Appendix S1. Instructions for retrieving and archiving the environmental covariates.

Supporting information for Scheuerell et al.

Contents

This is version 0.19.02.25.

1 Background

This appendix describes how to retrieve the environmental covariates used in our analyses. After reading in the raw data, summarizing them , and trimming them to the appropriate time frame, they table of covariates is written to a `.csv` file.

All of the analyses require the R software (v3.4.3 or later) for data retrieval and processing. We also need several packages that are not included with the base installation of **R**.

```
if(!require("here")) {
  install.packages("here")
  library("here")
}
## set data dir
datadir <- here("data")
if(!require("EGRET")) {
  install.packages("EGRET")
  library("EGRET")
}
if(!require("corrplot")) {
  install.packages("corrplot")
  library("corrplot")
}
if(!require("captioner")) {
  devtools::install_github("adletaw/captioner")
```

```
  library("captioner")
}
fig_cap <- captioner(prefix = "Figure S", suffix = ". ",
                     style = "b", style_prefix = TRUE,
                     auto_space = FALSE)
```

2 User inputs

We begin by specifying

1. the name of the data file that contains the observed total number of adult spawners (escapement) by year; and
2. the minimum age of an adult spawner.

```
## [n_yrs x 2] matrix of obs counts; 1st col is calendar yr
fn_esc <- "chin_esc.csv"

## min adult age
age_min <- 3
```

Next we retrieve the metadata for the covariates from a `.csv` file with the following columns:

1. `spp`: the species to which the covariate applies
2. `life_stage`: life stage at which the effect is thought to occur
3. `covariate`: type of covariate (e.g., flow, temperature)
4. `code`: USGS code for data type (i.e., usually 5-digit integer)
5. `long_name`: long name for the covariate
6. `short_name`: function name to derive the covariate
7. `lag_1`: years to lag begin date
8. `lag_2`: years to lag end date
9. `begin`: beginning date as 2-digit text (ie, mo-yr)
10. `end`: ending date as 2-digit text (ie, mo-yr)
11. `location`: description of location
12. `gage`: USGS gage number
13. `flag`: flag to include (1) or exclude (0) the covariate
14. `flow_scen`: flag to include (1) or exclude (0) flow scenario
15. `group`: integer indicator for life stage grouping

```
## covariate metadata
cov_meta_file <- "chin_cov_metadata.csv"
```

3 Function definitions

```
## over90k
## returns the total number of days that flow exceeded 90k cfs
over90k <- function(x) {
  return(sum(x >= 90000))
```

```
}

## min7mean
## returns the min of 7-day means over period
min7mean <- function(x) {
  return(round(min(filter(x, rep(1,7)/7, "convolution", sides = 1),
                   na.rm = TRUE), 0))
}

## max7mean
## returns the max of 7-day means over period
max7mean <- function(x) {
  return(round(max(filter(x, rep(1,7)/7, "convolution", sides = 1),
                   na.rm = TRUE), 0))
}

## med7mean
## returns the max of 7-day means over period
med7mean <- function(x) {
  return(round(median(filter(x, rep(1,7)/7, "convolution", sides = 1),
                   na.rm = TRUE), 0))
}

## rng7mean
## returns the max of 7-day means over period
rng7mean <- function(x) {
  return(round(diff(range(filter(x, rep(1,7)/7, "convolution", sides = 1),
                          na.rm = TRUE)), 0))
}

## range2
## returns the range as a scalar = max-min
range2 <- function(x) {
  return(diff(range(x)))
}
```

4 Loading the fish data

We begin by loading the spawner data so we can get the time frame of interest for the covariates.

```
## escapement
dat_esc <- read.csv(file.path(datadir,fn_esc))

## get first & last years
yr_frst <- min(dat_esc$year)
yr_last <- max(dat_esc$year)
```

5 Retrieve flow covariates

The analyses are based upon several environmental indicators related to river discharge. Load the metadata file containing all of the specifications for the covariates to be used.

```
cov_meta <- read.csv(file.path(datadir, cov_meta_file), stringsAsFactors = FALSE)
cov_meta$code <- gsub("\"","",cov_meta$code)
cov_meta$begin <- gsub("\"","",cov_meta$begin)
cov_meta$end <- gsub("\"","",cov_meta$end)
```

We need to define the beginning and ending dates for the covariates.

```
yr1 <- yr_frst
yr2 <- yr_last - age_min
## start date
startDate <- paste0(yr1 + min(cov_meta$lag_1),"-01-01")
## end date
endDate <- paste0(yr2 + max(cov_meta$lag_2),"-12-31")
```

We begin by getting the daily flow data from the US Geological Service National Water Information System for the complete time period of interest.

```
## metadata for flow covariates
flow_meta <- subset(cov_meta, covariate=="flow" & flag==1)
## data to get: flow (cfs)
parameterCD <- "00060"
## get all flow data for period of interest & gages
gages <- unique(flow_meta$gage)
n_gages <- length(gages)
tmp <- readNWISDaily(gages[1], parameterCD, startDate, endDate, convert=FALSE)
tmp$yr <- floor(tmp$DecYear)
flow_data <- tmp[,c("Date","waterYear","yr","Month","Day")]
colnames(flow_data) <- c("date","H2Oyr","yr","mon","day")
flow_data[,as.character(gages[1])] <- tmp$Q
if(n_gages > 1) {
  for(i in 2:n_gages) {
    tmp <- readNWISDaily(gages[i], parameterCD, startDate, endDate, convert=FALSE)
    flow_data[,as.character(gages[i])] <- tmp$Q
  }
}
```

```
## There are 6940 data points, and 6940 days.
```

