STA442 P3

Aichen Liu

10/11/2020

CO2

Introduction

Carbon dioxide(CO2) emission is one of the environmental problems people concern the most. In this research, we are going to explore the impact of two historical events: the fall of the Berlin wall in November 1989 and the global lockdown during the Covid-19 pandemic starting in February 2020 on CO2 emission. Some experts believe that such events would cause a dramatic fall in industrial production and economy and lead to a significant change in CO2 emissions. The data of atmosphere Carbon Dioxide concentrations from an observatory in Hawaii is used to address this research question.

There are two key variables in the dataset: day (The date the data is collected) and co2 (parts per million co2 in the atmosphere). Figure 1 visualizes the whole data and figure 2 visualizes data from the year 2015-2020. It is obvious that there exists a trend and seasonality in the data.

Figure 1: CO2 concentration

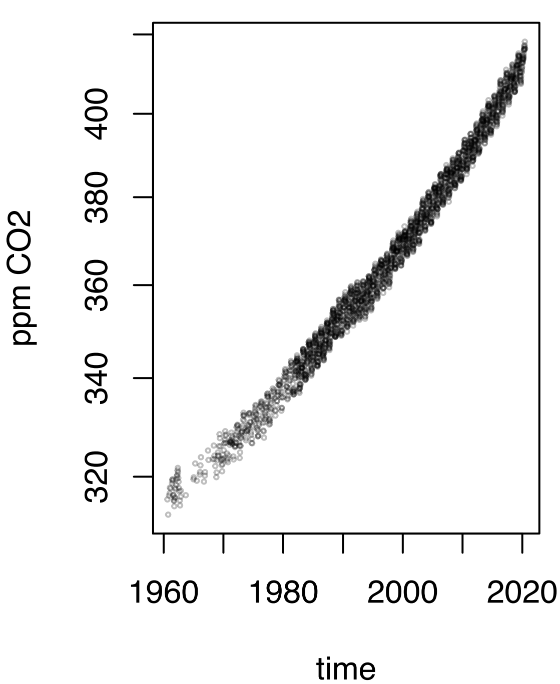
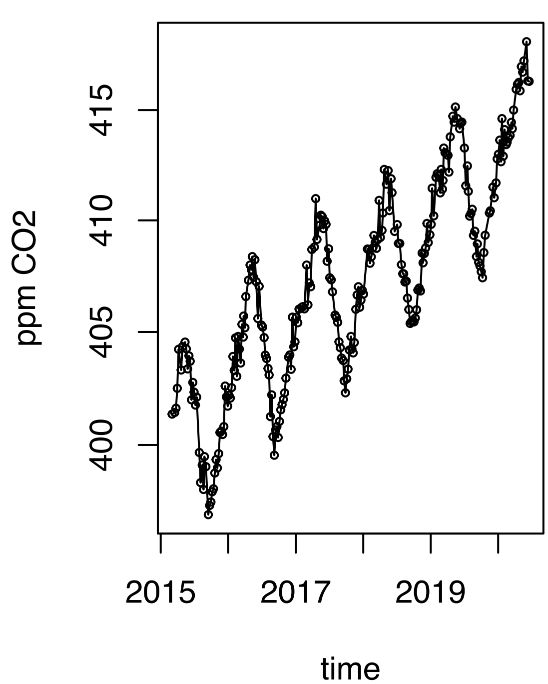


Figure 2: CO2 from 2015–2020



Model

he time-series data contains 2170 observations from 1960-08-18 to 2020-06-16. As indicated in Figure 1 and figure 2, the data doesn't follow any specific distributions. So we will use a non-parametric model. Since we are interested in the concentration of CO2, and the concentration of CO2 must be a positive continuous number, we assume y follows Gamma distribution then set variable co2 to be the dependent variable (y). To better fit the model, we created several new variables: TimeInla is the difference of years between the data collection date and 2000-01-01. cos12 is calculated from $cos(2\pi * TimeInla)$ cos6, is calculated from $cos(4\pi * TimeInla)$ sin12 is calculated from $sin(2\pi * TimeInla)$ sin6 is calculated from $sin(4\pi * TimeInla)$ TimeInla is set to be the random effect as a random slope, and sin6, sin12, cos6, and cos12 are used to smooth the data. Here is the mathematics equation of the model.

$$Y_i \sim Gamma(\theta_1, \frac{\lambda_i}{\theta_1})\, with\, E(Y) = \lambda_i\, and\, Var(Y) = \frac{\lambda_i^2}{\theta_1}$$

$$log(\lambda_i) = X_i\beta + U(t_i)$$

$where$

$$[U_1...U_T]^T \sim RW2(0, \sigma_U^2)$$

$Priors:$

$$P(\sigma_U > 0.1) = 0.5$$

$where$

$$U_{t+1} \mid U_k, k < t \sim N(-2U_t + U_{t-1}, \tau^2)$$

$$(U_{t+1} - U_t) - (U_t - U_{t-1}) \sim N(0, \tau^2)$$

$$U_{t+1} - 2U + U_{t-1} \sim N(0, \tau^2)$$

In this model, Yi is the concentration of CO2 in the atmosphere which follows a gamma distribution. Xi are sin12, sin6, cos12, cos6. $U(t_i)$ is the random effect (random slope), and it is a second-order random walk, which also can be interpreted as a random slope. β_i are the parameters.

Base on prior knowledge, the prior of the random effect (random slope) is set as p(σ_U > 0.1) = 0.5 which can be interpreted as the median of the variationσ_U is 0.1, the prior median is 0.1

Result

After smoothing the data, we can generate a plot combining the predicted trend and seasonality with the original data, shown in figure 3, and the random effect in Figure 4. To better observe the change in CO2 in the atmosphere, a plot of derivatives (change in CO2) is generated in figure 5; The red line indicates the 1989 November and the blue line indicates 2020 February. Since the data in figure 5 is too crowded, two detailed plots are generated to zoom in to 2017-2020 and 1987-1990.

Firstly, the impact of the fall of the Berlin wall in November 1989 on CO2 emission can be observed in figure 5 and figure 7. In figure 5 and figure 7, since the derivatives are all positive, the change in CO2 is positive. Thus we can conclude that the CO2 in the atmosphere is always increasing. In figure 5, the fluctuation on the left of the red line is almost flat, whereas on the right side it is slowly increasing. If we zoom in to that period and look at figure 7, we can see that preceding November 1989, the data already had a slightly increasing trend, and there are no drastic changes in fluctuation after November 1989. This observation implies that the CO2 in the atmosphere increased at an increasing speed. However, in figure 5, the increase in CO2 is within a reasonable range with no drastic changes. Hence, we can conclude that, even though the fall of the Berlin wall in November 1989 precedes a dramatic fall in industrial production in the Soviet Union and Eastern Europe, it has a mere impact on CO2 emissions.

Now consider the recent Covid-19 pandemic. In figure 5, on the left side of the blue line, the slope decreases, whereas the right side has no pattern. If we zoom in and look at figure 6, we can see that preceding February

2020 the fluctuation is in a decreasing trend, which means that the CO2 in the atmosphere increased at a decreasing speed. However, the pattern after February 2020 is ambiguous, since some of the lines go up some of them go down and some of them even become negative. To better observe the result, let's look at figure 3 and figure 4; Figure 3 indicates the predicted trend, the seasonality and their confidence intervals from July 2020 and onward, and figure 4 plots the random effect (random slope). The predicted value can be obtained by adding those two together. As indicated by the dashed line in figure 3 and figure 4, the predicted concentration of CO2 in the atmosphere can either increases or steady, and the random slope can either increases or decreases. Hence we don't have enough information to conclude how the global lockdown during the COVID-19 pandemic in February 2020 is going to affect the CO2 emission.

Summary

This research explored the impact of two historical events: the fall of the Berlin wall in November 1989 and the global lockdown during the Covid-19 pandemic starting in February 2020, on CO2 emission. Since both of the events preceding a dramatic fall in industrial production or a shutting down of the economy, I first hypothesize that these two events may lead to a slower speed of increase in CO2. However, based on the model above, the result is intriguing. It shows that the fall of the Berlin wall has a mere effect on the emission of CO2. On the other hand, the lockdown during the COVID-19 pandemic may have an effect on CO2 emissions. However, base on the information we have, we cannot determine how the global lockdown is going to affect the CO2 emission.

