

Kenai River Water Quality Assessment

Benjamin Meyer, Kenai Watershed Forum

2022-02-15

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: "2022-02-15"
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 scripts 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 Metals/Nutrients Lab Results (SGS Labs) [Download Original 2021 Metals/Nutrients Lab Results \(Click Link\)](#)

Show/Hide Code used to Prepare 2021 Metals/Nutrients Results

```
# 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
```

```
library(hms)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following object is masked from 'package:hms':
##
##      hms

## The following objects are masked from 'package:data.table':
##
##      hour, isoweek, mday, minute, month, quarter, second, wday, week,
##      yday, year

## The following objects are masked from 'package:base':
##
##      date, intersect, setdiff, union
```

```

# 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 t

### 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_res
  # 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 are

#### 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 matrix 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 names

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

# 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 (Dissolved/Soluble)
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 "left_join" are
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"))

# separate 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 cs

## 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/AQWMS_KWF_Baseline_2021.xlsx", sheet = "Sheet1",
  select("Monitoring Location Name", "Monitoring Location ID") %>%
  distinct()

# write 2021 sgs site names to an excel file
site_match_table_path <- "other/input/AQWMS/sgs_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 "sgs_site_names_matching_table_manual.xlsx"

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

## read in site names join table
sitenames21_match <- read_excel("other/input/AQWMS/sgs_site_names_matching_table_manual.xlsx",
  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 dfs
rm(sgs21_sitenames, aqwms21_sitenames, sitenames21_match)

# remove old "client_sample" column
sgs21 %<>% select(-client_sample)

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

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

# assign "Result Analytical Method ID" and "Result Analytical Method Context" to dataset using matching table

# 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/AQWMS_KWF_Baseline_2021.xlsx", sheet = "Methods")
  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-N02'
```

```
## Warning in read_fun(path = enc2native(normalizePath(path)), sheet_i = sheet, :
## Expecting logical in F1130 / R1130C6: got '4500-N03'
```

```

# join two tables above
epa_analysis_codes <- inner_join(aqwms_analytical_methods, analysis_code_matching_table
  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")

rm(analysis_code_matching_table, aqwms_analytical_methods, epa_analysis_codes)

##### Part F: Assign lab analysis in/out dates #####

# TBD if necessary

##### Miscellaneous Steps #####

# rectify select column classes

sgs21 %<>%
  transform(sample_date = mdy(sample_date))

```

A.1.1.1.2 Fecal Coliform Lab Results (Soldotna Wastewater Treatment Plant/Taurianen Engineering) [Download Original 2021 Fecal Coliform Lab Results \(Click Link\)](#)

Show/Hide Code used to Prepare 2021 Fecal Coliform Results

```

#####
##### Read in and Clean SWWTP / Taurianen FC Data #####
#####

##### Part A: SWWTP FC Data Read In #####
swwtp_spring21 <- read_excel("other/input/2021_wqx_data/spring_2021_wqx_data/SWWTP/KRWI
  clean_names() %>%

## fix site naming and terminology

```

```

# move info about duplicate sample and/or sample blank status into separate column
mutate(sample_condition = case_when(
  grepl("BLANK",sample_location_rm) ~ "Lab Blank",
  grepl("DUP",sample_location_rm) ~ "DUP",
  grepl("POSITIVE",sample_location_rm) ~ "Positive Control")) %>%
# remove "BLANK", and "POSITIVE" designation from sample_location column
mutate(sample_location_rm = (str_replace(sample_location_rm, "BLANK|POSITIVE", "")))

# remove "DUP" from site name column and trim white spaces in site name column
swwtp_spring21 %<>%
  mutate(sample_location_rm = str_remove(sample_location_rm,"DUP")) %>%
  mutate(sample_location_rm = str_trim(sample_location_rm,"right"))

# address different site naming systems
# use manually generated matching table
# read in matching table and match
swwtp_spring21_site_matching <- read_excel("other/input/AQWMS/swwtp_site_names_matching_table_mar

# join
swwtp_spring21 %<>%
  full_join(swwtp_spring21_site_matching) %>%
  select(-sample_location_rm)

rm(swwtp_spring21_site_matching)

## fix lab analysis times and dates
swwtp_spring21 %<>%
  # lab processing time/date
  mutate(analysis_time_in = as_hms(time_in),
    analysis_date_in = mdy(spring21_sample_date),
    analysis_time_out = as_hms(time_out),
    # see file "other/input/2021_wqx_data/spring_2021_wqx_data/SWWTP/KRWF Fecal 05-11-21.xls"
    analysis_date_out = ymd("2021-05-12")) %>%
  select(-time_in,-time_out) %>%
  transform(time_sampled = as_hms(time_sampled)) %>%
  # field sample date and time
  mutate(time_sampled = as_hms(time_sampled),
    sample_date = mdy(spring21_sample_date))

## rename existing column names and create new ones to match sgs21 data format at end of prior co
swwtp_spring21 %<>%

