PCB 5443 Exam 2

James Sturges

19 March 2023

This script uses long-term climate data from the Palmer Antarctic LTER site to look at the relationship between mean monthly temperature and chlorophyll a (Chl a) concentrations. We know that there are strong seasonal pulses in the Antarctic region and expect to see Chl a spike in the summer months when light availability is no longer a limiting nutrient. While this relationship is fairly obvious we want to know how well this relationship coincides with Adelie penguin reproductive success annually. Do greater amounts of primary productivity lead to greater prey resource availability for the penguins or are there other drivers limiting their reproductive success? Furthermore, what are the time lags in the process? Do high levels of primary productivity take time to translate to penguin reproduction?

Data Preparation Read in libraries for time series and cross-wavelet analyses

```
library(tidyverse)
library(ggpubr)
library(lubridate)

### TIME SERIES ANALYSES
library(forecast)
library(tseries)

### CROSS-WAVELET ANALYSES
library(WaveletComp)
library(dplyr)
library(matrixStats)
library(ggplot2)
```

Read in publicly available Palmer LTER data sets

```
weather = read.csv("table_189.csv")
#https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-pal.189.8

chla = read.csv("D197_PumpHouseChlorophyll_1991thru2021.csv")
#https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-pal.197.8

adel = read.csv("D93_AdeliePenguinReproductionSuccess.csv")
#https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-pal.93.7
```

Confine weather data set to 1991-2021 to align with the penguin reproduction study

```
weather = weather %>%
  filter(Year >= "1991")
```

Penguin reproduction data spans two calendar years (example 2000-2001 breeding season). I am using the end of the season to represent success for that year (values in 1991-1992 get classified as reproduction for the year 1992)

Manually fixes poor metadata

```
adel = adel %>%
  mutate(Year = case_when(
    startsWith(studyName, "PAL9192") ~ "1992",
    startsWith(studyName, "PAL9293") ~ "1993",
    startsWith(studyName, "PAL9394") ~ "1994",
    startsWith(studyName, "PAL9495") ~ "1995",
    startsWith(studyName, "PAL9596") ~ "1996",
    startsWith(studyName, "PAL9697") ~ "1997",
    startsWith(studyName, "PAL9798") ~ "1998",
    startsWith(studyName, "PAL9899") ~ "1999",
    startsWith(studyName, "PAL9900") ~ "2000",
    startsWith(studyName, "PAL0001") ~ "2001",
    startsWith(studyName, "PAL0102") ~ "2002",
    startsWith(studyName, "PAL0203") ~ "2003",
    startsWith(studyName, "PAL0304") ~ "2004",
    startsWith(studyName, "PAL0405") ~ "2005",
    startsWith(studyName, "PAL0506") ~ "2006",
    startsWith(studyName, "PAL0607") ~ "2007",
    startsWith(studyName, "PAL0708") ~ "2008",
    startsWith(studyName, "PAL0809") ~ "2009",
    startsWith(studyName, "PAL0910") ~ "2010",
    startsWith(studyName, "PAL1011") ~ "2011",
    startsWith(studyName, "PAL1112") ~ "2012",
    startsWith(studyName, "PAL1213") ~ "2013",
    startsWith(studyName, "PAL1314") ~ "2014",
    startsWith(studyName, "PAL1415") ~ "2015",
    startsWith(studyName, "PAL1516") ~ "2016",
    startsWith(studyName, "PAL1617") ~ "2017",
    startsWith(studyName, "PAL1718") ~ "2018",
    startsWith(studyName, "PAL1819") ~ "2019",
    startsWith(studyName, "PAL1920") ~ "2020",
    startsWith(studyName, "PAL2021") ~ "2021"))
```

Reclassifies variables for plotting purposes

```
chla = chla %>%
  mutate(Date = ymd(Date)) %>%
  mutate_at(vars(Date), funs(year, month, day))
```

```
## Warning: `funs()` was deprecated in dplyr 0.8.0.
## i Please use a list of either functions or lambdas:
##
## # Simple named list: list(mean = mean, median = median)
##
```

```
## # Auto named with `tibble::lst()`: tibble::lst(mean, median)
##
## # Using lambdas list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
```

```
chla$Chlorophylla = as.numeric(chla$Chlorophylla)
```

Takes an aggregate monthly Chl a value

```
chla_y_m = chla %>%
  select(!c(Date,day)) %>%
  group_by(year, month) %>%
  summarise(mean.chla = mean(Chlorophylla))
```