Figure 3: Predicted Value with Original Data

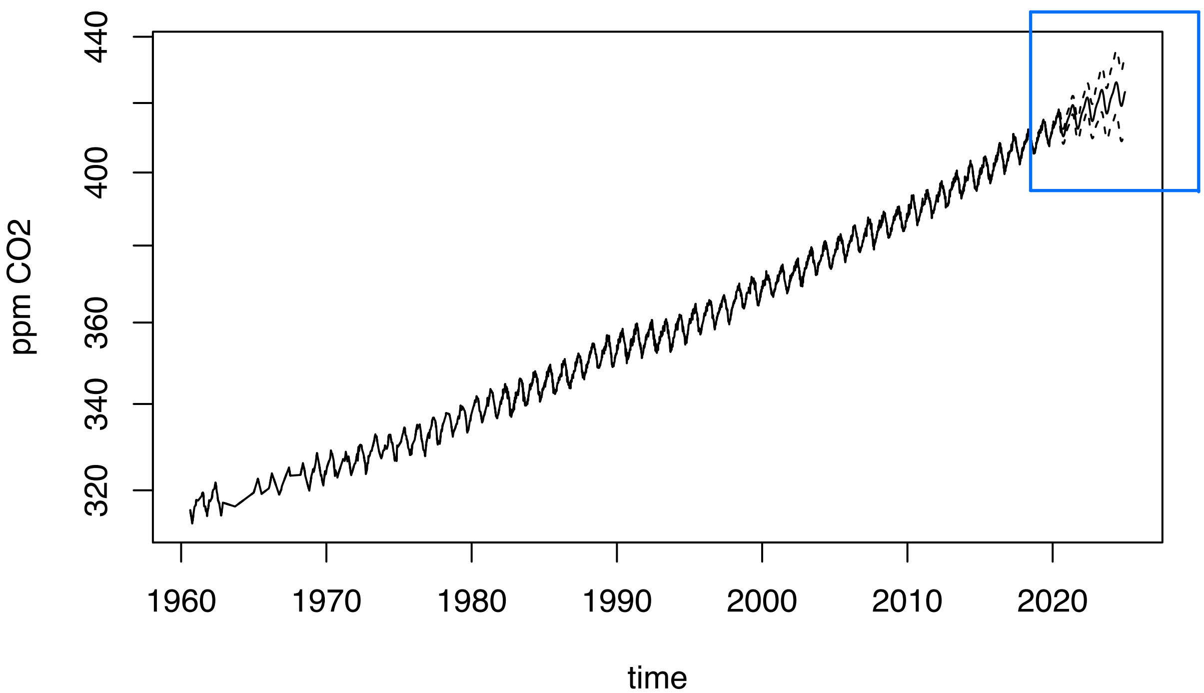


Figure 4: Random Slope

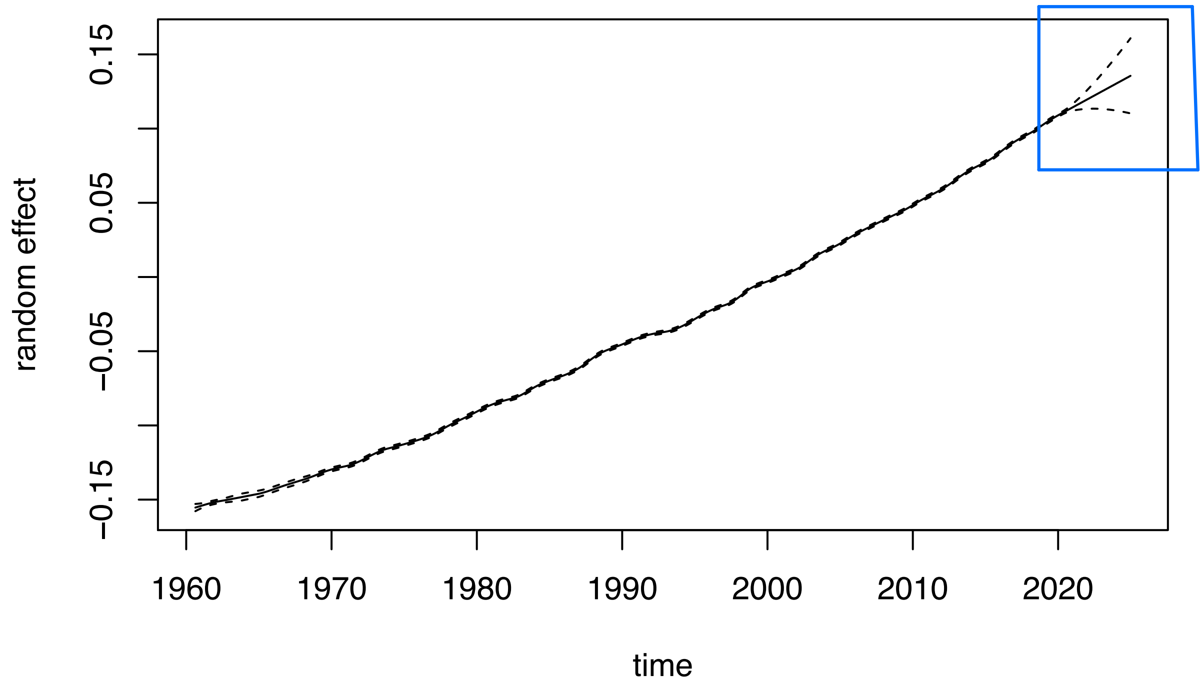


Figure 5: Rate of Change of CO2 (derivative)

