ContDataSumViz User Guide

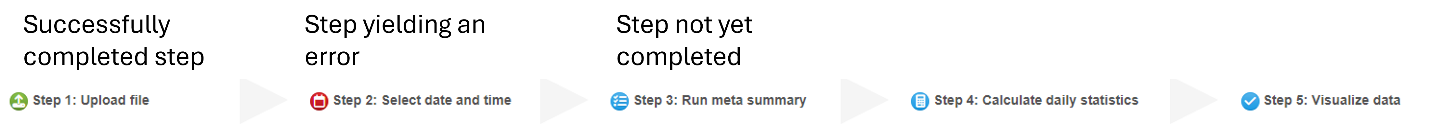
Data Visualization & Reporting for Continuous Monitoring Data

# Layout

This tool contains data summary and visualization modules for continuous monitoring data. Tool functionalities appear as the user provides required inputs. The landing tab walks the user through the process of uploading their data. Test data is available from the download button in the upper righthand corner of the application.

Placeholder: top of the page with Download test data button

Upload data progress is represented by the icons labeled with the upload data steps at the top of the page. Icons for steps that have not yet been completed are blue, successfully completed steps are green, and steps yielding errors are red.



Once the five steps of the upload data tab are successfully completed, additional tabs will display at the top of the page. The USGS & Daymet Exploration tab allows the user to download USGS gage data at their gage of choice and Daymet daily weather and climatology data at their coordinates of choice for the same time period as their uploaded data. The Discrete Data Exploration tab allows users to overlay discrete data points on a continuous time series of the same parameter. The Continuous Exploration tab provides daily summary statistics and visualizations for any uploaded parameters as well as specific statistics for temperature and hydrology parameters.

## Start over with a new dataset

To analyze a new dataset, refresh the page to reset the inputs and upload the new input data file in step 1.

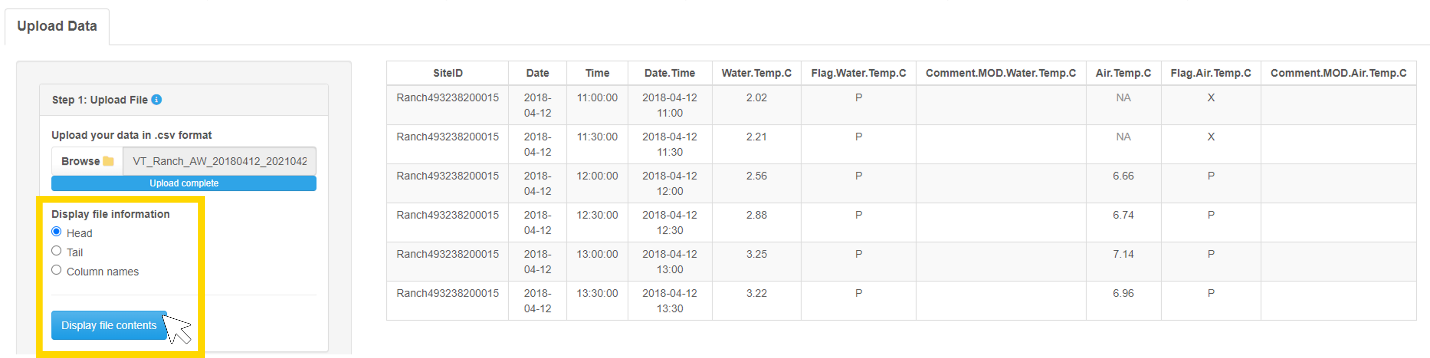
# Upload Data

## Step 1: Upload file

Ensure that your dataset is saved as a .csv file on your computer. To upload the dataset, select the browse button in the lefthand sidebar. Navigate to your file in the popup file explorer window and select open.

Placeholder

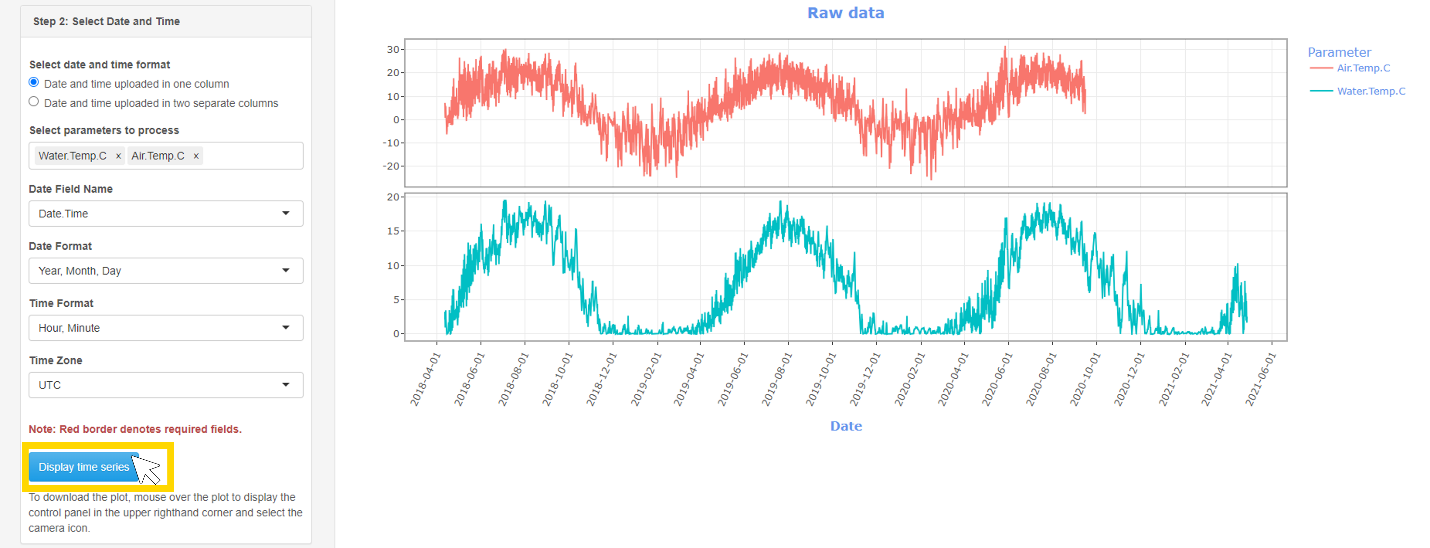
Select the file information you would like to display—head (the first six rows), tail (the last six rows), or column names—then select display file contents. The date format displayed may be different from the display in Excel, because Excel often adds date formatting without alerting the user. If you open the file in a simple text editor like Notepad, the date format should match the application display.



