

Processing Water Quality Portal data in R

EPA Water Modeling Webinar Series
June 21, 2022

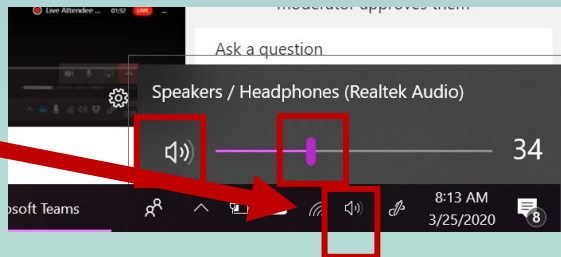
General Housekeeping

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

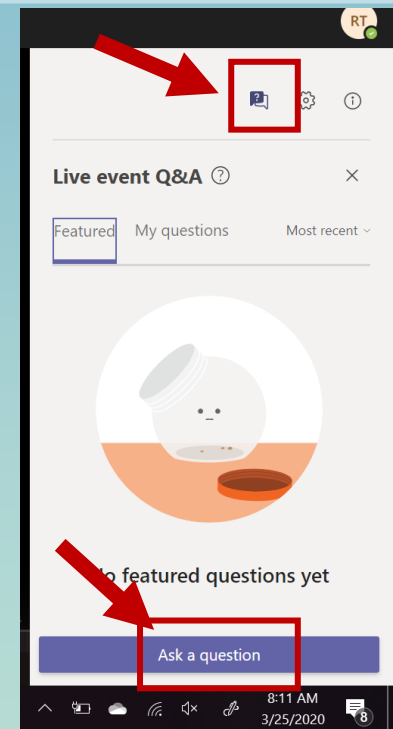


Computer
Audio
(bottom right)



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.
- 30 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

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



Water Quality Data

- **Biological, chemical, physical constituents**
- **Historically:**
 - Collected in person, analyzed in the lab
 - Organizations maintained datasets individually
- **Current:**
 - Much still collected and analyzed manually
 - Increased automation for some constituents
 - Greater capacity for centralized data storage and maintenance
 - Expectations for more consistent SAPs and QA/QC



Considerations: Assessment, TMDLs

- **Sampling frequency**
- **Exceedance frequency**
- **Sampling location**
 - **Lakes:** pelagic/littoral, riverine/transitional/lacustrine, depth
 - **Streams:** spatial representativeness (e.g., reach), relation to tributaries, diversions, dams



Data Sources

Water Quality

- USGS NWIS (R package: dataRetrieval)
- EPA STORET (R package: dataRetrieval)
- Agencies/other entities
- National Lakes Assessment, National Rivers & Streams Assessment
- LAGOS-NE (soon, LAGOS-US) (R package: LAGOSNE)



NWIS and STORET: Water Quality Portal

- USGS
- EPA
- NWQMC
- >400 state, federal, tribal, local agencies



NATIONAL WATER
QUALITY MONITORING
COUNCIL

[Home](#)

