

# **Processing USGS streamflow data in R**

**EPA Water Modeling Webinar Series**  
**July 19, 2022**

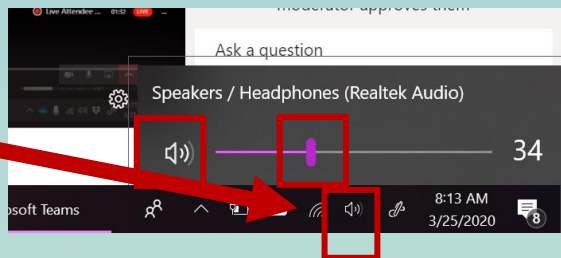
# General Housekeeping

## Meeting Audio

- ✓ Attendees must use **Device Audio** (computer, headset, tablet, smart phone)
- ✓ Presenter Audio Only (like a TV broadcast)
- ✓ If you have trouble with audio, ensure your system/device audio is turned up and not muted

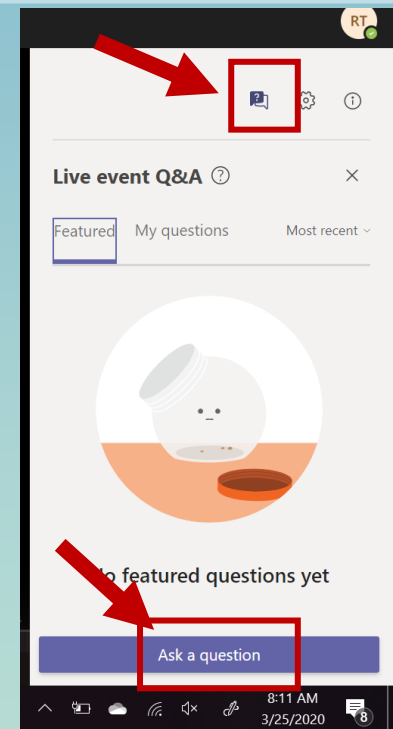


Computer  
Audio  
(bottom right)



## Meeting Q & A's

- ✓ Ask a question under Live Event Q&A panel on the right
- ✓ Click “Ask a question”
- ✓ Moderator will respond



# Water Quality Modeling Webinar Series

- Purpose: To help water quality professionals better understand water quality modeling and how models can be used to solve the problems facing water quality regulators.
- 31 webinars to date
- Webinars recorded and posted:  
<https://www.epa.gov/waterdata/surface-water-quality-modeling-training>

# Audience

- Water quality professionals
- Clean Water Act (CWA) regulators: TMDL, standards, wetlands, assessment, permitting, etc.
- Scientists, engineers, managers, students, attorneys
- Assumptions for audience members:
  - Have an understanding of basic hydrology, water quality, and land use principles, such as eutrophication, flow calculations, erosion processes, etc.

# Speakers

**Kateri Salk**

Tetra Tech

[kateri.salkgundersen@tetrattech.com](mailto:kateri.salkgundersen@tetrattech.com)

# Workshop Goals

1. Incorporate reproducible and efficient workflows into data workflows for water quality modeling
2. Acquire, process, and explore common types of data used in water quality modeling
3. Create and modify coding scripts for future use



# Webinar Series Overview (Topics)

- Processing Water Quality Portal data in R
- Processing USGS Streamflow data in R
- Processing NOAA Weather data in R
- Processing gridded weather data in ArcGIS and Python
- Watershed delineation in ArcGIS and Python



