Example code

Matt Callahan & Jordan Watson

7/6/2021

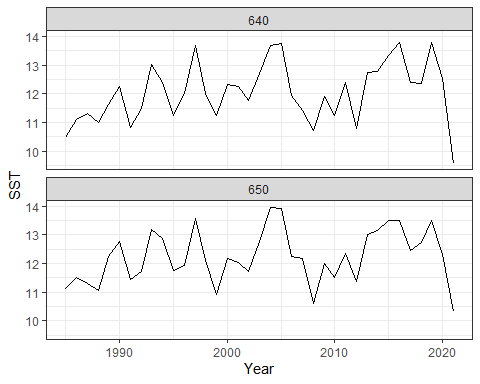
**R Code Input - 1**

#Install packages  
library(httr) # For pulling data via a URL  
library(tidyverse) # Data manipulation  
library(lubridate) # Date formatting  
library(odbc) # For connecting to oracle database  
  
#Web service query of sea surface temperature for NMFS Areas 640.  
  
head(httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&start\_date=19850101&end\_date=20220101'),   
 type = "application/json") %>%   
 bind\_rows)

## # A tibble: 6 x 5  
## MEANSST NMFSAREA READ\_DATE YEAR JULIAN  
## <dbl> <chr> <chr> <chr> <chr>   
## 1 6.06 640 1985-01-01T12:00:00Z 1985 001   
## 2 5.89 640 1985-01-02T12:00:00Z 1985 002   
## 3 5.78 640 1985-01-03T12:00:00Z 1985 003   
## 4 5.76 640 1985-01-04T12:00:00Z 1985 004   
## 5 5.78 640 1985-01-05T12:00:00Z 1985 005   
## 6 5.92 640 1985-01-06T12:00:00Z 1985 006

**R Code Input - 2**

#Web service query of sea surface temperature for NMFS Areas 640 and 650, averaged for June, July, and August  
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640,650&start\_date=19850101&end\_date=20220101'),   
 type = "application/json") %>%   
 bind\_rows %>%   
 mutate(MONTH=month(as\_date(READ\_DATE))) %>% # Extract month   
 filter(MONTH==6 | MONTH==7 | MONTH==8) %>% # Filter summer months  
 group\_by(YEAR,NMFSAREA)%>%  
 summarize(SST=mean(MEANSST))%>% # Average by year and area.  
 ggplot(aes(as.numeric(YEAR),SST)) +   
 geom\_line() +   
 facet\_wrap(~NMFSAREA, nrow=2) +  
 xlab("Year") +  
 theme\_bw()



**R Code Output - 2.** Web service query of sea surface temperature for NMFS Areas 640 and 650, averaged for June, July, and August.

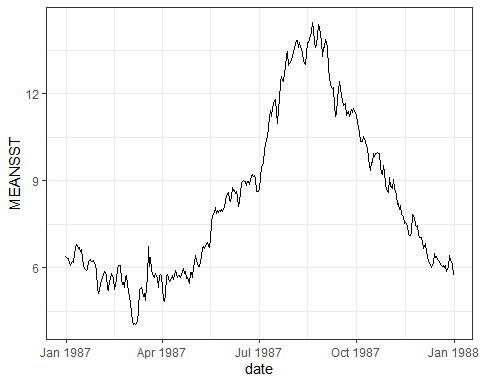
**R Code Input - 3**

#The full time series yields more than 13,000 rows of data per area (i.e., daily data from 1985-01-01 to present).   
  
data <- httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&start\_date=19850101&end\_date=20220101'),   
 type = "application/json") %>%   
 bind\_rows  
  
str(data)

## tibble [13,333 x 5] (S3: tbl\_df/tbl/data.frame)  
## $ MEANSST : num [1:13333] 6.06 5.89 5.78 5.76 5.78 5.92 6 5.92 5.85 5.83 ...  
## $ NMFSAREA : chr [1:13333] "640" "640" "640" "640" ...  
## $ READ\_DATE: chr [1:13333] "1985-01-01T12:00:00Z" "1985-01-02T12:00:00Z" "1985-01-03T12:00:00Z" "1985-01-04T12:00:00Z" ...  
## $ YEAR : chr [1:13333] "1985" "1985" "1985" "1985" ...  
## $ JULIAN : chr [1:13333] "001" "002" "003" "004" ...