Selects only the climate variables we need

```
weather = weather %>%
  select(c(Year, Month, Average.Temperature..C.))
```

Fixes inconsistent column naming scheme

```
chla_y_m = chla_y_m %>%
  rename(Year = year, Month = month)
```

Joins climate and chl a dat to single dataframe

```
chla_temp = weather %>%
  left_join(chla_y_m)
```

I hate working with dates

```
chla_temp = chla_temp %>%
  mutate(day1 = "01")

 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "12","Dec"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "11","Nov"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "10","Oct"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "9","09"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "8","08"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "7","07"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "6","06"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "5","05"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "4","04"))
```

```
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "3","03"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "2","02"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "1","01"))
  chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "Dec","12"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "Nov","11"))
 chla_temp = chla_temp %>%
   mutate(Month = str_replace(Month, "Oct","10"))

  chla_temp = chla_temp %>%
   mutate(Monthlab = Month)

  chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "12","Dec"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "11","Nov"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "10","Oct"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "09","Sep"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "08","Aug"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "07","Jul"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "06","Jun"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "05","May"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "04","Apr"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "03","Mar"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "02","Feb"))
 chla_temp = chla_temp %>%
   mutate(Monthlab = str_replace(Monthlab, "01","Jan"))
  chla_temp = chla_temp

 chla_temp = chla_temp %>%
  mutate(yyyymmdd = paste(Year,Month,day1,sep = ""))

chla_temp$yyyymmdd <- ymd(chla_temp$yyyymmdd)
```

Creates a unique ID of Month and Year

```
chla_temp = chla_temp %>%
  mutate(y_m = paste(Monthlab, Year, sep = " "))
```

Estimates mean and standard deviation of chl a data

```r
mean(chla_temp$mean.chla, na.rm = T)
```

```
## [1] 1.996992
```

```r
sd(chla_temp$mean.chla, na.rm = T)
```

```
## [1] 3.649089
```

Transforms temperature data for plotting purposes

```r
chla_temp = chla_temp %>%
  mutate(postemp = Average.Temperature..C. + 12.43)
```

Analyzing Trends Plots mean monthly chl a concentrations and temperature from 1991-2021

```r
chla_1 = ggplot()+
  geom_line(data = chla_temp, aes(x = factor(y_m, levels = unique(y_m)), y = postemp*1), color = "blue"
  geom_line(data = chla_temp, aes(x = factor(y_m, levels = unique(y_m)), y = mean.chla), color = "black
  geom_point(data = chla_temp, aes(x = factor(y_m, levels = unique(y_m)), y = mean.chla, na.rm = T), co
  geom_text(data=subset(chla_temp, mean.chla > 9.3),
            aes(y_m,mean.chla,label=y_m, vjust = -.8, hjust = .9, angle = 30), size = 3, color = "red")
  # geom_point(data = chla_temp, aes(x = factor(y_m, levels = unique(y_m)), y = Average.Temperature..C.
  scale_y_continuous(name = "Mean Chl a Concentration [  ]", sec.axis = sec_axis(~./1, name = "Temperatu
  theme_classic() +
  scale_x_discrete(expand = c(0, 0)) +
  theme(axis.ticks = element_blank(), axis.text.x = element_blank()) +
  labs(x = "Time (1991 - 2021)")


chla_1
```