# Follow-Up from Water Quality Session

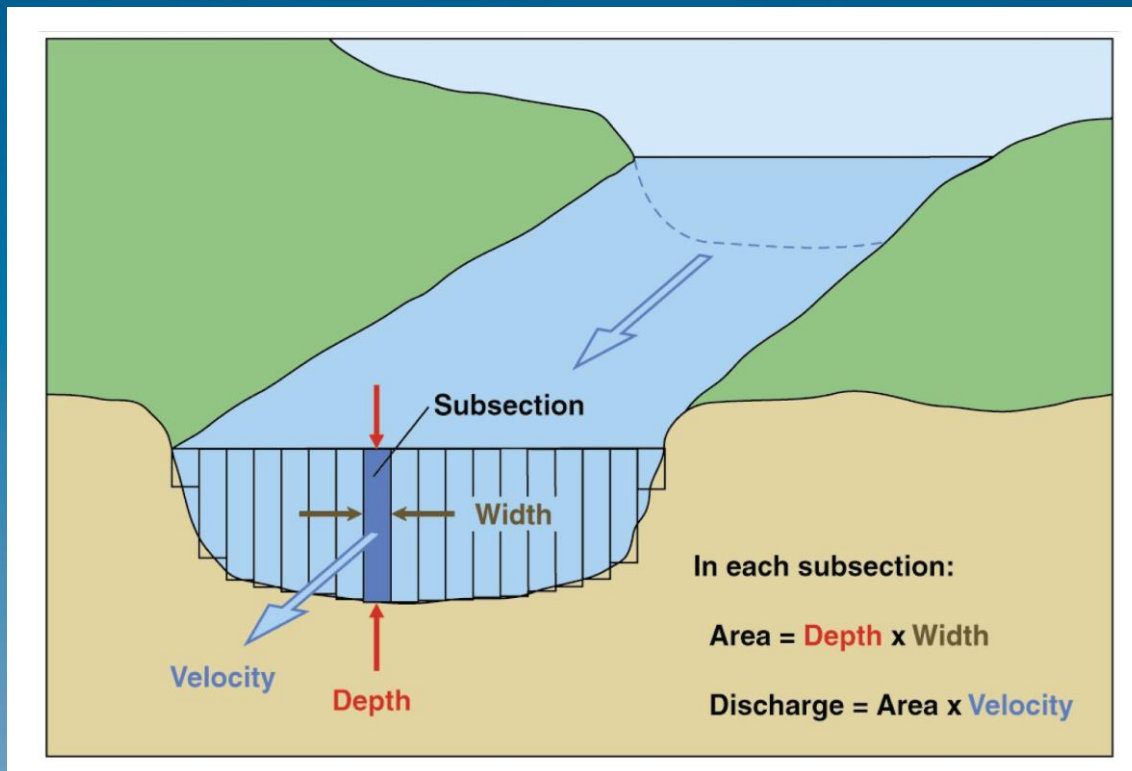
- Webinar intended as a basic introduction to using dataRetrieval to acquire and process water quality data
- Several packages can further automate data QA
- e.g., NADA package can run robust ROS on nondetects
- **Recommendations**
  - Automation should never be a substitute for understanding your data
  - Always check multiple fields in WQP data downloads; reporting is not always consistent!