**R Code Input - 4**

#SST in NMFS area 640 in 1987.   
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&start\_date=19870101&end\_date=19880101'),   
 type = "application/json") %>%   
 bind\_rows %>%   
 mutate(date=as\_date(READ\_DATE)) %>%   
 ggplot(aes(date,MEANSST)) +   
 geom\_line()+  
 theme\_bw()



**R Code Input - 5**

#You can query a specific date with "read\_date". For example SST in MFS 640 on Y2K.  
  
#Query the day after your date of interest because omitting the time component in read\_date misses that day's reading.  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&read\_date=20000101'),   
 type = "application/json") %>%   
 bind\_rows

## # A tibble: 1 x 5  
## MEANSST NMFSAREA READ\_DATE YEAR JULIAN  
## <dbl> <chr> <chr> <chr> <chr>   
## 1 4.94 640 1999-12-31T12:00:00Z 1999 365

**R Code Input - 6**

#You can specify a number of days prior to any date using a "days\_back" parameter specification. For example the three days before Y2K.   
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&read\_date=20000101&days\_back=2'),   
 type = "application/json") %>%   
 bind\_rows

## # A tibble: 3 x 5  
## MEANSST NMFSAREA READ\_DATE YEAR JULIAN  
## <dbl> <chr> <chr> <chr> <chr>   
## 1 5.43 640 1999-12-29T12:00:00Z 1999 363   
## 2 5.13 640 1999-12-30T12:00:00Z 1999 364   
## 3 4.94 640 1999-12-31T12:00:00Z 1999 365

**R Code Input - 7**

#If "read\_date" is not specified, "days\_back" returns the most recent SSTs. Here are SSTs for the last three days in NMFS 640.  
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&days\_back=2'),   
 type = "application/json") %>%   
 bind\_rows

## # A tibble: 3 x 5  
## MEANSST NMFSAREA READ\_DATE YEAR JULIAN  
## <dbl> <chr> <chr> <chr> <chr>   
## 1 10.8 640 2021-07-02T12:00:00Z 2021 183   
## 2 11.0 640 2021-07-03T12:00:00Z 2021 184   
## 3 11.1 640 2021-07-04T12:00:00Z 2021 185

**R Code Input - 8**

#To query multiple areas, separate the values by a comma. For example to query NMFS areas 640 and 650 (Southeast Alaska outside waters).  
  
httr::content(  
 httr::GET("https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640,650"),   
 type = "application/json")%>%   
 bind\_rows

## # A tibble: 2 x 5  
## MEANSST NMFSAREA READ\_DATE YEAR JULIAN  
## <dbl> <chr> <chr> <chr> <chr>   
## 1 11.1 640 2021-07-04T12:00:00Z 2021 185   
## 2 11.6 650 2021-07-04T12:00:00Z 2021 185

**R Code Input - 9**

#View strata included in the lookup table  
lkp <- readRDS("Data/crwsst\_spatial\_lookup\_table.RDS")   
  
unique(lkp$Ecosystem\_sub)

## [1] NA "Northern Bering Sea"   
## [3] "Western Gulf of Alaska" "Eastern Gulf of Alaska"   
## [5] "Southeastern Bering Sea" "Eastern Aleutians"   
## [7] "Central Aleutians" "Western Aleutians"

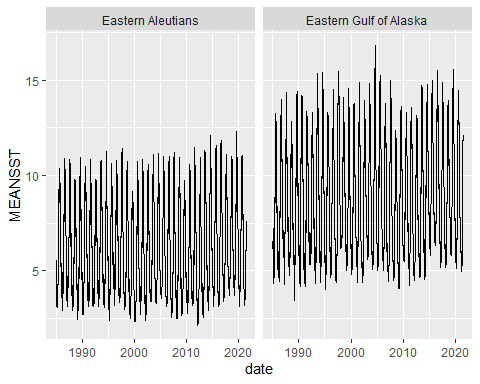
**R Code Input - 10**

#To query the data for the "Southeastern Bering Sea", for example, add "ecosystem\_sub=Southeastern%20Bering%20Sea", where spaces are filled by "%20".  
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/ecosystem\_sub\_crw\_avg\_sst?ecosystem\_sub=Southeastern%20Bering%20Sea'),   
 type = "application/json") %>%  
 bind\_rows

## # A tibble: 1 x 5  
## MEANSST ECOSYSTEM\_SUB READ\_DATE YEAR JULIAN  
## <dbl> <chr> <chr> <chr> <chr>   
## 1 7.83 Southeastern Bering Sea 2021-07-04T12:00:00Z 2021 185