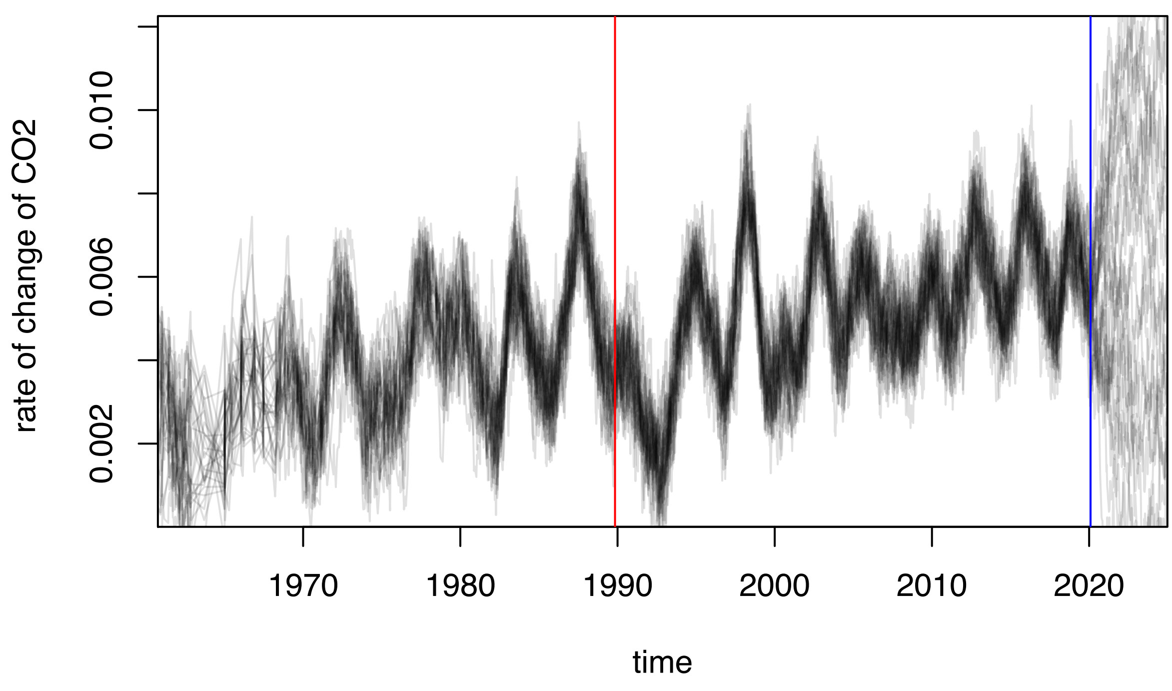


Figure 6: Detailed Rate of Change of CO2 (derivative) from 2017–202

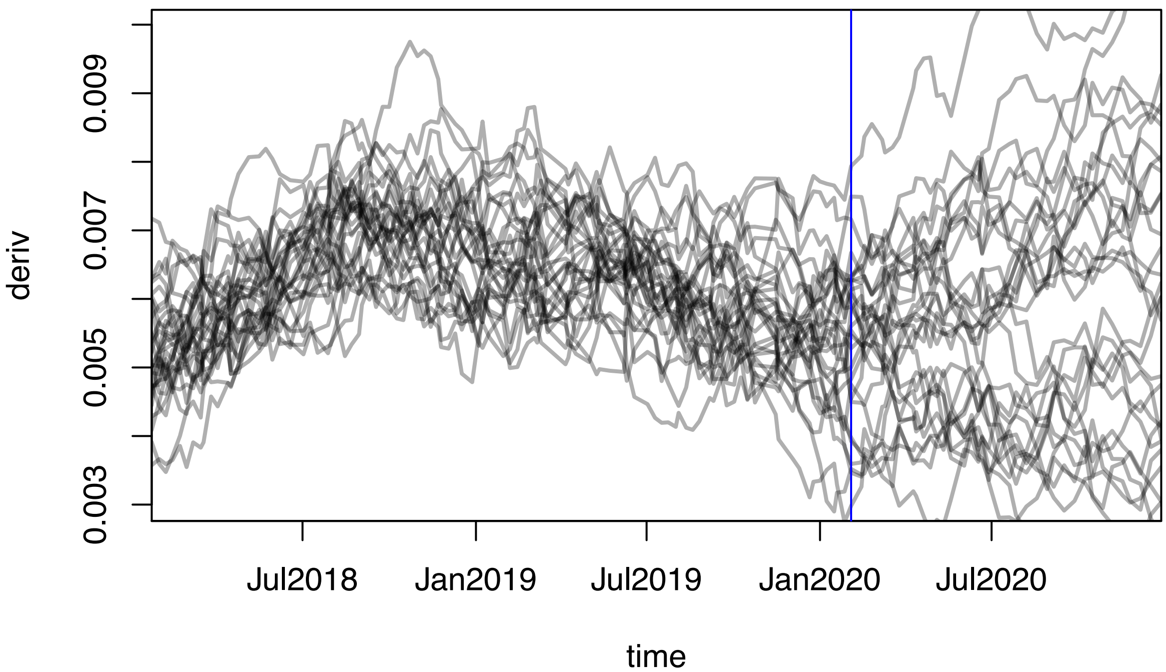
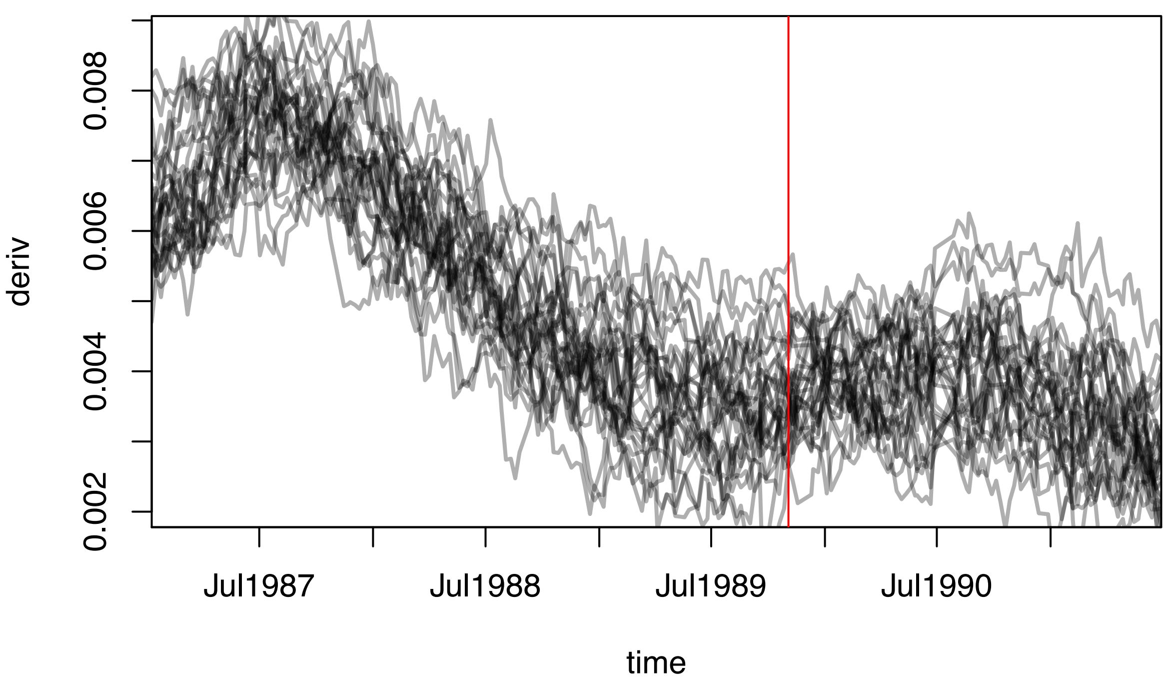


Figure 7: Detailed Rate of Change of CO2 (derivative) from 1987–199



Death

Introduction

Covid-19, a new infectious disease, causes respiratory afflictions that have affected people all around the world. The government believes that the first wave of the COVID-19 epidemic from March 2020 to May 2020 primarily affected the elderly, because they are susceptible to the virus. However, the second wave of the epidemic is mainly caused by young people such as undergraduate students. In this research, we are going to use Daily Mortality Data in Quebec to compare deaths among the people over 70 in the spring and the people under 50 in recent months to historical averages. In order to prove whether the government's beliefs are correct.

There are some key variables in the data set: Age: A categorical variable indicating which age group the sample is in. There are three levels, 0-49 years old, 50-69 years old, 70 years old and over. dead: mortality counts for a certain day and certain year group. time: The information collection date.

Figure 8 and figure 9 gives a visualization of the data. Figure 8 illustrates an overall look at the data and figure 9 categorizes the data for each year. It is obvious that the cases of deaths are significantly higher in 2020 compared with other years.

Figure 8: Mortality in Quebec

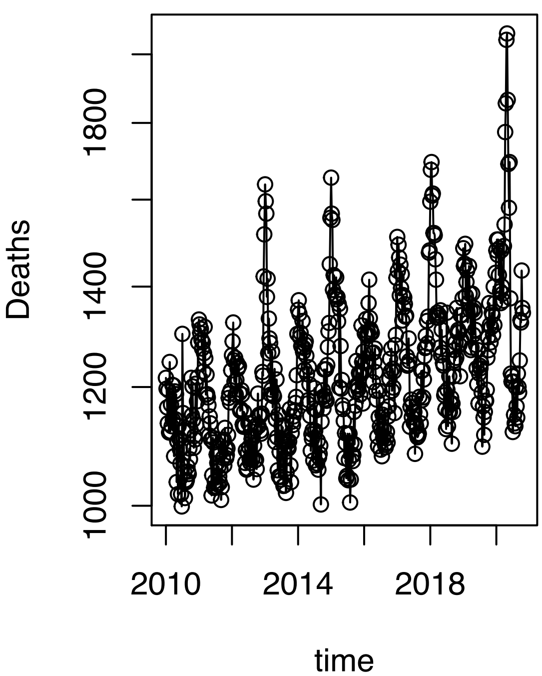
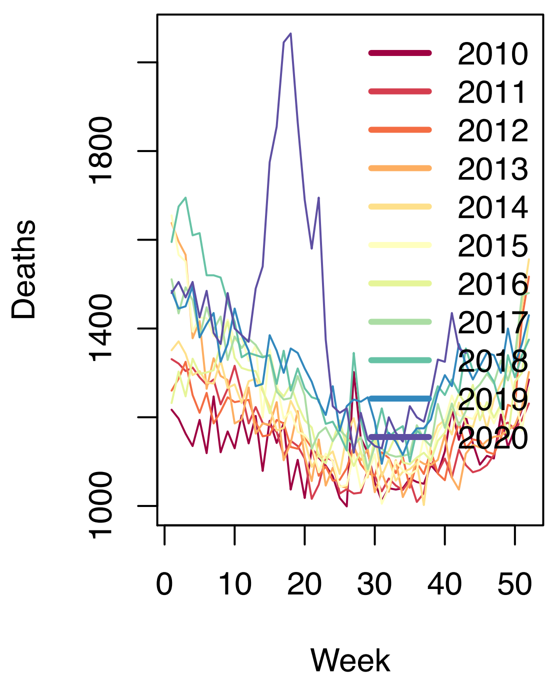


Figure 9: Mortality for Each Year



Model

The time-series data contains seven variables and 2252 observations from 2010-01-01 to 2020-10-21. As shown in figure 8, the time series data doesn't follow a specific distribution, but it is necessary to include a random intercept for time therefore a semi-parametric time trend is used. Some new variables are created to better fit the model. sin6, cos6, sin12, cos12 are sine and cosine values to smooth the data. timeIid is the numerical representation of data. timeForInla is the difference of years between the data collection date and 2015/1/1. Since we want to compare the number of deaths and deaths must be positive integers, the variable dead is set to be the dependent variable and follows a Poisson distribution. Whereas timeIid is set to be the random intercept and timeForInla is set to be the random slope. Since we are interested in two groups of people: people over 70 years old and people under 50 years old, two models, model-1 and model-2 are fitted one for each subgroup. Here is the mathematics equation for both models.

$$Y_i \sim Poisson(\lambda_i)$$

$$log(\lambda_i) = X_i\beta + U(t_i) + V_i$$

$where$

$$[U_1...U_T]^T \sim RW2(0, \sigma_U^2)$$

$$V_i \sim N(0, \sigma_V^2)$$

$Priors:$

$$P(\sigma_U > 0.01) = 0.5$$

$$P(\sigma_V > log(1.2)) = 0.5$$

$where$

$$U_{t+1} \mid U_k, k < t \sim N(-2U_t + U_{t-1}, \tau^2)$$

$$(U_{t+1} - U_t) - (U_t - U_{t-1}) \sim N(0, \tau^2)$$

$$U_{t+1} - 2U + U_{t-1} \sim N(0, \tau^2)$$

$$\sigma_V = \frac{1}{\sqrt{\tau}}\ where\ \pi(\tau) = \frac{\lambda}{2}\tau^{-3/2}exp(-\lambda\tau^{-1/2})\ \tau > 0$$

The only difference between the two models is their dependent variable. For model-1 the dependent variableY_i represents the mortality counts for people who are over 70 years old, but for model-2, it represents the mortality counts for people who are less than 50 years old. X_i are sin12, sin6, cos12, cos6.$U(t_i)$ is the random slope and it follows a second-order random walk. V_i is a random intercept. β_i are the parameters.

