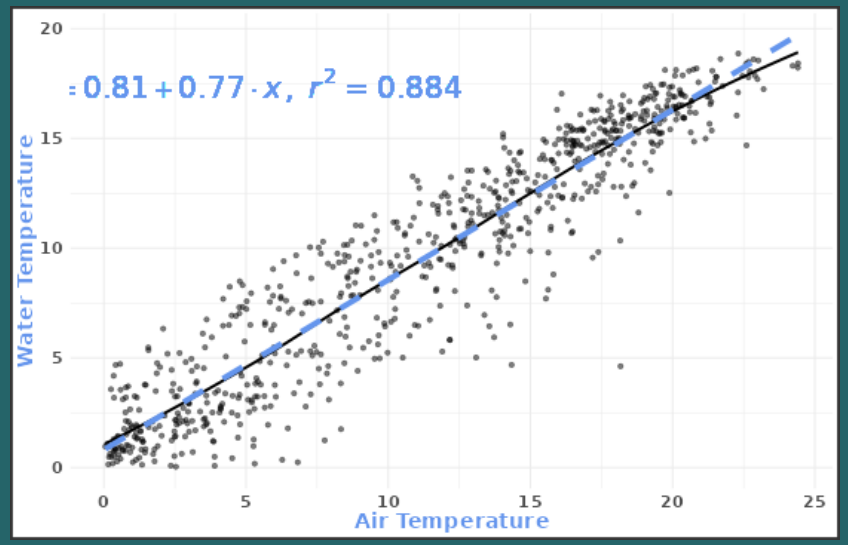
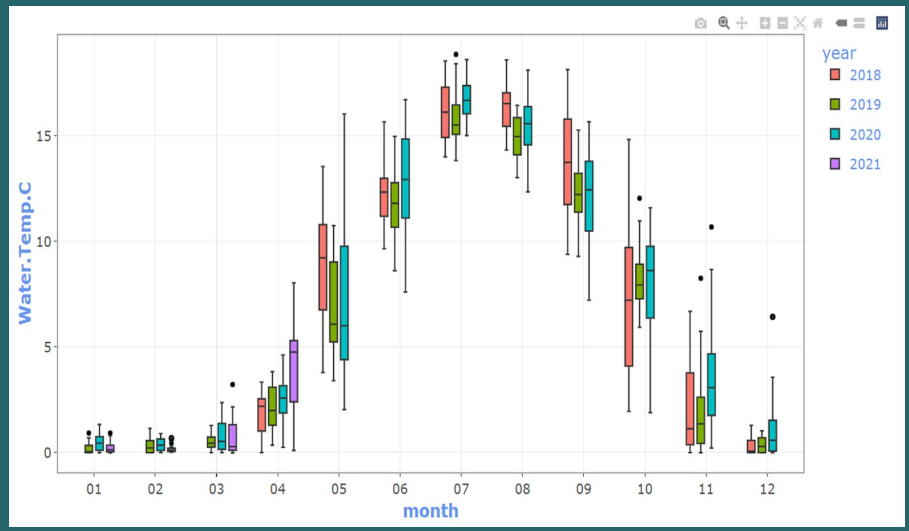
ContDataSumViz: An open source R Shiny web application for summarizing and visualizing continuous water quality sensor data





Laura Naslund (EPA), Michael Pennino (EPA), Jen Stamp (Tetra Tech), Erik Leppo (Tetra Tech), Britta Bierwagen (EPA)

2025 National Training Workshop on Water Quality Data, Assessment, and Plans. Shepherdstown, WV. June 4, 2025
* Breakout Session - Continuous Monitoring: Approaches to Managing and Analyzing the Data.

Disclaimers

The views expressed in this presentation are those of the author and do not necessarily reflect views or policies of the U.S. Environmental Protection Agency or other collaborating agencies.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use but is for descriptive purposes only.

This document does not constitute an endorsement of a particular procedure or method.

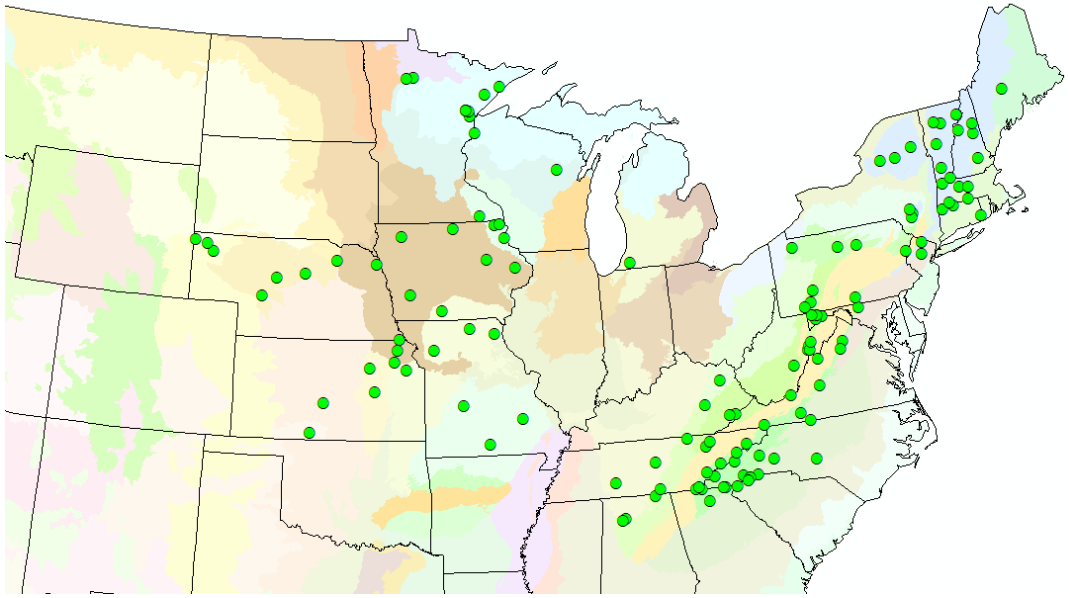
Today's talk

- Background (why we developed the tool)
- Features of the tool
- Status of tool development

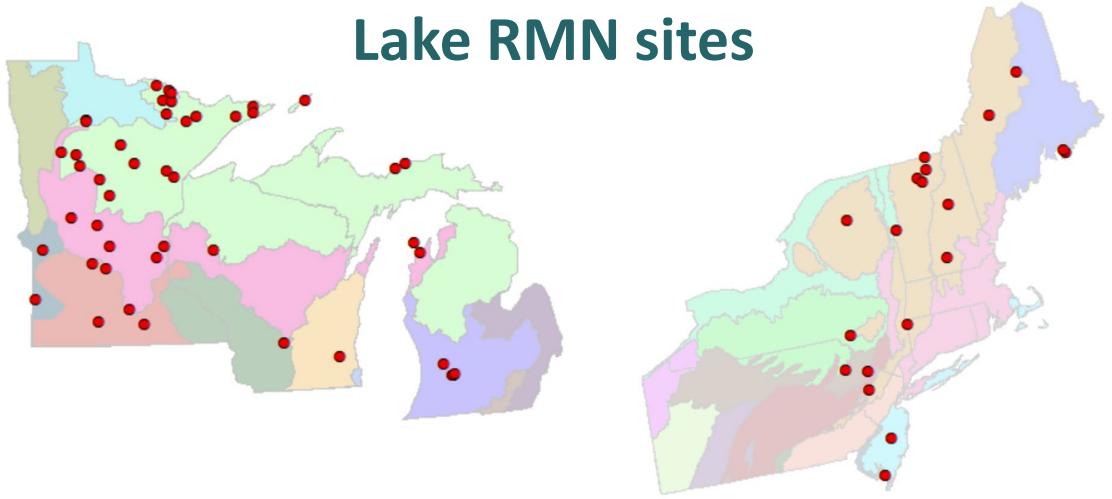
Background

- Northeast stream Regional Monitoring Network (RMN) partners **started collecting continuous sensor data in 2012.**
- Most RMN partners **did not have a data management system that could accommodate continuous data.**
- Most **did not have a formal system in place for QC'ing continuous sensor data.**
- The goal of this project was **to provide a free, open-source R-based option** to help address these needs.