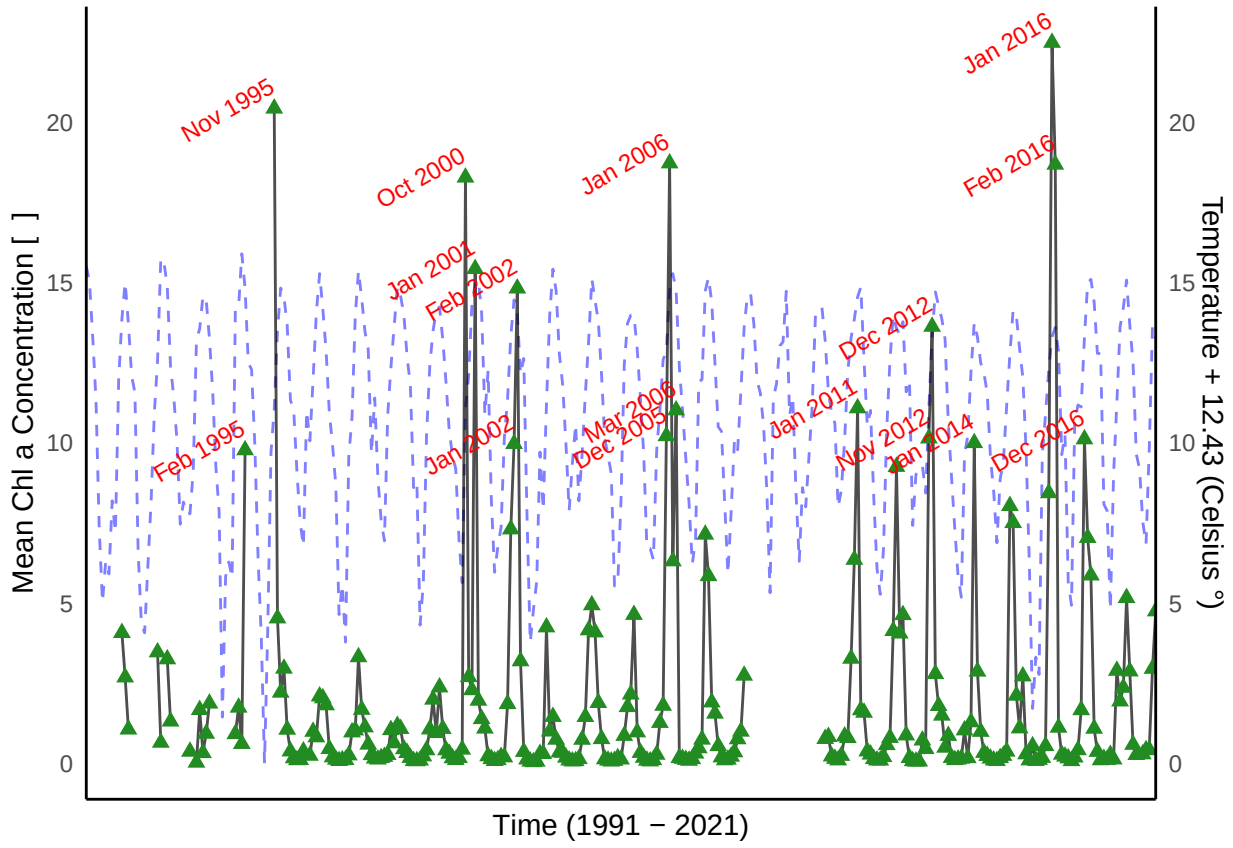


Figure 1. Mean monthly chlorophyll a concentrations (green triangles) and mean monthly temperature (blue line) in Palmer Station Antarctica from 1991 to 2021. Temperatures are transformed with an addition of 12.43 degrees so that the lowest temperature recorded is scaled to zero. Mean monthly chlorophyll a values that were greater than 2 standard deviations (3.65) above the mean (1.99) and indicated with a red label describing the month and year of significantly higher chlorophyll a levels. Chlorphyll a values were not recorded during 2008.

```
adel = adel %>%
  mutate(hatchtime1 = Chick.1.Hatch.Date - Egg.1.Lay.Date,
         hatchtime2 = Chick.2.Hatch.Date - Egg.1.Lay.Date)

adel = adel %>%
  select(studyName,Egg.1.Lay.Date,Egg.2.Lay.Date,Chick.1.Hatch.Date,Chick.2.Hatch.Date,hatchtime1,hatch

adel = adel %>%
  filter(hatchtime1 %in% (20:50),
         hatchtime2 %in% (22:48))

adel.hatch.t = adel %>%
  group_by(Year) %>%
  summarise(mean.hatch.y1 = mean(hatchtime1),
            mean.hatch.y2 = mean(hatchtime2),
            sd.h1 = sd(hatchtime1),
            sd.h2 = sd(hatchtime2))
```

```r
sig.yrs = read.csv("sig.chla.yrs.csv")

# sig.yrs = sig.yrs %>%
#   rename(Year = ï..Year)
```