Base on prior knowledge, the prior of the random slope is set as p($\sigma_U > 0.01$) = 0.5 which can be interpretated as the prior median is 0.01. The prior for the random intercept is set as p($\sigma_V >$ log(1.2)) = 0.5, which means the prior median in this case is log(1.2).

Result

After fitting two models some plots are generated to better illustrate the results. Firstly, focus on deaths among people who are over 70. In figure 10 the red dots represent the actual number of deaths among people who are over 70 for each year, the black line is the predicted trend of deaths and the dashed line is the credible interval. As indicated, the actual number of deaths in previous years aligns with the predicted trend, whereas in 2020 the actual deaths excess a lot by the predicted value. Figure 11 illustrates the random slope and its credible interval, in 2020 the CI for the predicted slope is almost all positive but it can either be increasing or decreasing. Hence we know that the number of deaths among people over 70 would be increasing but we are not certain about the speed of the increase. To better observed the pattern for 2020, detailed plots are generated using a subset of the data. In figure 12 and figure 13, the red dots are the actual number of deaths, and the black lines are the predicted number of deaths using historical data. These two plots show that base on historical averages, the number of deaths among people over 70 is supposed to be around 1000, but under the effect of the COVID-19, the actual deaths excess the predicted values a lot. By subtracting the predicted number of deaths from actual deaths, we get an excessive number of deaths and the plot is shown in figure 14. As indicated by figure 14, the original excess deaths preceding 2020 March is around zero, then it increases drastically in March and reaches the peak value 800 in 2020 May, but then it returns to zero around 2020 July. Furthermore, in figure 15, the 80 percent quantile corresponds to around 5000 deaths in March to June 2020, but the 80 percent quantile for the recent weeks is around 250. To summarize all the information, we know that the number of deaths among people over 70 is above historical averages from March 2020 to June 2020, and it goes back to the pre-covid level after June 2020. Therefore we know that the first wave of the COVID-19 epidemic had a great impact on the elderly, but the second wave did not affect the elderly so much.

In the second model, we focus on the number of deaths among young people who are under 50 years old during the second wave of the COVID-19. Figure 16 illustrates the predicted number of deaths for young people in black lines and its credible interval in dashed lines. Compare the predicted values with the real data (in red dots), the red dots are randomly scattered around the predicted trend with no specific patterns observed, and the actual number of deaths does not align with the historical averages. Hence, the number of deaths among young people during the second wave of the COVID-19 epidemic is not very different than in previous years. On the other hand, figure 17 indicates the random slope of the model. The overall trend of the slope is decreasing, the slope itself is positive before 2014 then it becomes negative afterward, meaning the number of deaths among young people increased at a decreasing speed before 2014, but the number of deaths decreases afterward. Focus on 2020, the predicted slope is also decreasing which aligns with the historical pattern. To zoom in on the year 2020, in Figures 18 and 19, the black lines represent the predicted number of deaths among young people based on historical data. As indicated, the predicted number of deaths in 2020 is supposed to be around 50, the actual deaths in red dots are in fact fluctuate around the predicted values. In figure 19, from March to June 2020, there is no drastic change in the number of deaths, and the historical data perfectly predicted the actual number of deaths in that period but during 2020 September and onward, there is a small increase in deaths. Figure 20 helps us look into the small increase during September. The y-axis represents excess deaths, and the x-axis is the months of 2020. The excess deaths of young people fluctuate around 0, the small increase of excess death in September is within a reasonable range. Furthermore, figure 21 gives that the 80 percent quantile during March to June is around 150 whereas the 80 percent quantile for the recent weeks is around 9, which means that the number of deaths for young people even decrease during the second wave of the COVID-19 epidemic compare to the first wave. To sum up, the number of death for young people doesn't exceed the historical average in both two waves of the COVID-19 epidemic. Therefore we know that young people who are under 50 years old were not significantly affected by the first wave of the COVID-19 epidemic, and the second wave is not caused by these young people either.

Summary

The certain government believes that the first wave of the COVID-19 pandemic during springtime primarily affected the elderly whereas the second wave of the pandemic which began in September is caused by young people. In this report, two models are fitted to determine whether the government's beliefs are true. Base on Daily Mortality Counts Data in Quebec, we get the following result. By comparing the actual number of deaths to historical averages, we observed that during the first wave of the pandemic, the excess deaths for elderly people are way above average, whereas the deaths for young people is in line with previous years. Hence we can conclude that the first wave indeed primarily affected the elderly. On the other hand, during the second wave, excess death for elderly people goes back to pre-covid level, in the meanwhile, there is no significant increase in deaths in people under 50. Therefore, we can conclude that the second wave of the COVID-19 epidemic may not be caused by those young people.

