Data Pull User Guide

Tetra Tech

## Required Packages

The required packages for the data\_pull.R code are *TADA* and *tidyverse*. In order to install the *TADA* package, you will also need the *remotes* package. The *remotes* package is not required otherwise. Here are the lines to install these packages:

# install.packages('tidyverse')  
library(tidyverse)  
  
# install.packages("remotes",  
# repos = "http://cran.us.r-project.org")  
library(remotes)  
  
# remotes::install\_github("USEPA/TADA",  
# ref = "develop",  
# dependencies = TRUE)  
library(TADA)

## Define Date Parameters

The only necessary input for the data pull is the start and end dates. Change the dates using the format ‘YYYY-MM-DD’ for startDate and endDate for whatever is required.

startDate <- '2018-02-01'  
endDate <- '2023-02-01'

## Data Pull/Download

The data download section is broken into five different pulls. Each different pull is for a specific site type that we identified as relevant come straight from the [Water Quality Portal](https://www.waterqualitydata.us/) (see ‘Site Type’ under the ‘Basic’ option). The actual download is performed using the TADA\_BigDataRetrieval function. Documentation on this function can be found by running the code below:

?TADA\_BigDataRetrieval

The inputs into the TADA\_BigDataRetrieval are the dates defined in the previous section, the specific site type we’re doing the pull for, the state code, and applying auto clean. The TADA\_BigDataRetrieval function will not return a TADA object without the applyautoclean = T.

#Estuary  
est\_data\_pull <- TADA\_BigDataRetrieval(startDate = startDate,  
 endDate = endDate,  
 siteType = 'Estuary',  
 statecode = 'AK',  
 applyautoclean = T)

The auto-cleaning applies the functions TADA\_ConvertSpecialChars, TADA\_ConvertResultUnit, TADA\_ConvertDepthUnits, and TADA\_IDCensoredData. For more information on the auto-cleaning functions run the following lines of code to review the documentation:

?TADA\_ConvertSpecialChars  
?TADA\_ConvertResultUnits  
?TADA\_ConvertDepthUnits  
?TADA\_IDCensoredData

The output of the TADA\_BigDataRetrieval function is a WQP data table with an additional 27 columns that contain the TADA-specific edits and flags.

## Exporting as CSV

The resulting tables from TADA\_BigDataRetrieval can either be combined using rbind into one large table for export and/or be exported individually by site type. Exporting to csv is performed using write\_csv from the *dplyr* package (loaded in with *tidyverse*). Change the export file location if needed.

write\_csv(est\_data\_pull, 'Data/data\_pull/data\_pull\_estuary.csv')