BDA - Project

Tomi Räsänen - 879626 & Erik Husgafvel - 528867

Contents

1. Introduction

One of the biggest challenges of humankind in the 2020s is figuring out ways to slow down the growth of greenhouse gas emissions and stop global warming (due to human activities) under 2 °C. The increasing trend of global temperature is easily seen in Figure 1 [cite NASA] in which the global surface temperature is illustrated relative to 1951-1980 average temperatures. Warming can also be seen with one's own eyes by observing the winters that are warming year by year, by noticing that the number of devastating hurricanes is increased, and by finding out the increased rate of ice melting in glacier's summer.

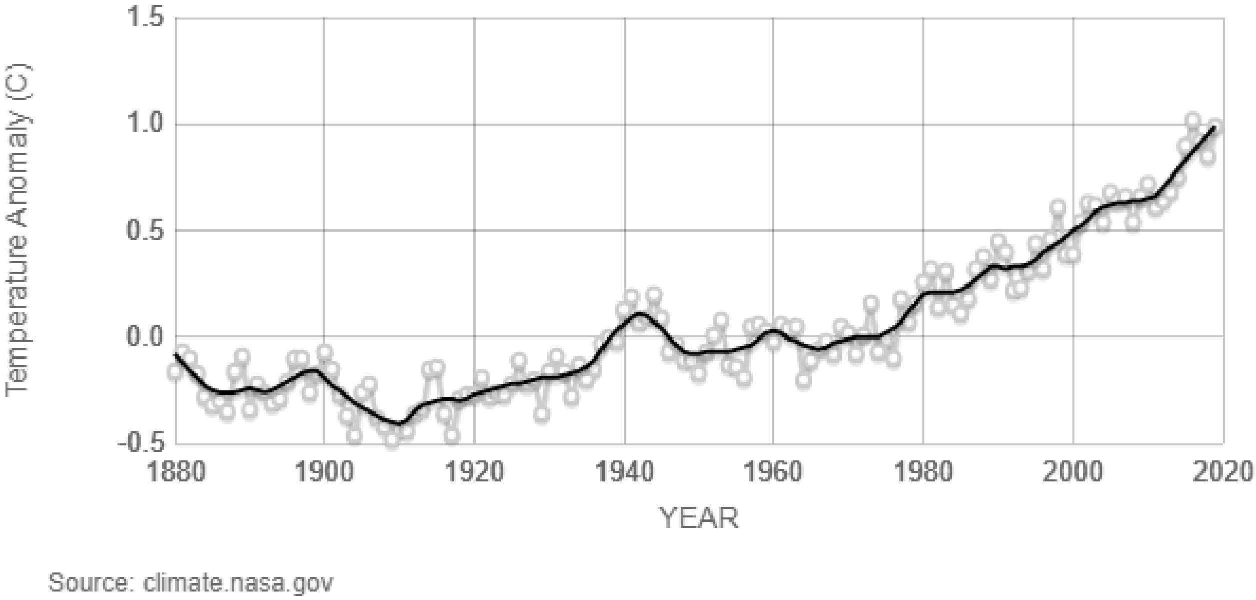


Figure 1: Global Land-Ocean Temperature Index

In response to that warming, many countries have declared a climate emergency to emphasize the criticality of the situation. In addition, young people have organized climate demonstrations around the world, politicians are talking more and more about climate change, and presidents and prime ministers are negotiating agreements and commitments to solve this, one of humanity's greatest, problem. But what if, despite attempts of negotiation, the necessary CO_2 reduction decisions are not achieved?

In this project, our goal is to model the historical emission trends of selected countries as well as attempts to model their future emissions. We are examining a scenario in which emissions continue to develop at a historical rate, and the necessary reductions are not achieved. In our modeling, the other parameters e.g. population growth and technical conditions, are similar to historical data in our modeling.

