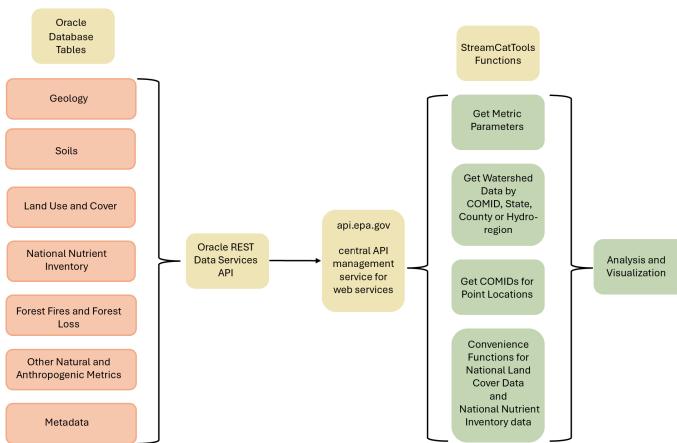
**StreamCatTools** provides functions for easily working with, visualizing and analyzing StreamCat(Hill, Weber, Leibowitz, Olsen, & Thornbrugh, 2016) and LakeCat(Hill, Weber, Debbout, Leibowitz, & Olsen, 2018) watershed metrics within **R**. The StreamCat and LakeCat datasets provide hundreds of landscape metrics for both the local catchment (e.g. landscape draining to a particular stream reach) 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) for the contiguous United States (CONUS). **StreamCatTools** functions wrap the application programming interface (API) for the StreamCat and LakeCat data and facilitate simple, straightforward access and use of these datasets within R.

## 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 curated watershed data for CONUS that has had stringent quality control applied. The data encompasses hundreds of watershed metrics for every stream reach and lake feature represented in the NHDPlusV21(McKay et al., 2012). StreamCatTools fills the 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 Hydro-region, 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).

## StreamCatTools Functionality

**StreamCatTools** provides a simple streamlined set of functions to easily query and ingest watershed landscape metrics into an R session. Figure 1 shows the overall framework of the StreamCat database, application programming interface, and functionality in the package that simplifies data access in R using web services for the StreamCat and LakeCat datasets.



**Figure 1:** The StreamCat framework: a backend Oracle database, web services exposed through api.gov, and functions in StreamCatTools which simplify accesss and analyzing the data in R via the API.

The core functionality of **StreamCatTools** simplifies the calls to the API for StreamCat and LakeCat data within the R programming language to allow a user to: 1) ask for details on available StreamCat and LakeCat metrics 2) Reference sites with location information to NHDPlusV2 catchments for stream and lake features in order to ask for StreamCat and LakeCat metrics and 3) Get StreamCat and LakeCat watershed metrics by COMID, by state, by county, by hydroregion, or for all of CONUS. Additional functionality is provided in convenience functions get data all data for particular years for particular data sources such as the National Land Cover Database (NLCD) ((USGS), 2024) and the National Nutrient Inventory (NNI) (Brehob et al., 2025). The core functions in **StreamCatTools** leverage the `httr2` library `httr2`(Wickham, 2025) for a modern, pipeable API for working with web APIs.

#### Installing `StreamCatTools`

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

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

```
library(StreamCatTools)
```

As we detail in the [package introduction page](#) we can list metric parameters and find out more about StreamCat and LakeCat data available in **StreamCatTools** using using the `sc_get_params` function. We can also see what metrics are available for what areas of interest and what years using the `sc_get_params` function (which returns a tibble of information about StreamCat metrics). Additional functions provided for getting metadata on the underlying StreamCat and LakeCat data include the `sc_fullname` and `lc_fullname` functions and the `sc_get_params` and `lc_get_params` functions. We can also filter metric names and information by the metric year(s), the indicator categories for metrics, the metric data set names, or the Areas of Interest the metrics are available for using the `sc_get_metric_names` or `lc_get_metric_names` functions.

The primary package functionality is in the `sc_get_data` and `lc_get_data` functions which allow users to extract whatever suite of metrics they are interested in from StreamCat or LakeCat using a provided set of unique identifiers of COMIDs for NHDPlusV2, or request a set of metrics for a given state(s), county(ies), hydroregion(s), or all of CONUS. Additionally, convenience functions are provided for accessing the NLCD and NNI datasets using `sc_nlcd` and `lc_nlcd` and `sc_nni` and `lc_nni`.