

Effects of Temperature Changes in the Indian Ocean on Australian Fires

Indian Fellas

30/01/2022

Fevzi Kaya - Halil Serdar

Project Proposal

As it is known, Australian forest fires occur at certain times each year. There is even talk of the ‘Bushfire Season’ in Australia. We will try to explain this disaster, in which millions of living species were affected, with climatic changes. The world is in a state of equilibrium, but this balance is breaking down day by day. This balance was also valid between South Africa and Australia, connected by the Indian Ocean. While there are fires on one side, there can be flood disasters on the other.

Project Goal

In this project:

We will examine the relationship between Indian Ocean temperature change and Australian bushfires. We will examine the Indian Ocean Dipole (temperature difference between the South African coast and the Australian coast). The 2019 Australian fire destroyed over 15 million hectares.

Data Documents

We couldn’t upload it to Github because the data is too large.

Data were obtained from the following sites.

You can view the Indian Ocean data **here**.

Australia data was created by collecting from the internet.

—

To access all data collectively **Drive Link**.

Variables we intend to use:

Surface temperature of the water, longitude, latitude, date, burned area

The codes of the operations performed are shown below.

```
library(tidyverse)
library(readr)
library(magrittr)
library(ggplot2)
library(readxl)
library(hrbrthemes)
library(babynames)
library(ggrepel)
library(tidyr)
library(ggthemes)
library(leaflet)
library(leaflet.extras)
```

```
indian_ocean_df <- read_csv("data/indian_ocean/indian_ocean_df.csv",
                           show_col_types = FALSE)

# I chose the columns that I could possibly use.
indian_ocean_df <- indian_ocean_df[,c('STATION', 'DATE', 'LATITUDE', 'LONGITUDE', 'AIR_TEMP',
                                       'DEW_PT_TEMP', 'SEA_SURF_TEMP')]
```

```
# I prefer the column names to be lowercase.
colnames(indian_ocean_df) <- tolower(colnames(indian_ocean_df))

# There are data at different times of the same day. We need monthly data.
# First I extracted the time data to be able to use group_by, just set it to be day/month.
# Thus, I can easily group_by. Because I will sort it.

indian_ocean_df['date'] = format(as.Date(as.POSIXct(indian_ocean_df$date)), "%Y%m")
indian_ocean_df$date <- as.integer(indian_ocean_df$date)
```

Let's start cleaning "NaN" data.

```
# [air_temp] column cleaning
sum(is.na(indian_ocean_df['air_temp']))
```

```
## [1] 1322311
```

```
# It is not healthy for the air_temp column of 1.3M rows to be "NaN" when I have 1.7M
# rows in total. This number is too high. So I'm removing the air_temp column.
indian_ocean_df <- subset(indian_ocean_df, select = -air_temp)
```

```
# [dew_pt_temp] column cleaning
sum(is.na(indian_ocean_df$dew_pt_temp))
```

```
## [1] 1416491
```

```
# It is not healthy for the dew_pt_temp column of 1.3M rows
# to be "NaN" when I have 1.7M rows in total. This number is too high.
# So I'm removing the air_temp column.
indian_ocean_df <- subset(indian_ocean_df, select = -dew_pt_temp)
```

```
# [sea_surf_temp] column cleaning
sum(is.na(indian_ocean_df$sea_surf_temp))
```

```
## [1] 266722
```

```
# It consists of 266,722 rows of "NaN" data in total. Considering that I have
# a large data with 1.7M rows, it seems to be healthier to remove
# these rows instead of filling them with average values.
# I think the data I have is sufficient.
indian_ocean_df <- indian_ocean_df[!(is.na(indian_ocean_df['sea_surf_temp'])),]
```

Classification of Regions

```
# I need to separate the coastal zones. In the [station] column,  
# the names of the regions are given, but they do not mean anything.  
# Because while giving the data, the explanations of these columns  
# were not made. So the regions I will separate them by longitude.  
# The [longitude] column has height data. If [longitude] > 90°,  
# I will consider it as the Australian coast, if [longitude] < 60°, I will  
# consider the South African coast.
```

```
indian_ocean_df %<>% mutate(indian_ocean_df,  
  station = ifelse(longitude > 105, "australian_coast",  
    ifelse(longitude < 50, "southAfrica_coast", "middle_zone"))  
)
```