Stream RMN sites with continuous sensors

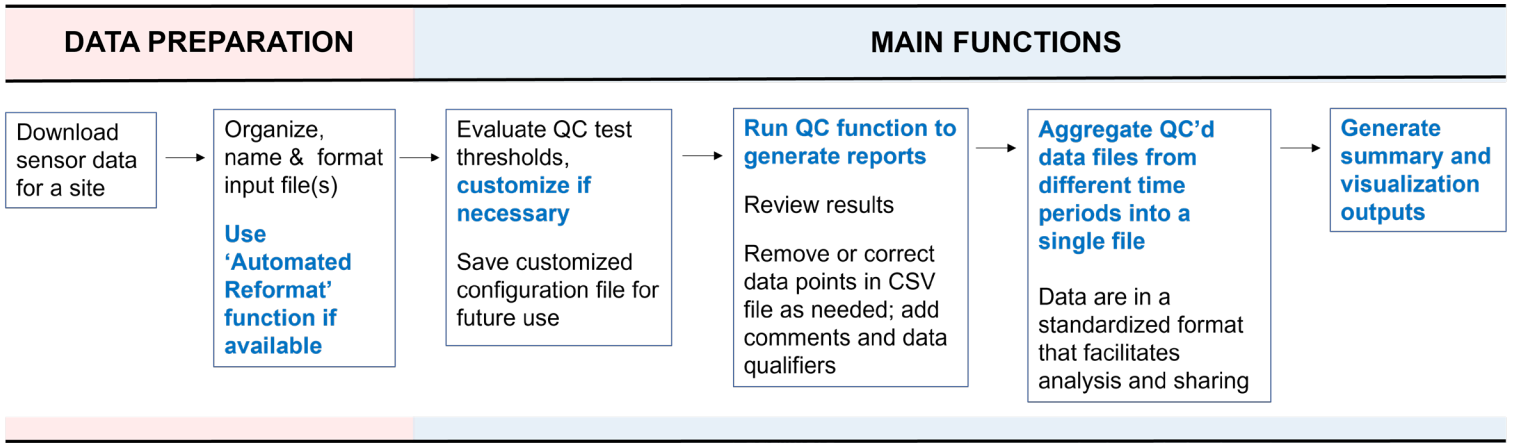


Lake RMN sites



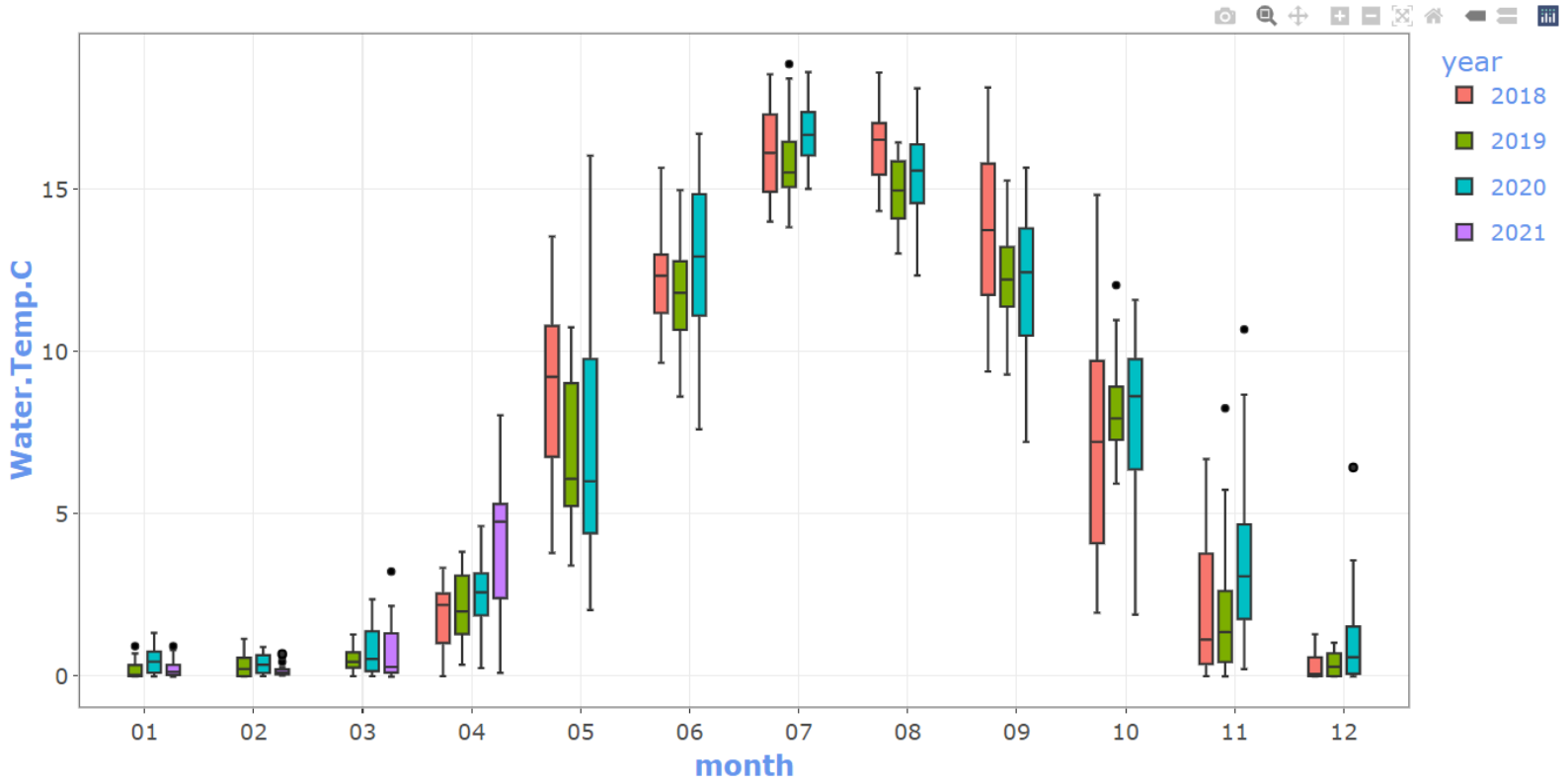
Background – ContDataQC

- R Package and Shiny App: Created by EPA ORD and Tetra Tech
- Generates QC reports to detect anomalies and erroneous data values
- Used for any sensors; customization is possible
- Parameters: Temperature, water level, discharge, conductivity, dissolved oxygen, pH, turbidity, chlorophyll-a, salinity
- R Shiny App: https://shiny.epa.gov/ContDataQC/
- GitHub: https://github.com/USEPA/ContDataQC



Background – ContDataSumViz

By 2021, there was a growing **need for a tool to help generate more summary stats and visualizations** for the RMN data (for **exploratory** purposes as well as for **reporting**).



ContDataSumViz

- **Purpose: summarize and visualize QC'd continuous sensor data**
  - Provides a streamlined, reproducible, simplified approach for summarizing and visualizing large and complex and temporal datasets
- Basic functions for all parameters
  - Summary statistics, time series plots, box plots, cumulative distribution functions, raster graphs
  - Correlations between different parameters
- Temperature functions
  - Thermal sensitivity, growing degree days, etc.
- Hydrology functions
  - Flashiness index, hydrologic alteration indicators, etc.

ContDataSumViz – Landing Page

Upload & Visualize Data

Tracks progress

Step 1: Upload file → Step 2: Select date and time → Step 3: Run meta summary → Step 4: Calculate daily statistics → Step 5: Visualize data

Upload Data | User Guide

Step 1: Upload File

Upload your data in .csv format

Browse | VT_Ranch_AW_20180412_20210427.csv

Upload complete

Display file information

- (•) First six rows
- ( ) Last six rows
- ( ) Column names

Display file contents

Step 2: Select Date and Time

Select date and time format

- (•) Date and time uploaded in one column
- ( ) Date and time uploaded in two separate columns