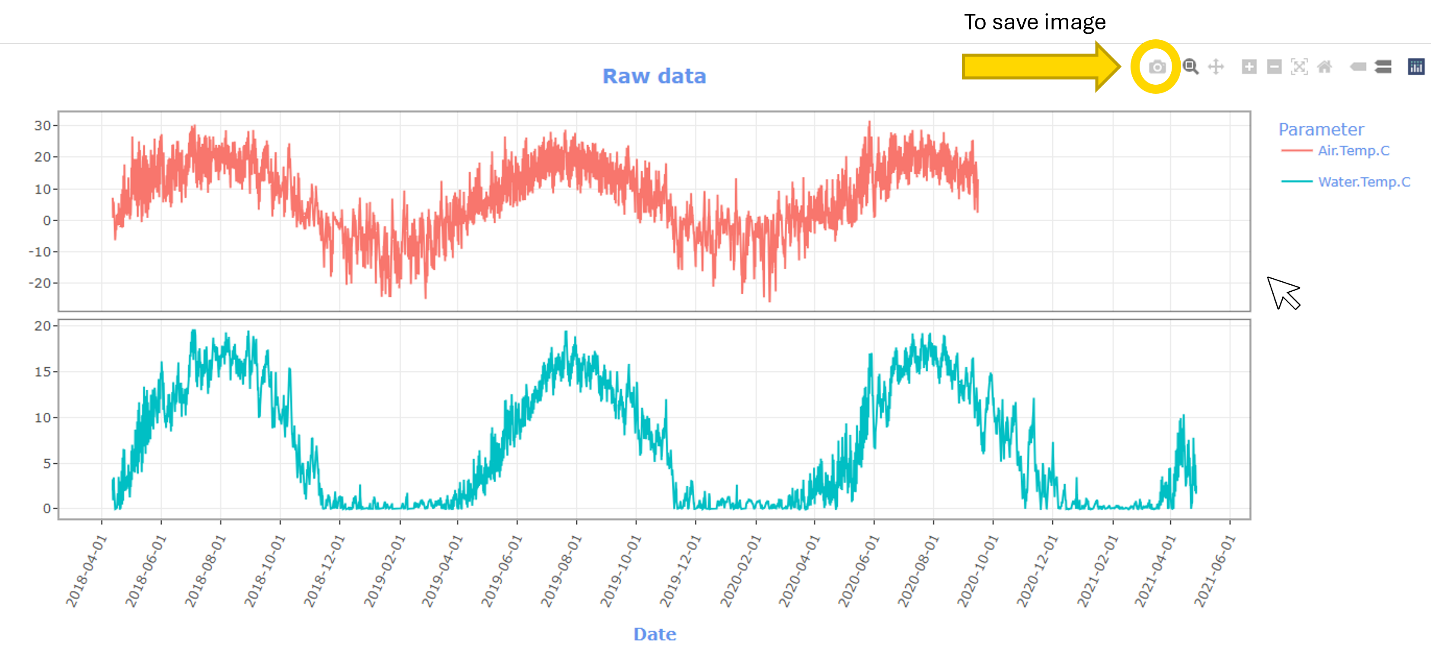
## Step 2: Select date and time

Selecting the display file contents button in step 1 will display step 2 in the lefthand sidebar. Select whether date and time are contained in a single column or two columns in the uploaded data file. Select the parameters to process, date/time field name, date format, time format, and time zone using the dropdowns below. For more information about the time zone options, including their deviation from UTC in standard time and daylight savings time visit: <https://en.wikipedia.org/wiki/List_of_tz_database_time_zones>

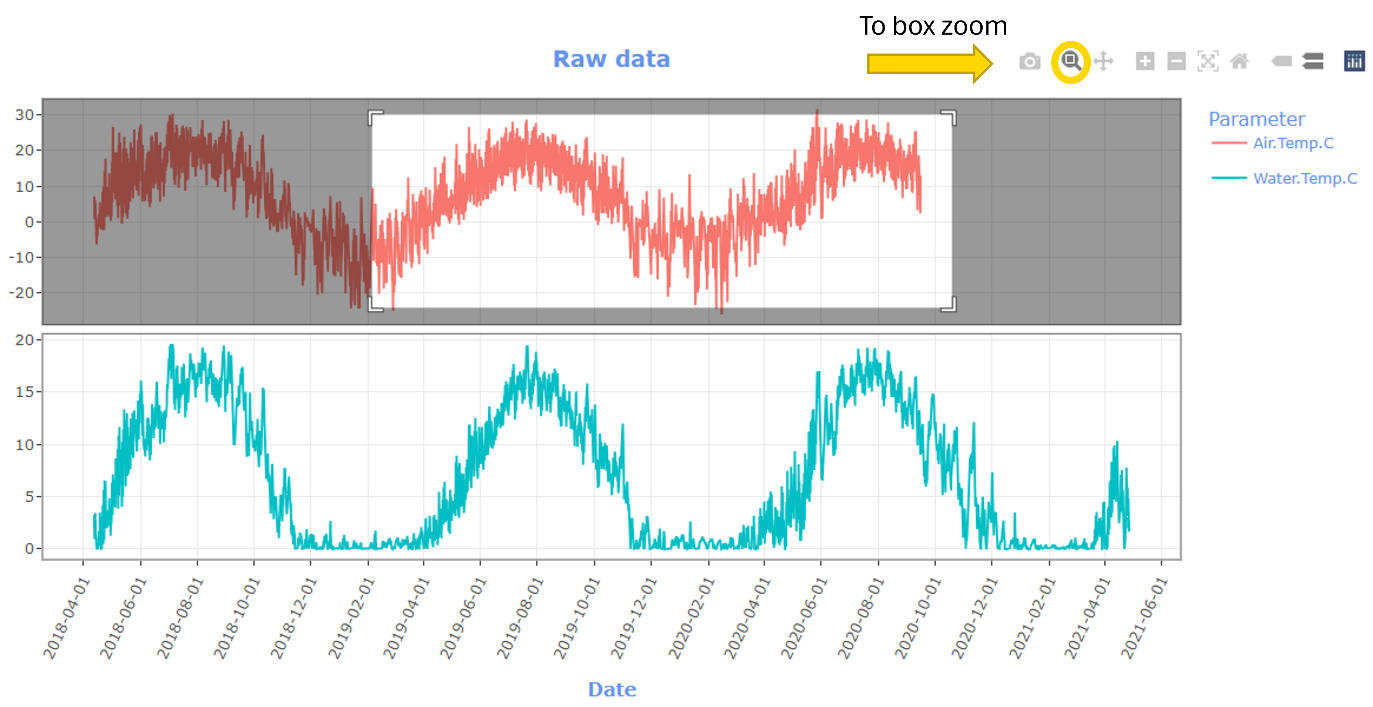
If the date and time format parameters match the uploaded file, the time series will display after clicking the display time series button. If there is a mismatch, an error will return upon selecting the display time series button. Correct the error(s) in the parameter selections, then click Display time series, which will display the plot and step 3.



Mousing over the time series will display the plotted values. To save the plot as an image, mouse over the plot to display the control panel in the upper righthand corner of the plot. Select the camera icon to open the file explorer window, then navigate to the location where you would like to save the plot and select save.

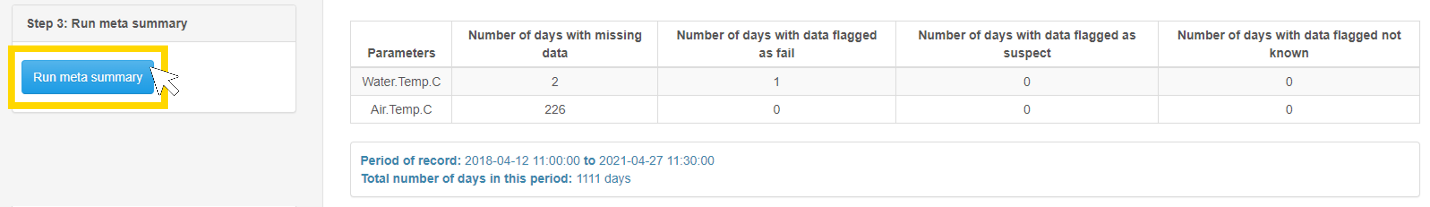


Zooming is available through the plus and minus buttons in the plot control panel, or by selecting the magnifying glass icon, which opens a box zoom that zooms to a user-drawn window. To reset the zoom, either double left click with the magnifying glass selected or select the home icon in the control panel.



## Step 3: Run meta summary

Select the run meta summary button to display a summary of missing and flagged data for each parameter, the period of record, and number of days in the period.



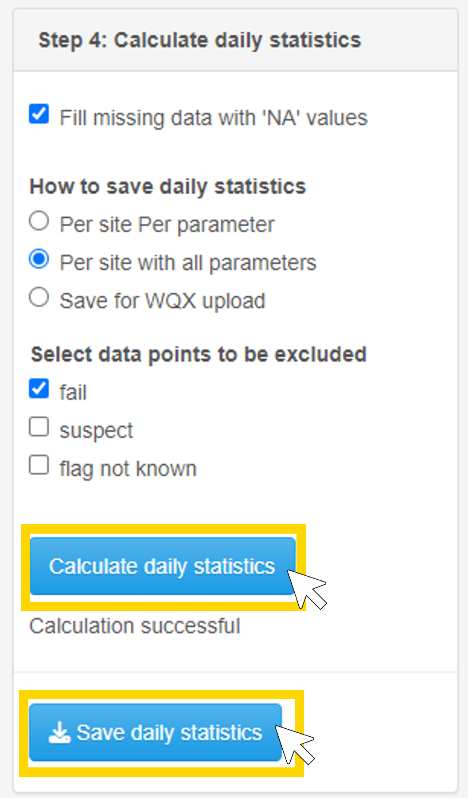
## Step 4: Calculate daily statistics

Select the options for calculating daily statistics in the side bar. To save different daily statistics summaries, select an option for How to save daily statistics, select Calculate daily statistics and then select Save daily statistics. After selecting new daily statistics options, you must select Calculate daily statistics before saving them.