# Stream and River Discharge

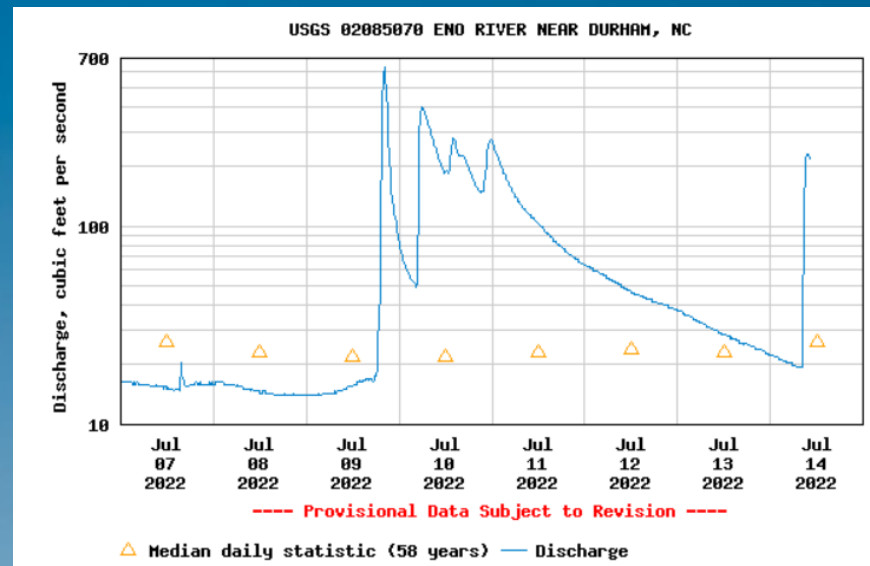
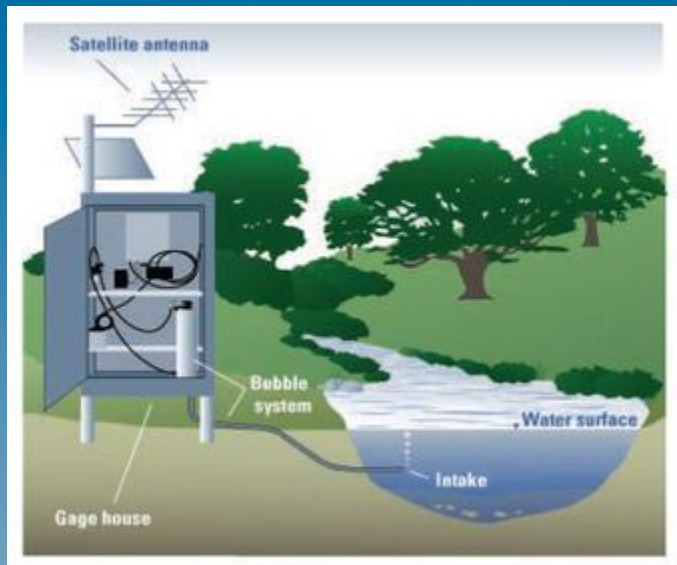
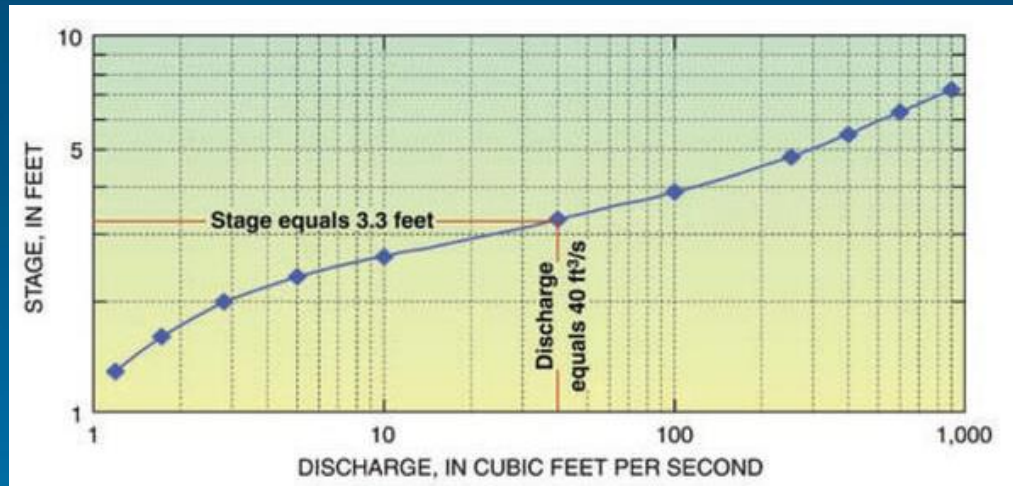
- Key physical characteristic in hydrology
- Volume/time, typically cfs for USGS data