Select parameters to process

Water.Temp.C × Air.Temp.C ×

Date Field Name

Date.Time

Date Format

Year, Month, Day

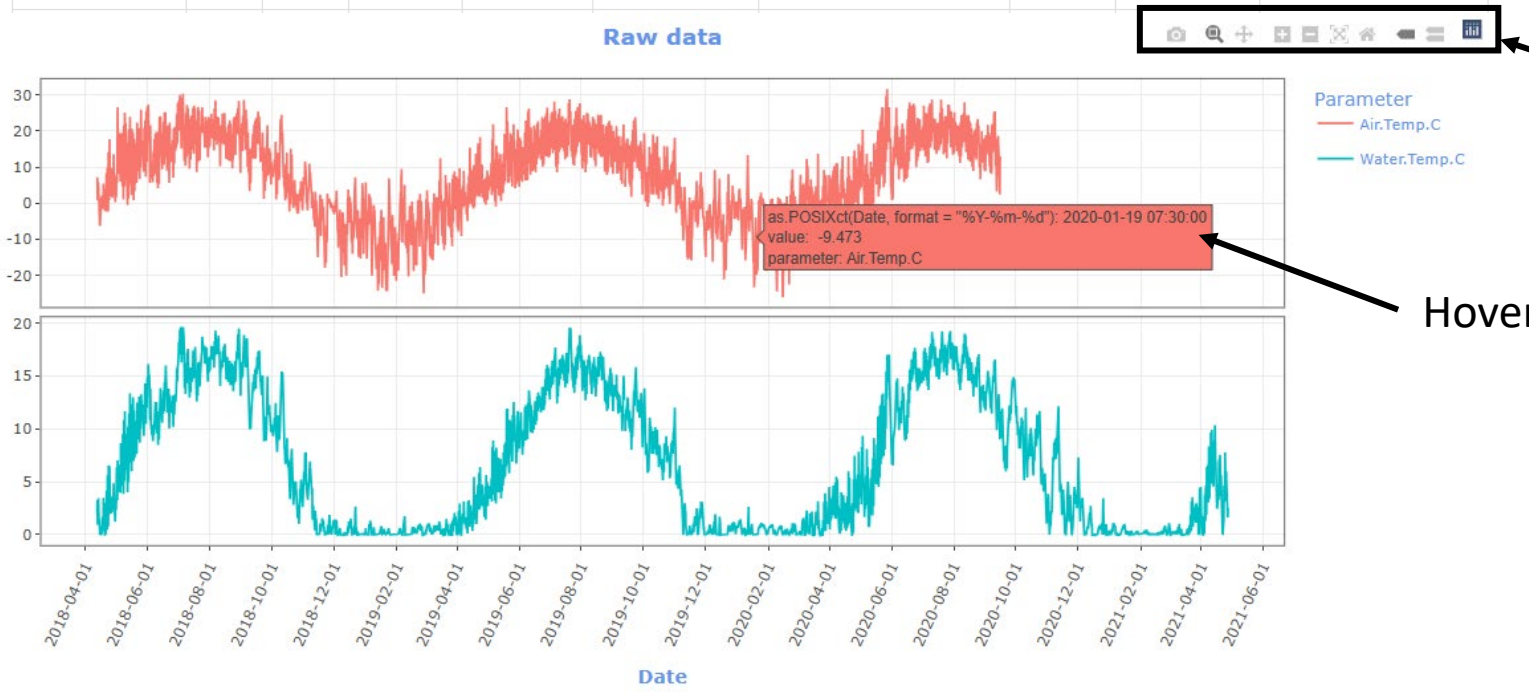
Time Format

Hour, Minute

Time Zone

UTC

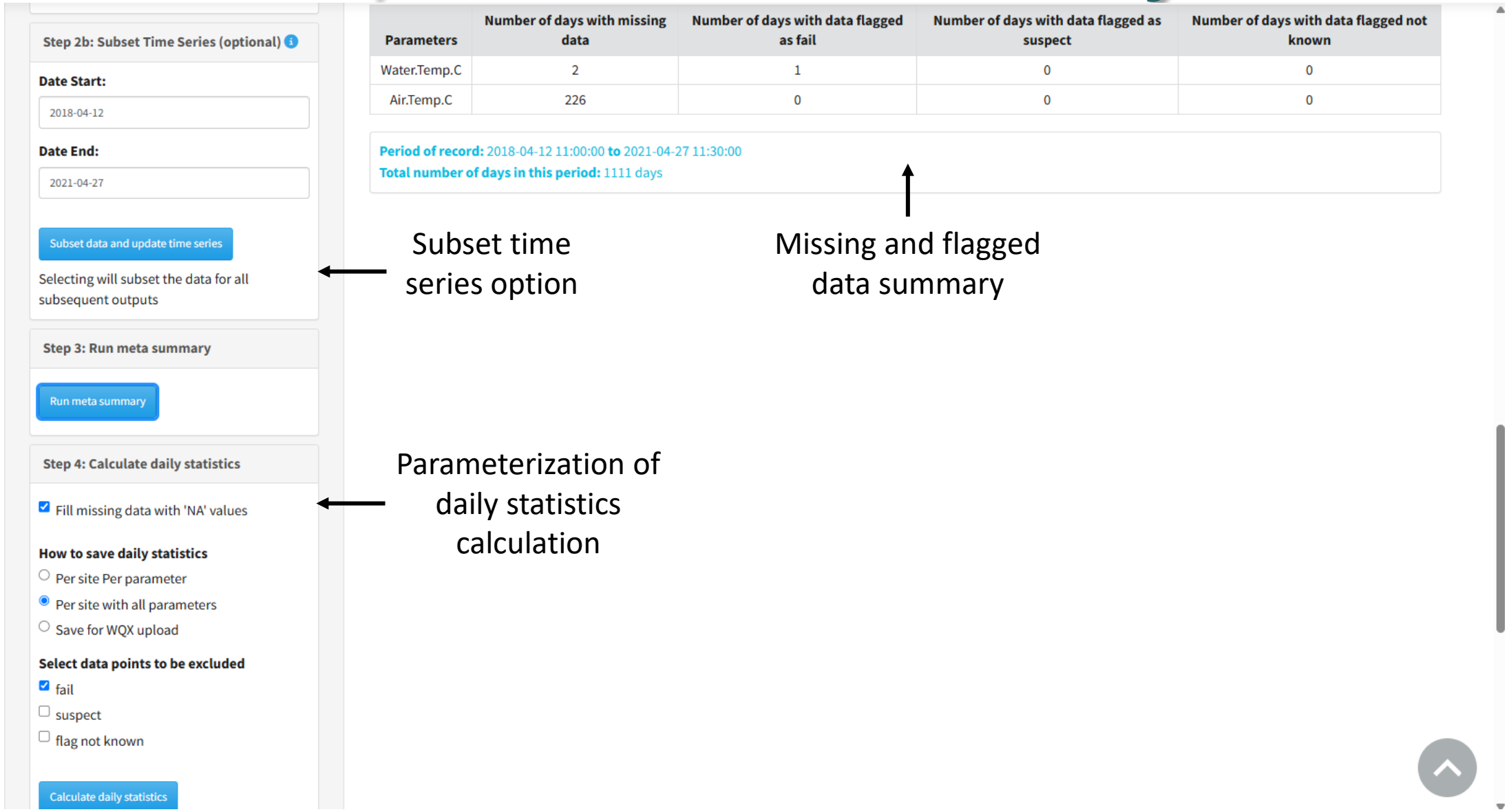
SiteID	Date	Time	Date.Time	Water.Temp.C	Flag.Water.Temp.C	Comment.MOD.Water.Temp.C	Air.Temp.C	Flag.Air.Temp.C	Comment.MOD.Air.Temp.C
Ranch493238200015	2018-04-12	11:00:00	2018-04-12 11:00	2.02	P		NA	X	
Ranch493238200015	2018-04-12	11:30:00	2018-04-12 11:30	2.21	P		NA	X	
Ranch493238200015	2018-04-12	12:00:00	2018-04-12 12:00	2.56	P		6.66	P	
Ranch493238200015	2018-04-12	12:30:00	2018-04-12 12:30	2.88	P		6.74	P	
Ranch493238200015	2018-04-12	13:00:00	2018-04-12 13:00	3.25	P		7.14	P	



Interactive plot controls

Hover text

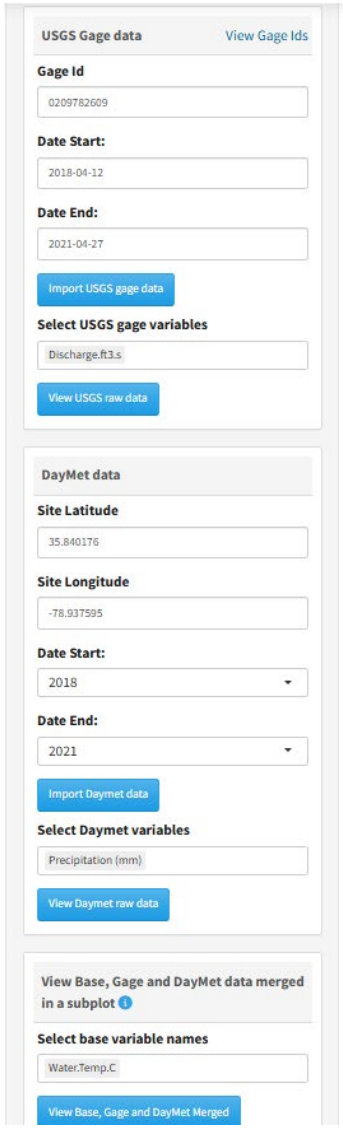
Summarize, Subset & Calculate Daily Statistics

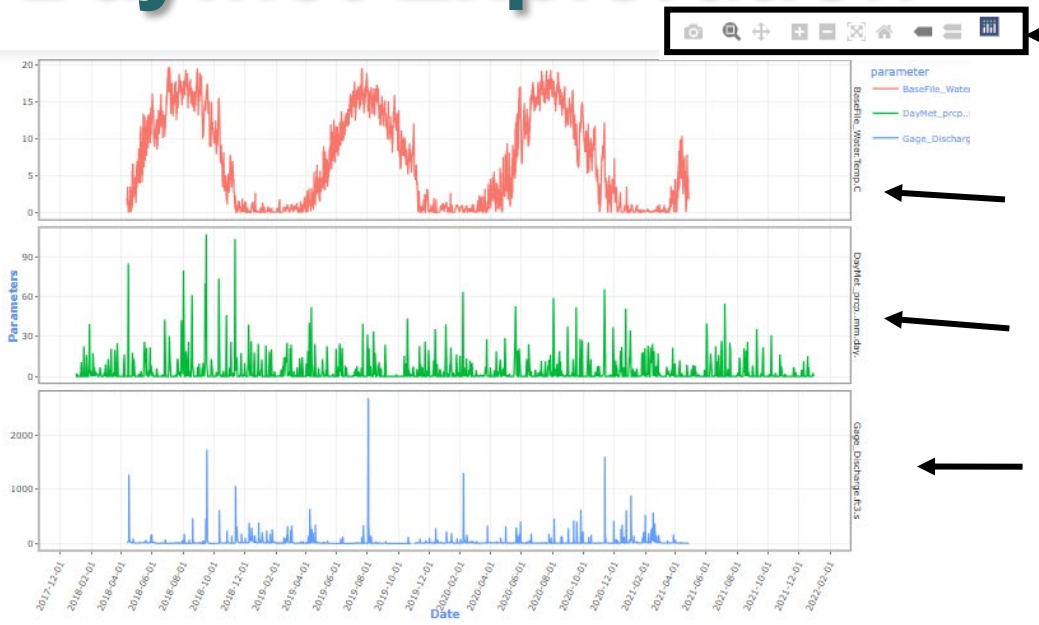


USGS & Daymet Exploration

Interactive plot controls

Options to select date range of data download





User Data

Daymet Climate Data

USGS Flow Data

Discrete Data Exploration

Upload Data | User Guide | USGS & Daymet Exploration | Discrete Data Exploration | Continuous Data Exploration