```

```

rename(lab_sample = dish_number,
       result = colony_count_100m_l) %>%
mutate(note = paste0("Lab analysis volume = ",ml," mL"),
       matrix = "Water (Surface, Eff., Ground)",
       analysis = "9222 D ~ Membrane filtration test for fecal coliforms",
       analyte = "Fecal Coliform",
       unit = "cfu/100ml",
       # reporting limit value from 2019 QAPP, pg 17
       reporting_limit = 1.0,
       lab_name = "Soldotna Wastewater Treatment Plant",
       mdl = "",
       note = "",
       qualifier = "",
       epa_analysis_id = "9222D",
       context_code = "APHA" ) %>%
clean_names() %>%
select(-ml,-colony_count) %>%
# transform to prep for bind with sgs21
transform(lab_sample = as.character(lab_sample),
          mdl = as.double(mdl))

## join SGS 2021 data to Spring 2021 Fecal Coliform data from SWWTP
dat <- bind_rows(sgs21,swwtp_spring21)

##### Part B: Taurianen FC Data Read In (Summer 2021) #####
taur_summer21 <- read_excel("other/input/2021_wqx_data/summer_2021_wqx_data/Taurianen/1
clean_names() %>%
select(-qc1,-data_entry,-qc2) %>%

## move info about duplicate sample and/or sample blank status into separate column
mutate(sample_condition = case_when(
  grepl("DUP",sample_location) ~ "DUP")) %>%
# remove "DUP" designation from sample_location column
mutate(sample_location = (str_replace(sample_location, "_DUP", ""))) %>%
# trim white spaces in site name column
mutate(sample_location = str_trim(sample_location,"right")) %>%

```

```
## add known info about times/dates, correct formats and names
mutate(sample_date = mdy(summer21_sample_date),
       analysis_date_in = mdy(summer21_sample_date),
       analysis_time_in = as_hms(time_relinquished),
       analysis_date_out = mdy("8/28/2021"),
       analysis_time_out = as_hms(time_tested),
       sample_date = mdy(summer21_sample_date),
       # drop old columns
       .keep = "unused") %>%
select(-date_of_testing) %>%
transform(time_sampled = as_hms(time_sampled)) %>%

## add lab name
mutate(lab_name = "Taurianen Engineering")

## fix site naming and terminology

# generate spreadsheet of unique site names from taurianen dataset
taur_summer21_sites <- data.frame(unique(taur_summer21$sample_location)) %>%
  rename(sample_location = unique.taur_summer21.sample_location.)

# export site names list to spreadsheet
write.xlsx(taur_summer21_sites, "other/input/AQWMS/taurianen_site_names_matching_table.xlsx")

# manually edit a new spreadsheet such that taurianen site names are paired iwth AQWMS site names
# read in manually edited site names sheet
taur_summer21_sites <- read_excel("other/input/AQWMS/taurianen_site_names_matching_table_manual_e

# join AQWMS site names to taurianen data

z <- left_join(taur_summer21,sitenames21_match,by = "sample_location")
```

A.1.1.1.3 Total Dissolved Solids Lab Results (Soldotna Wastewater Treatment Plant (SWWTP)) Download Original 2021 Total Dissolved Solids Lab Results (Click Link)

Show/Hide Code used to Prepare 2021 Total Dissolved Solids Results

```
# SWWTP data  
## Reformat TSS data to match AQWMS template
```

A.2 2021 Baseline Water Quality Data