Per site Per parameter: generates a zipped folder with different files for each parameter

Per site with all parameters: generates a single file with summaries for all parameters

Save for WQX upload: generates daily summaries in Water Quality eXchange (WQX) format



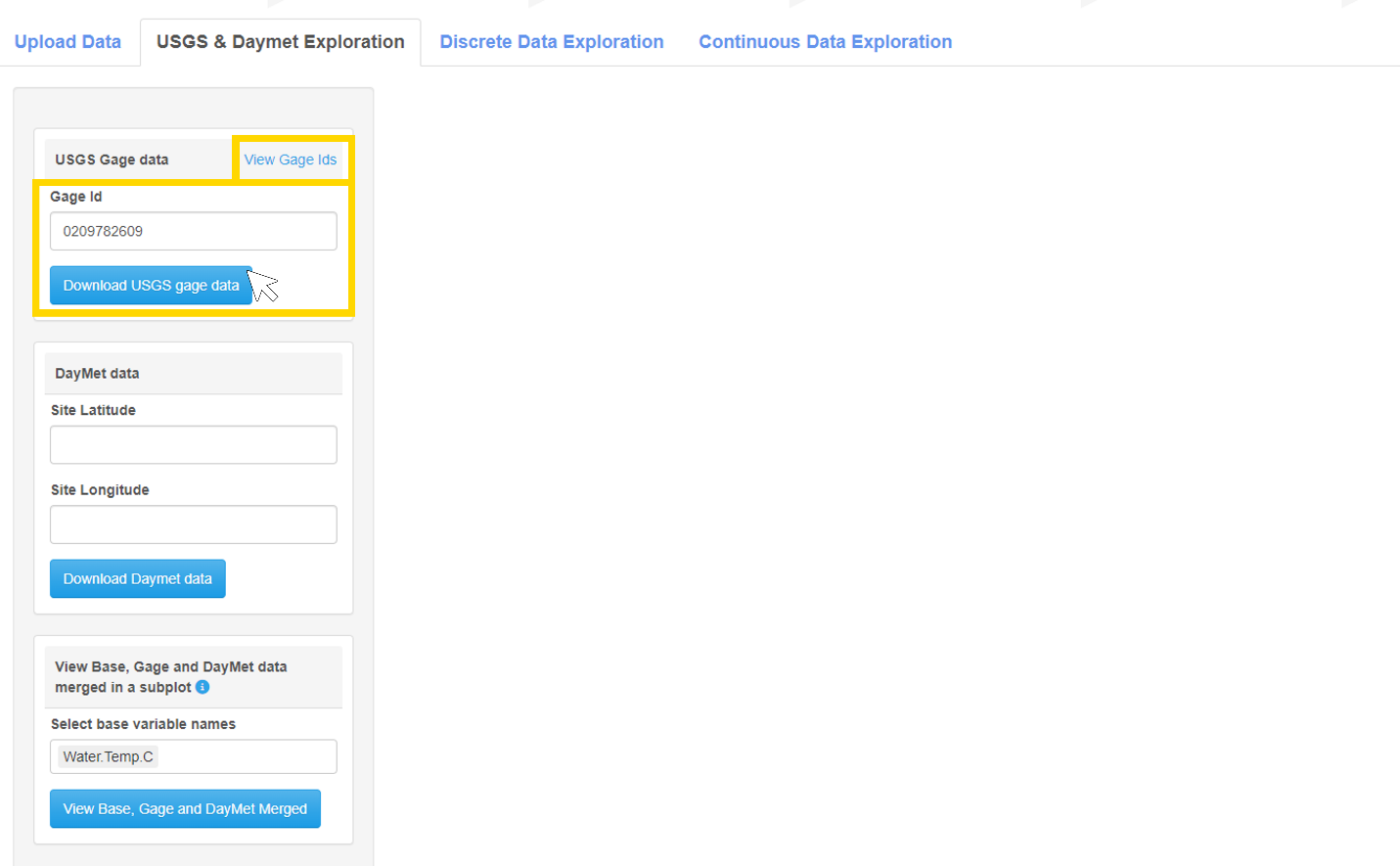
## Step 5: Visualize data

After successful calculation of daily statistics in step 4, step 5 will appear on the upload tab instructing the user to proceed to the other tabs at the top of the page to visualize their data.

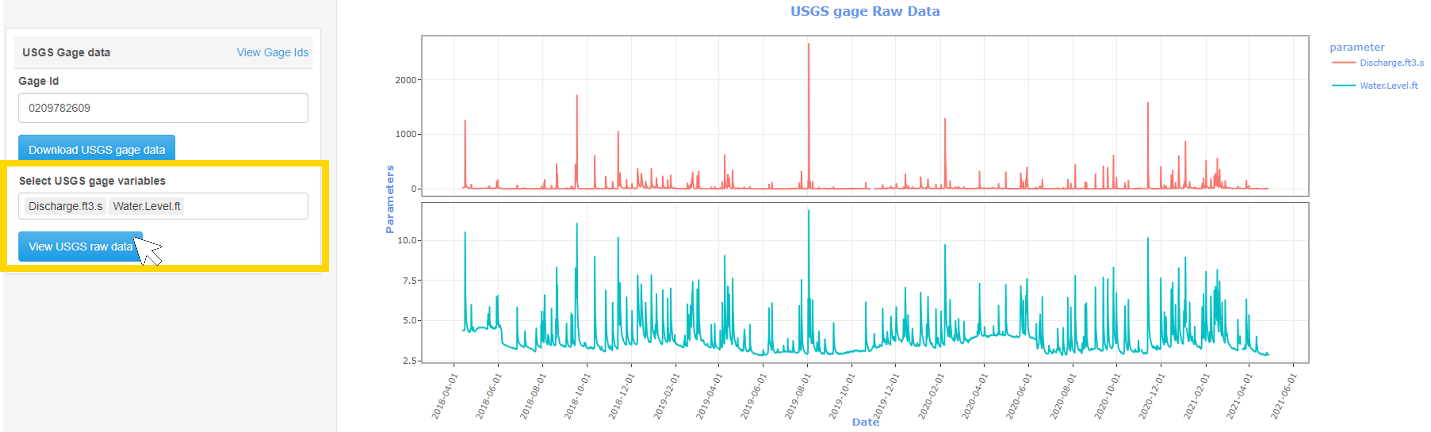
# USGS & Daymet Exploration

## USGS gage data

Input the id of the desired USGS gage in the gage id box and select Download USGS gage data. To find a gage id, select the link in the upper right corner of the sidebar to visit the National Water Information System Web Interface. USGS gage ids are also available from the [USGS National Water Dashboard](https://dashboard.waterdata.usgs.gov/app/nwd/en/). The larger the requested data range, the longer it will take to download the data. If you receive an error message proceed to the USGS gage data error messages section of the user guide for additional guidance. When the USGS data have been successfully downloaded, the Select USGS gage variables box will appear. Select the desired variables from the dropdown and click View raw USGS data to generate a timeseries.



The plot controls to zoom, pan, and save the USGS gage data plot are the same as for the raw data time series in the Upload Data tab.



#### USGS gage data error messages

* “missing value where TRUE/FALSE needed”: this error generally displays when an invalid gage id is provided. Check that the id you are inputting is an id in the National Water Information System (NWIS) database.
* “arguments implying different number of rows: 1, 0”: this error can be displayed when the input date range is entirely outside of the available range for the gage. Check the site page for your selected gage to confirm that the available date range does not overlap that of the uploaded data. Select a new gage that has an overlapping time period with the uploaded data.
* Other errors: If the USGS web service times out, it is possible to see other errors. This may happen because too much data is being requested or the service itself is down. Try reducing the date range requested from the service. If that does not resolve the error, try downloading the data again in a few hours or days to rule out web service down time. See [USGS documentation](https://waterservices.usgs.gov/docs/instantaneous-values/instantaneous-values-details/) for more information on the USGS gage web service. If none of these solutions work, please submit an issue on the ContDataSumViz repository [link] with details about how this module failed (gage id requested and date range of the uploaded data).