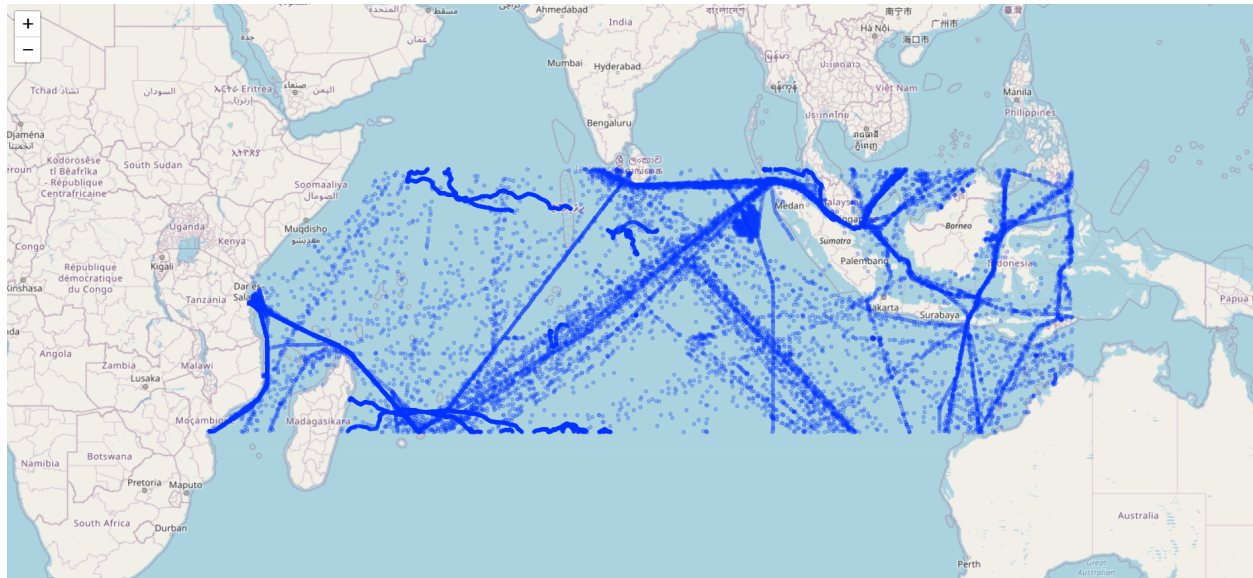
```
# I want to get a specific latitude range. We have -60° to 10°.  
# The lower part, namely -60° latitude, is very south.  
# To observe the Indian Ocean dipole, I will set this range between -20° and 10°.  
indian_ocean_df <- filter(indian_ocean_df, latitude >= -20)
```

```
# Date is taken from 2005 to 2020  
indian_ocean_df <- filter(indian_ocean_df, (date >= 201601 & date < 202000))
```