Upload discrete data in .csv format

Browse | CRBuoy discrete different

Upload complete

Select continuous parameters to process

temp.c ×

Select Date and Time for discrete data

Select date and time format

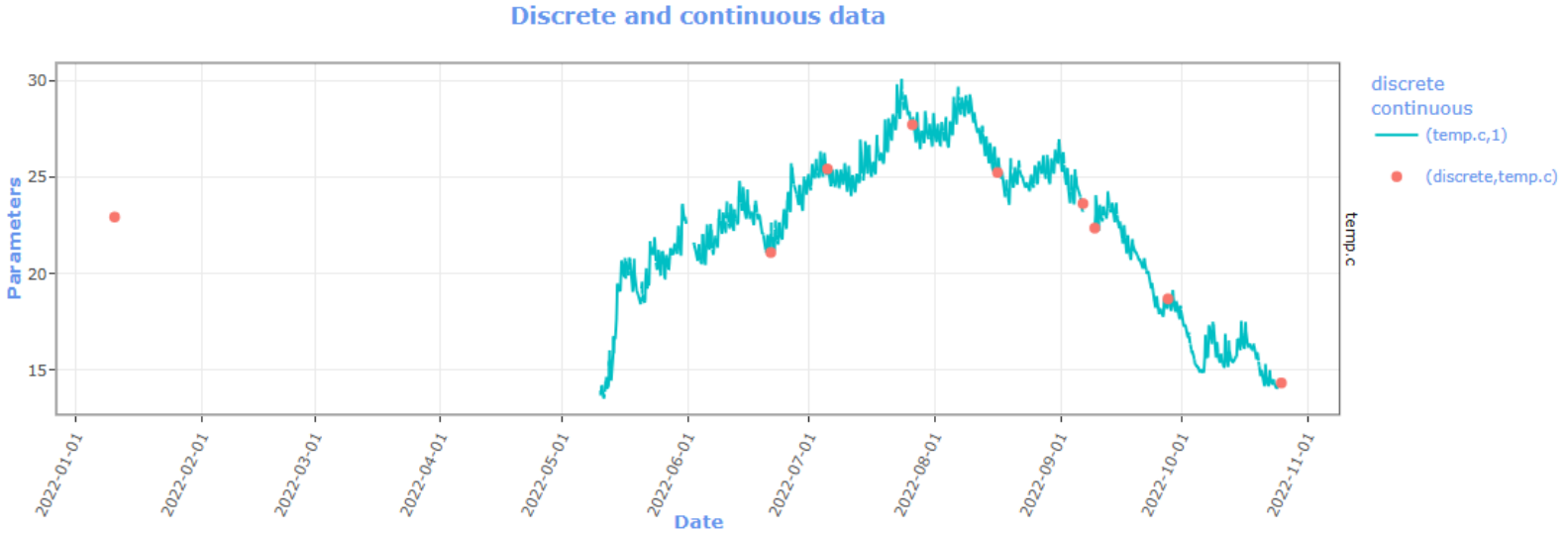
- (•) Date and time uploaded in one column
- ( ) Date and time uploaded in two separate columns

Select parameters to process

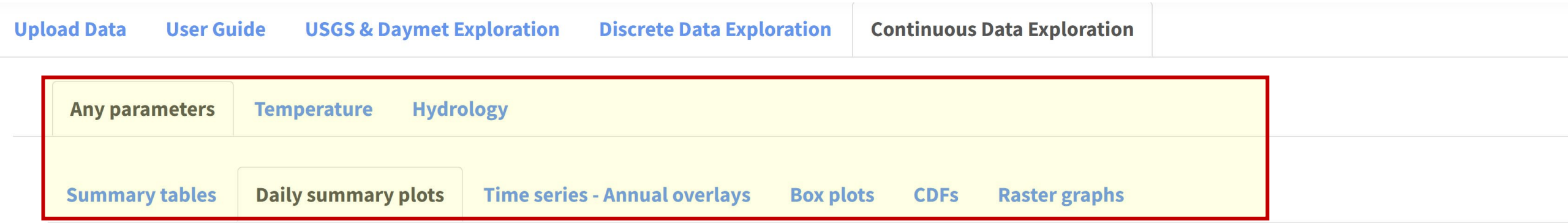
temp.c ×

Date Field Name

Site	Sonde	Date...Time.edt	temp.c	spcond..ms.cm.	ph	do..mg.l.	do....	turbidity..fnu.	chlorophyll..rfu.	phycocy
Charles	18	1/10/2022 11:00	22.94	1.14	6.63	7.89	92.50	2.52	5.50	2.
Charles	18	6/21/2022 11:16	21.11	1.54	7.80	7.85	89.10	2.32	1.92	0.
Charles	17	7/5/2022 10:47	25.43	1.53	7.75	8.78	107.50	1.06	2.45	0.
Charles	18	7/26/2022 10:40	27.73	2.43	7.86	6.93	88.60	-2.01	2.48	0.
Charles	14	8/16/2022 10:32	25.25	3.54	7.01	8.12	99.80	1.44	1.60	0.
Charles	18	9/6/2022 12:32	23.64	3.90	7.68	7.76	92.40	2.50	2.65	0.



Continuous Data Exploration – Any parameters



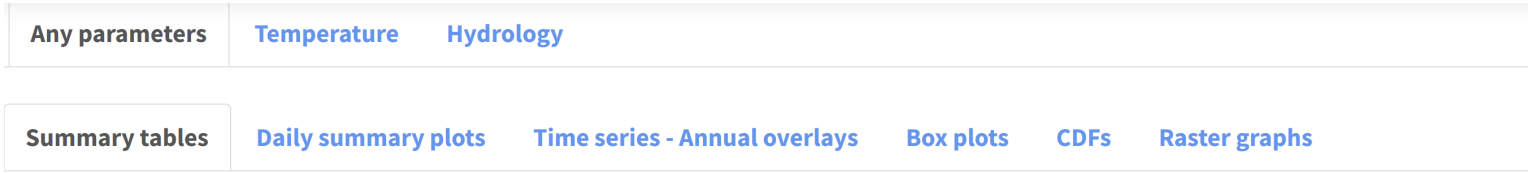
Functions available for any parameters

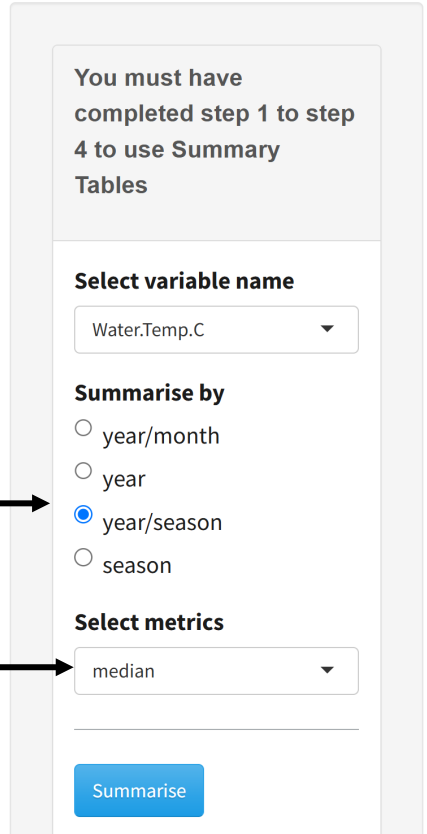
- Summary statistic tables
- Time series plots
- Time series – annual overlays
- Box plots
- CDFs (Cumulative Distribution Functions)
- Raster graphs