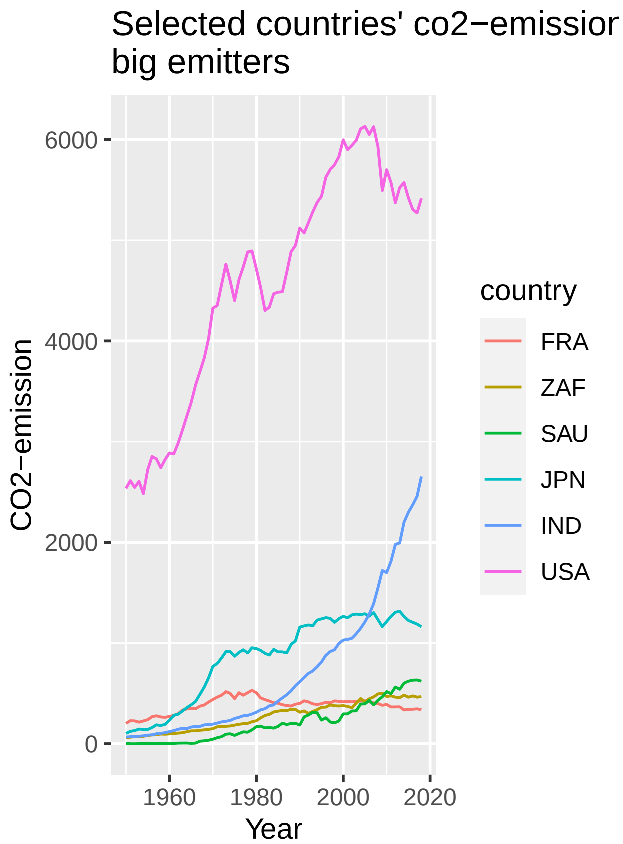
2. Data description

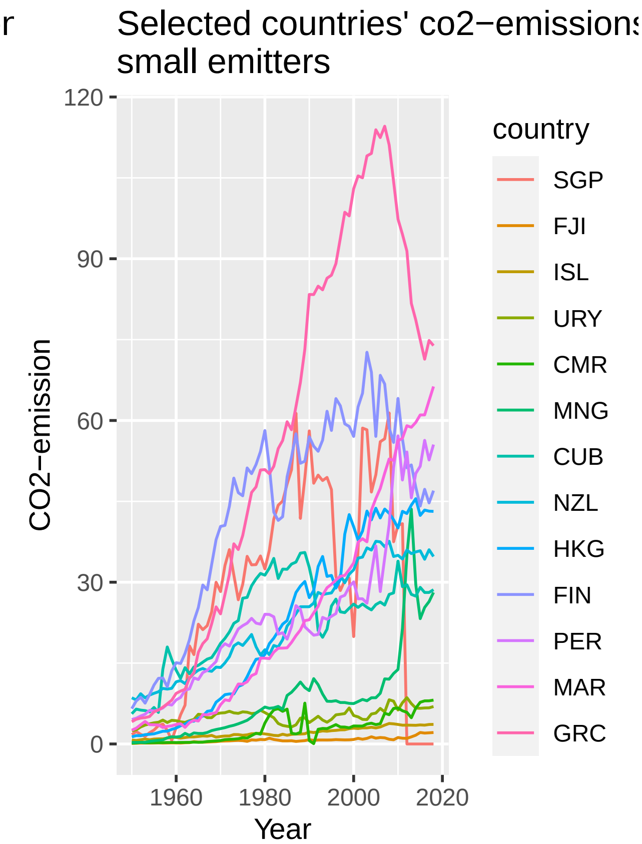
Our CO_2 data was obtained from *Our World in Data* (OWID) web page [cite OWID_net] and the actual *CSV* file from OWID GitHub page [cite OWID_git]. As mentioned earlier, climate change is a hot topic in the daily news, and there is a lot of studies and research concerning how CO_2 emissions are influencing global warming. The data set was also used, for example, when researchers studied the climate impact of the different policy recommendation which targeted to reduce greenhouse gases from the atmosphere.

In our modeling, we are going to select 19 different countries from that OWID data set and examine CO_2 data between the years 1950-2018. The countries we have chosen cover the whole globe and are roughly evenly distributed across continents. There are both small and large polluters among the countries.

```r
library(rstan)
library(ggplot2)
library(reshape2)
library(gridExtra)
#library(aaltobda)
#data("factory")
```

```r
# Read data to data frame
data <- read.csv("./data_co2.csv")
# Our exploratory data analysis discovered some missing values. In order to
# avoid errors, we'll replace them with zeros.
data[is.na(data)] <- 0


# We discovered that the CO2-emissions difference between our selected countries is so vast
# that it's better to split the data into two different plots.

df_data1 <- data[, (data[dim(data)[1], ]) >= 100]
# Sorting the columns in ascending order based on the last row value
df_data1 <- df_data1[,order(df_data1[69,])]
df_data1_2 <- data.frame(years=seq(1950,2018), df_data1)

# Plot the result
df_plot1 <- melt(data = df_data1_2, id.vars = "years", variable.name = "country")

df_data2 <- data[, (data[dim(data)[1], ]) < 100]
# Sorting the columns in ascending order based on the last row value
df_data2 <- df_data2[,order(df_data2[69,])]
df_data2_2 <- data.frame(years=seq(1950,2018), df_data2)

# Plot the result
df_plot2 <- melt(data = df_data2_2, id.vars = "years", variable.name = "country")


# Plot the countries
plot1 <- ggplot(df_plot1, aes(x=years, y=value, colour=country), environment = ) +
        geom_line() +
        ggtitle("Selected countries' co2-emissions: \nbig emitters") +
        xlab("Year") +
        ylab("CO2-emission")

plot2 <- ggplot(df_plot2, aes(x=years, y=value, colour=country)) +
        geom_line() +
        ggtitle("Selected countries' co2-emissions: \nsmall emitters") +
        xlab("Year") +
        ylab("CO2-emission")

grid.arrange(plot1, plot2, ncol=2)
```





Example code for hierarchical model from assignment

```
# # Setting seed to get same "random" results
# SEED <- 12345
#
# # Printing out our hierarchical model
# writeLines(readLines("assignment9_hierarchical_model.stan"))
#
#
# num_of_chains = 10
# hier_model <- rstan::stan_model(file = "assignment9_hierarchical_model.stan")
# hier_data <- list(N = nrow(factory),
#                   J = ncol(factory),
#                   y = factory)
# hier_fit <- rstan::sampling(object = hier_model,
#                             data = hier_data,
#                             iter = 5000,
#                             warmup = 1000,
#                             chains = num_of_chains,
#                             verbose = FALSE,
#                             refresh = 0,
#                             show_messages = FALSE,
#                             seed = SEED)
```

References