## DayMet data

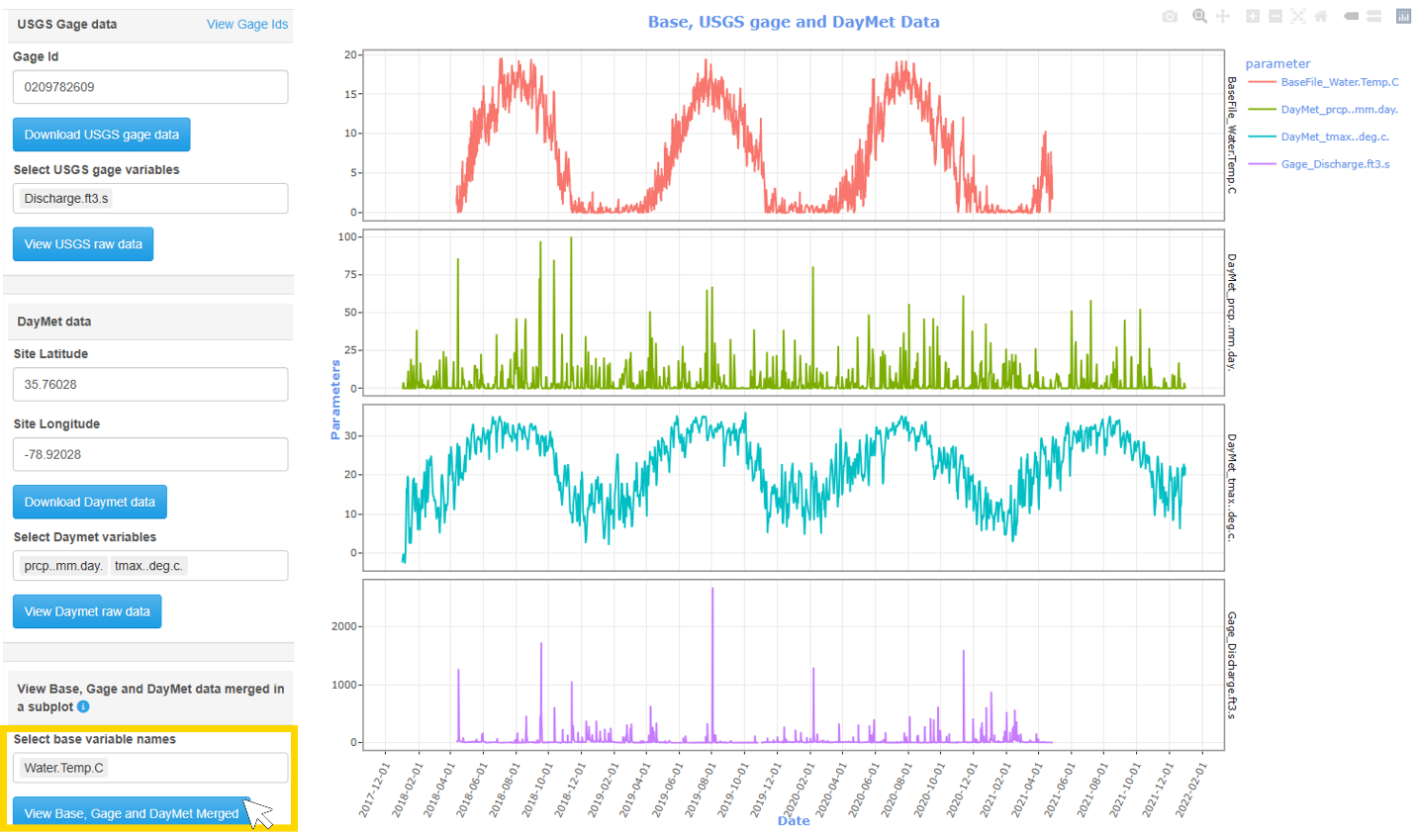
To download daily weather and climatology data from Daymet, input the latitude and longitude of your requested site in decimal degrees. When Daymet data have been successfully downloaded, Select Daymet variables will appear. Select desired Daymet parameters (see table below) and select the View Daymet raw data button to display the time series

**Table 1**. Daymet parameters available for download in ContDataSumViz. Additional metadata concerning these variables are available on the [Daymet website](https://daac.ornl.gov/DAYMET/guides/Daymet_Daily_V4R1.html).

|  |  |
| --- | --- |
| **Parameter name** | **Meaning** |
| Precipitation (mm) | Total accumulated precipitation in water-equivalent depth in millimeters |
| Shortwave radiation (W m^-2) | Incident shortwave radiation in watts per square meter |
| Snow water equivalent (kg m^-2) | Snow water equivalent (estimate of the water contained in snowpack) in kilograms per square meter |
| Maximum air temperature (degrees C) | Maximum air temperature in a day in degrees C |
| Minimum air temperature (degrees C) | Minimum air temperature in a day in degrees C |
| Water vapor pressure (Pa) | Water vapor pressure in pascals |

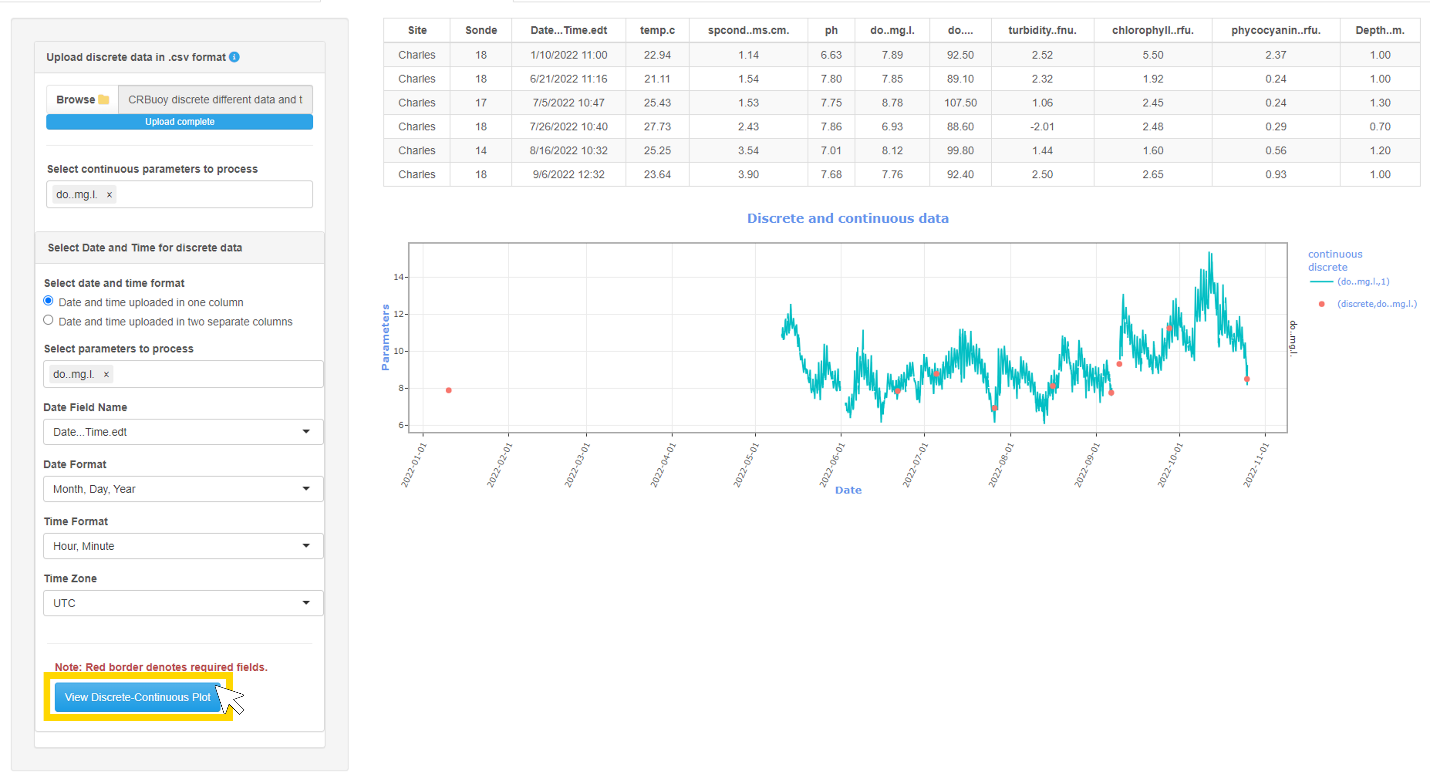
## Merged base data, USGS data, and Daymet data

To display a plot with uploaded, USGS gage, and Daymet data, select the desired parameters from the Select base variable names dropdown and click View Base, Gage and DayMet Merged. The plots will populate with the USGS gage and Daymet variables previously selected in the sidebar.



# Discrete Data Exploration

This module allows users to plot discrete data on top of the uploaded continuous data. To begin, upload the .csv file with your discrete data. Select date and time parameters to match the uploaded data, as done in the Upload Data tab. The name of the parameter in the discrete file must match a parameter in the continuous data file, otherwise the application will display an error. Upon uploading the discrete data file, the application will display the first six rows in the main panel. Select parameters for the date and time format. Then select View Discrete-Continuous Plot to view the time series. The plot controls for zoom, pan, and save the continuous-discrete data plot are the same as for the raw data time series in the Upload Data tab.