**R Code Input - 11**

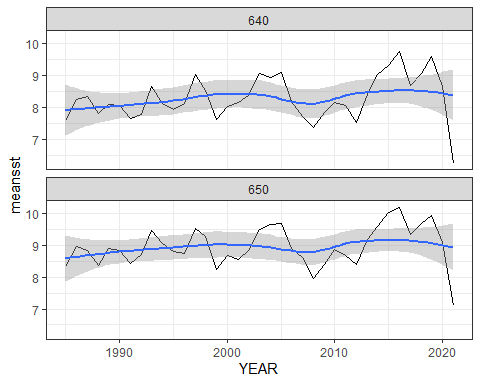
#Eastern GOA and Eastern Aleutians SST from 1985 - Present."  
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/ecosystem\_sub\_crw\_avg\_sst?ecosystem\_sub=Eastern%20Gulf%20of%20Alaska,Eastern%20Aleutians&start\_date=19850101&end\_date=20220101'),   
 type = "application/json") %>%   
 bind\_rows %>%   
 mutate(date=as\_date(READ\_DATE)) %>%   
 ggplot(aes(date,MEANSST)) +   
 geom\_line() +   
 facet\_wrap(~ECOSYSTEM\_SUB)



Eastern GOA and Eastern Aleutians SST from 1985 - Present.

**R Code Input - 12**

#"Annual average SST for NMFS areas 640 and 650."  
  
httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640,650&start\_date=19850101&end\_date=20220101'),   
 type = "application/json") %>%   
 bind\_rows %>%   
 mutate(date=as\_date(READ\_DATE),  
 YEAR=as.numeric(YEAR)) %>%   
 group\_by(YEAR,NMFSAREA) %>%   
 summarise(meansst=mean(MEANSST)) %>%   
 ggplot(aes(YEAR,meansst)) +   
 geom\_line() +   
 geom\_smooth() +  
 facet\_wrap(~NMFSAREA, nrow=2)+  
 theme\_bw()



Annual average SST for NMFS areas 640 and 650.

**R Code Input - 13**

#Marine heatwave calculation (Schlegel et al. 2018) for NMFS region 640.  
#Step 1: Get SST data using httr  
  
updateddata <- httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/nmfs\_area\_crw\_avg\_sst?nmfs\_area=640&start\_date=19850101&end\_date=20211231'),   
 type = "application/json") %>%   
 bind\_rows %>%   
 mutate(date=as\_date(READ\_DATE)) %>%   
 data.frame %>%   
 dplyr::select(date,  
 meansst=MEANSST,  
 NMFSAREA) #simplify data frame for clarity.

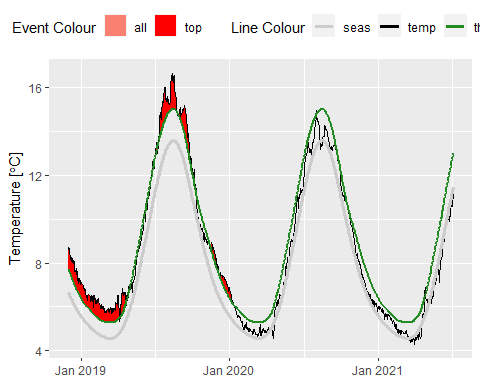
**R Code Input - 14**

#Step 2: Calculate marine heatwave indices using the mhw package.  
library(heatwaveR)  
mhw <- detect\_event(  
 ts2clm(updateddata %>%   
 rename(t=date,  
 temp=meansst) %>%   
 arrange(t),  
 climatologyPeriod = c("1985-01-01", "2014-12-31"))) #Specify baseline period

**R Code Input - 15**

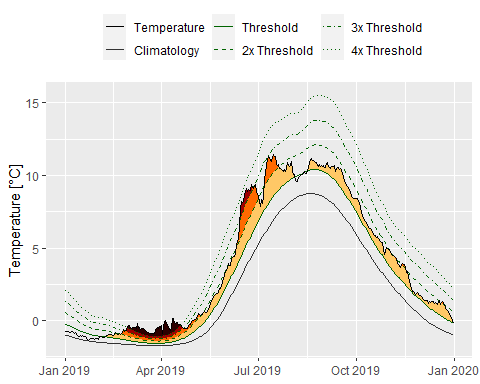
#Create "flame" plots.  
# Plotting code directly from heatwaveR vignette.  
mhw\_clim <- mhw$climatology %>%   
 filter(t>=as.Date("2018-12-01")) #Extract the MHW data  
  