[Explore WQP Sites](#) ▾

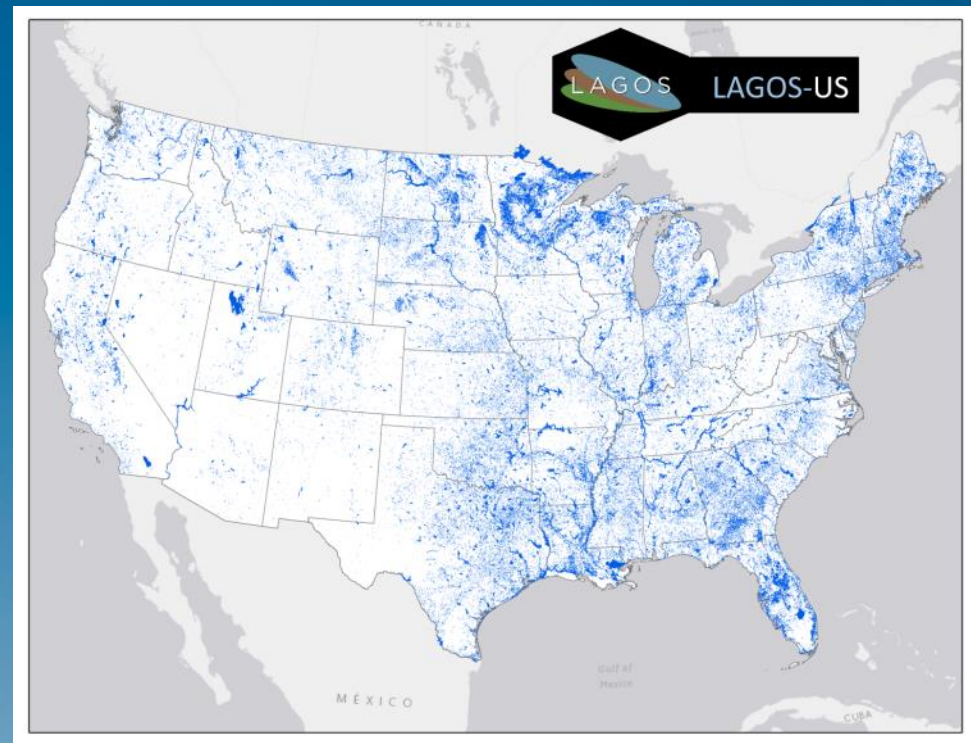
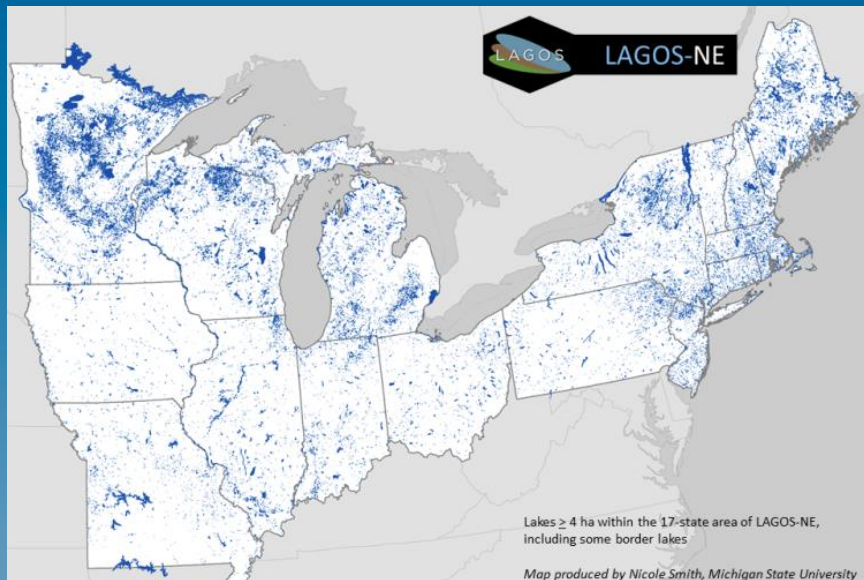
[Help & About](#) ▾