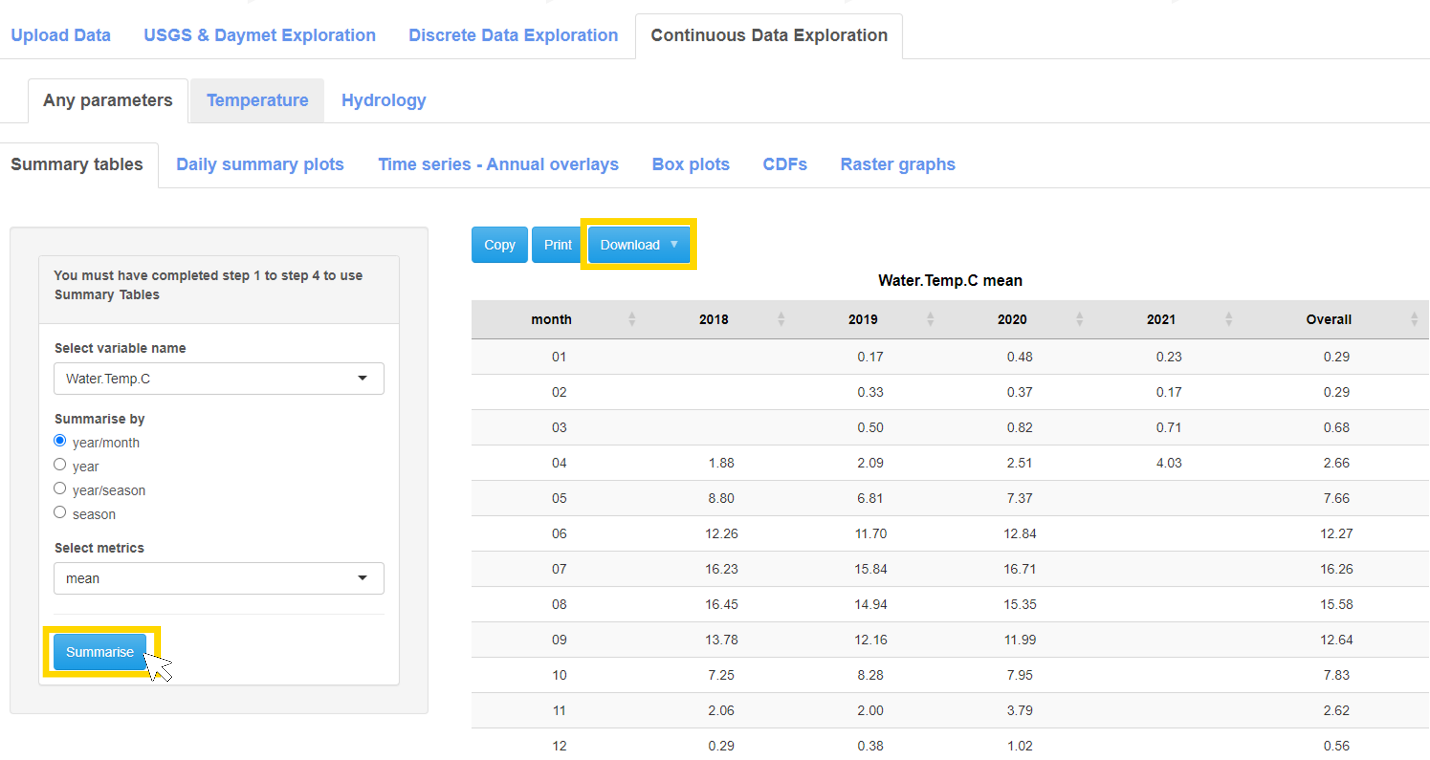
# Continuous Data Exploration

## Any parameters

Modules under this tab can calculate summaries and develop visualizations for any uploaded continuous parameter.

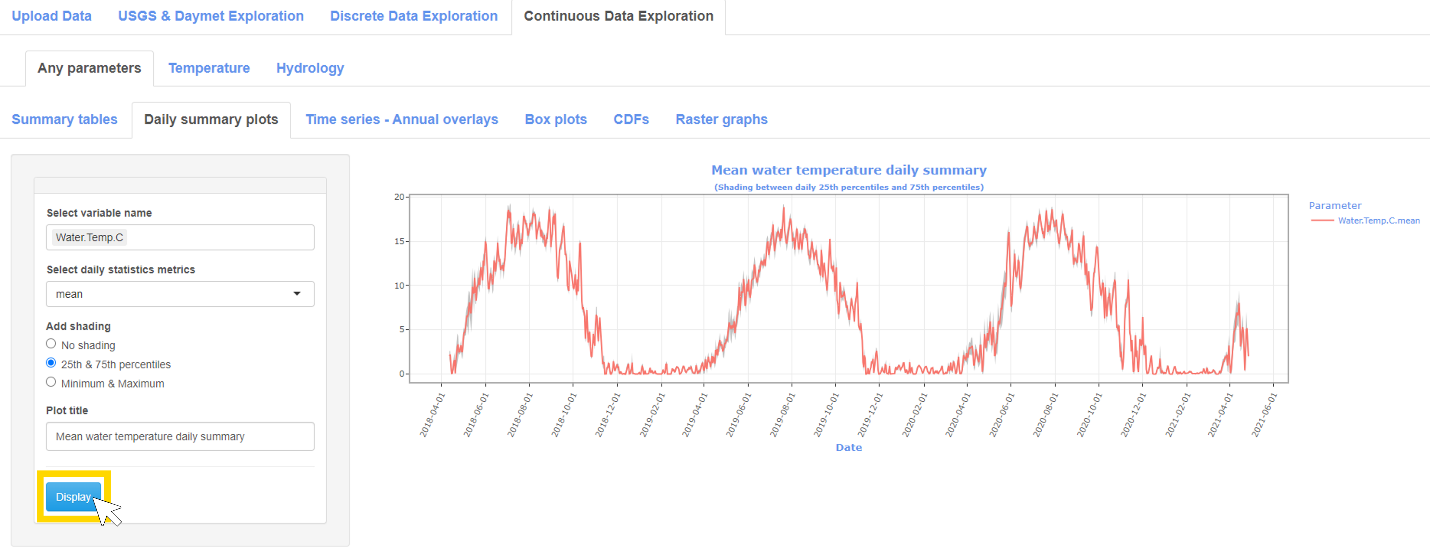
### Summary tables

This module allows users to calculate summary statistics for different variables, time periods, and metrics. Users can select any of the variables in their uploaded file from the Select variable name dropdown. Radio buttons display the time period summary options: year/month, year, year/season, and season. The metrics options available in the dropdown are mean, median, minimum, maximum, range, standard deviation, variance, coefficient of variation, and number of observations. Options to copy, print, and download the summary table are displayed above the table. Downloads are available as csv, Excel, and PDF files.



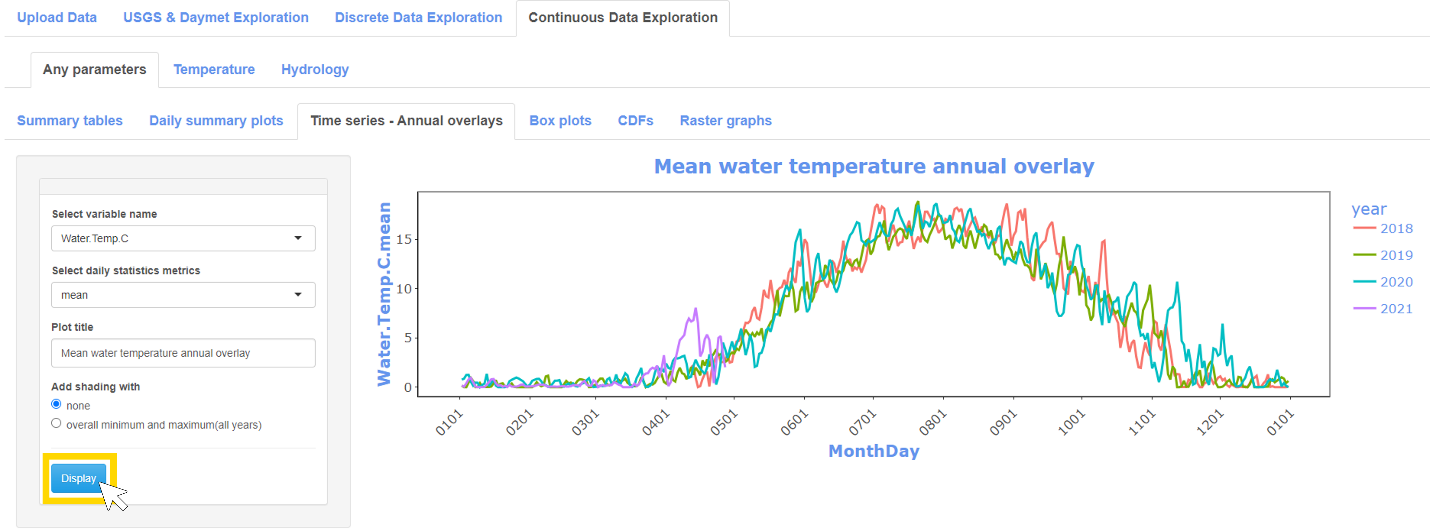
### Daily summary plots

This module displays time series of daily summaries for selected variables and offers the option of adding shading between the 25th and 75th percentiles or the minimum and maximum values. The Plot title text box allows users to alter the time series title. Available summary statistics are the same options as for the Summary tables tab above.