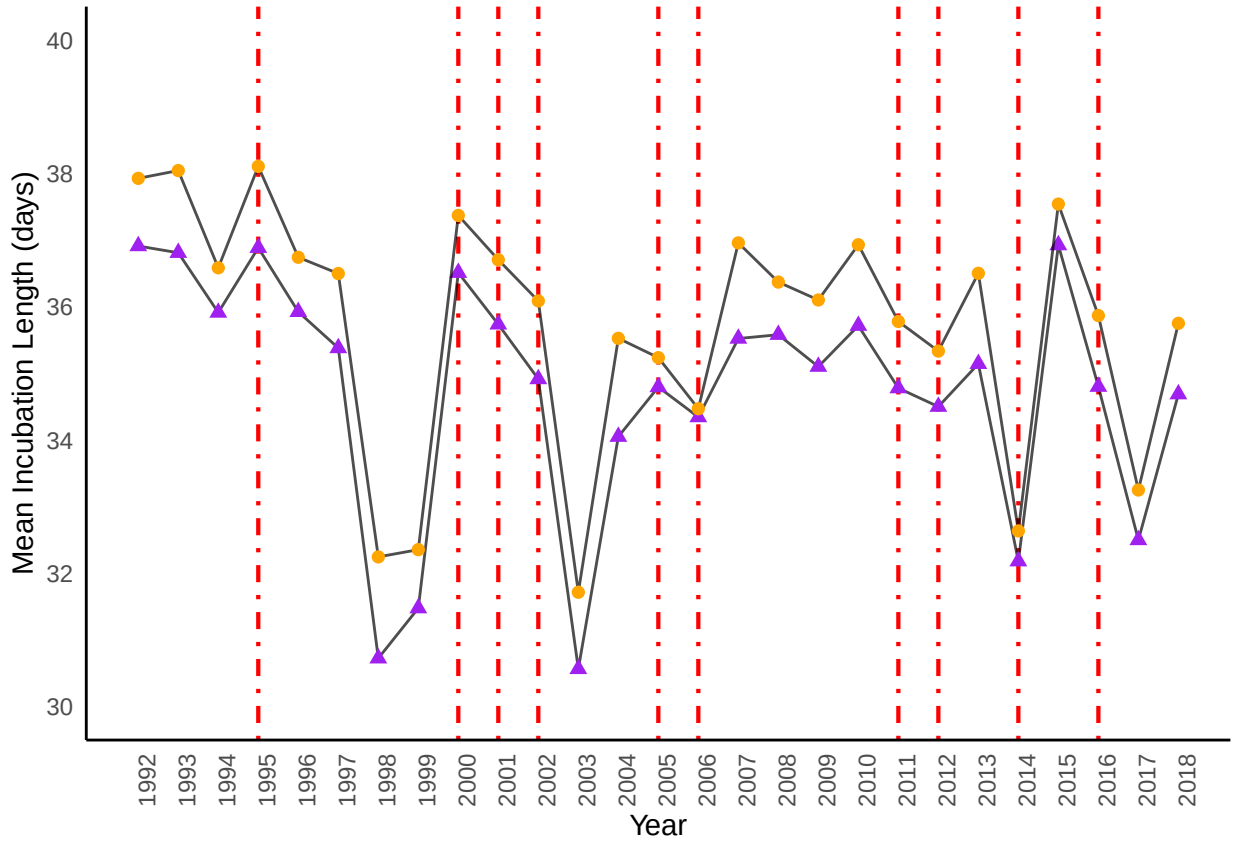
```r
adel.hatch.t$Year = as.numeric(adel.hatch.t$Year)
sig.yrs$Year = as.numeric(sig.yrs$Year)

adel.hatch.t = adel.hatch.t %>%
  left_join(sig.yrs)
```

```
## Joining, by = "Year"
```

```r
peng.eggs = adel.hatch.t %>%
  ggplot() +
  geom_vline(aes(xintercept = 1995), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2000), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2001), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2002), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2006), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2005), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2011), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2012), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2014), linetype=4, size=0.75, color="red")+
  geom_vline(aes(xintercept = 2016), linetype=4, size=0.75, color="red")+
  geom_line(aes(x = Year, y = mean.hatch.y1), color = "black", group = 1, alpha = 0.7) +
  geom_point(aes(x = Year, y = mean.hatch.y1), color = "purple", size = 2, shape = 17)+
  geom_line(aes(x = Year, y = mean.hatch.y2), color = "black", group = 1, alpha = 0.7) +
  geom_point(aes(x = Year, y = mean.hatch.y2), color = "orange", size = 2, shape = 16) +
  scale_x_continuous(breaks=seq(1992, 2018, 1)) +
  theme_classic()+
  scale_y_continuous(limits = c(30, 40), breaks = seq(30, 40, 2)) +
  theme(axis.ticks = element_blank(), axis.text.x = element_text(angle = 90))+
  labs(x = "Year", y = "Mean Incubation Length (days)")
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
```

```r
peng.eggs
```




```
  # geom_text(data=subset(year, mean.chla > 9.3),
  #           aes(y_m,mean.chla,label=y_m, vjust = -.8, hjust = .9, angle = 30), size = 3, color = "red
```

Figure 2)

