Fish In Hot Water: Made for Chart Challenge

Ellie White

2023-04-03

Set up

Load libraries

```
# Load libraries
library(tidyverse)
library(readr)
library(scales)
library(gganimate)
library(scico)
```

Load files

Copy-pasted data from paper to a csv. Paper is here: https://afspubs.onlinelibrary.wiley.com/doi/full/10.1002/mcf2.10076

```
fish_data <- read_csv("in/fish_data.csv")
```

```
## Rows: 12 Columns: 11
## -- Column specification ---------------------------------------------------
## Delimiter: ","
## chr (3): species, period, variable
## dbl (8): RCP_26, RCP_26_SE, RCP_45, RCP_45_SE, RCP_60, RCP_60_SE, RCP_85, RC...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

This extra data came from the author in personal communications 04/03/2023

I modified the spreadsheet to be more R friendly

It's a time series record spanning 1950-2099 of fish spawning onset and cessation dates (modeled values)

```
fish_data_ts <- read_csv("in/fish_data_from_nack.csv")
```

```
## Rows: 1200 Columns: 8
## -- Column specification ---------------------------------------------------
## Delimiter: ","
## chr (4): species, condition, start_date, end_date
```

```
## dbl (4): year, start_date_julianday, end_date_julianday, duration
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```r
# pulled the 1951 values as origin dates and saved to this file for easy reference
# note that the year date is set to 2015 to be consistent with the rest of the reference dates but the
fish_data_ts[fish_data_ts$year == 1951, ]
```

```
## # A tibble: 8 x 8
##    species       condition  year start_date_jul~1 end_d~2 durat~3 start~4 end_d~5
##    <chr>         <chr>     <dbl>            <dbl>   <dbl>   <dbl> <chr>   <chr>
## 1 American Shad RCP_26     1951              116     181      65 26-Apr  30-Jun
## 2 Striped Bass  RCP_26     1951              129     189      60 9-May   8-Jul
## 3 American Shad RCP_45     1951              117     181      64 27-Apr  30-Jun
## 4 Striped Bass  RCP_45     1951              130     188      58 10-May  7-Jul
## 5 American Shad RCP_60     1951              114     180      66 24-Apr  29-Jun
## 6 Striped Bass  RCP_60     1951              126     189      63 6-May   8-Jul
## 7 American Shad RCP_85     1951              118     180      62 28-Apr  29-Jun
## 8 Striped Bass  RCP_85     1951              130     188      58 10-May  7-Jul
## # ... with abbreviated variable names 1: start_date_julianday,
## #   2: end_date_julianday, 3: duration, 4: start_date, 5: end_date
```

```r
fish_data_origin_dates <- read_csv("in/fish_data_origin_dates.csv")
```

```
## Rows: 4 Columns: 6
## -- Column specification ---------------------------------------------------
## Delimiter: ","
## chr (6): species, variable, RCP_26, RCP_45, RCP_60, RCP_85
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