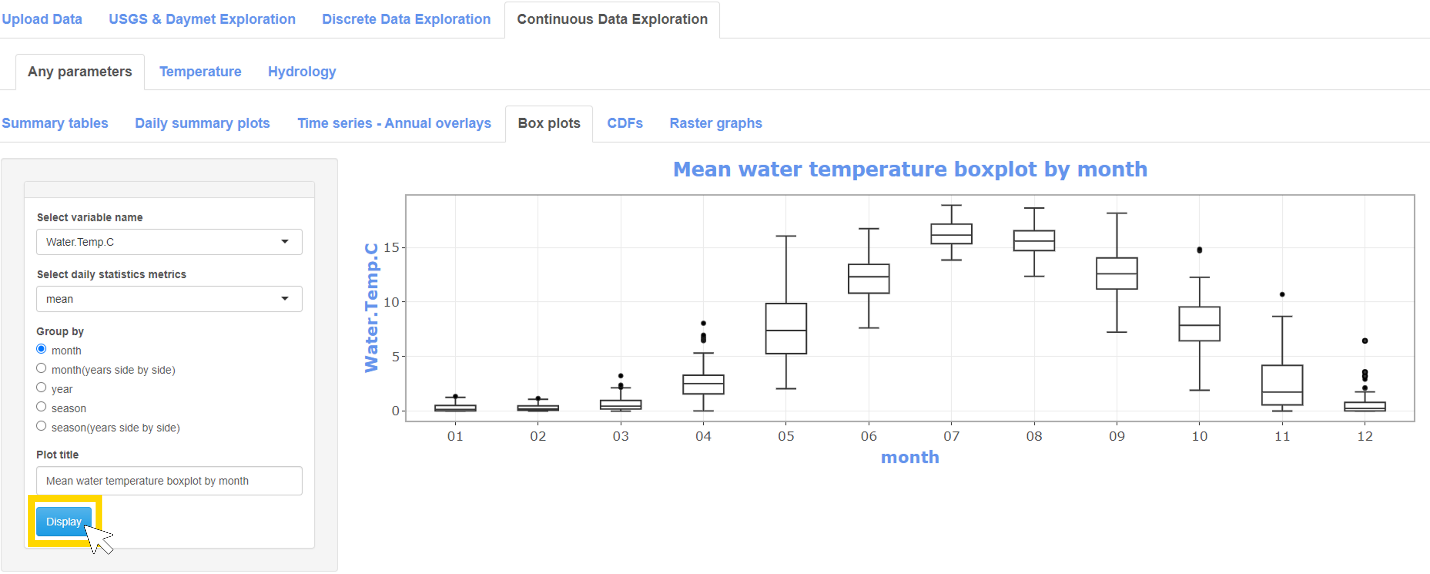
### Time series – Annual overlays

This module overlays time series of daily summary metrics for each year in the uploaded data. The available summary metrics are the same as the previous Any parameters tabs. Shading for overall minimum and maximum across all years can be added to the plot.



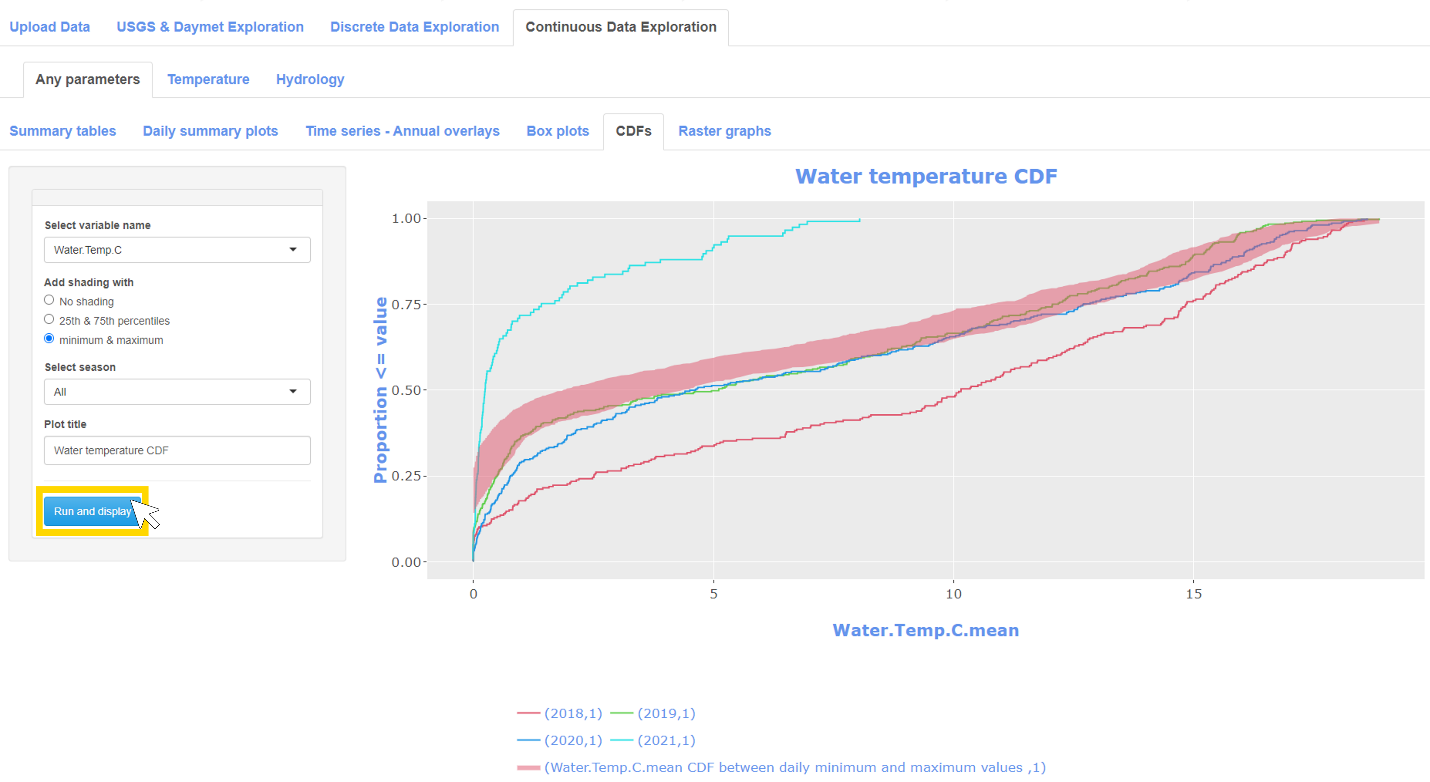
### Box plots

This module prepares boxplots of daily summary statistics grouped by different time periods. The daily statistics metrics options are the same as the above Any parameters tabs. The available summary time periods are: 1) month, 2) month (years side by side), 3) year, 4) season, 5) season (years side by side).