Cross Wavelet Analysis Temp vs Chl a In order to complete a cross wavelet analysis of Temperature and Chl a we need to remove the gaps from the Chl a monthly data set

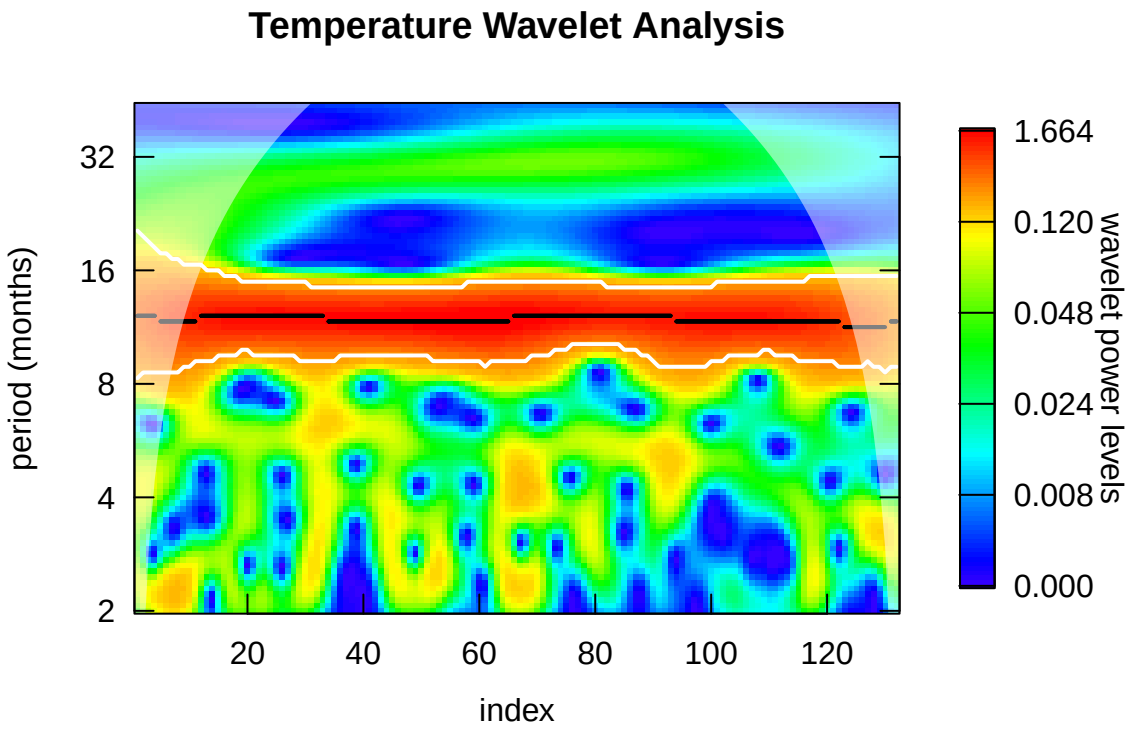
```
chla_temp_9607 = chla_temp %>%
  filter(Year %in% (1997:2007))
```

```
temp.waves <- analyze.wavelet(chla_temp_9607, "Average.Temperature..C.",make.pval = TRUE, n.sim = 100)
```

```
## Smoothing the time series...
## Starting wavelet transformation...
## ... and simulations...
##   |                                                                      |
```

```
wt.image(temp.waves, main = "Temperature Wavelet Analysis",
         periodlab = "period (months)",
         label.time.axis = T, show.date = T, date.format = "%Y/%m/%d",
         color.key = "quantile",legend.params = list(label.digits = 3, lab = "wavelet power levels", ma
```

```
## Warning in wt.image(temp.waves, main = "Temperature Wavelet Analysis", periodlab = "period (months)"
## Please check your calendar dates, format and time zone: dates may not be in an unambiguous format or
```