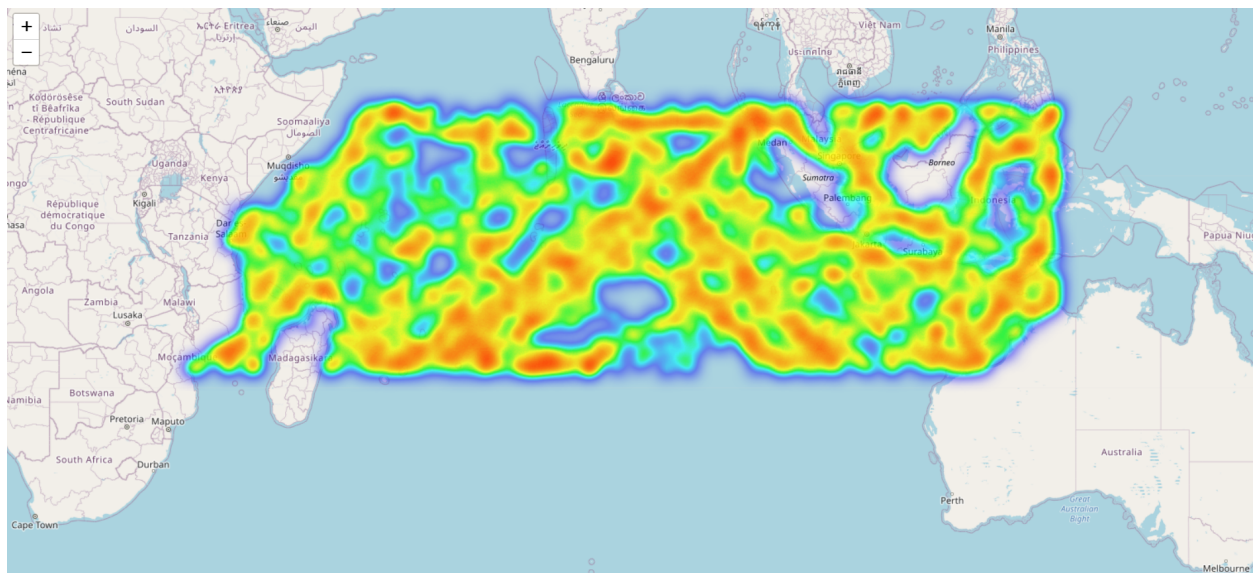
```
heatMap_df <- indian_ocean_df %>%  
  group_by(date, latitude, longitude) %>%  
  summarise_at(vars(sea_surf_temp), list(monthly_average_temperature = mean))
```

Investigated area

```
l <- leaflet(heatMap_df) %>%  
  addCircles(lng = ~longitude, lat = ~latitude, weight = 0.5,  
    radius = ~20000) %>%  
  addTiles()  
htmlwidgets::saveWidget(l, 'html/investigated_area.html')
```



```
l <- leaflet(heatMap_df) %>%
  addTiles() %>%
  addHeatmap(lng = ~longitude, lat = ~latitude,
    max=50, radius=10, blur=18, intensity = ~monthly_average_temperature)
htmlwidgets::saveWidget(l, 'html/2016_2020_heatMap.html')
```



```
# I created the regions as different dataframes.
middle_indian_ocean_df <- filter(indian_ocean_df, station == 'middle_zone')
south_africa_coast_df <- filter(indian_ocean_df, station == 'southAfrica_coast')
australian_coast_df <- filter(indian_ocean_df, station == 'australian_coast')
```

```
# There is more than one data belonging to different time of the same day.
# I found the monthly average temperature values from the daily temperature data.
# I just need date and average temperature data. I removed the other columns.
```

```

middle_indian_ocean_df <- middle_indian_ocean_df %>%
  group_by(date) %>%
  summarise_at(vars(sea_surf_temp), list(monthly_average_temperature = mean))
##
south_africa_coast_df <- south_africa_coast_df %>%
  group_by(date) %>%
  summarise_at(vars(sea_surf_temp), list(monthly_average_temperature = mean))
##
australian_coast_df <- australian_coast_df %>%
  group_by(date) %>%
  summarise_at(vars(sea_surf_temp), list(monthly_average_temperature = mean))

```

```

# Lost years will be filled with average values.
temp_df <- data.frame(date = c(201908, 201903, 201812, 201710, 201706))
temp_df["monthly_average_temperature"] <-
  mean(south_africa_coast_df$monthly_average_temperature)
south_africa_coast_df <- rbind(temp_df, south_africa_coast_df)