LAGOS-NE (soon, US)

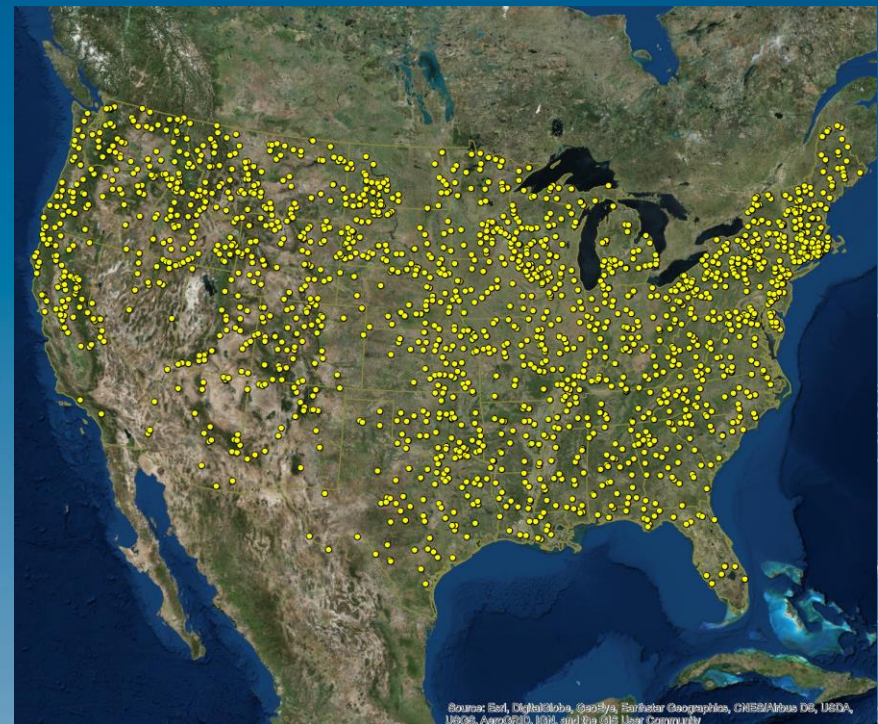
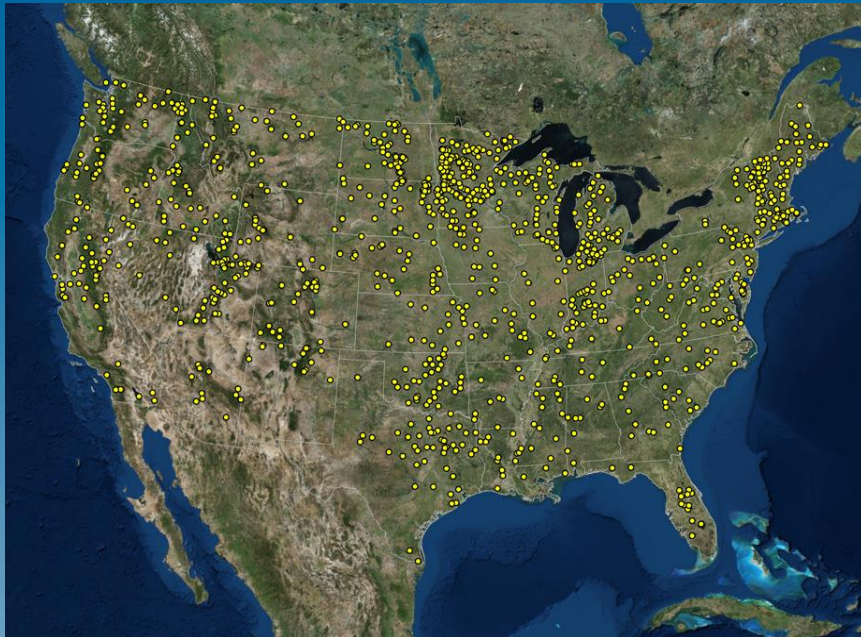
<https://lagoslakes.org/>

