Kenai River Water Quality Assessment

Benjamin Meyer, Kenai Watershed Forum

2022-02-10

Contents

Part I

Part I

No text here

Preface

Preface

Acknowledgements

Acknowledgement

Abstract

Abstract here

Units

Units

Acronyms

Acronyms

Chapter 1

Introduction

Introduction/history

Chapter 2

Study Area

2.1 Description

2.2 Figures/maps

2.3 Sampling sites descriptions/photos

2.3.1 Tributary Sites

2.3.2 Main Stem Sites

Part II

Part II

Chapter 3

Methods

Chapter 4

Results

4.0.1 Parameters and graphs

Explain how to interpret figures

4.0.1.1 Parameters

For each parameter, basic time series figures in context of EPA and AEDC limits similar to Guerron Orejuela 2016. Update references and interpretation.

(code note: use "parameters" option in front matter of each page to switch, https://bookdown.org/yihui/rmarkdown/params-knit.html)

Chapter 5

Data QA/QC

Data QA/QC report basics

Chapter 6

Summary

References

List pasted from Zotero library, and link to online library:

https://www.zotero.org/groups/4571804/kenai-river-water-quality-assessment

author: “Benjamin Meyer, Kenai Watershed Forum”
date: “2022-02-10”
output: html_document:
code_folding: hide

Appendix A

Appendix: Data Uplift to EPA WQX

A.1 Introduction

Prior to analysis and interpretation of water quality data, we will ensure that all data that meets QA/QC standards outlined in the project Quality Assurance Project Plan (QAPP) (**?**) is accessible in the appropriate repository. Water quality data for this project is ultimately destined for the EPA Water Quality Exchange (EPA WQX), formerly EPA STORET.

Section B10 of the 2020 QAPP describes data management details and responsible parties for each step of the data pipeline from observation to repository.

A.1.1 2021 Water Quality Data

Water quality data generated from the Kenai River Baseline Water Quality Monitoring (KRBWQM) program was submitted to the Soldotna office of the Alaska Department of Environmental Conservation (ADEC) in January 2022 using the project-specific AQWMS Template provided by ADEC.

A.1.1.1 2021 Water Quality Data AQWMS Formatting

The code script below assembles water quality data from the three analytic laboratories that partnered with Kenai Watershed Forum for this project in 2021:

- SGS Laboratories (Anchorage, AK)

- Soldotna Wastewater Treatment Plant (Soldotna, AK)
- Taurianen Engineering (Soldotna, AK)

A.1.1.1.1 Spring 2021 Download Original Spring 2021 SGS Lab Results Files: (Click Link)