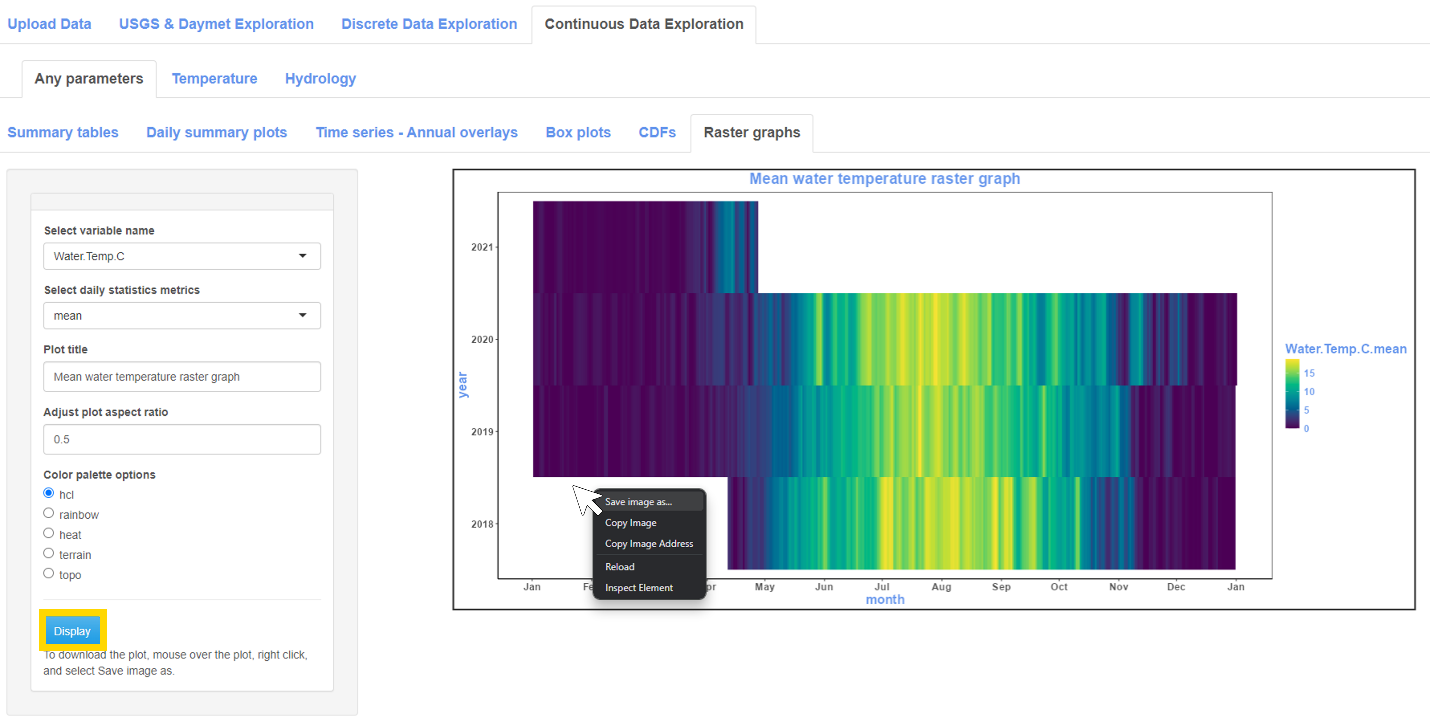
### CDFs

This module calculates and displays the empirical cumulative distribution function of the selected variable for each year subset by the selected season.



### Raster graphs

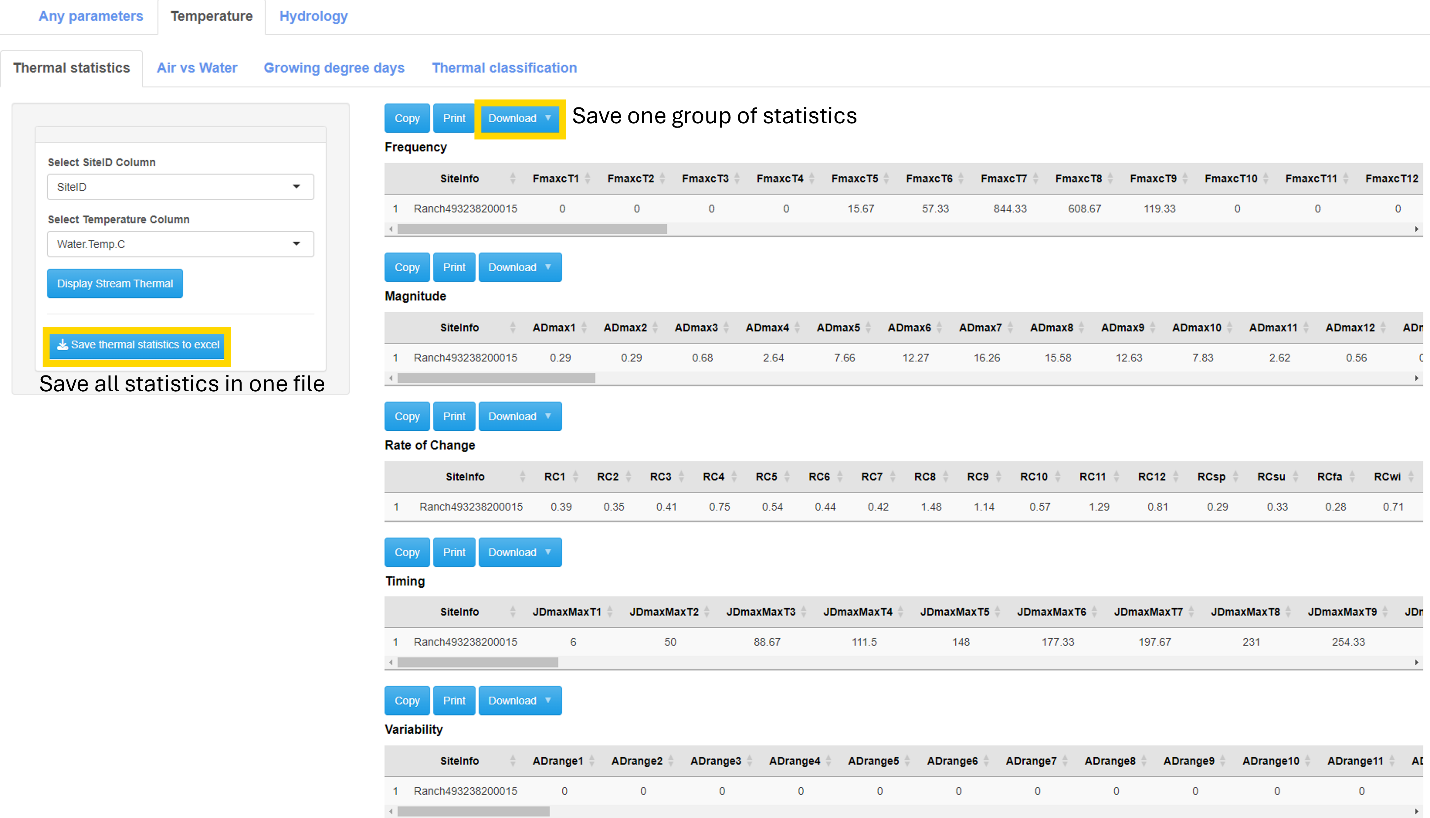
This module displays raster graphs depicting daily summary statistics with rows representing years, columns representing days, and colors representing summary values. The Adjust plot aspect ratio parameter in the sidebar of this tab allows the user to adjust the plot aspect ratio. To save the plot, mouse over the plot, right click, and select Save image as.