Get data ready for plotting

```r
# change to factors
fish_data <- fish_data |>
  mutate(species = factor(species, levels = c("American Shad", "Striped Bass")),
         variable = factor(variable, levels = c("Onset", "Cessation", "Duration")),
         period = factor(period, levels = c("Historical", "Future")))

# take out duration and confidence intervals
fish_data <- fish_data[fish_data$variable %in% c("Onset", "Cessation"), ]
fish_data_uncertainty <- select(fish_data, -c("RCP_26", "RCP_45", "RCP_60", "RCP_85")) # save this to b
fish_data <- select(fish_data, -c("RCP_26_SE", "RCP_45_SE", "RCP_60_SE", "RCP_85_SE"))

# prep origin dates
fish_data_origin_dates <- gather(fish_data_origin_dates, condition, origin_date, RCP_26:RCP_85)
fish_data_origin_dates$origin_date <- as.Date(fish_data_origin_dates$origin_date, format = "%m/%d/%Y")

# add in origin dates
```

```
fish_data_long <- gather(fish_data, condition, value, RCP_26:RCP_85)
fish_data_long <- full_join(fish_data_long, fish_data_origin_dates, by = c("species", "variable", "cond
fish_data_long$end_date <- fish_data_long$origin_date + fish_data_long$value

# prep uncertainty in dates to be brought in later
fish_data_uncertainty_long <- gather(fish_data_uncertainty, condition, confidence_interval, RCP_26_SE:R
fish_data_uncertainty_long$condition <- substr(fish_data_uncertainty_long$condition, 1, 6)

# make into wide format
fish_data_long <- fish_data_long |>
  select(-c(origin_date, value))
fish_data_wide <- spread(fish_data_long, variable, end_date)
names(fish_data_wide)[names(fish_data_wide)=="Onset"] <- "onset"
names(fish_data_wide)[names(fish_data_wide)=="Cessation"] <- "cessation"

# bring in uncertainty that was prepped before
fish_data_uncertainty_wide <- spread(fish_data_uncertainty_long, variable, confidence_interval)
names(fish_data_uncertainty_wide)[names(fish_data_uncertainty_wide)=="Onset"] <- "onset_se"
names(fish_data_uncertainty_wide)[names(fish_data_uncertainty_wide)=="Cessation"] <- "cessation_se"

fish_data_wide <- full_join(fish_data_wide, fish_data_uncertainty_wide, by = c("species", "period", "con
```

Now to add extra things for plotting purposes only:

```
# add in y location for segment plot
y_location <- tibble(condition = c("RCP_26", "RCP_45", "RCP_60", "RCP_85"),
                     y = rep(c(1:4)*2 + 5))
fish_data_wide <- full_join(fish_data_wide, y_location, by = c("condition"))

# offset the y locations and mess with spacing for historical vs. future so they slightly overlap but d
fish_data_wide <- fish_data_wide |>
  mutate(y_offset = ifelse(period == "Historical", y - 0.5, y))
```

Set up main plot

Theme:

```
theme_usgs <- function(legend.position = "right"){
  theme(
    plot.title = element_text(vjust = 3, size = 14, face = "bold", family="sans"),
    plot.subtitle = element_text(vjust = 3, size = 12,family="sans"),
    panel.border = element_rect(colour = "black", fill = NA, linewidth = 0.1),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.background = element_rect(fill = "white"),
    legend.background = element_blank(),
    legend.justification=c(0, 0),
    legend.position = legend.position,
    legend.key = element_blank(),
    legend.title = element_blank(),
    legend.text = element_text(size = 10),
    axis.title.x = element_text(size = 10, family="sans"),
```

```
        axis.title.y = element_text(vjust = 1, angle = 90, size = 9, family="sans"),
        axis.text.x = element_text(size = 10, vjust = -0.25, colour = "black",
                                   family="sans", margin=margin(10,5,20,5,"pt")),
        axis.text.y = element_text(size = 10, hjust = 1, colour = "black",
                                   family="sans", margin=margin(5,10,10,5,"pt")),
        axis.ticks = element_line(colour = "black", linewidth = 0.1),
        axis.ticks.length = unit(-0.25 , "cm")
  )
}
```