Click Black Arrow to Show/Hide Code used to Prepare 2021 (Spring & Summer) SGS Results for AQWMS Template

```
# clear environment
rm(list=ls())

# load packages
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
library(openxlsx)
library(data.table)
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
library(stringr)
library(magrittr)
```

```
##
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':
##
##     set_names

## The following object is masked from 'package:tidyr':
##
##     extract
```

```
library(janitor)
```

```
##
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```
# Assign 2021 Field Sample Dates

# Spring 2021 sampling date
spring21_sample_date <- "5/11/2021"

# Summer 2021 Sampling Date
summer21_sample_date <-"8/27/2021"

##########################################################################################################
########################################## Read in  and Clean SGS/ALS Data ###############################
##########################################################################################################


############################## Part A: SGS Data Read In ##############################

## Reformat SGS data downloaded from their server client (SGS Engage) to match AQWMS template

### spring 2021 SGS data

#### read in results downloaded from SGS
```

```
spring_sgs21 <- read_excel("other/input/2021_wqx_data/spring_2021_wqx_data/SGS/Spring 2
  mutate(`Lab Sample` = as.double(`Lab Sample`)) %>%
  # assign field sample date
  mutate(sample_date = spring21_sample_date) %>%
  clean_names()

### summer 2021 SGS data

#### read in results downloaded from SGS
summer_sgs21 <- read_excel("other/input/2021_wqx_data/summer_2021_wqx_data/SGS/SGS_resu
  # filter out extraneous info
  filter(!`Client Sample` %in%c("Sample Comments","Client Sample Id"),
         !`Matrix` %in% "200.7 Total Ca, Fe, Mg were analyzed by ALS of Kelso, WA.",
         !is.na(`Lab Sample`)) %>%
  # remove blank columns
  select(-starts_with("...")) %>%
  # assign field sample date
  mutate(sample_date = summer21_sample_date) %>%
  clean_names()
```

```
## New names:
## * `` -> ...9
## * `` -> ...10
## * `` -> ...11
## * `` -> ...12
## * `` -> ...13
```

```
# transform column types to prep for join
summer_sgs21 %<>%
  mutate(lab_sample = as.double(lab_sample),
         reporting_limit = as.double(reporting_limit))

# append spring and summer 2021 results
sgs21 <- bind_rows(spring_sgs21,summer_sgs21) %>%
  # add lab name
  mutate(lab_name = "SGS North America, Anchorage, Alaska") %>%
  # make lab sample column character
  transform(lab_sample = as.character(lab_sample))

rm(spring_sgs21,summer_sgs21)
```

```
###################### Part B: ALS Data Read In ##############################

## SGS subcontracted analyses of Ca, Fe, and Mg to ALS laboratories (Kelso, WA). These results ar

#### read in spring 2021 results from ALS
spring_als21 <- read_excel("other/input/2021_wqx_data/spring_2021_wqx_data/SGS/Spring 2021 Total
  # assign field sample date
  mutate(sample_date = spring21_sample_date) %>%
  clean_names()

#### read in summer 2021 results from ALS
summer_als21 <- read_excel("other/input/2021_wqx_data/summer_2021_wqx_data/SGS/ALS_Data_Results_S
  # assign field sample date
  mutate(sample_date = summer21_sample_date) %>%
  clean_names()

# join spring and summer 2021 datasets
als21 <- bind_rows(spring_als21,summer_als21)

# remove old dataframes
rm(spring_als21,summer_als21)

# prep ALS data to be joined with SGS data
als21 %<>%
  rename(client_sample = client_id,
         lab_sample = lab_id,
         reporting_limit = mrl,
         analysis = method) %>%
  select(-qc1,-data_entry,-qc2) %>%
  # transform column types as needed for join in next step
  transform(lab_sample = as.character(lab_sample),
            analysis = as.character(analysis),
            result = as.character(result)) %>%
  # add lab name
  mutate(lab_name = "ALS Environmental - Kelso Laboratory",
         # assign matric type
         matrix = "Water (Surface, Eff., Ground)")

# join SGS data with ALS data
sgs21 <- bind_rows(sgs21,als21)

rm(als21)
```

```
################ Part C: Address spelling/format issues and inconsistent sample/site nam

# upon visual inspection, we can see that the location names in the AQWMS template dif


# move info about duplicate sample and/or sample blank status into separate column
sgs21 %<>%
  mutate(sample_condition = case_when(
    grepl("Blank",client_sample) ~ "Blank",
    grepl("DUP",client_sample) ~ "DUP")) %>%
  # remove "DUP" designation from client_sample column
  mutate(client_sample = str_replace(client_sample, "DUP", ""))


# remove from "client_sample" names the text containing the suffixes Diss/Dis (Dissolv
sgs21 %<>%
  mutate(client_sample = (str_replace(client_sample, "Diss|Dis|DUP", ""))) %>%

  # remove "Diss" suffix and "EP" prefix from "analysis" column
  mutate(analysis = str_replace(analysis, "Diss", "")) %>%
  # note trailing space after "EP200.8 "
  mutate(analysis = str_replace(analysis,"EP200.8 ","200.8")) %>%

  # address the one stubborn site name still containing "Diss"
  mutate(client_sample = case_when(
    client_sample == "RM0-No Name Creek  Diss" ~ "RM0-No Name Creek",