Figure 10: Predicted Number of Deaths with Original data

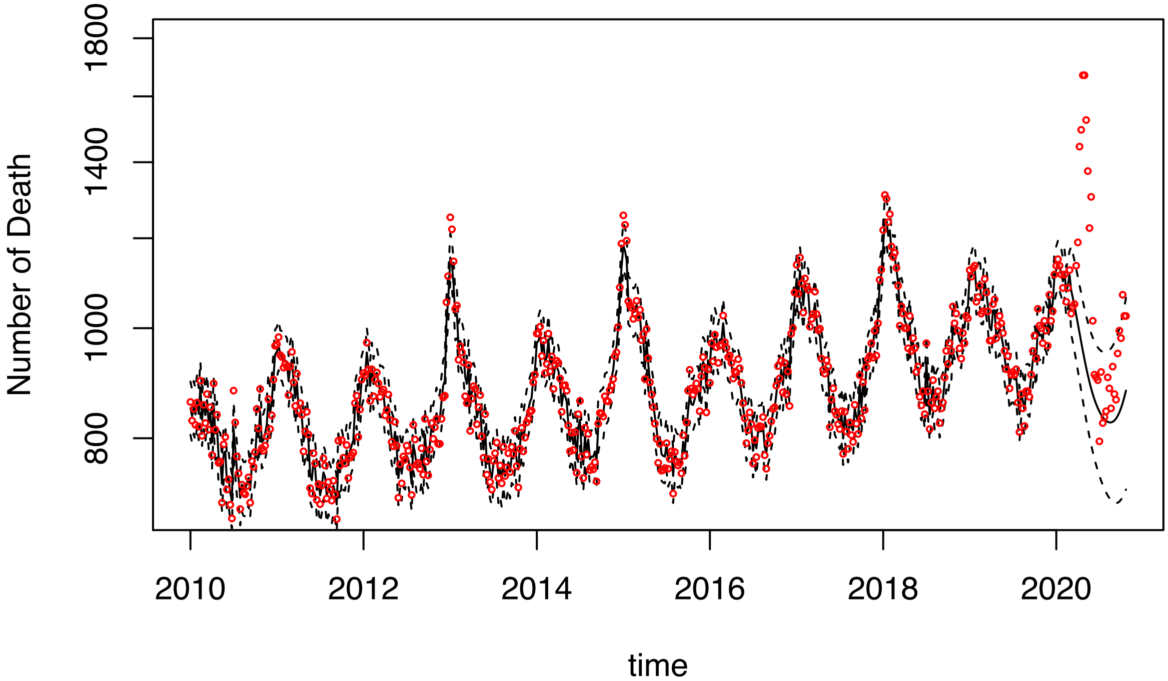


Figure 11: Time Effect

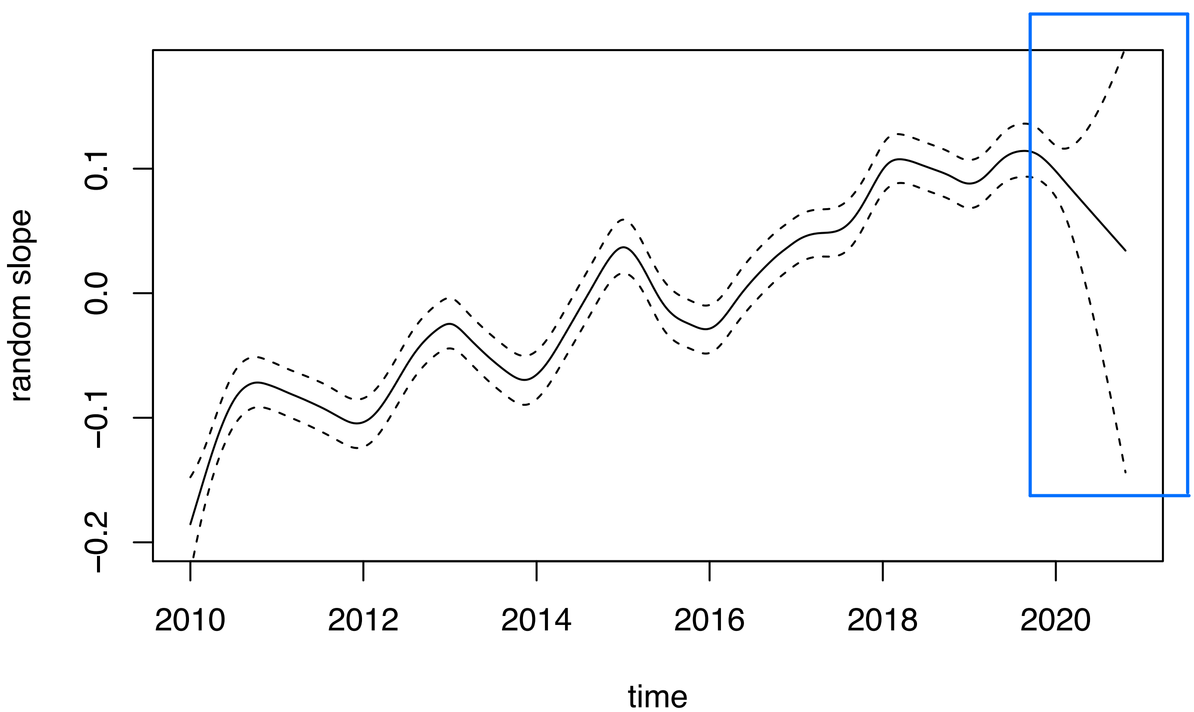


Figure 12: Sample Deaths VS. Real data

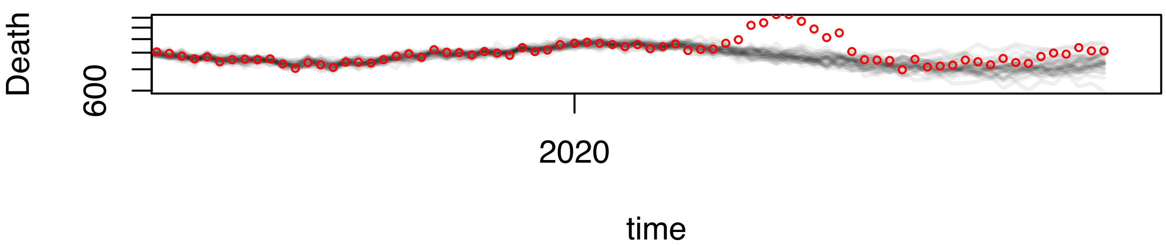


Figure 13: Post COVID–19 Forecast VS. Real data

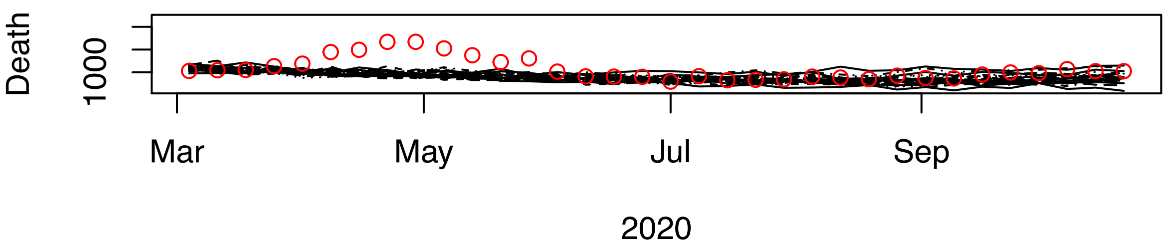


Figure 14: Excess Deaths

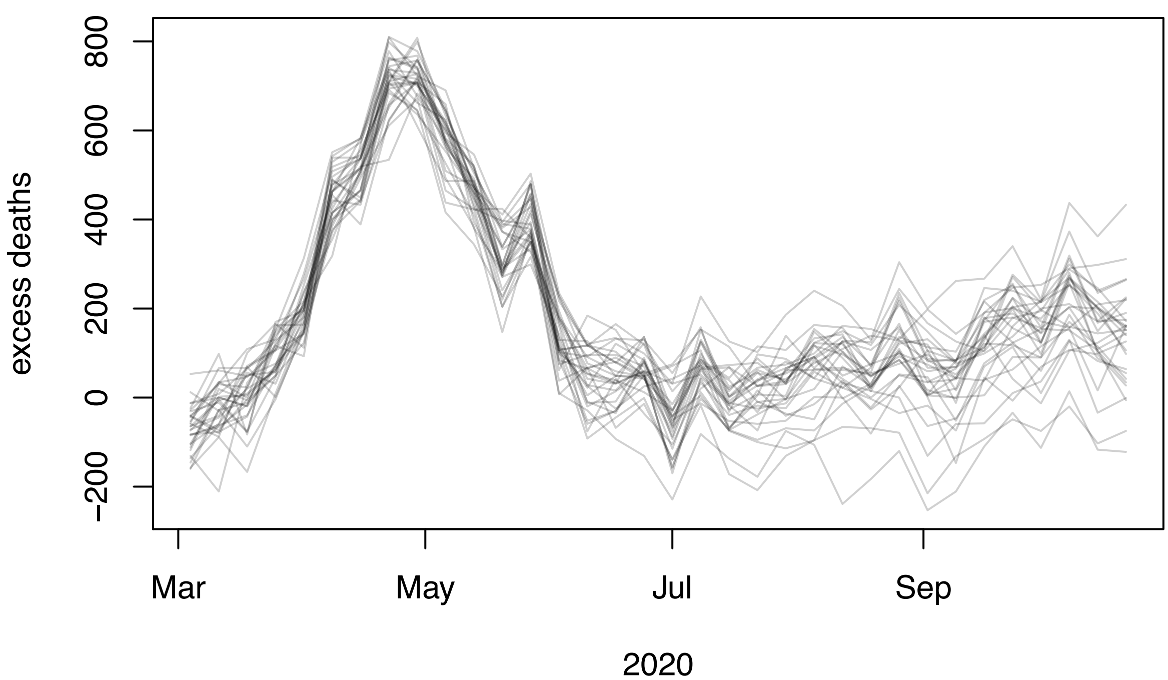
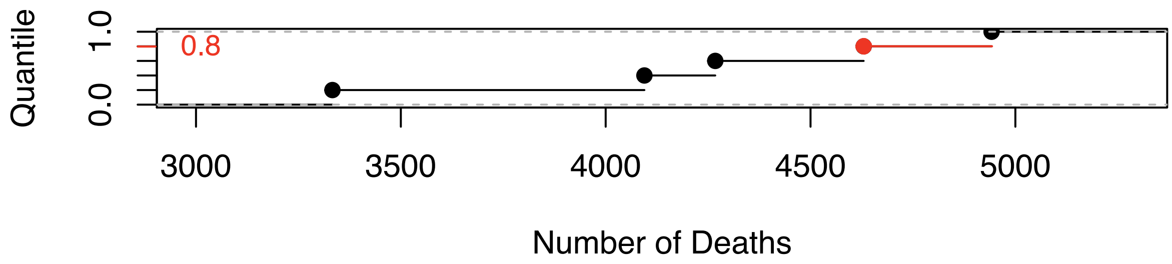