<https://github.com/cont-limno/LAGOSNE>



NLA and NRSA

- National & Regional focus (includes AK & HI)
- Statistically representative
- Not linked to R packages



Setting Up Reproducible Workflows



NATIONAL WATER
QUALITY MONITORING
COUNCIL

[Home](#)

[Explore WQP Sites](#) ▾

[Help & About](#) ▾

- **Example: acquiring and processing publicly available water quality data from the Water Quality Portal (WQP)**
- **Highlight of:**
 - Non-reproducible workflow
 - Reproducible workflow in R



WQP: Non-Reproducible Workflow

Select Location Parameters

Specify location parameters to describe the spatial extent of the desired dataset. All fields are optional.

Country ⁱ

All Countries ▼

State ⁱ

All States ▼

County ⁱ

All Counties ▼

Point Location ⁱ

Within

miles of
Latitude

0

Longitude

0

Use my location

Bounding Box ⁱ

North:

90

South:

-90

East:

180

West:

-180

Site Type ⁱ

All Site Types ▼

Organization ID ⁱ

All Organization IDs ▼

Site ID ⁱ

All Site IDs ▼

HUC ⁱ

All HUCs

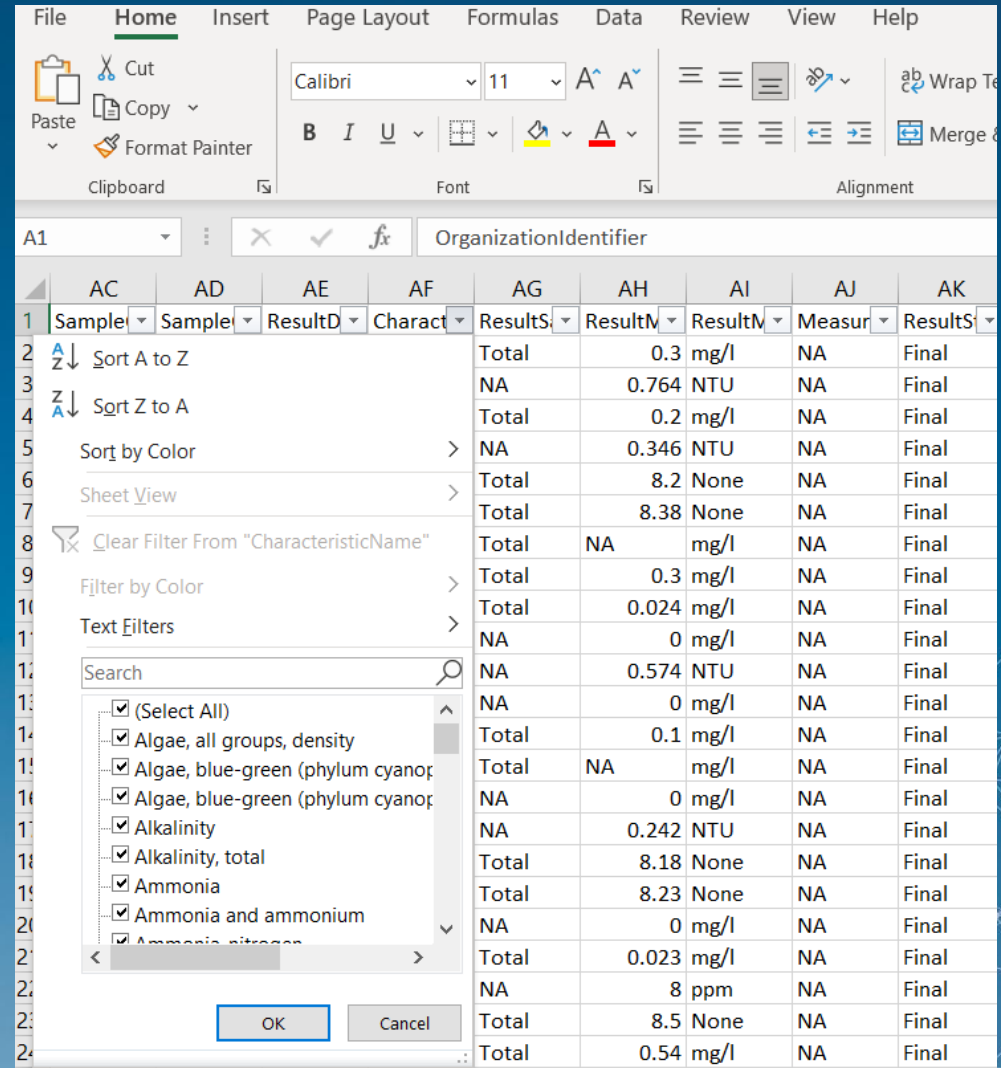
Show upstream downstream mapper

BETA

Upstream/Downstream information can be used to help you determine where to collect water quality data based on the flow of water. Select a feature source, and optionally search for a location. Then, click a feature to select upstream/downstream navigation and enter a desired distance. These search criteria will be populated in your query.

WQP: Non-Reproducible Workflow

- In Excel:
 - Filtering for constituents of interest
 - Manual handling of samples below detection
 - Pivot tables
 - Manual graph construction
- Updating/Expanding
 - Re-download data
 - Copy-paste
 - Consistency not guaranteed



	AC	AD	AE	AF	AG	AH	AI	AJ	AK
	Sample	Sample	ResultD	Character	ResultS	ResultM	ResultM	Measur	ResultS
1					Total	0.3	mg/l	NA	Final
2					NA	0.764	NTU	NA	Final
3					Total	0.2	mg/l	NA	Final
4					NA	0.346	NTU	NA	Final
5					Total	8.2	None	NA	Final
6					Total	8.38	None	NA	Final
7					Total	NA	mg/l	NA	Final
8					Total	0.3	mg/l	NA	Final
9					Total	0.024	mg/l	NA	Final
10					NA	0	mg/l	NA	Final
11					NA	0.574	NTU	NA	Final
12					NA	0	mg/l	NA	Final
13					Total	0.1	mg/l	NA	Final
14					Total	NA	mg/l	NA	Final
15					NA	0	mg/l	NA	Final
16					NA	0.242	NTU	NA	Final
17					Total	8.18	None	NA	Final
18					Total	8.23	None	NA	Final
19					NA	0	mg/l	NA	Final
20					Total	0.023	mg/l	NA	Final
21					NA	8	ppm	NA	Final
22					Total	8.5	None	NA	Final
23					Total	0.54	mg/l	NA	Final

WQP: Reproducible Example

- In R:

- Pull from WQP with one line of code
- Filter using a customizable list of constituents
- Employ logical statements to process non-detects, etc.
- Code data summaries and exploratory visualization