```

```

changeOf_temperature <- australian_coast_df$monthly_average_temperature -
  south_africa_coast_df$monthly_average_temperature

changeOf_temperature_df <- data.frame(changeOf_temperature)
changeOf_temperature_df["date"] <- australian_coast_df$date

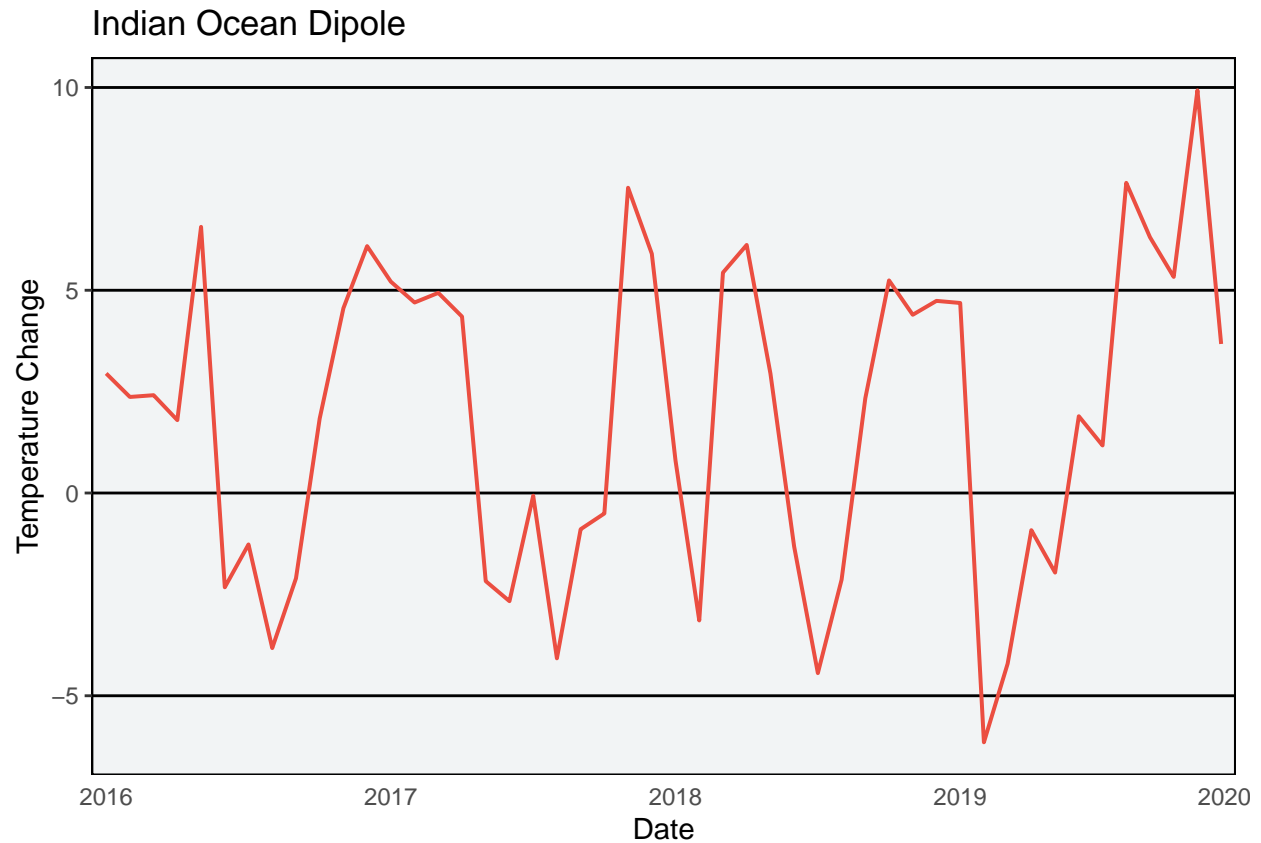
changeOf_temperature_df["date"] <- as.character(changeOf_temperature_df$date)

```

```

ggplot(changeOf_temperature_df, aes(x=date, y=changeOf_temperature, group=1)) +
  geom_line(color="#eb4e41", linetype = "solid", size=0.7) +
  theme_excel() +
  theme(panel.background = element_rect(fill = "#f2f4f5", colour = NA),
        axis.ticks.x=element_blank()) +
  labs(x="Date", y="Temperature Change") +
  scale_x_discrete(labels=c("201601"="2016", "201602"="", "201603"="", "201604"="",
    "201605"="", "201606"="", "201607"="", "201608"="", "201609"="",
    "201610"="", "201611"="", "201612"="", "201701"="2017", "201702"="",
    "201703"="", "201704"="", "201705"="", "201706"="", "201707"="",
    "201708"="", "201709"="", "201710"="", "201711"="", "201712"="",
    "201801"="2018", "201802"="", "201803"="", "201804"="",
    "201805"="", "201806"="", "201807"="", "201808"="", "201809"="",
    "201810"="", "201811"="", "201812"="", "201901"="2019", "201902"="",
    "201903"="", "201904"="", "201905"="", "201906"="", "201907"="",
    "201908"="", "201909"="", "201910"="", "201911"="", "201912"=" 2020")
  ) +
  ggtitle("Indian Ocean Dipole")