Figure 15: Quantile Plot from 2020–03 to 2020–06



Quantile Plot for Most Recent Weeks

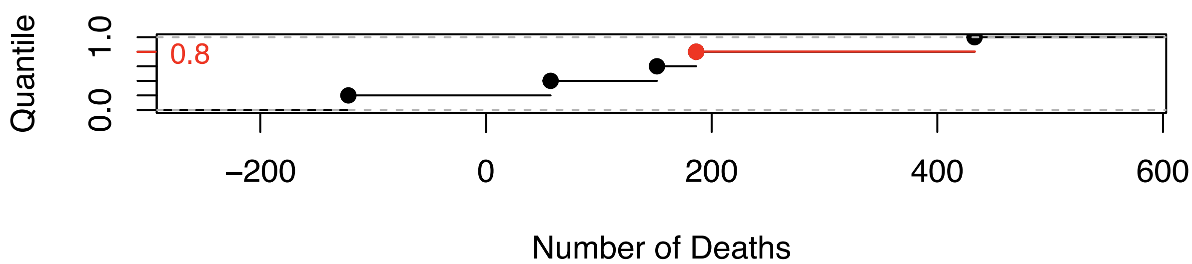


Figure 16: Predicted Number of Deaths with Original data

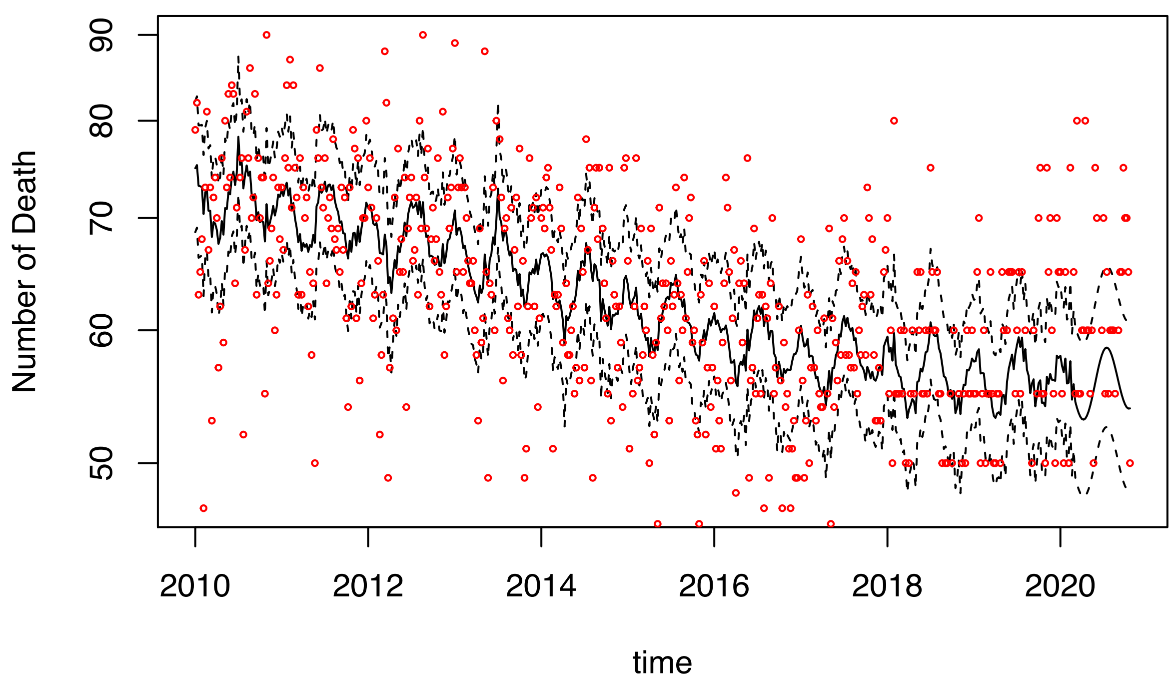


Figure 17: Time Effect

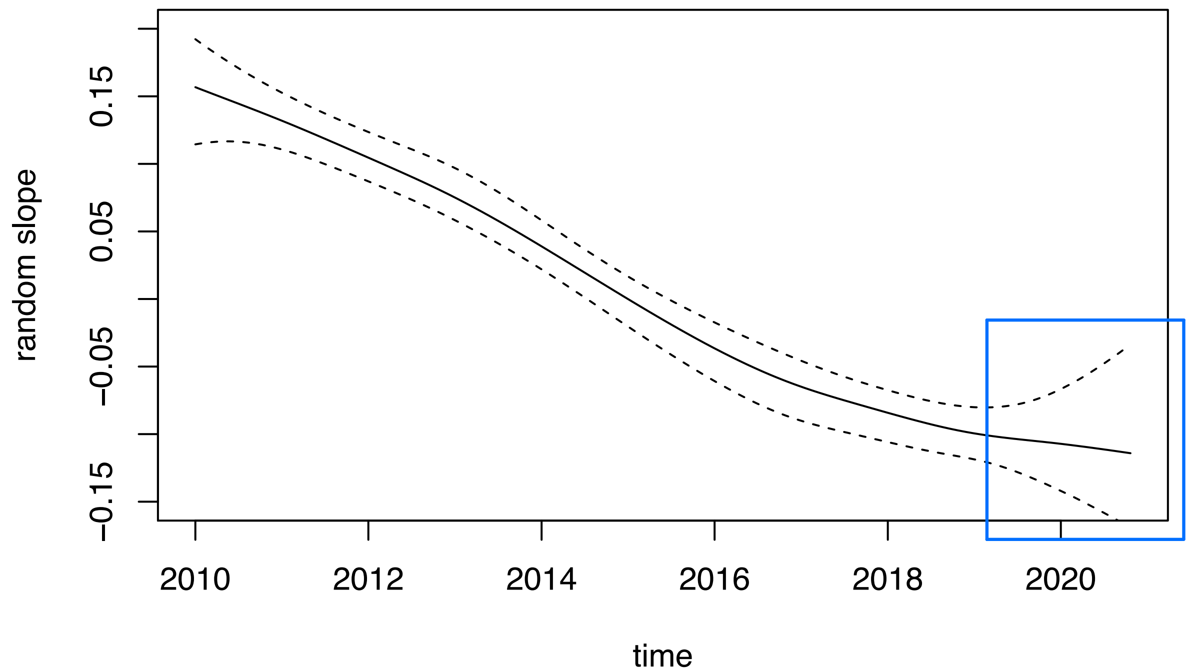


Figure 18: Sample Deaths VS. Real data

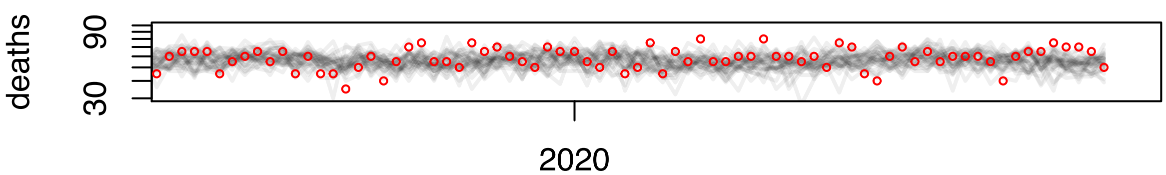
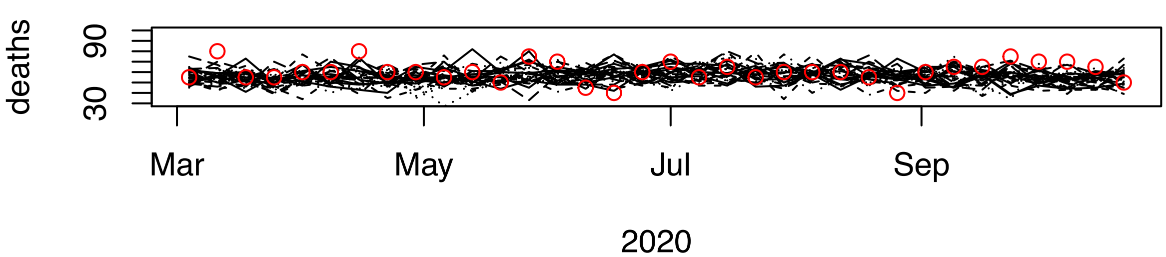