Source: USGS

# Stream and River Discharge

Source: USGS



# Data Sources

## Hydrology/Streamflow

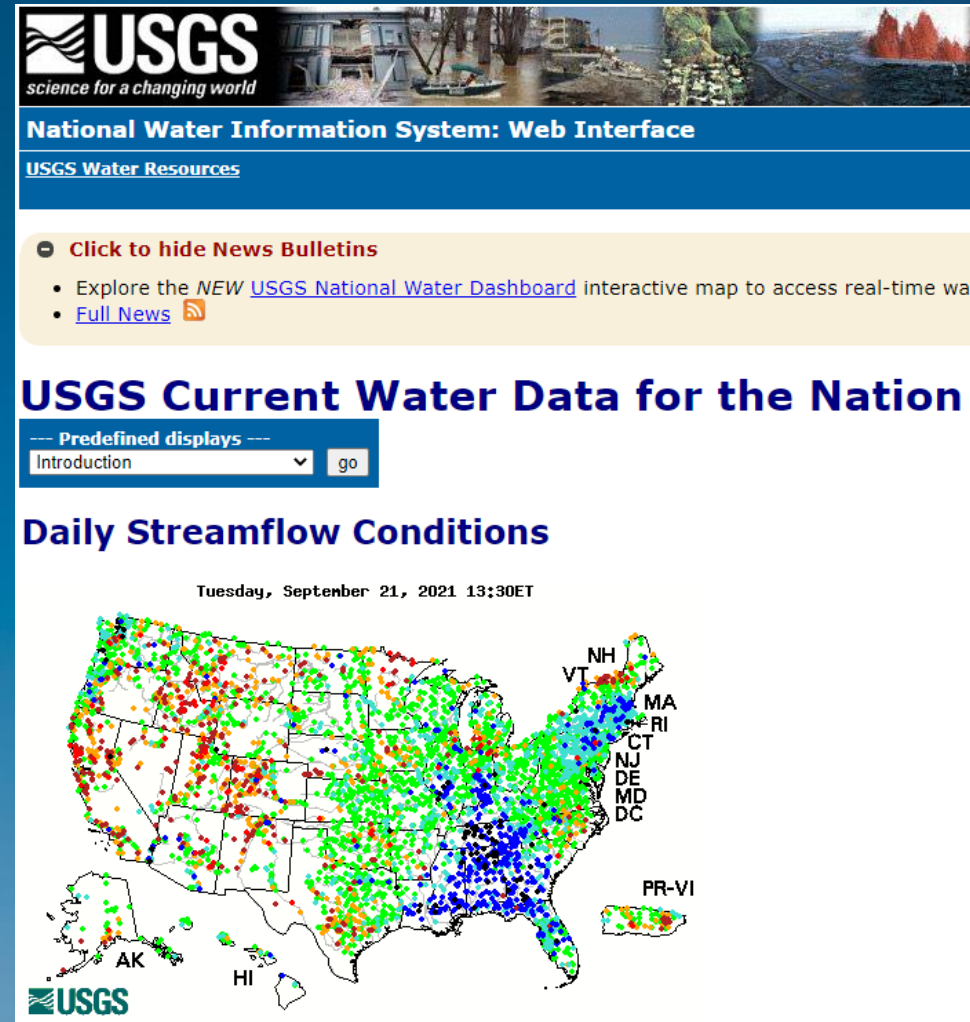
- USGS gages (R package: dataRetrieval)
- Agencies/other entities
- StreamCat
- StreamStats (R package: streamstats)
- NHDPlus (R package: nhdplusTools)



# USGS Gage Data

<https://cran.r-project.org/web/packages/dataRetrieval/vignettes/dataRetrieval.html>

<https://github.com/USGS-R/dataRetrieval>



The screenshot shows the USGS National Water Information System Web Interface. At the top is the USGS logo with the tagline "science for a changing world". Below the logo is a banner image showing a flooded area with a boat. The main heading is "National Water Information System: Web Interface". Underneath is a link to "USGS Water Resources". A yellow box contains a link to "Click to hide News Bulletins" and a list of links: "Explore the NEW USGS National Water Dashboard interactive map to access real-time wa" and "Full News". Below this is a section titled "USGS Current Water Data for the Nation". It features a dropdown menu for "Predefined displays" with "Introduction" selected and a "go" button. The main section is titled "Daily Streamflow Conditions" and shows a map of the United States with numerous colored dots representing streamflow gauges. The map is dated "Tuesday, September 21, 2021 13:30ET". State abbreviations are visible on the map, including NH, VT, MA, RI, CT, NJ, DE, MD, DC, AK, HI, and PR-VI. The USGS logo is in the bottom left corner of the map area.



## dataRetrieval


Retrieval functions for USGS and EPA hydrologic and water quality data.