Produce plots

```
# Main base plot, trying out blurring edges
library(ggnewscale)
ggplot() +
  geom_segment(data = fish_data_wide[fish_data_wide$period == "Historical", ], aes(x = onset, xend = ces
  geom_segment(data = fish_data_wide[fish_data_wide$period == "Historical", ], aes(x = onset-onset_se, :
  new_scale_colour() +
  geom_segment(data = fish_data_wide[fish_data_wide$period == "Future", ], aes(x = onset, xend = cessati
  geom_segment(data = fish_data_wide[fish_data_wide$period == "Future", ], aes(x = onset-onset_se, xend
  scale_x_date(limits = c(as.Date("2015-01-01"), as.Date("2015-12-31")), date_breaks = "1 month", date_
  scale_y_continuous(limits = c(0, 13)) +
  coord_polar(theta = "x", direction = 1, start = -1.57*1.5) + # start is in radians, 90 Deg is Jan
  facet_wrap(~species) +
  labs(x = "",
       y = "",
       title = "FISH IN HOT WATER",
       subtitle = "Under projected climate change scenarios, the American Shad and Striped Bass of the
       # caption = "Data Source: Nack, C. et. al. (2019). https://doi.org/10.1002/mcf2.10076
       # Plot made by Ellie White, ewhite@usgs.gov 04/02/2023"
       ) +
  theme_bw()+
  theme(plot.title = element_text(vjust = 3, size = 14, face = "bold", family="sans"),
        plot.subtitle = element_text(vjust = 3, size = 12,family="sans"),
        axis.text.y = element_blank(),
        axis.ticks = element_blank(),
        panel.border = element_blank(),
        strip.background = element_blank(),
        strip.text.x = element_blank())
```
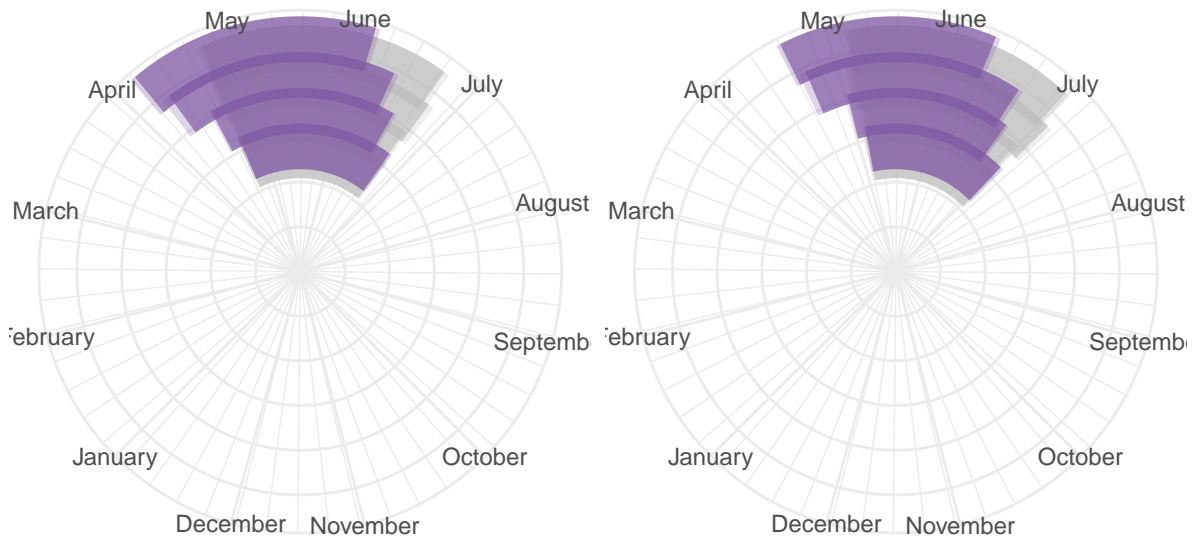
```
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
## Scale for y is already present.
## Adding another scale for y, which will replace the existing scale.
```

FISH IN HOT WATER

Under projected climate change scenarios, the American Shad and Striped Bass of t


```
ggsave("out/26_local-change_ewhite_base.png", width = 16, height = 9, units = "in", dpi = 1200)
```