Choose desired statistic for plotting

- Mean
- Median
- Min
- Max
- Range
- Standard deviation
- And more…

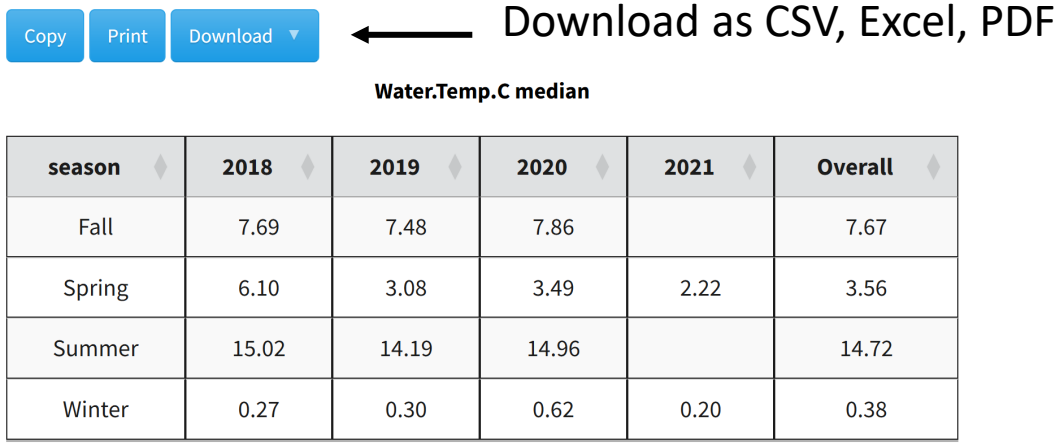
Any parameters – Summary tables





Select summary time period

Select summary metrics



Download as CSV, Excel, PDF

Water.Temp.C median

season	2018	2019	2020	2021	Overall
Fall	7.69	7.48	7.86		7.67
Spring	6.10	3.08	3.49	2.22	3.56
Summer	15.02	14.19	14.96		14.72
Winter	0.27	0.30	0.62	0.20	0.38

Any parameters – Daily summary plots

Upload Data | User Guide | USGS & Daymet Exploration | Discrete Data Exploration | **Continuous Data Exploration**

Any parameters | Temperature | Hydrology

Summary tables | **Daily summary plots** | Time series - Annual overlays | Box plots | CDFs | Raster graphs

Select variable name

Water.Temp.C

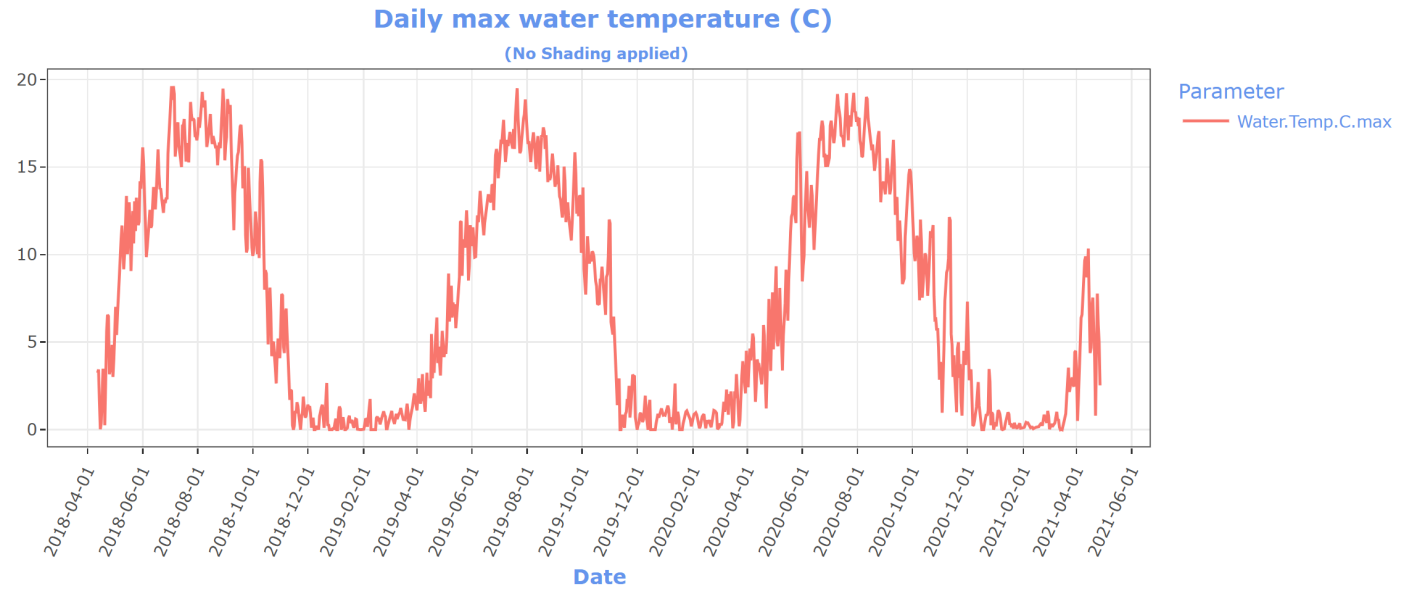
Select daily statistics metrics

max

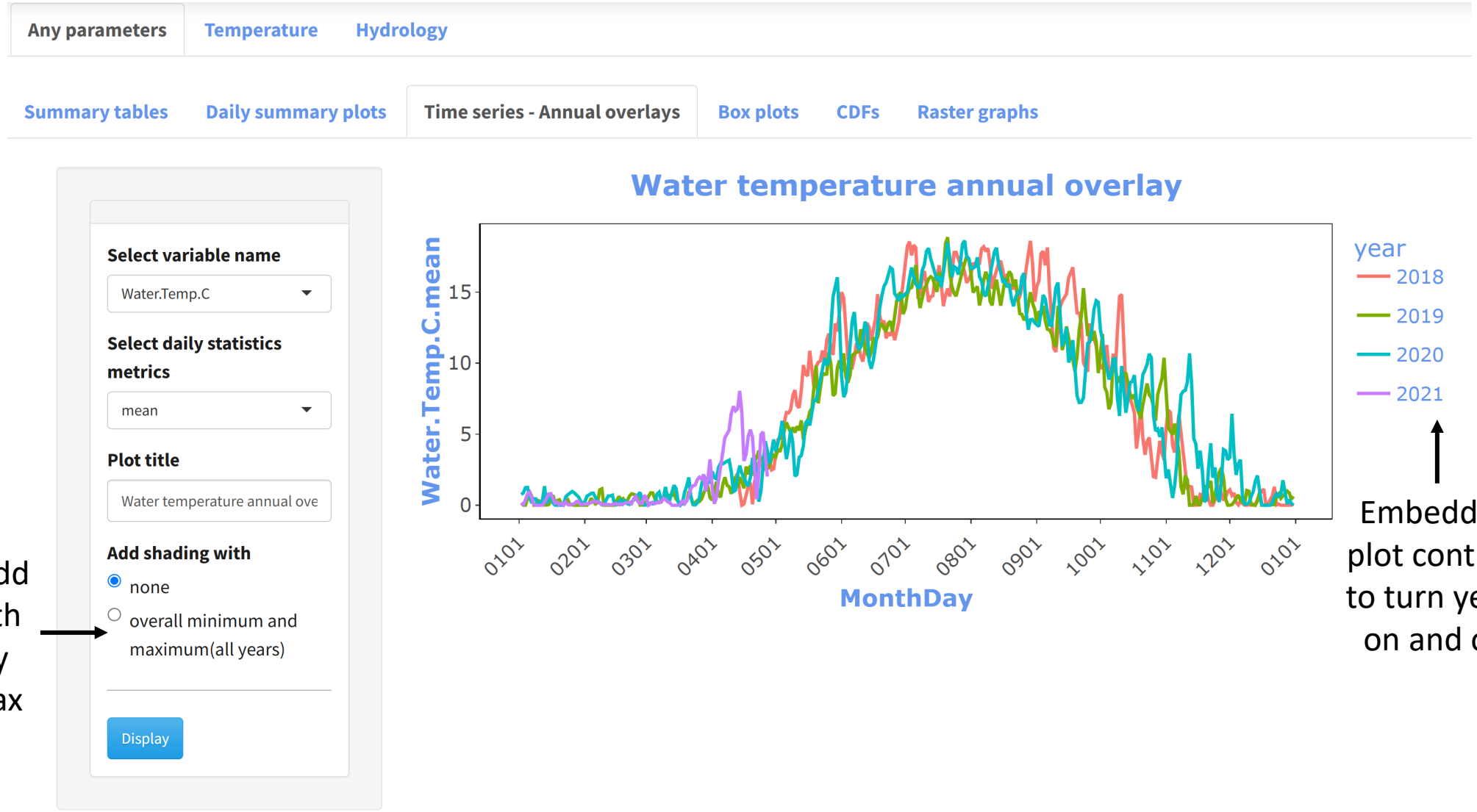
Plot title

Daily max water temperature (

Display



Any parameters – Time series annual overlays



Option to add shading with mean daily min and max across all years

Embedded plot controls to turn years on and off

Any parameters – Boxplots

Any parameters | Temperature | Hydrology

Summary tables | Daily summary plots | Time series - Annual overlays | Box plots | CDFs | Raster graphs