```

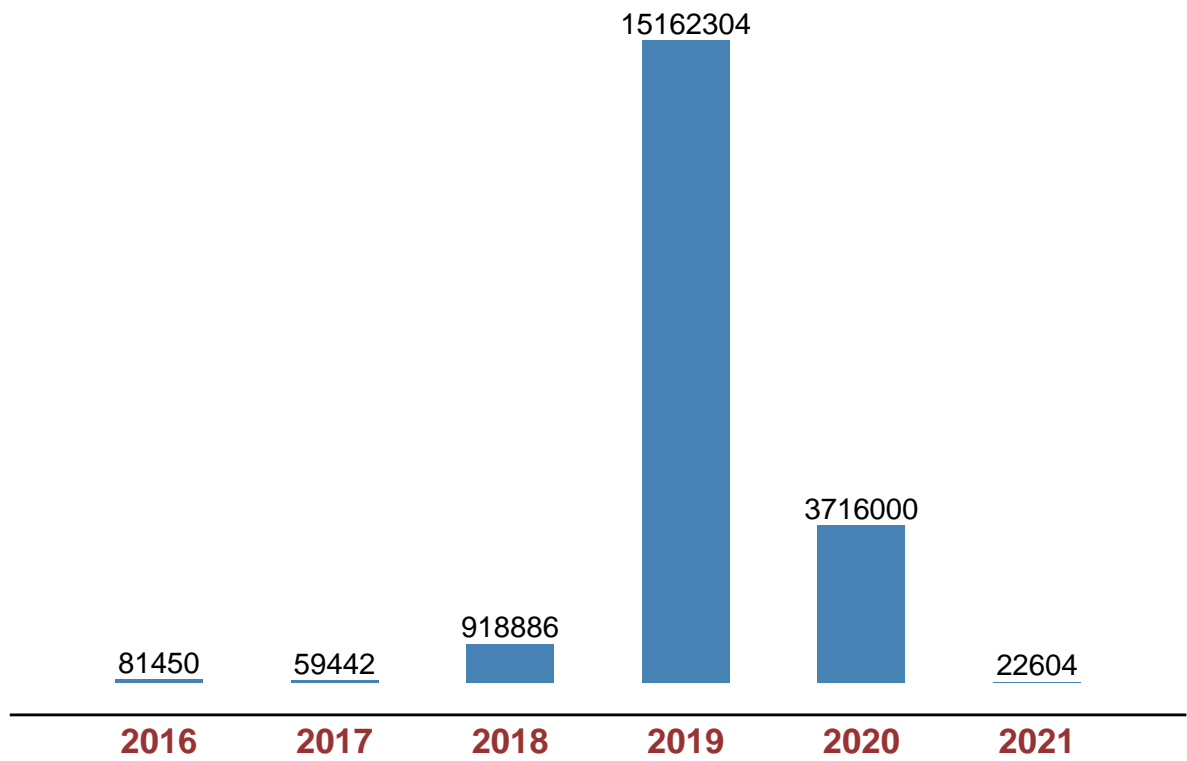


```
# Australian fire data
burned_area_df <- read_excel("data/australia/burnedforestarea.xlsx")
burned_area_df <- filter(burned_area_df, Tarih >= 2016)

# A graph showing the difference between fire in 2019 compared to other years
burned_area <- ggplot(data=burned_area_df, aes(x=Tarih, y=`Yanan Alan (Hektar)`)) +
  geom_bar(stat="identity", fill="steelblue", width=0.5) +
  theme_classic() +
  labs(x="", y="",
        title = "AUSTRALIAN FIRE TOTAL AMOUNT OF FIELDS BURNED IN HECTARES") +
  theme(plot.title = element_text(hjust=0.5, size=13),
        axis.ticks.y = element_blank(), axis.text.y = element_blank(),
        axis.line.y = element_blank(), axis.ticks.x=element_blank(),
        axis.text.x = element_text(face="bold", color="#993333",
                                    size=13)) +
  scale_x_discrete(limits=2016:2021) +
  geom_text(aes(label = `Yanan Alan (Hektar)`), vjust = -0.3)

burned_area
```