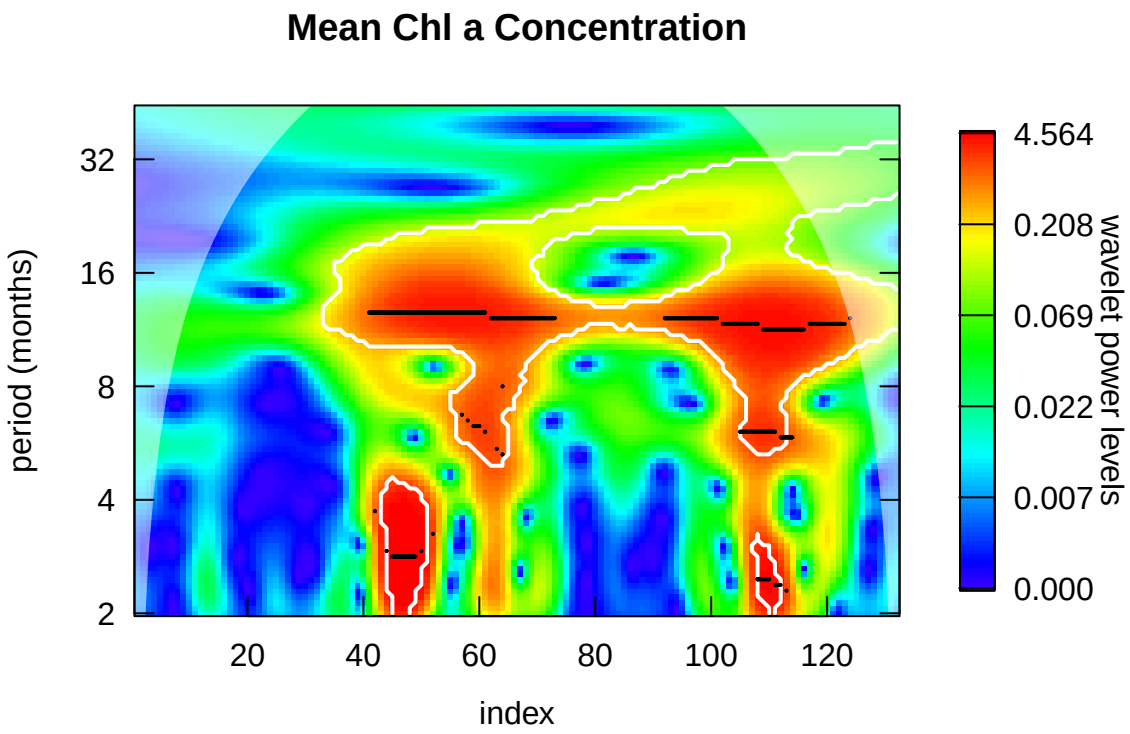



```
chla.waves <- analyze.wavelet(chla_temp_9607, "mean.chla",make.pval = TRUE, n.sim = 100)
```

```
## Smoothing the time series...
## Starting wavelet transformation...
## ... and simulations...
##   |                                                                      |
```

```
wt.image(chla.waves, main = "Mean Chl a Concentration",
         periodlab = "period (months)",
         label.time.axis = T, show.date = T, date.format = "%Y/%m/%d",
         color.key = "quantile",legend.params = list(label.digits = 3, lab = "wavelet power levels", ma
```

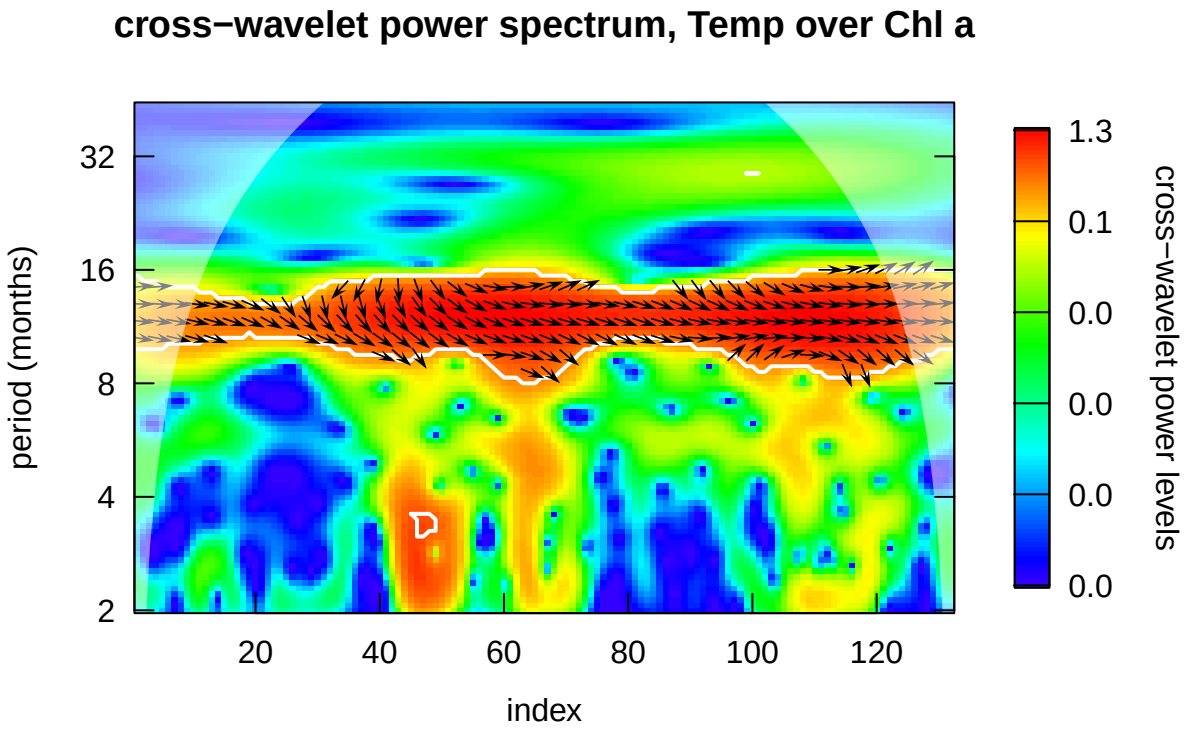
```
## Warning in wt.image(chla.waves, main = "Mean Chl a Concentration", periodlab = "period (months)", :
## Please check your calendar dates, format and time zone: dates may not be in an unambiguous format or
```