Figure 19: Post COVID−19 Forecast VS. Real data



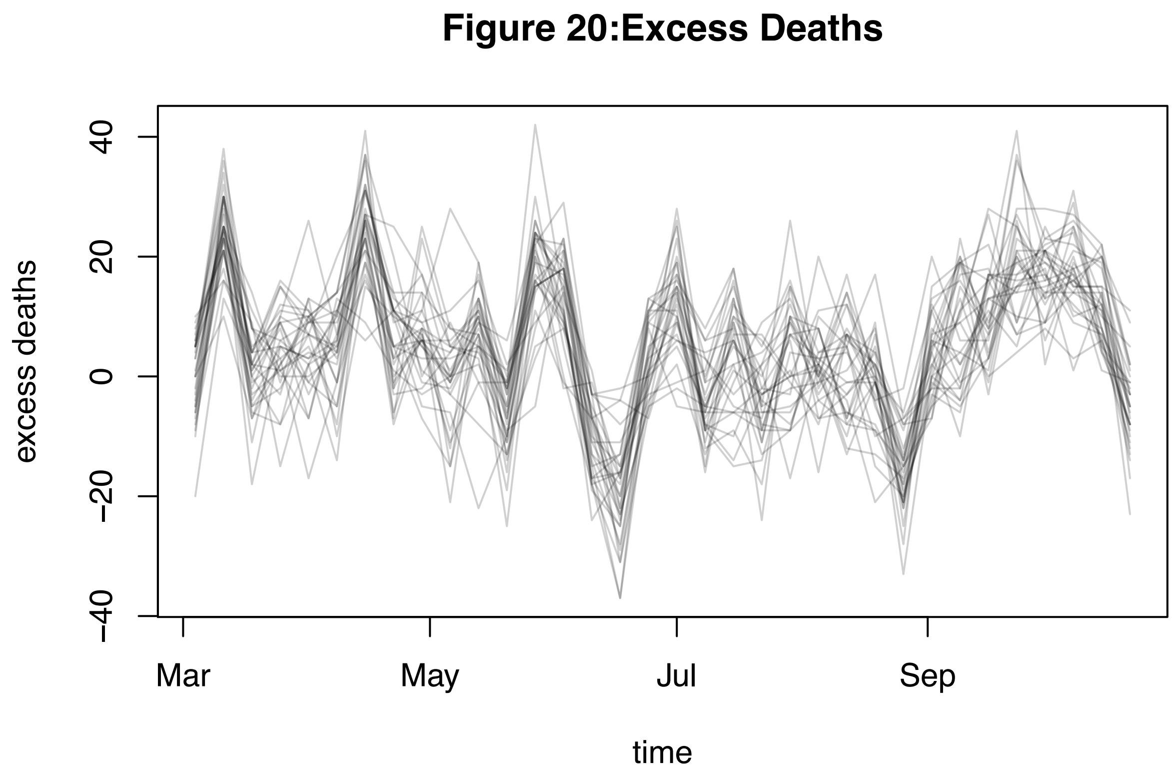


Figure 20:Excess Deaths
excess deaths
40
20
0
−20
−40
Mar
May
Jul
Sep
time

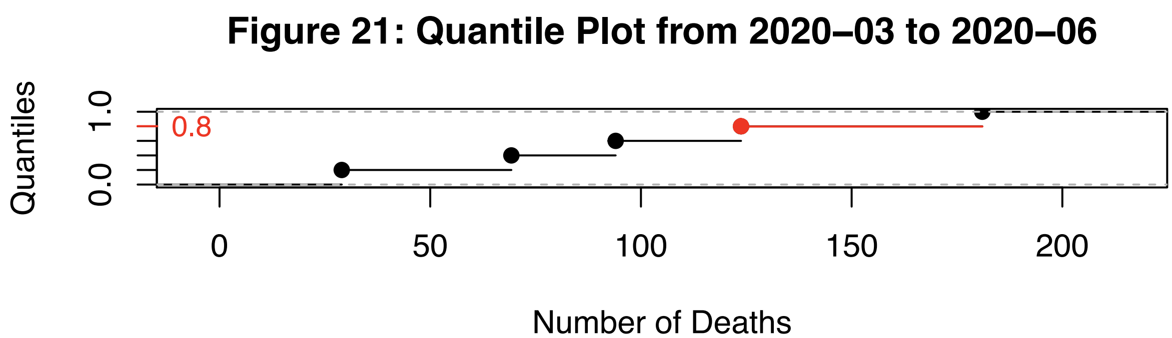
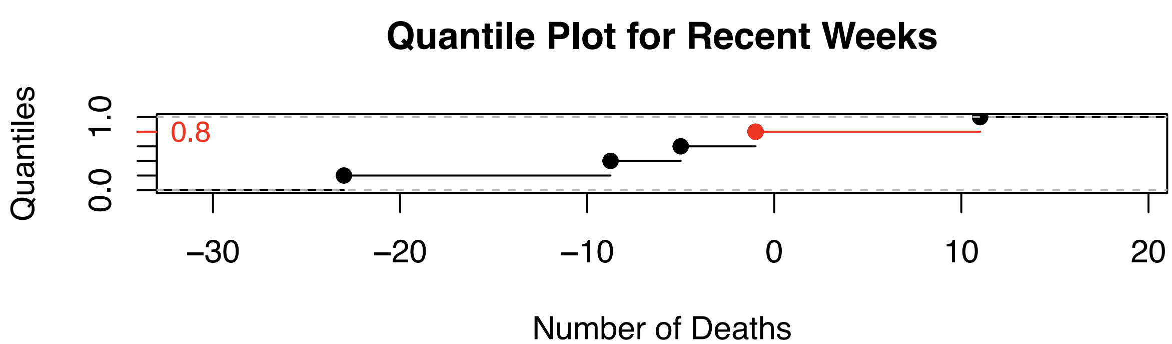


Figure 21: Quantile Plot from 2020−03 to 2020−06
Quantiles
1.0
0.0
0.8
0
50
100
150
200
Number of Deaths


Quantile Plot for Recent Weeks
Quantiles
1.0
0.0
0.8
−30
−20
−10
0
10
20
Number of Deaths