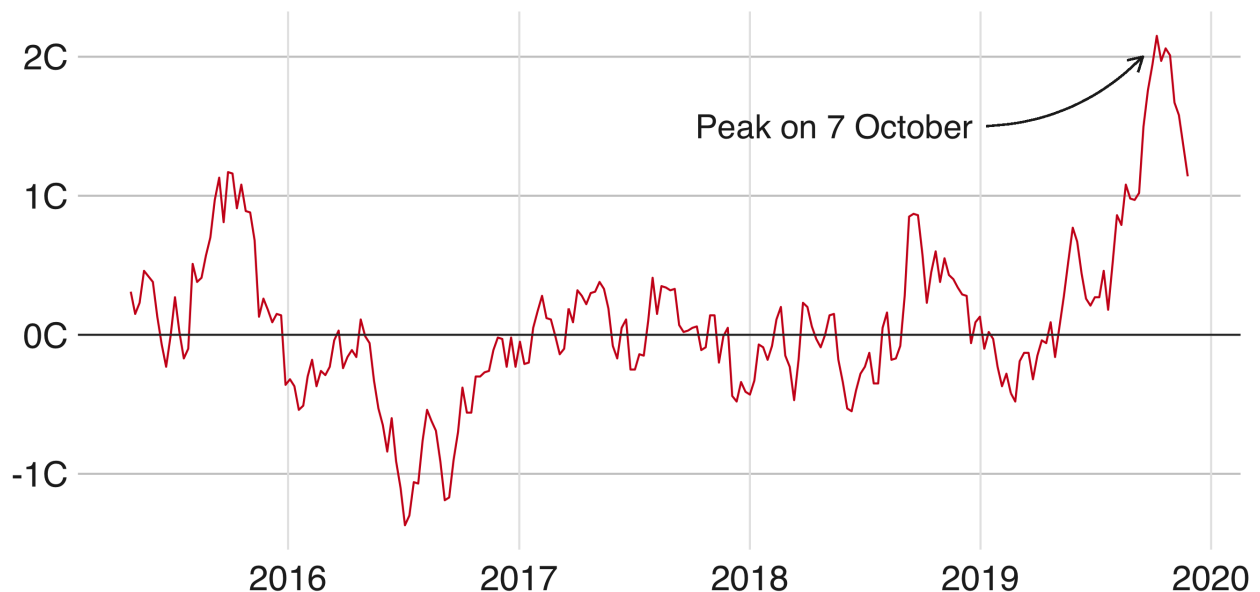
AUSTRALIAN FIRE TOTAL AMOUNT OF FIELDS BURNED IN HECTARES



Result and Discussion

There is a difference between the temperature values found by climate scientists and what we find, and this is natural. Because, as it is written in the source image below, they examined the temperature difference between the two points they chose. On the other hand, we examined a large region and actually got the result as we expected. As can be seen, the temperature, which normally hovers between (-5, +5) values, reached +10 degrees towards the end of 2019. This coincides with the fire season. As a result, we were able to establish the relationship we wanted.

This year's dipole has been unusually strong Indian Ocean Dipole index (degrees C)



Note: Index refers to the temperature difference between two points in the Indian Ocean

Source: Australia's Bureau of Meteorology

BBC

Contributions of Each Team Members

- We had weekly meetings on Discord.
- All work has done together.
- Each member contributed 50%

Sources

- <https://www.bbc.com/news/science-environment-50602971>
- <https://www.ncei.noaa.gov/access/search/index>
- https://data.ceda.ac.uk/badc/cru/data/cru_cy/cru_cy_4.04/data