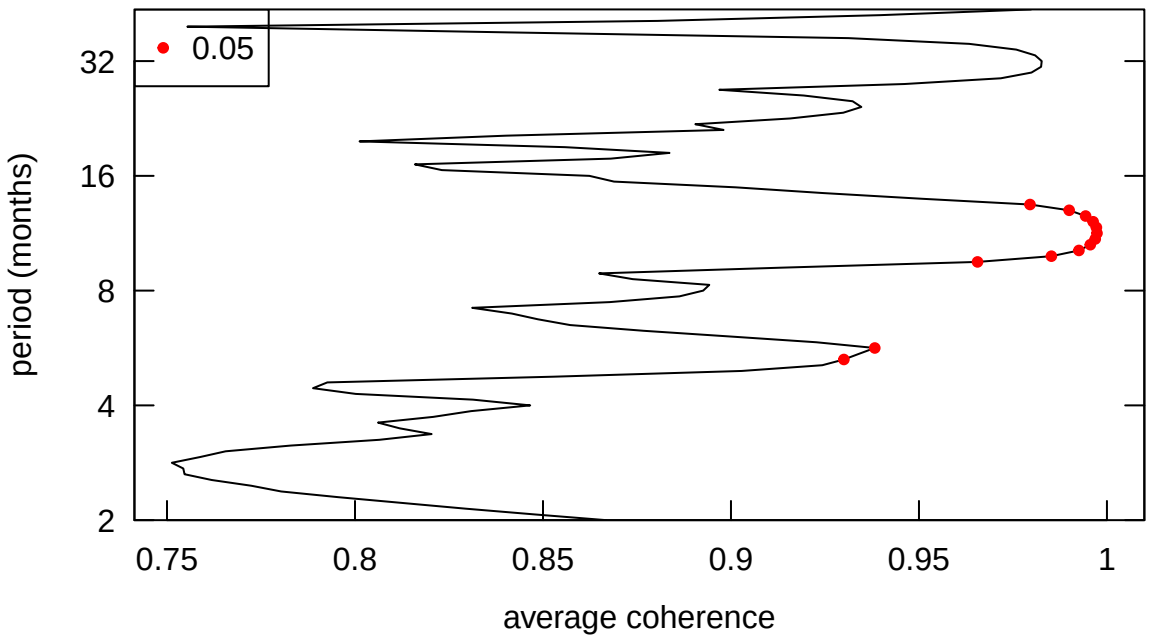

```
cross.waves <- analyze.coherency(chla_temp_9607, c("Average.Temperature..C.","mean.chla"), n.sim = 100)
```

```
## Smoothing the time series...
## Starting wavelet transformation and coherency computation...
## ... and simulations...
##   |                                                                      |
```

```
wc.image(cross.waves, main = "cross-wavelet power spectrum, Temp over Chl a",
         legend.params = list(lab = "cross-wavelet power levels"),
         periodlab = "period (months)")
```




```
# plot of average coherence:
wc.avg(cross.waves, which.avg = "wc",
       siglvl = 0.05, sigcol = 'red',
       legend.coords = "topleft",
       periodlab = "period (months)")
```