Select summary time period →

Select variable name

Water.Temp.C

Select daily statistics metrics

mean

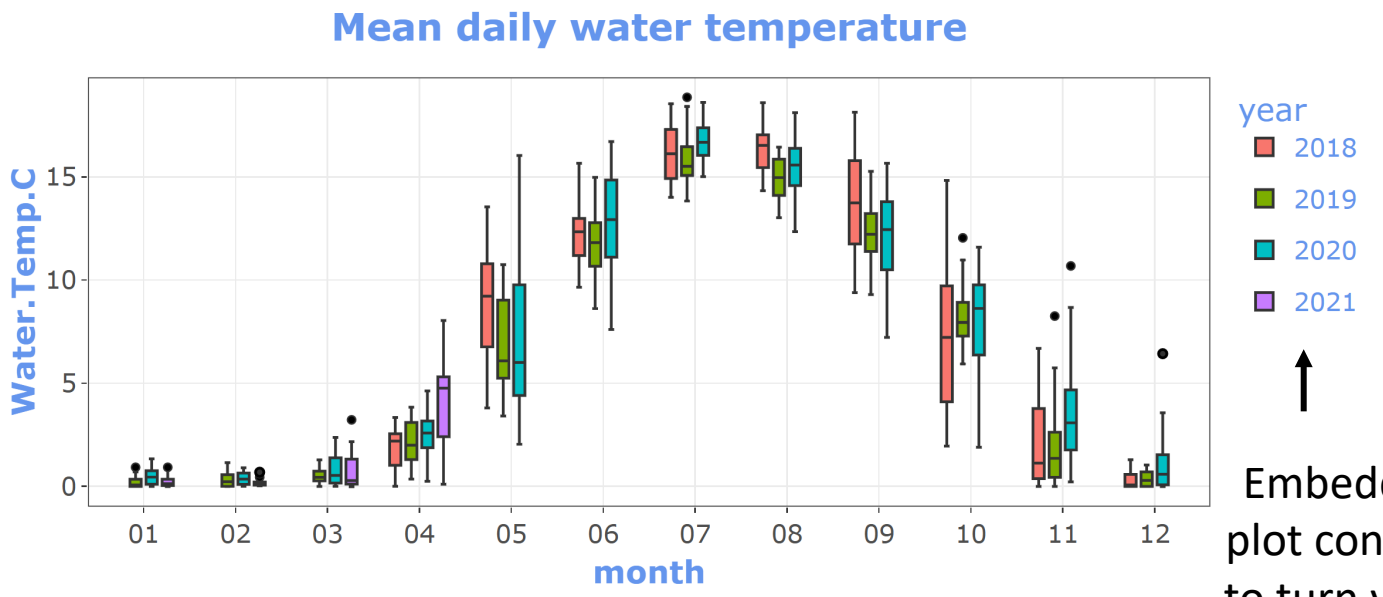
Group by

- ○ month
- ◉ month(years side by side)
- ○ year
- ○ season
- ○ season(years side by side)

Plot title

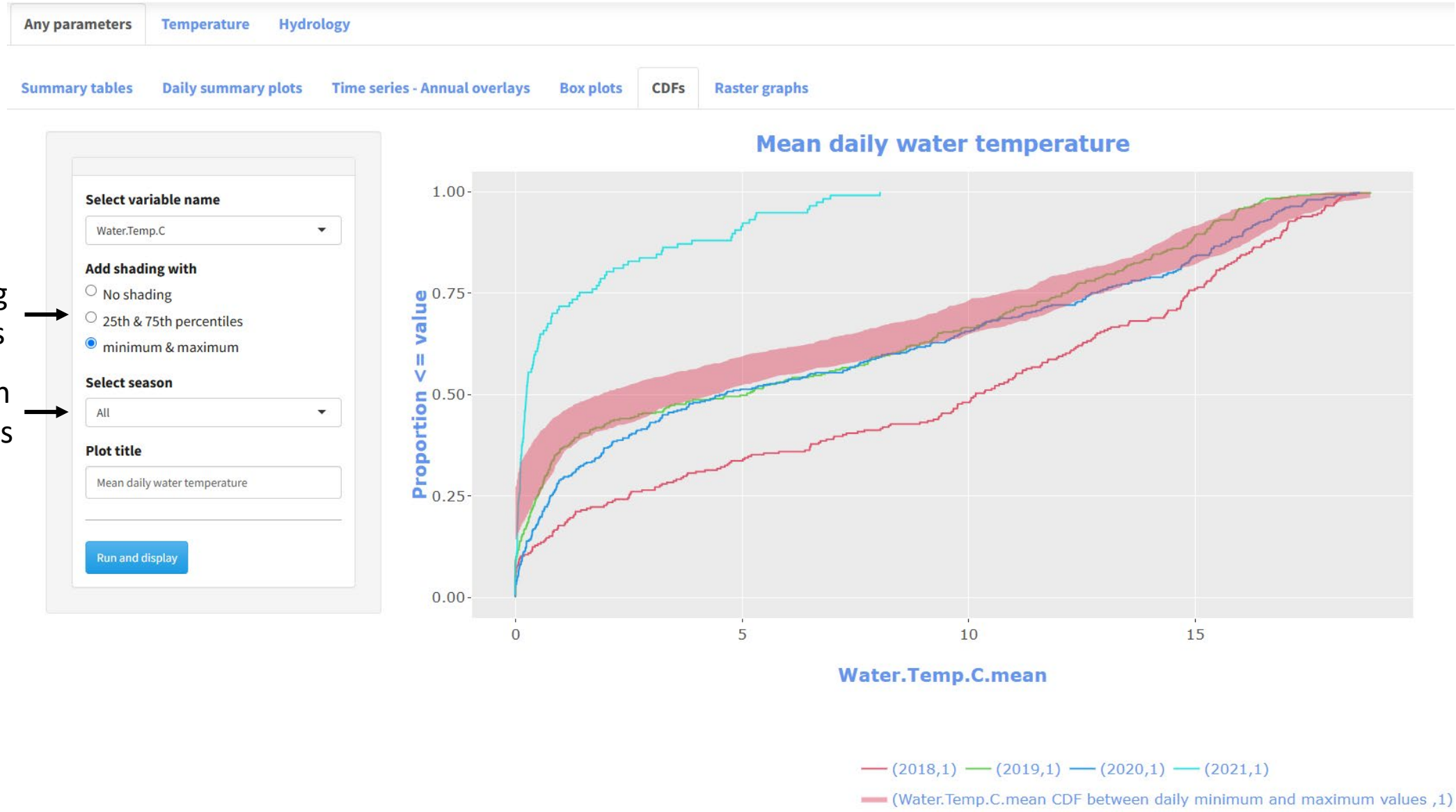
Mean daily water temperature

Display



Embedded plot controls to turn years on and off

Any parameters – CDFs



Shading options →

Season options →

Any parameters – Raster graphs

Any parameters | Temperature | Hydrology

Summary tables | Daily summary plots | Time series - Annual overlays | Box plots | CDFs | Raster graphs

Select variable name

Water.Temp.C

Select daily statistics metrics

max

Plot title

Water temperature max

Adjust plot aspect ratio

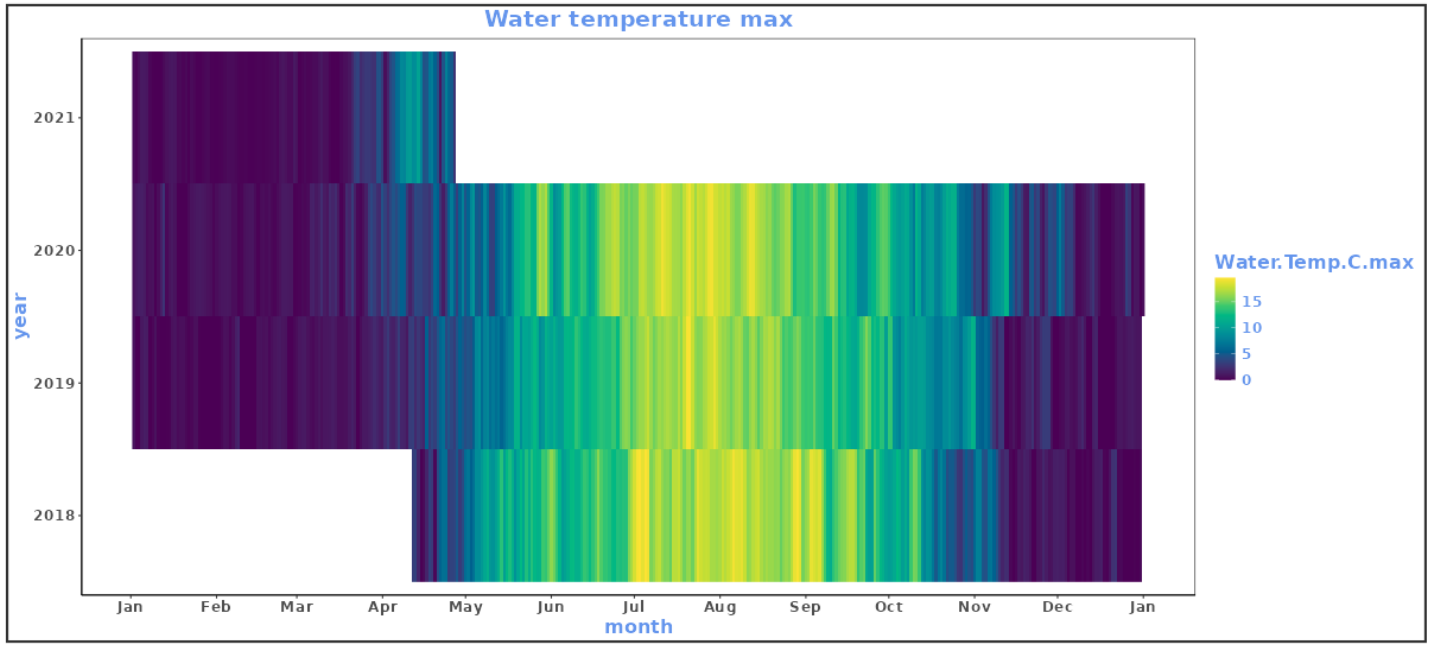
0.5

Color palette options

- (•) hcl
- ( ) rainbow
- ( ) heat
- ( ) terrain
- ( ) topo

Display

To download the plot, mouse over the plot, right click, and select Save image as.