For complete tutorial information, see:

<https://usgs-r.github.io/dataRetrieval>

<https://owi.usgs.gov/R/dataRetrieval.html>

# StreamCat


United States  
Environmental Protection  
Agency

Search EPA.gov

Environmental Topics ▾
Laws & Regulations ▾
Report a Violation ▾
About EPA ▾

National Aquatic Resource Surveys
CONTACT US

National Aquatic Resource Surveys Home
Background
Indicators
Manuals
Map of Sampled Sites
**NARS Data**
Journal Articles
Applying the Data
Related Studies and Tools

## StreamCat Dataset

StreamCat currently contains over 600 metrics that include local catchment (Cat), watershed (Ws), and special metrics. The special metrics were derived through modeling or by combining other StreamCat metrics. These variables include predicted water temperature, predicted biological condition, and the indexes of catchment and watershed integrity. See [Geospatial Framework and Terms](#) below for definitions of catchment and watershed as used with the StreamCat Dataset.

These metrics are available for ~2.65 million stream segments and their associated catchments across the conterminous US. StreamCat metrics represent both natural (e.g., soils and geology) and anthropogenic (e.g, urban areas and agriculture) landscape information.

**Looking for data on lakes?**

Data for 378,088 lakes also available from the [LakeCat Dataset](#)

## StreamCatTools


lifecycle experimental R-CMD-check failing

`StreamCatTools` is a draft R package for accessing StreamCat data via the in-development StreamCat API and for working with site data in conjunction with StreamCat and NHDPlus. The StreamCat API is currently behind the agency firewall so the functions in the package will not currently work for those outside the EPA. `StreamCatTools` is in pre-alpha development - this is a work in progress, look for a working package release in the next year.



# StreamStats

<https://github.com/markwh/streamstats>




SCIENCE  
Topics, centers,  
missions

PRODUCTS  
Maps, data,  
publications

NEWS  
Releases,  
I'm a reporter

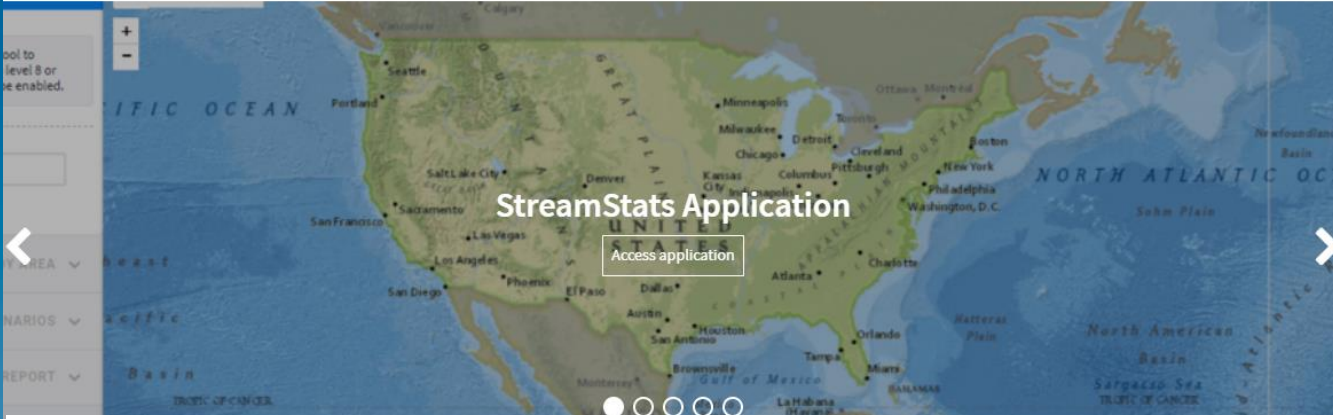
CONNECT  
Contact, chat,  
social media

ABOUT  
Organization,  
jobs, budget



Water Resources

## StreamStats: Streamflow Statistics and Spatial Analysis Tools for Water-Resources Applications



### streamstats : An R package for using the USGS Streamstats API

See API documentation over [here](#)

Downloading and using (in R):

```
devtools::install_github("markwh/streamstats")
library(streamstats)
```

# nhdplusTools

<https://github.com/USGS-R/nhdplusTools>

## nhdplusTools

R-CMD-check passing codecov 94% downloads 15K CRAN 0.4.3



🔗 **nhdplusTools: Tools for Accessing and Working with the NHDPlus**

### Terminology:

The following definitions have been used as much as possible throughout the package.