Cross Wavelet Analysis Chl a vs Incubation

```
adel.hatch.t.9607 = adel.hatch.t %>%
  filter(Year %in% (1997:2007))


chla_temp_9607_mon = chla_temp_9607 %>%
  filter(Monthlab %in% ("Jan"))

complete_time_series_df = chla_temp_9607_mon %>%
  left_join(adel.hatch.t.9607)
```

```
## Joining, by = "Year"
```

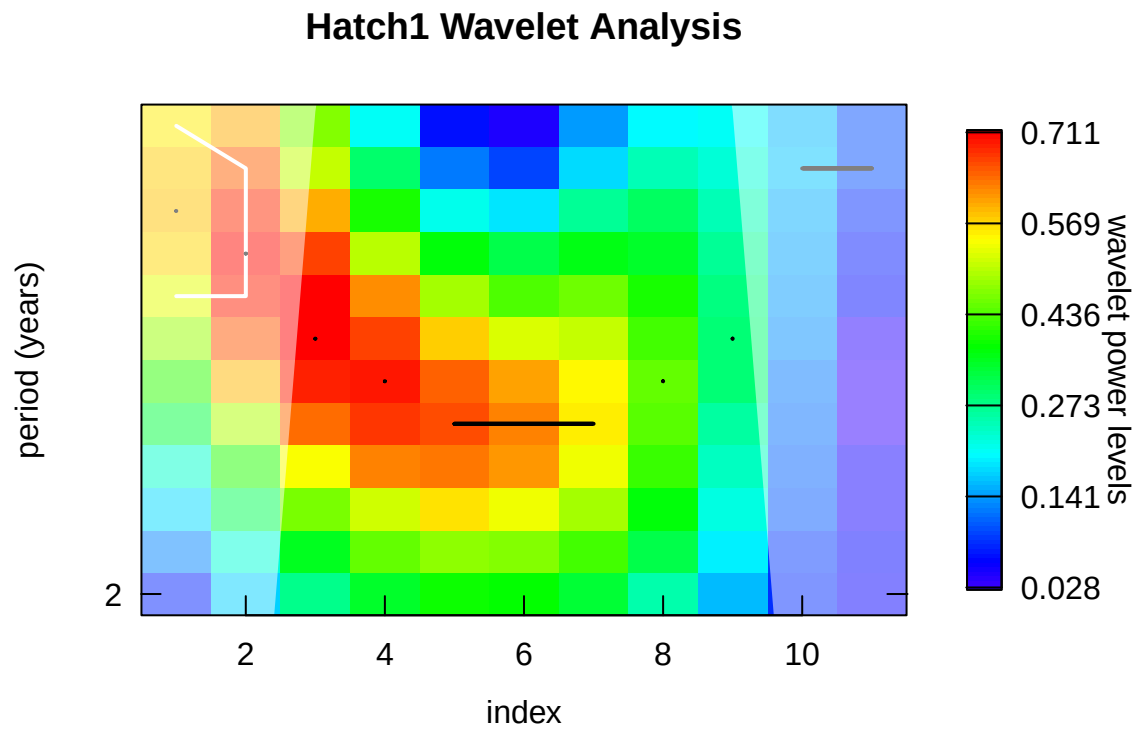
```
hatch1.waves <- analyze.wavelet(complete_time_series_df, "mean.hatch.y1",make.pval = TRUE, n.sim = 100)
```

```
## Smoothing the time series...
## Starting wavelet transformation...
## ... and simulations...
##   |                                                                      |
```

```
wt.image(hatch1.waves, main = "Hatch1 Wavelet Analysis",
         periodlab = "period (years)",
```

```
        label.time.axis = T, show.date = T, date.format = "%Y/%m/%d",
        color.key = "quantile",legend.params = list(label.digits = 3, lab = "wavelet power levels", ma
```

```
## Warning in wt.image(hatch1.waves, main = "Hatch1 Wavelet Analysis", periodlab = "period (years)", :
## Please check your calendar dates, format and time zone: dates may not be in an unambiguous format or
```



cross wavelet for hatch and chl a in progress

```
# cross.waves.2 <- analyze.coherency(complete_time_series_df, c("mean.chla", "mean.hatch.y1"), n.sim =
# wc.image(cross.waves.2, main = "cross-wavelet power spectrum, Hatch over Chl a",
#          legend.params = list(lab = "cross-wavelet power levels"),
#          periodlab = "period (years)")
#
# # plot of average coherence:
# wc.avg(cross.waves.2, which.avg = "wc",
#        siglvl = 0.05, sigcol = 'red',
#        legend.coords = "topleft",
#        periodlab = "period (year)")
```