Tests and Checks

```r
# is the mean what is reported for onset and cessation deviations in the linear modeling?
fish_data_ts_summarized <- fish_data_ts_long |>
  group_by(species, condition, period, spawning) |>
  summarize(mean_value = mean(value),
            se_value = sd(value)/sqrt(length(value))) # calculate standard error
```

```
## 'summarise()' has grouped output by 'species', 'condition', 'period'. You can
## override using the '.groups' argument.
```

```r
fish_data_summarized <- full_join(fish_data_long, fish_data_ts_summarized, by = join_by(species == spec
  mutate(date_diff = end_date - mean_value)

# It's not the mean value of the data, it's slightly off!
fish_data_summarized_summarized <- fish_data_summarized |>
  group_by(period, variable) |>
  summarize(mean(mean_value))
```

```
## 'summarise()' has grouped output by 'period'. You can override using the
## '.groups' argument.
```

```
# cessation
fish_data_summarized_summarized$`mean(mean_value)`[1] - fish_data_summarized_summarized$`mean(mean_valu
```

Time difference of 13.89881 days

```
# onset
fish_data_summarized_summarized$`mean(mean_value)`[2] - fish_data_summarized_summarized$`mean(mean_valu
```

Time difference of 9.543446 days

```
# never mind, we aren't going to recreate the author's analysis, going to take the paper values as they
```

Produce Gif

```
# make the yearly data wide so onset and cessation columns are side by side
fish_data_ts_wide <- spread(fish_data_ts_long, spawning, value)

# filter by RCP
fish_data_ts_plot <- fish_data_ts_wide |>
  filter(condition == "RCP_85") |>
  filter(year != 1950) |>
  mutate(year = as.integer(year))

animation <- ggplot(data = fish_data_ts_plot) +
  geom_segment(aes(x = Onset, xend = Cessation, y = year, yend = year, col = Onset), linewidth = 1 , sho
  scale_color_scico(palette = "tokyo") +
  scale_x_date(limits = c(as.Date("2015-01-01"), as.Date("2015-12-31")), date_breaks = "1 month", date_
  scale_y_continuous(limits = c(1900, 2120)) +
  coord_polar(theta = "x", direction = 1, start = -1.57*1.5) + # start is in radians, 90 Deg is Jan
  facet_wrap(~species) +
  labs(x = "",
       y = "",
       title = "FISH IN HOT WATER",
       subtitle = "RCP 8.5: Business-as-Usual.
       Year: {frame_time}
       "
       # caption = "Data Source: Nack, C. et. al. (2019). https://doi.org/10.1002/mcf2.10076
       # Plot made by Ellie White, ewhite@usgs.gov 04/02/2023"
       ) +
  theme_bw()+
  theme(plot.title = element_text(vjust = 3, size = 14, face = "bold", family="sans"),
        plot.subtitle = element_text(vjust = 3, size = 12,family="sans"),
        # axis.text.y = element_blank(),
        axis.ticks = element_blank(),
        panel.border = element_blank(),
        strip.background = element_blank(),
        strip.text.x = element_blank()) +
  # Here comes the gganimate specific bits
  transition_time(year) +
  shadow_mark(past = TRUE, future = FALSE, alpha = 0.3)
```

```
a <- animate(animation, end_pause = 30, width = 16, height = 9, units = "in", res = 300)
anim_save("out/26_local-change_ewhite.gif", a)
```

Supporting information

Key takeaways of this viz (1-2 sentences each)

1. The American Shad and the Striped Bass are migratory species needing both freshwater and marine habitats to complete their life cycle. This makes them particularly vulnerable to human activities. The Hudson River Shad has declined in stock so much that all its fisheries were closed in 2010. The Striped Bass, while declining in relative abundance, still remains the most important game fish in the Hudson River.

Data source(s)

Paper is here: https://afspubs.onlinelibrary.wiley.com/doi/full/10.1002/mcf2.10076

Citation: Nack, C. C., Swaney, D. P., & Limburg, K. E. (2019). Historical and projected changes in spawning Phenologies of American Shad and Striped bass in the Hudson River Estuary. Marine and Coastal Fisheries, 11(3), 271-284.

DOI: https://doi.org/10.1002/mcf2.10076

Process

1) produced `out/25_local-change_ewhite_base.png` with ggplot

2) made markups in PowerPoint

3) final plot is called `out/25_local-change_ewhite_final.png`