Terms for rivers:

**Flowline:** The NHD name for a hydrographic representation of a flowing body of water. Flowline is generally used when referring to geometry.

**Flowpath:** The HY\_Features name for a hydrologic feature that is the primary path water follows through a catchment; either from headwater to outlet or inlet to outlet. Flowpath is used when describing aspects of the abstract flowpath featuretype, generally in relation to a flowpath's relationship to a catchment.

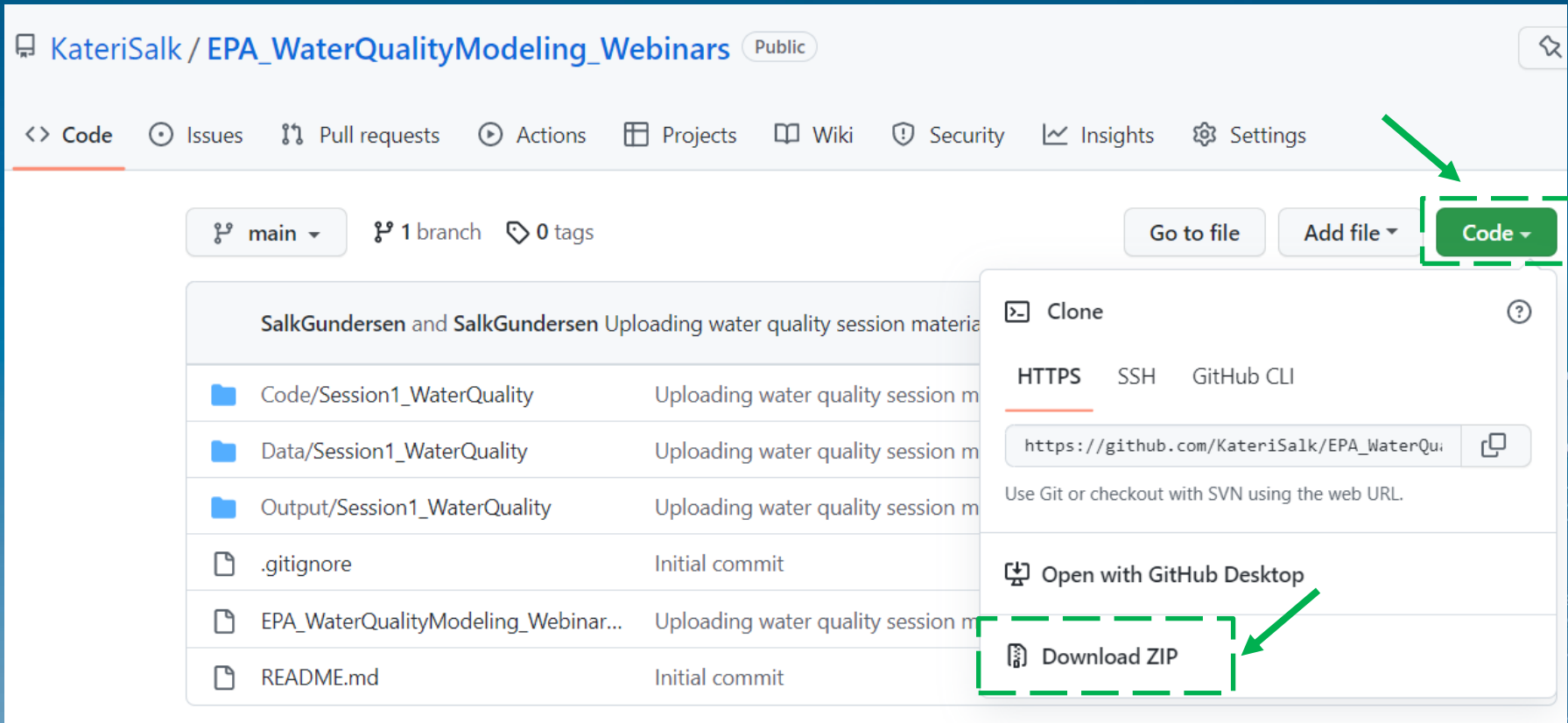
Terms used for hydrologic units:

**Catchment:** The most abstract unit of hydrology in HY\_Features is the catchment. It is a physiographic unit with zero or one inlets and one outlet. It does not inherently have any conceptual realizations. Rather, a given catchment can be realized in a number of ways; flowpath, divide, and networks of flowpaths and divides are the primary realizations.