ggplot(data = mhw\_clim,   
 aes(x = t)) +  
 geom\_flame(aes(y = temp, y2 = thresh, fill = "all"),  
 show.legend = T) +  
 geom\_flame(data = mhw\_clim,   
 aes(y = temp,y2 = thresh, fill = "top"),   
 show.legend = T) +  
 geom\_line(aes(y = temp, colour = "temp")) +  
 geom\_line(aes(y = thresh, colour = "thresh"), size = 1.0) +  
 geom\_line(aes(y = seas, colour = "seas"), size = 1.2) +  
 scale\_colour\_manual(name = "Line Colour",  
 values = c("temp" = "black",   
 "thresh" = "forestgreen",   
 "seas" = "grey80")) +  
 scale\_fill\_manual(name = "Event Colour",   
 values = c("all" = "salmon",   
 "top" = "red")) +  
 scale\_x\_date(date\_labels = "%b %Y") +  
 guides(colour = guide\_legend(  
 override.aes = list(fill = NA))) +  
 labs(y = expression(paste("Temperature [", degree, "C]")), x = NULL) +   
 theme(legend.position="top")

## Warning: Removed 1 rows containing missing values (geom\_flame).  
  
## Warning: Removed 1 rows containing missing values (geom\_flame).



**R Code Input - 16**

# A better illustration of heatwave categories from the Northern Bering Sea  
  
# Here we'll use an example where we do not save the SST data as a separate object first, simply embedding it into the ts2clm() function.  
  
clim\_cat <- (detect\_event(  
 ts2clm(  
 httr::content(  
 httr::GET('https://apex.psmfc.org/akfin/data\_marts/akmp/ecosystem\_sub\_crw\_avg\_sst?ecosystem\_sub=Northern%20Bering%20Sea&start\_date=19850101&end\_date=20211231'),  
 type = "application/json") %>%   
 bind\_rows %>%   
 mutate(date=as\_date(READ\_DATE)) %>%   
 data.frame %>%   
 dplyr::select(t=date,temp=MEANSST) %>%   
 arrange(t),   
 climatologyPeriod = c("1985-01-01", "2014-12-31"))))$climatology %>% #Specify the baseline period.  
 dplyr::mutate(diff = thresh - seas,  
 thresh\_2x = thresh + diff,  
 thresh\_3x = thresh\_2x + diff,  
 thresh\_4x = thresh\_3x + diff) %>%   
 filter(t>=as.Date("2019-01-01") & t<=as.Date("2019-12-31")) # Select the time period to display.  
  
# Plotting code directly from heatwaveR vignette.  
# Set line colours  
lineColCat <- c(  
 "Temperature" = "black",  
 "Climatology" = "gray20",  
 "Threshold" = "darkgreen",  
 "2x Threshold" = "darkgreen",  
 "3x Threshold" = "darkgreen",  
 "4x Threshold" = "darkgreen"  
)  
  
# Set category fill colours  
fillColCat <- c(  
 "Moderate" = "#ffc866",  
 "Strong" = "#ff6900",  
 "Severe" = "#9e0000",  
 "Extreme" = "#2d0000"  
)  
  
ggplot(data = clim\_cat,   
 aes(x = t, y = temp)) +  
 geom\_flame(aes(y2 = thresh, fill = "Moderate")) +  
 geom\_flame(aes(y2 = thresh\_2x, fill = "Strong")) +  
 geom\_flame(aes(y2 = thresh\_3x, fill = "Severe")) +  
 geom\_flame(aes(y2 = thresh\_4x, fill = "Extreme")) +  
 geom\_line(aes(y = thresh\_2x, col = "2x Threshold"),  
 size = 0.7, linetype = "dashed") +  
 geom\_line(aes(y = thresh\_3x, col = "3x Threshold"),   
 size = 0.7, linetype = "dotdash") +  
 geom\_line(aes(y = thresh\_4x, col = "4x Threshold"),   
 size = 0.7, linetype = "dotted") +  
 geom\_line(aes(y = seas,col = "Climatology"), size = 0.7) +  
 geom\_line(aes(y = thresh,col = "Threshold"), size = 0.7) +  
 geom\_line(aes(y = temp,col = "Temperature"), size = 0.6) +  
 scale\_colour\_manual(name = NULL,   
 values = lineColCat,  
 breaks = c("Temperature",   
 "Climatology",  
 "Threshold",  
 "2x Threshold",   
 "3x Threshold",  
 "4x Threshold")) +  
 scale\_fill\_manual(name = NULL,   
 values = fillColCat,   
 guide = FALSE) +  
 scale\_x\_date(date\_labels = "%b %Y") +  
 guides(colour = guide\_legend(  
 override.aes = list(linetype = c("solid",   
 "solid",   
 "solid",  
 "dashed",   
 "dotdash",   
 "dotted"),  
 size = c(0.6, 0.7, 0.7,0.7, 0.7, 0.7)))) +  
 labs(y = "Temperature [°C]", x = NULL) +  
 theme(legend.position="top")