    TRUE ~ client_sample))


# We need to remove white spaces, apostrophes, and dashes. Join functions such as "lef

sgs21 %<>%
    # remove excess white spaces
  mutate(client_sample = str_trim(client_sample,"both")) %>%
  mutate(client_sample = str_squish(client_sample)) %>%

  # make remaining white spaces underscores
  mutate(client_sample = gsub("\\s+","_",client_sample)) %>%

  # remove apostrophes
  mutate(client_sample = gsub("\\'","",client_sample)) %>%
```

```
  # replace dashes with underscores
  mutate(client_sample = gsub("\\-","_",client_sample)) %>%

  # replace multiple underscores with single
  mutate(client_sample = gsub("\\__","_",client_sample)) %>%
  mutate(client_sample = gsub("\\___","_",client_sample))

# apply note regarding trip blanks (for BTEX organics)
# assigned in sequence as encountered on chain of custody
sgs21 %<>%
  mutate(note = case_when(
    grepl("Trip_Blank_1", client_sample) ~ "KWF Crew, RM1.5_Kenai_City_Dock",
    grepl("Trip_Blank_2", client_sample) ~ "USFWS Crew, RM6.5_Cunningham_Park",
    grepl("Trip_Blank_3", client_sample) ~ "DEC Crew, RM40_Bings_Landing",
    grepl("Trip_Blank_4", client_sample) ~ "DEC Crew, RM43_Upstream_of_Dow_Island"))


# seperate result qualifiers (U, J, B) in to a new column
sgs21 %<>%
  mutate(qualifier = str_extract(result,"[aA-zZ]+")) %>%
  mutate(result = str_remove(result,"[aA-zZ]+")) %>%

  # again replace multiple underscores with single
  mutate(client_sample = gsub("\\__","_",client_sample))









############### Part D: Prepare SGS/ALS Location/Site Names ############################

# In preparation for a join to AQWMS table, we will manually generate a match table csv file that

## generate list of unique site names from 2021 SGS data
sgs21_sitenames <- data.table(unique(sgs21$client_sample))


# generate list of unique site names from 2021 AQWMS template
aqwms21_sitenames <- read_excel("other/input/AQWMS/AWQMS_KWF_Baseline_2021.xlsx", sheet = "Monito
  select("Monitoring Location Name", "Monitoring Location ID") %>%
  distinct()

# write 2021 sgs site names to an excel file
```

```
site_match_table_path <- "other/input/AQWMS/site_names_matching_table.xlsx"
write.xlsx(sgs21_sitenames, site_match_table_path)

# create an excel file with two sheets: a.) SGS site names, and b.) AQWMS site names
wb <- loadWorkbook(site_match_table_path)
addWorksheet(wb,"Sheet2")
writeData(wb,"Sheet2",aqwms21_sitenames)
saveWorkbook(wb,site_match_table_path,overwrite = TRUE)


# Using these two tables, we will manually create a new file titled "site_names_matchi


# append "Monitoring Location Name" and "Monitoring Location ID" info from WQX to spri

## read in site names join table
sitenames21_match <- read_excel("other/input/AQWMS/site_names_matching_table_manual_ed
  select(`Monitoring Location Name`,`Monitoring Location ID`,sgs_sitenames) %>%
  rename(client_sample = sgs_sitenames)
```

```
## New names:
## * `` -> ...4
## * `` -> ...5
## * `` -> ...6
```

```
# append monitoring location names
sgs21 %<>%
  left_join(sitenames21_match, by = "client_sample") %>%
  clean_names()

# remove
rm(sgs21_sitenames,aqwms21_sitenames,sitenames21_match)
```

```
######################### Part E: "Result Analytical Method Context" name rectification

# In the AQWMS template, the EPA names that will go in the column "Result Analytical M

# assign "Result Analytical Method ID" and "Result Analytical Method Context" to datas

# read in matching table
```

```
analysis_code_matching_table <- read_excel("other/input/AQWMS/analysis_code_matching_table.xlsx")
  select(-Comments,-`EPA Name`) %>%
  clean_names() %>%
  rename(analysis = sgs_analysis_code) %>%
  # remove "EP" prefix from method "EP200.8"
  mutate(analysis = str_replace(analysis,"EP200.8","200.8"))


# read in AQWMS Analytical Methods list
aqwms_analytical_methods <- read_excel("other/input/AQWMS/AWQMS_KWF_Baseline_2021.xlsx", sheet =
  select("ID","Context Code") %>%
  clean_names() %>%
  rename(epa_analysis_id = id)
```

```
## Warning in read_fun(path = enc2native(normalizePath(path)), sheet_i = sheet, :
## Expecting logical in F1126 / R1126C6: got '4500-NO2'

## Warning in read_fun(path = enc2native(normalizePath(path)), sheet_i = sheet, :
## Expecting logical in F1130 / R1130C6: got '4500-NO3'
```

```
# join two tables above
epa_analysis_codes <- inner_join(aqwms_analytical_methods,analysis_code_matching_table, by = "epa
  filter(!context_code %in% c("USEPA Rev 5.4",
                              "APHA (1997)",
                              "APHA (1999)"))

# join EPA analysis IDs and context codes to overall dataset
sgs21 %<>%
  left_join(epa_analysis_codes, by = "analysis")

#"EP200.8 Diss" remains in analysis column, wth ...

rm(analysis_code_matching_table,aqwms_analytical_methods,epa_analysis_codes)
```

```
# Taurianen data
## Reformat Taurianen data to match AQWMS template
```

```
# Soldotna Wastewater Treatment Plant
## Reformat SWWTP data to match AQWMS template
```

```
# check out example from other WQP download
z1 <- read.csv("other/input/wqp_data/narrowresult.csv")

### 1st step:::: make sure that monitoring list locations matches those currently match those in
```

A.2 2021 Baseline Water Quality Data