- Updating/Expanding

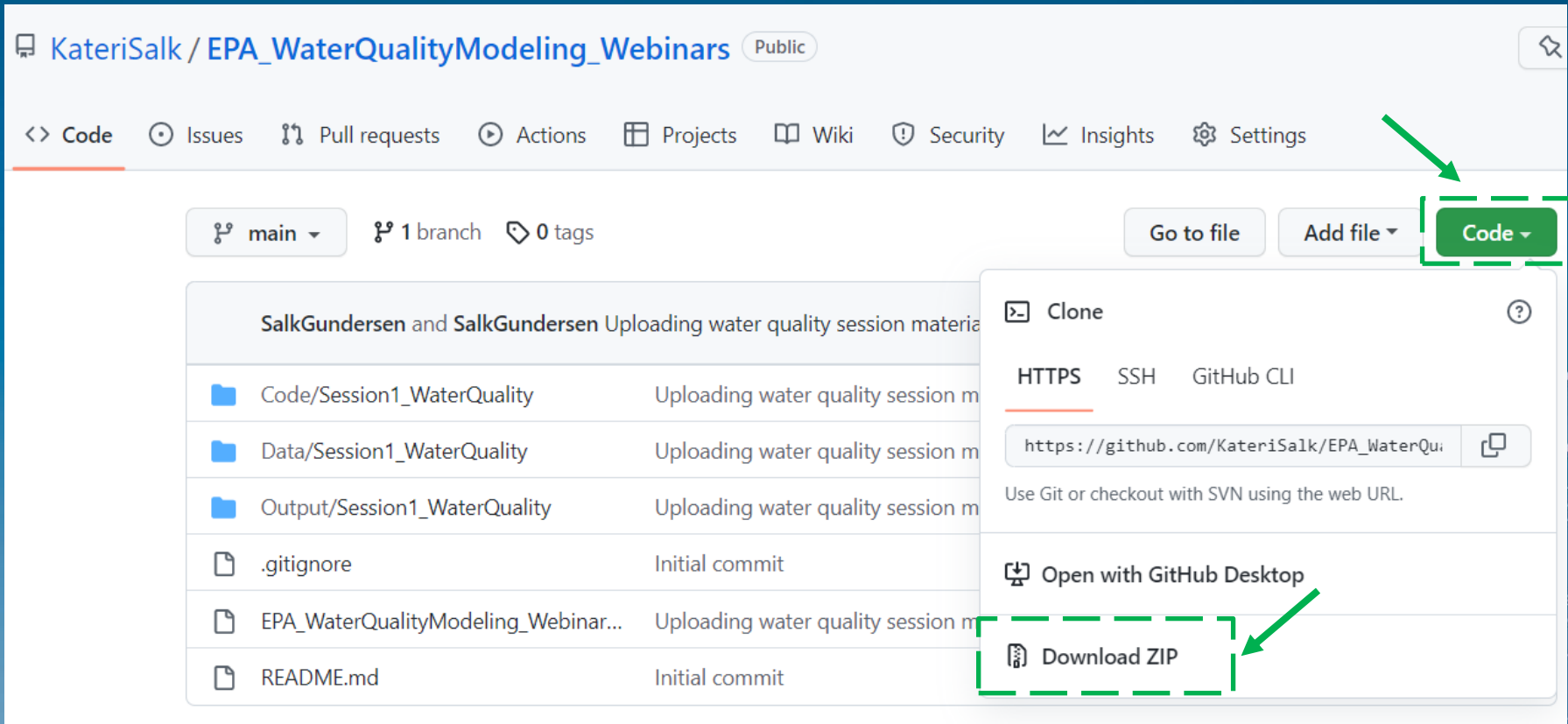
- Re-run script at the click of a button

```
23 # {r}
24 AZdata <- readWQPdata(statecode = "Arizona",
25                       siteType = "Lake, Reservoir, Impoundment",
26                       sampleMedia = c("water", "water"))
27
28 AZsites <- whatWQPsites(statecode = "Arizona",
29                         siteType = "Lake, Reservoir, Impoundment")
30
```

```
38 AZdata_filtered <- AZdata %>%
39   filter(CharacteristicName %in%
40     c("Inorganic nitrogen (nitrate and nitrite)",
41       "Turbidity",
42       "pH",
43       "Phosphate-phosphorus",
44       "Total suspended solids",
45       "Dissolved oxygen (DO)",
46       "Kjeldahl nitrogen",
47       "Organic carbon",
48       "Depth, Secchi disk depth",
49       "Phosphorus",
50       "Temperature, water",
51       "Chlorophyll a, corrected for pheophytin",
52       "Algae, blue-green (phylum cyanophyta) percent",
53       "Alkalinity, total",
54       "Ammonia-nitrogen",
55       "Dissolved oxygen saturation",
56       "Nitrogen",
57       "Algae, all groups, density",
58       "Nitrate",
59       "Orthophosphate",
60       "Oxygen",
61
```

Webinar Resources

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

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 Session1_WaterQuality folder
 - Click on WaterQualitySession.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 USGS streamflow data in R July 19, 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>