**Oracle database queries**

**R Code Input - 17**

#Connect to the AKFIN database with R  
  
# Load the AKFIN database user name and password from an external file.  
params <- read\_csv("markdown\_odbc\_params.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## uid = col\_character(),  
## pass = col\_character()  
## )

# Connect to the AKFIN database  
con <- dbConnect(odbc::odbc(), "akfin", UID=params$uid, PWD=params$pass)

**SQL Code Input - 1**

-- Spatial lookup table on the AKFIN Oracle database  
  
select \* from afsc.erddap\_crw\_sst\_spatial\_lookup   
where rownum<=5

5 records

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | STATE\_FED | STAT\_AREA | DEPTH | LONGITUDE | LATITUDE | NMFSAREA | BSIERP\_ID | BSIERP\_NAME | ECOSYSTEM | ECOSYSTEM\_SUB | CRAB |
| 58040 | FED | 525902 | -126 | -152.175 | 59.175 | 630 | NA | NA | Gulf of Alaska | Western Gulf of Alaska | NA |
| 58041 | FED | 525902 | -130 | -152.125 | 59.175 | 630 | NA | NA | Gulf of Alaska | Western Gulf of Alaska | NA |
| 58042 | FED | 525902 | -119 | -152.075 | 59.175 | 630 | NA | NA | Gulf of Alaska | Western Gulf of Alaska | NA |
| 58043 | FED | 525902 | -108 | -152.025 | 59.175 | 630 | NA | NA | Gulf of Alaska | Western Gulf of Alaska | NA |
| 58044 | STATE | 515906 | -115 | -151.975 | 59.175 | 630 | NA | NA | Gulf of Alaska | Western Gulf of Alaska | NA |

**SQL Code Input - 2**

-- CRW SST from the AKFIN Oracle database  
  
select \* from afsc.erddap\_crw\_sst   
where rownum<=5

5 records

|  |  |  |
| --- | --- | --- |
| READ\_DATE | CRW\_ID | TEMP |
| 1985-04-03 12:00:00 | 66728 | -1.72 |
| 1985-04-03 12:00:00 | 66729 | -1.72 |
| 1985-04-03 12:00:00 | 66730 | -1.71 |
| 1985-04-03 12:00:00 | 66731 | -1.70 |
| 1985-04-03 12:00:00 | 66732 | -1.69 |

**SQL Code Input - 3**

-- query SST (“TEMP”) data that fall within a crab management area and add a field for “Year”  
  
select read\_date,  
 temp,  
 to\_char(read\_date,'YYYY') as Year,  
 crab  
from afsc.erddap\_crw\_sst a  
INNER JOIN (select \* from afsc.erddap\_crw\_sst\_spatial\_lookup  
where crab <> 'NA') b  
ON a.crw\_id =b.id  
where rownum<=5

5 records

|  |  |  |  |
| --- | --- | --- | --- |
| READ\_DATE | TEMP | YEAR | CRAB |
| 2017-01-05 12:00:00 | 3.83 | 2017 | stmatts |
| 2017-01-06 12:00:00 | 3.72 | 2017 | stmatts |
| 2017-01-07 12:00:00 | 3.64 | 2017 | stmatts |
| 2017-01-08 12:00:00 | 4.36 | 2017 | stmatts |
| 2017-01-13 12:00:00 | 3.67 | 2017 | stmatts |

**R / SQL Code Input - 1**

#Plotted query of Bristol Bay crab management area SST data averaged daily and plotted with default smoothing.  
  
dbFetch(dbSendQuery(con,  
 paste0("select read\_date,  
 round(avg(temp),2) as sst,  
 crab  
 from afsc.erddap\_crw\_sst a  
 INNER JOIN (select \*   
 from afsc.erddap\_crw\_sst\_spatial\_lookup  
 where crab = 'bb') b  
 ON a.crw\_id = b.id   
 group by   
 crab,  
 read\_date"))) %>%   
 ggplot(aes(READ\_DATE,SST)) +   
 geom\_line() +   
 geom\_smooth()

## `geom\_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'