**Catchment divide:** NHD "catchment" polygons are more accurately described as "catchment divide" features. Because of the overlap with the HY\_Features abstract "catchment" feature type, "catchment divide" is used for polygon representations of catchments.

# Webinar Resources

[github.com/KateriSalk/EPA\\_WaterQualityModeling\\_Webinars](https://github.com/KateriSalk/EPA_WaterQualityModeling_Webinars)



KateriSalk / EPA\_WaterQualityModeling\_Webinars Public

<> Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags

Go to file Add file Code

SalkGundersen and SalkGundersen Uploading water quality session materia

Code/Session1_WaterQuality	Uploading water quality session m
Data/Session1_WaterQuality	Uploading water quality session m
Output/Session1_WaterQuality	Uploading water quality session m
.gitignore	Initial commit
EPA_WaterQualityModeling_Webinar...	Uploading water quality session m
README.md	Initial commit

Clone

HTTPS SSH GitHub CLI

[https://github.com/KateriSalk/EPA\\_WaterQui](https://github.com/KateriSalk/EPA_WaterQui)

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP



# Coding Portion for Today's Webinar

- **Open EPA\_WaterQualityModeling\_Webinars.Rproj**
  - This will open R and RStudio
  - Project file allows you to navigate within a project folder
- **From the “Files” tab inside RStudio:**
  - Navigate to the Code folder
  - Navigate to the Session2\_Discharge folder
  - Click on DischargeSession.Rmd



# R Setup

- See “Setup” folder in the repository
- Use the R\_Setup file if you don’t have R and Rstudio installed
- Watch along for today’s session if software aren’t installed yet

## EPA Water Modeling Webinar Series: R Coding Setup

### R

If you already have R installed, please ensure you are using version 4.1.1 or higher. If not, follow the instructions below. If you do not have R installed, follow the instructions below.

1. Open a web browser and navigate to the Duke CRAN mirror: <http://archive.linux.duke.edu/cran/>
2. Please select your operating system (Mac, Windows, Linux)

For Windows Users: install the **base package**

For Mac Users: install the R-4.1.1.pkg file. Note: you must have OS X 10.13 (High Sierra) installed or higher. You must also install XQuartz, which can be found in the Latest Release paragraphs on the same page.

### RStudio

If you already have RStudio installed, please ensure you are using version 1.4.1717 or higher (if you are using a relatively recent version, you likely won’t run into any problems). To check that you are using the most up-to-date version, go to the Help menu in RStudio and select “Check for Updates.” You will be instructed to install any updates that are available.

If you do not have RStudio installed on your computer, please do the following:

1. Open a web browser and navigate to <http://www.rstudio.com/products/rstudio/download/>
2. Select and download the appropriate installer for your operating system (Windows, Mac, Linux)
3. Open the installer and follow the onscreen directions

### Other R resources

If you would like additional resources for using R and RStudio, R for Data Science is a great resource for developing R skills.

# Questions?

# Contact Information

Kateri Salk

kateri.salkgundersen@tetrattech.com

Brian Pickard

brian.pickard@tetrattech.com

**[https://www.epa.gov/waterdata/surface-  
water-quality-modeling](https://www.epa.gov/waterdata/surface-water-quality-modeling)**

# **Next Topic: Processing NOAA Weather Data in R August 16, 2022**

**EPA Water Modeling Workgroup  
Webpage:**

**[https://www.epa.gov/waterdata/surface-  
water-quality-modeling-training](https://www.epa.gov/waterdata/surface-water-quality-modeling-training)**

# Subscribe to EPA Office of Water Updates

**Get on our email list here:**

<https://www.epa.gov/aboutepa/about-office-water>