ContDataSumViz: Temperature & Hydrology

Temperature

- Output from the StreamThermal R package (over 200 metrics) (https://github.com/tsangyp/StreamThermal; Tsang et al. 2016)
- Thermal sensitivity (plots air vs. water temperature; calculates correlation coefficient)
- Thermal class based on mean July/August temperature
- Growing degree days (heat accumulation over a threshold temperature)
- Temperature not to exceed (e.g., 4T3 - temperature not to be exceeded for ≥ 4 hours in a 24-hour period on > 3 days)

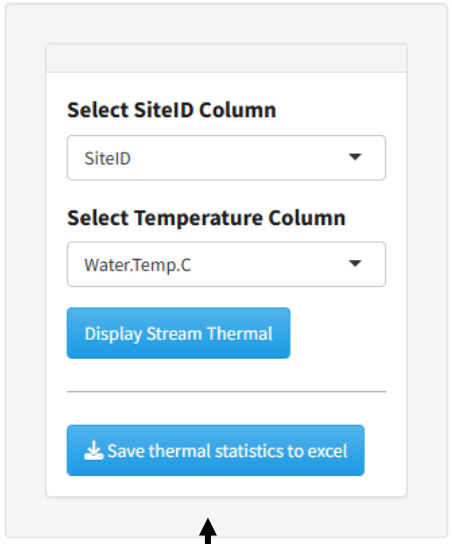
Hydrology

- Output from Indicators of Hydrologic Alteration (IHA) R package (ecologically relevant flow regime characteristics)
- Flashiness index (Richard-Baker Index – based on change in mean daily discharge)

Temperature – Thermal statistics

Any parameters | Temperature | Hydrology

Thermal statistics | Air vs Water | Growing degree days | Thermal classification | Temperature not to exceed



Save to single PDF with metadata

Copy Print Download

Download individual tables as CSV, Excel, PDF

Frequency

	SiteInfo	FmaxcT1	FmaxcT2	FmaxcT3	FmaxcT4	FmaxcT5	FmaxcT6	FmaxcT7	FmaxcT8
1	Ranch493238200015	0	0	0	0	15.67	57.33	844.33	608.67

Copy Print Download

Magnitude

	SiteInfo	ADmax1	ADmax2	ADmax3	ADmax4	ADmax5	ADmax6	ADmax7	ADmax8	A
1	Ranch493238200015	0.29	0.29	0.68	2.64	7.66	12.27	16.26	15.58	

Copy Print Download

Rate of Change

	SiteInfo	RC1	RC2	RC3	RC4	RC5	RC6	RC7	RC8	RC9	RC10	RC11	RC12
1	Ranch493238200015	0.39	0.35	0.41	0.75	0.54	0.44	0.42	1.48	1.14	0.57	1.29	0.81

Temperature – Air vs Water

Any parameters | Temperature | Hydrology

Thermal statistics | Air vs Water | Growing degree days | Thermal classification | Temperature not to exceed

Select Air Temperature Column

Air.Temp.C

Select Water Temperature Column

Water.Temp.C

Limit the data points with air temperature

- No
- Yes

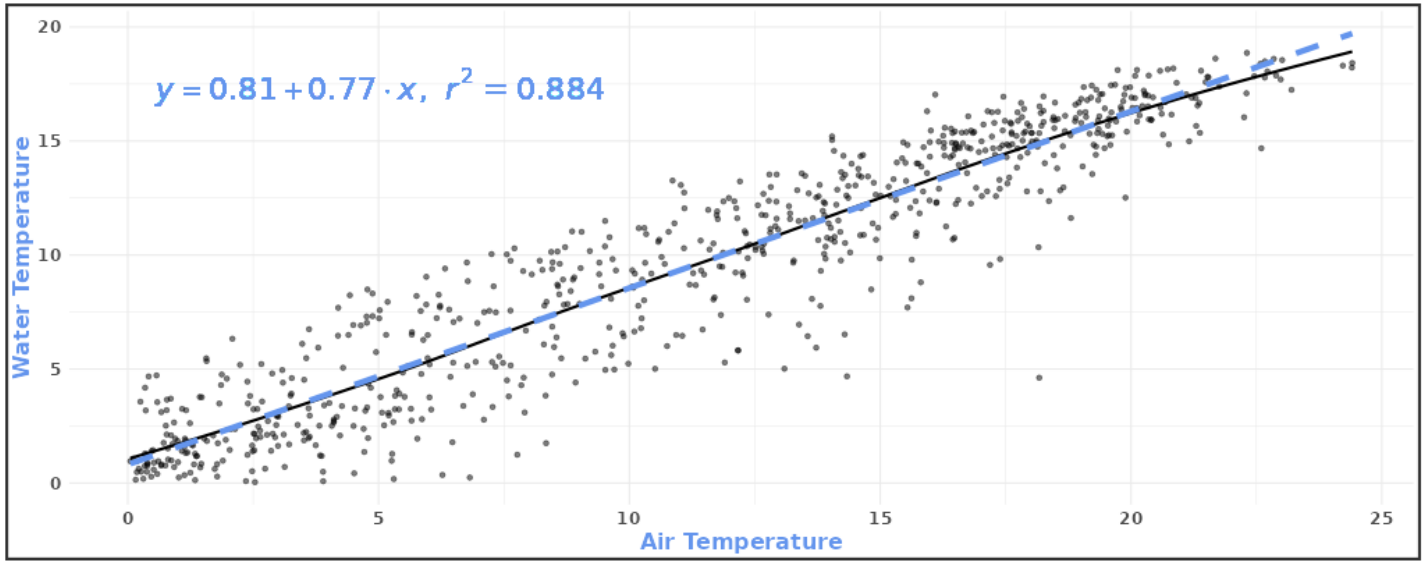
air temperature less than this value will be excluded

0

Display thermal sensitivity

To download the plot, mouse over the plot, right click, and select Save image as.

Option to exclude points < threshold



Temperature – Not to Exceed

Thermal statistics | Air vs Water | Growing degree days | Thermal classification | **Temperature not to exceed**