## Temperature

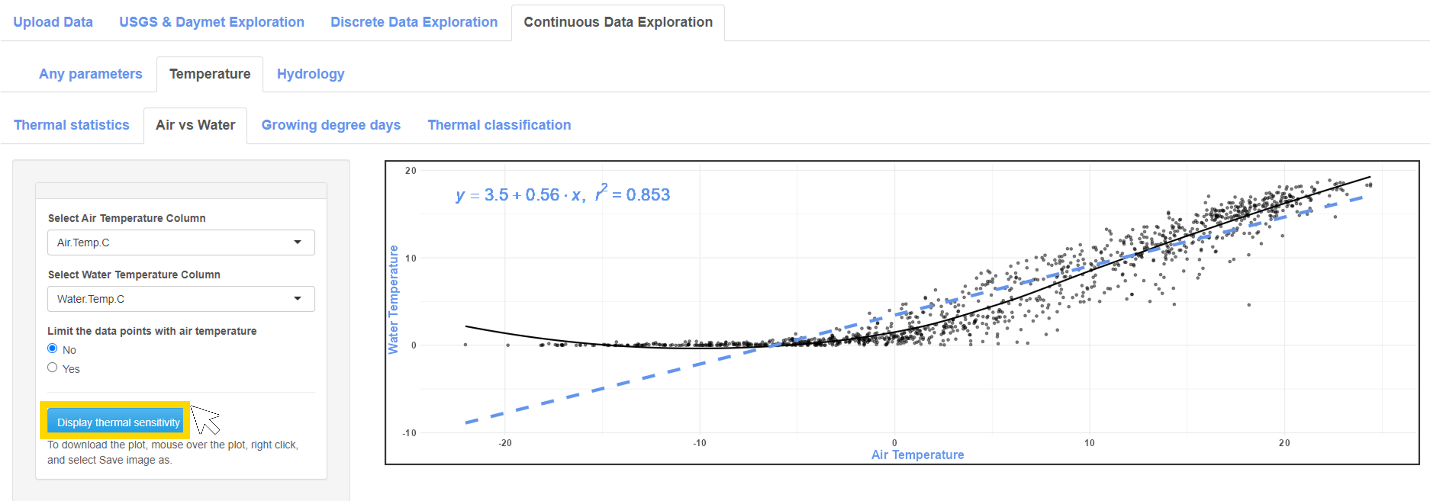
### Thermal statistics

This module calculates stream thermal regime statistics from [Tsang et al. 2016](https://doi.org/10.1080/03632415.2016.1210517) covering temperature magnitude, variability, frequency, timing, and rate of change. To run this module, select the columns representing SiteID and water temperature, then select Display Stream Thermal. When the statistics have successfully calculated, the summary tables will appear in the main panel on the right and the Save thermal statistics in excel button will appear in the sidebar on the left, which generates an Excel workbook containing all the thermal statistics tables. Individual tables can be downloaded using the download button above the corresponding table. These tables can be downloaded as csv, Excel, and PDF files.



### Air vs Water

To save this plot, mouse over the plot, right click, and select Save image as. The blue dashed line indicates the linear best fit which corresponds to the regression parameters in blue in the upper lefthand corner of the plot. The black line represents a local polynomial regression fit of the data points in black. The sidebar contains an option to exclude values with air temperatures less than a specified value. The default value provided (when Limit the data points with air temperature = Yes) is 0˚C. At this temperature ice formation generally breaks down the relationship between air and water temperature.



### Growing degree days

This module is still in development, but R code to calculate growing degree days is available from the Download R script link.

The growing degree days module in the Temperature tab under the Continuous Data Exploration tab in ContDataSumViz. Contains text "COMING LATER.. COMING LATER…
When resources permit, we will add in a function to calculate Growing Degree Days (GDD), which are used to estimate the growth and development of insects during the growing season. The basic concept is that development will only occur if the temperature exceeds some minimum development threshold, or base temperature, which varies depending on the type of organisms being studied.

Although we aren’t able to provide a GDD calculator at this time, one of the RMN partners, Tim Martin from Minnesota DNR (tim.martin@state.mn.us), has generously shared code that people with R software can use in the meantime. Click below to download the R script." Below is a clickable link labeled "Download R script"

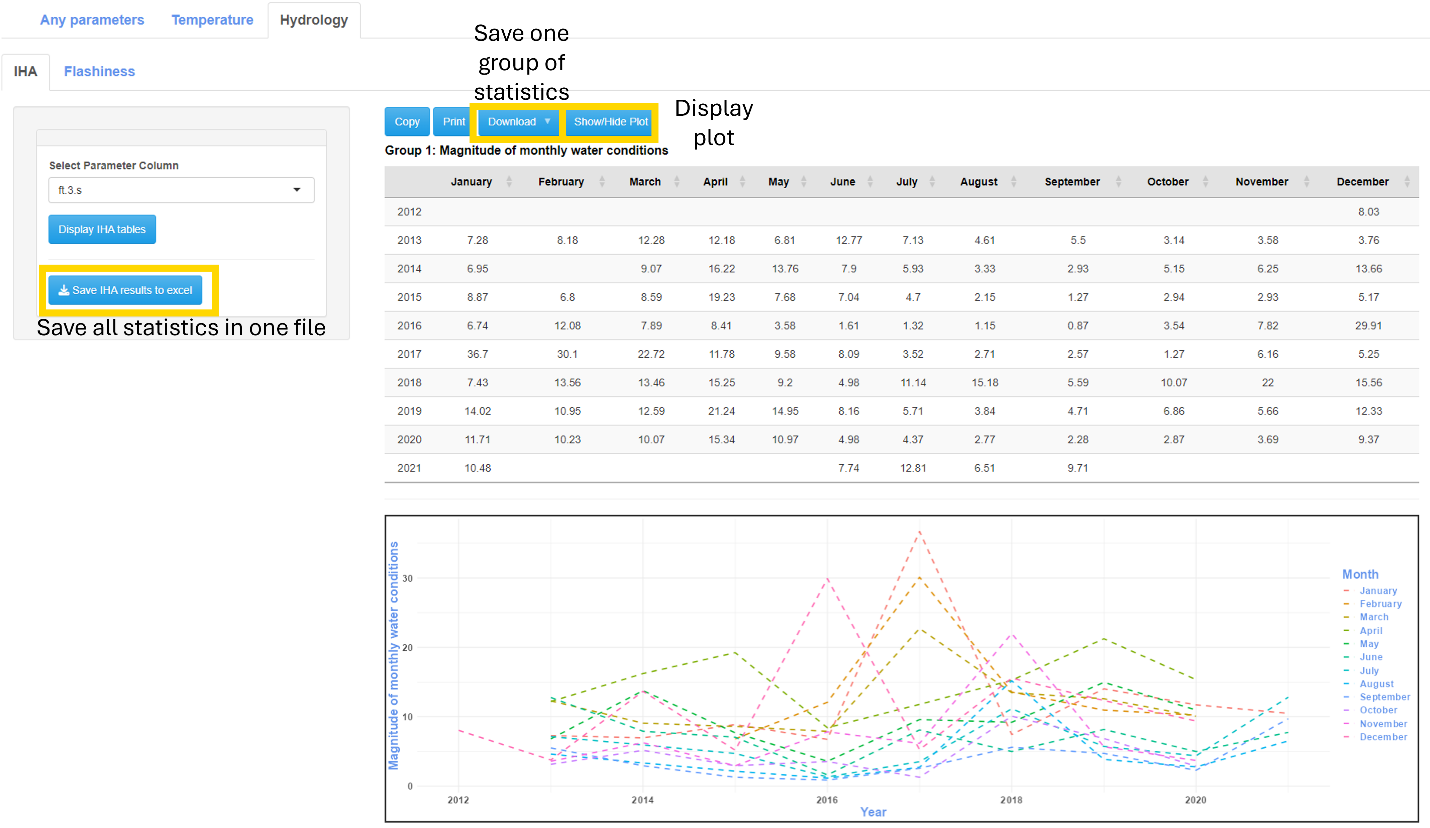
### Thermal classification

PENDING instruction from Michael

## Hydrology

### IHA

This module calculates tables and figures for five groups of Indicators of Hydrologic Alteration (IHA): Group 1 Magnitude of monthly water conditions, Group 2 Twelve magnitude/duration statistics, Group 3 Julian day of annual minimum and maximum discharge, Group 4 Frequency and duration of high and low pulses, Group 5 Rate and frequency of water condition changes. To download all statistics group in a single Excel workbook, select Save IHA results to excel in the lefthand sidebar. To download individual tables, select the download button above the corresponding table. To display the corresponding plot, select Show/Hide Plot. To download a plot, mouse over the plot, right click, and select Save image as.



### Flashiness

This module to calculate the Richards-Baker flashiness index is coming later.

The Flashiness module in the Hydrology tab under the Continuous Data Exploration tab of the ContDataSumViz tool. 

Contains a message:
"COMING LATER…
The Richards-Baker flashiness index (RBI) (Baker et al. 2004) reflects the frequency and rapidity of short-term changes in streamflow. It measures oscillations in discharge relative to total discharge, such that flashier streams receive higher scores. Results are scaled from 0 to 1 (most flashy).

The calculation is based on mean daily flows and is calculated by dividing the sum of the absolute values of day-to-day changes in mean daily flow by total discharge during the specified time period.

The Shiny app calculation is based on mean daily values and calendar year. Those settings can be changed if you are using the R package instead of Shiny app. For more information, contact Erik Leppo (Erik.Leppo@tetratech.com).

The RBI is intended to be used with discharge data but we’re experimenting with using it with sensor depth and water level data as well (since discharge data aren’t available for some of the RMN sites).

Citation:
Baker, D.B., Richards, R.P., Loftus, T.T. and J.K. Kramer. 2004. A New Flashiness Index: Characteristics and Applications to Midwestern Rivers and Streams. Journal of the American Water Resources Association 40(2): 503-522.
https://doi.org/10.1111/j.1752-1688.2004.tb01046.x"