Now we can extract the specific flow covariates that relate to each of the hypotheses about the affected life stage.

```
cov_flow <- matrix(NA,length(seq(yr1,yr2)),dim(flow_meta)[1]+1)
n_mods <- dim(cov_flow)[2] - 1
cov_flow[,1] <- seq(yr_frst,yr2)
for(i in 1:dim(flow_meta)[1]) {
  fn <- get(flow_meta[i,"short_name"])
```

```
  cnt <- 1
  for(t in yr1:yr2) {
    beg <- paste0(t+flow_meta[i,"lag_1"],"-",flow_meta[i,"begin"])
    end <- paste0(t+flow_meta[i,"lag_2"],"-",flow_meta[i,"end"])
    tmp <- flow_data[flow_data$date>=beg & flow_data$date<=end,
                     as.character(flow_meta[i,"gage"])]
    cov_flow[cnt,i+1] <- fn(tmp)
    cnt <- cnt+1
  }
}
## write flow covariates to file
write.csv(cov_flow, row.names = FALSE,
          file = file.path(datadir,
                           "Willamette_Chin_SR_mainstem_flow_covariates.csv"))
```

6 Correlation of covariates

Below is a graphical representation of the pairwise correlation between each of the 32 flow covariates.

```
par(mai=rep(0.1, 4), omi=rep(0.1, 4))
corrplot(cor(cov_flow[,-1]), method="ellipse", type="lower",
         tl.col = "black", tl.srt = 0, tl.cex = 0.5, tl.offset = 0.7,
         cl.cex = 0.8, cl.offset = 0.9, cl.ratio = 0.2)
```

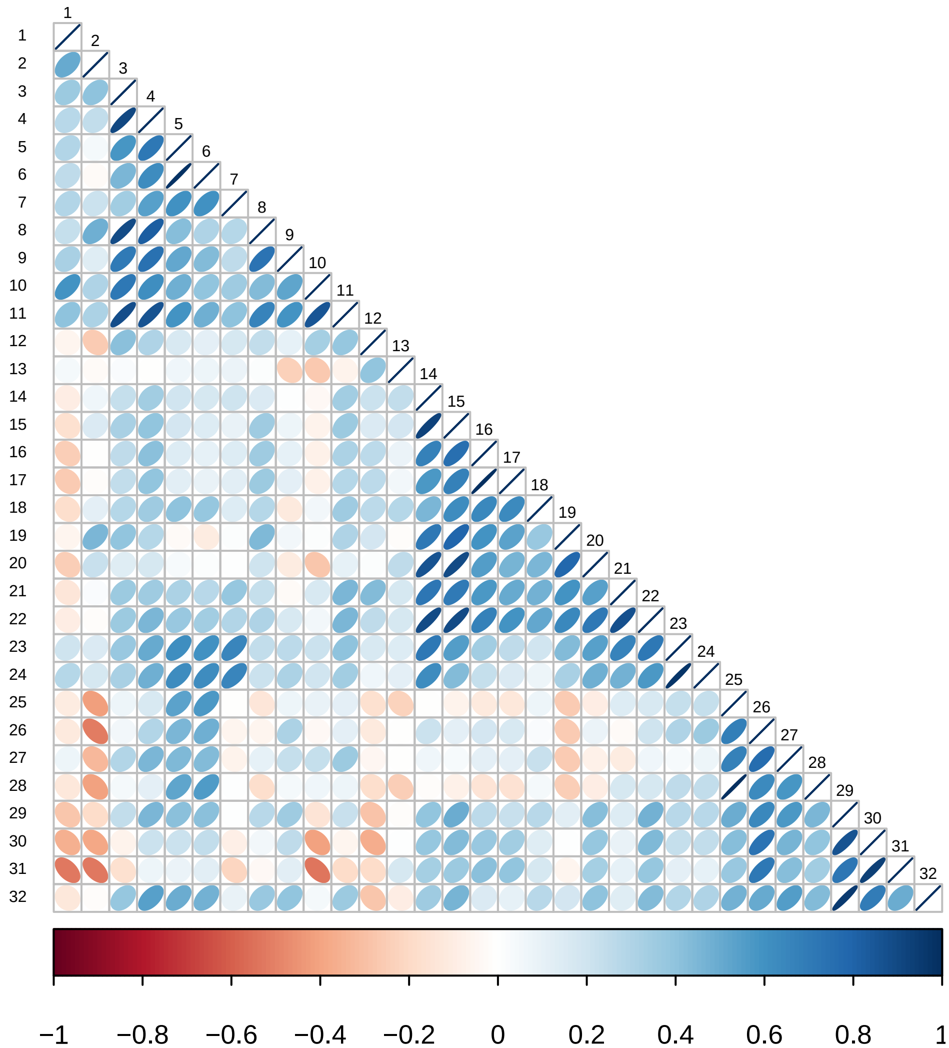

Figure S1. Pairwise correlation for each of the 32 flow covariates considered in the analysis.