Select Temperature Column

SiteID

Select number of hours for the time window calculation

4

Select number of days for the time window calculation

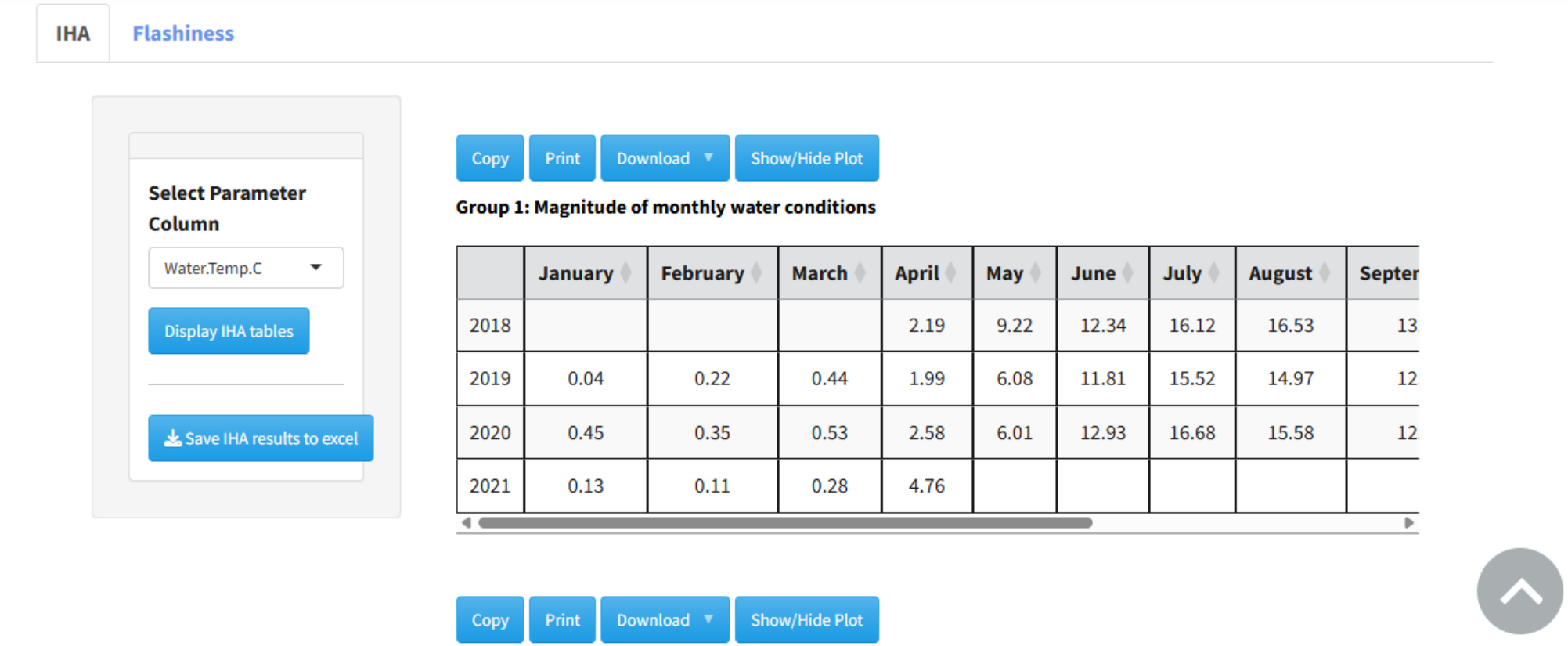
3

Copy | Print | Download

Year	4T3	Number of days
2018	Ranch493238200015	264
2019	Ranch493238200015	365
2020	Ranch493238200015	366
2021	Ranch493238200015	117

Hydrology – IHA

Indicators of Hydrologic Alteration



Hydrology – Flashiness

Any parameters | Temperature | **Hydrology**

IHA | **Flashiness**

Select Variable

Water.Temp.C

Display RB Index

Copy | Print | Download

Year	RB_Index
2018	0.088
2019	0.093
2020	0.11
2021	0.332

Status of ContDataSumViz

- User testing and updates to application are complete.
- App is currently going through EPA review and clearance.
- App will be published with a journal article later this year.

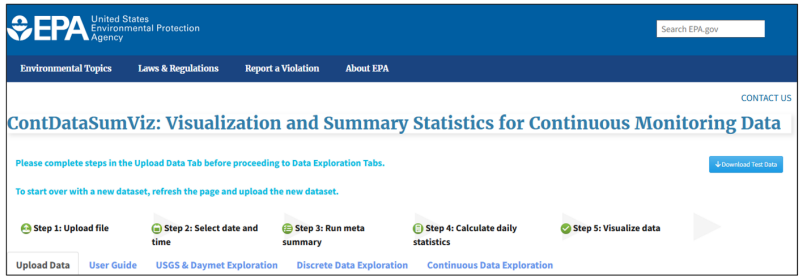
Acknowledgments

- Helped in App Design
  - Tom Faber (EPA Region 1, Retired) & Leah Ettema (EPA Region 3)
- Helped in App Development
  - Yadong Xu, EPA Environmental Monitoring and Visualization Lab (EMVL)
  - Nilima Gandhi & Gopakumar Nair, Science Applications International Corporation (SAIC)
- App Reviewers
  - User testers: Erik Heitshusen (EPA Region 3), Meredith Zeigler (New Mexico Environment Department), Nicole Sadecky (EPA Region 3), Joel Owen (EPA Region 4)
  - EPA reviewers of app and draft publication: Leah Ettema (Region 3), Tori McLeod (Region 1), Tom Johnson (ORD)
- Funding
  - EPA EMVL award
  - EPA innovation grant (Lead: Tom Faber, EPA Region 1)
- RMN partners who provided data and helped with testing and development

Questions? Feedback?

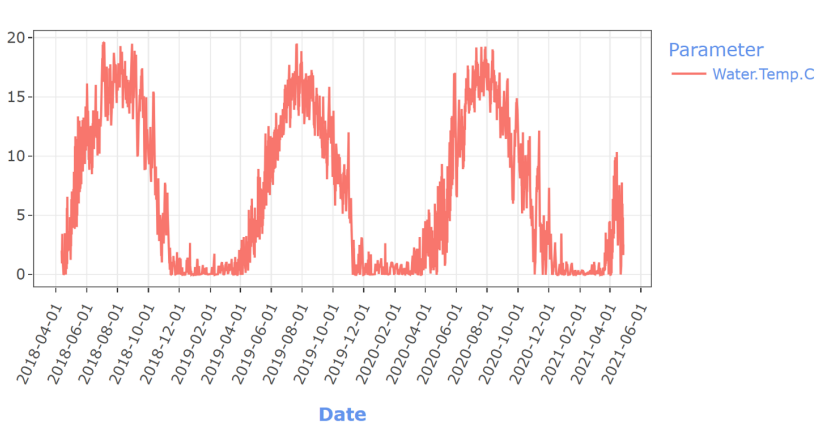
How can we improve the tools?
What additional features would you like to see?

pennino.michael@epa.gov
naslund.laura@epa.gov



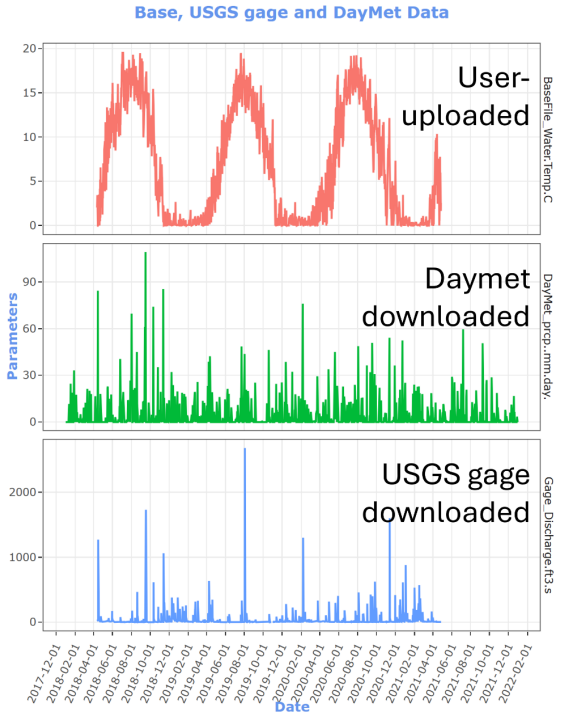
Upload & format Data

- Upload file
- Format date & time
- Display raw time series

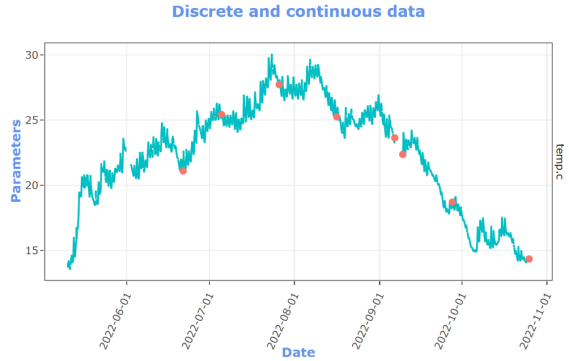


- Subset time series (optional)
- Create metadata table
- Calculate daily statistics
- Save daily statistics

Download & view context data



Overlay discrete data



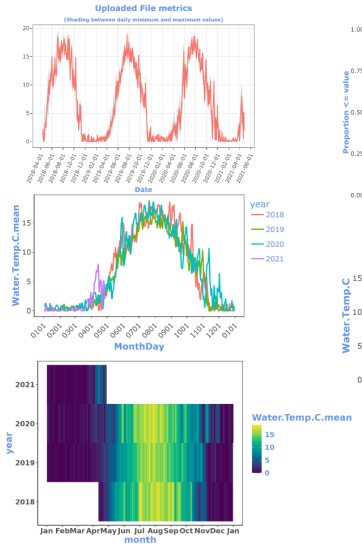
Summarize & visualize continuous data

Any parameters – daily summaries

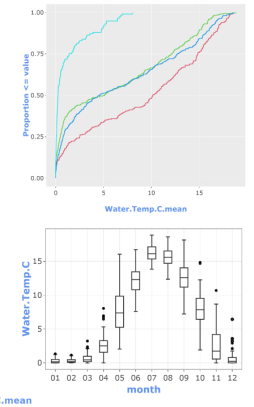
Summary table

- Time frame (year, year/month, season, year/season)
- Metrics (e.g., mean, median, min, max, sd)

Temporal patterns

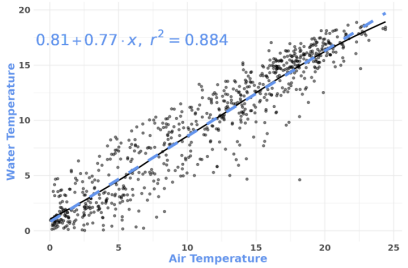


Distributions



Temperature metrics

- StreamThermal statistics
- Growing degree days
- Thermal classification
- Temperature not to exceed



Hydrology metrics

- Indicators of Hydrologic Alteration (IHA) statistics
- Richards-Baker flashiness index