Appendix

```
knitr::opts_chunk$set(echo = TRUE)
#install.packages("INLA", repos=c(getOption("repos"), INLA="https://inla.r-inla-download.org/R/stable"), dep=TRUE)
#BiocManager::install("Biobase")
library('INLA', verbose=FALSE)
library(ecdfHT)
library(reshape2)
library(RColorBrewer)
cUrl = paste0("http://scrippsco2.ucsd.edu/assets/data/atmospheric/",
"stations/flask_co2/daily/daily_flask_co2_mlo.csv")
cFile = basename(cUrl)
if (!file.exists(cFile)) download.file(cUrl, cFile)
co2s = read.table(cFile, header = FALSE, sep = ",",
skip = 69, stringsAsFactors = FALSE, col.names = c("day",
"time", "junk1", "junk2", "Nflasks", "quality",
"co2"))
co2s$date = strptime(paste(co2s$day, co2s$time), format = "%Y-%m-%d %H:%M",
tz = "UTC")
# remove low-quality measurements
co2s = co2s[co2s$quality == 0, ]
par(mfrow=c(1,2))
plot(co2s$date, co2s$co2, log = "y", cex = 0.3, col = "#00000040",
xlab = "time", ylab = "ppm CO2",main="Figure 1: CO2 concentration")
plot(co2s[co2s$date > ISOdate(2015, 3, 1, tz = "UTC"),
c("date", "co2")], log = "y", type = "o", xlab = "time",
ylab = "ppm CO2", cex = 0.5,main="Figure 2: CO2 from 2015-2020")
co2s$day = as.Date(co2s$date)
toAdd = data.frame(day = seq(max(co2s$day) + 3, as.Date("2025/1/1"),
by = "10 days"), co2 = NA)

co2ext = rbind(co2s[, colnames(toAdd)], toAdd)
timeOrigin = as.Date("2000/1/1")
co2ext$timeInla = round(as.numeric(co2ext$day - timeOrigin)/365.25,
2)
co2ext$cos12 = cos(2 * pi * co2ext$timeInla)
co2ext$sin12 = sin(2 * pi * co2ext$timeInla)
co2ext$cos6 = cos(2 * 2 * pi * co2ext$timeInla)
co2ext$sin6 = sin(2 * 2 * pi * co2ext$timeInla)
mm = get("inla.models", INLA:::inla.get.inlaEnv())
if(class(mm) == 'function') mm = mm()
mm$latent$rw2$min.diff = NULL
assign("inla.models", mm, INLA:::inla.get.inlaEnv())

co2res = inla(co2 ~ sin12 + cos12 + sin6 + cos6 +
f(timeInla, model = 'rw2',
prior='pc.prec', param = c(0.1, 0.5)),
data = co2ext, family='gamma',
control.family = list(hyper=list(prec=list(
prior='pc.prec', param=c(0.1, 0.5)))),
# add this line if your computer has trouble
control.inla = list(strategy='gaussian'),
control.predictor = list(compute=TRUE, link=1),
control.compute = list(config=TRUE),
verbose=FALSE)

qCols = c('0.5quant','0.025quant','0.975quant')
sampleList = INLA::inla.posterior.sample(30, co2res,
selection = list(timeInla = 0))
sampleMean = do.call(cbind, Biobase::subListExtract(sampleList,
"latent"))
sampleDeriv = apply(sampleMean, 2, diff)/diff(co2res$summary.random$timeInla$ID)


matplot(co2ext$day, co2res$summary.fitted.values[,
qCols], type = "l", col = "black", lty = c(1, 2,
2), log = "y", xlab = "time", ylab = "ppm CO2")
title("Figure 3: Predicted Value with Original Data")

Stime = timeOrigin + round(365.25 * co2res$summary.random$timeInla$ID)
matplot(Stime, co2res$summary.random$timeInla[, qCols],
type = "l", col = "black", lty = c(1, 2, 2), xlab = "time",
```

```
ylab = "random effect")
title("Figure 4: Random Slope")

matplot(Stime[-1], sampleDeriv, type = "l", lty = 1,
xaxs = "i", col = "#00000020", xlab = "time", ylab = "rate of change of CO2 ",
ylim = quantile(sampleDeriv, c(0.01, 0.995)))
title("Figure 5: Rate of Change of CO2 (derivative)")
forX = as.Date(c("2018/1/1", "2021/1/1"))
forX = seq(forX[1], forX[2], by = "6 months")
toPlot = which(Stime > min(forX) & Stime < max(forX))
abline(v=Stime[679],col = "red")
abline(v=Stime[2162],col = "blue")

matplot(Stime[toPlot], sampleDeriv[toPlot, ], type = "l",
lty = 1, lwd = 2, xaxs = "i", col = "#00000050",
xlab = "time", ylab = "deriv", xaxt = "n", ylim = quantile(sampleDeriv[toPlot,
], c(0.01, 0.995)))
axis(1, as.numeric(forX), format(forX, "%b%Y"))
abline(v=Stime[2162],col = "blue")

title("Figure 6: Detailed Rate of Change of CO2 (derivative) from 2017-2020")
forXXX = as.Date(c("1987/1/1", "1991/12/1"))
forXXX = seq(forXXX[1], forXXX[2], by = "6 months")
toPlotXXX = which(Stime > min(forXXX) & Stime < max(forXXX))
matplot(Stime[toPlotXXX], sampleDeriv[toPlotXXX, ], type = "l",
lty = 1, lwd = 2, xaxs = "i", col = "#00000050",
xlab = "time", ylab = "deriv", xaxt = "n", ylim = quantile(sampleDeriv[toPlotXXX,
], c(0.01, 0.995)))
axis(1, as.numeric(forXXX), format(forXXX, "%b%Y"))
abline(v=Stime[679],col = "red")
title("Figure 7: Detailed Rate of Change of CO2 (derivative) from 1987-1990")
xWide = read.table(paste0("https://www.stat.gouv.qc.ca/statistiques/",
"population-demographie/deces-mortalite/", "WeeklyDeaths_QC_2010-2020_AgeGr.csv"),
sep = ";", skip = 7, col.names = c("year", "junk",
"age", paste0("w", 1:53)))
xWide = xWide[grep("^[[:digit:]]+$", xWide$year), ]
x = reshape2::melt(xWide, id.vars = c("year", "age"),
measure.vars = grep("^w[[:digit:]]+$", colnames(xWide)))
x$dead = as.numeric(gsub("[[:space:]]", "", x$value))
x$week = as.numeric(gsub("w", "", x$variable))
x$year = as.numeric(x$year)
x = x[order(x$year, x$week, x$age), ]

newYearsDay = as.Date(ISOdate(x$year, 1, 1))
x$time = newYearsDay + 7 * (x$week - 1)
x = x[!is.na(x$dead), ]
x = x[x$week < 53, ]
par(mfrow=c(1,2))
plot(x[x$age == "Total", c("time", "dead")], type = "o",
log = "y", ylab = "Deaths")
title("Figure 8: Mortality in Quebec ")

xWide2 = reshape2::dcast(x, week + age ~ year, value.var = "dead")
Syear = grep("[[:digit:]]", colnames(xWide2), value = TRUE)
Scol = RColorBrewer::brewer.pal(length(Syear), "Spectral")
matplot(xWide2[xWide2$age == "Total", Syear], type = "l",
lty = 1, col = Scol, ylab = "Deaths", xlab = "Week")
legend("topright", col = Scol, legend = Syear, bty = "n",
lty = 1, lwd = 3)
title("Figure 9: Mortality for Each Year")
dateCutoff = as.Date("2020/3/1")
xPreCovid = x[x$time < dateCutoff, ]
xPostCovid = x[x$time >= dateCutoff, ]
toForecast = expand.grid(age = unique(x$age), time = unique(xPostCovid$time),
dead = NA)
xForInla = rbind(xPreCovid[, colnames(toForecast)],
toForecast)
xForInla = xForInla[order(xForInla$time, xForInla$age),
]
xForInla$timeNumeric = as.numeric(xForInla$time)
```

```
xForInla$timeForInla = (xForInla$timeNumeric - as.numeric(as.Date("2015/1/1")))/365.25
xForInla$timeIid = xForInla$timeNumeric
xForInla$sin12 = sin(2 * pi * xForInla$timeNumeric/365.25)
xForInla$sin6 = sin(2 * pi * xForInla$timeNumeric *
2/365.25)
xForInla$cos12 = cos(2 * pi * xForInla$timeNumeric/365.25)
xForInla$cos6 = cos(2 * pi * xForInla$timeNumeric *
2/365.25)
xForInlaTotal= xForInla[xForInla$age == '70 years old and over', ]
res = inla(dead ~ sin12 + sin6 + cos12 + cos6 +
f(timeIid, prior='pc.prec', param= c(log(1.2), 0.5)) +
f(timeForInla, model = 'rw2', prior='pc.prec', param= c(0.01, 0.5)),
data=xForInlaTotal,
control.predictor = list(compute=TRUE, link=1),
control.compute = list(config=TRUE),
# control.inla = list(fast=FALSE, strategy='laplace'),
family='poisson')
qCols = paste0(c(0.5, 0.025, 0.975), "quant")
#rbind(res$summary.fixed[, qCols])
#Pmisc::priorPostSd(res)$summary[,qCols]
matplot(xForInlaTotal$time, res$summary.fitted.values[,
qCols], type = "l", ylim = c(690, 1800), lty = c(1,
2, 2), col = "black", log = "y", xlab = "time",ylab = "Number of Death")
points(x[x$age == "70 years old and over", c("time", "dead")], cex = 0.4,
col = "red")
title("Figure 10: Predicted Number of Deaths with Original data")
matplot(xForInlaTotal$time, res$summary.random$timeForInla[,
c("0.5quant", "0.975quant", "0.025quant")], type = "l",
lty = c(1, 2, 2), col = "black", ylim = c(-2, 1.8) *
0.1, xlab = "time",ylab = "random slope")
title("Figure 11: Time Effect")
sampleList = INLA::inla.posterior.sample(30, res, selection = list(Predictor = 0))
sampleIntensity = exp(do.call(cbind, Biobase::subListExtract(sampleList,
"latent")))
sampleDeaths = matrix(rpois(length(sampleIntensity),
sampleIntensity), nrow(sampleIntensity), ncol(sampleIntensity))
#matplot(xForInlaTotal$time, sampleDeaths, col = "#00000010",
#lwd = 2, lty = 1, type = "l", log = "y")
#points(x[x$age == "70 years old and over", c("time", "dead")], col = "red",
#cex = 0.5,ylab = "time", xlab = "Sample Deaths")
#title("Figure 12: Sample Death VS. Real data")
par(mfrow=c(2,1))
matplot(xForInlaTotal$time, sampleDeaths, col = "#00000010",
lwd = 2, lty = 1, type = "l", log = "y", xlim = as.Date(c("2019/6/1",
"2020/11/1")), ylim = c(0.6, 1.6) * 1000, ylab = "Death",xlab = "time")
title("Figure 12: Sample Deaths VS. Real data")
points(x[x$age == "70 years old and over", c("time", "dead")], col = "red",
cex = 0.5)

xPostCovidTotal = xPostCovid[xPostCovid$age == "70 years old and over",
]
xPostCovidForecast = sampleDeaths[match(xPostCovidTotal$time,
xForInlaTotal$time), ]
excessDeaths = xPostCovidTotal$dead - xPostCovidForecast
matplot(xPostCovidTotal$time, xPostCovidForecast, type = "l",
ylim = c(600, 2200), col = "black",xlab = "2020",ylab = "Death")
title("Figure 13: Post COVID-19 Forecast VS. Real data")
points(xPostCovidTotal[, c("time", "dead")], col = "red")
matplot(xPostCovidTotal$time, excessDeaths, type = "l",
lty = 1, col = "#00000030", xlab = "2020", ylab = "excess deaths")
title("Figure 14: Excess Deaths")
excessDeathsSub = excessDeaths[xPostCovidTotal$time >
as.Date("2020/03/01") & xPostCovidTotal$time <
as.Date("2020/06/01"), ]
excessDeathsInPeriod = apply(excessDeathsSub, 2, sum)
#round(quantile(excessDeathsInPeriod))
#round(quantile(excessDeaths[nrow(excessDeaths), ]))
par(mfrow=c(2,1))
plot(ecdf(quantile(excessDeathsInPeriod)),main="",ylab = "Quantile", xlab = "Number of Deaths")
title("Figure 15: Quantile Plot from 2020-03 to 2020-06")
```

```
plot(ecdf(quantile(excessDeaths[nrow(excessDeaths), ])),main="",ylab = "Quantile", xlab = "Number of Deaths")
title("Quantile Plot for Most Recent Weeks")
xForInlaTotal= xForInla[xForInla$age == '0-49 years old', ]
res = inla(dead ~ sin12 + sin6 + cos12 + cos6 +
f(timeIid, prior='pc.prec', param= c(log(1.2), 0.5)) +
f(timeForInla, model = 'rw2', prior='pc.prec', param= c(0.01, 0.5)),
data=xForInlaTotal,
control.predictor = list(compute=TRUE, link=1),
control.compute = list(config=TRUE),
# control.inla = list(fast=FALSE, strategy='laplace'),
family='poisson')

qCols = paste0(c(0.5, 0.025, 0.975), "quant")
#rbind(res$summary.fixed[, qCols], Pmisc::priorPostSd(res)$summary[,
#qCols])
matplot(xForInlaTotal$time, res$summary.fitted.values[,
qCols], type = "l", ylim = c(47, 90), lty = c(1,
2, 2), col = "black", log = "y", xlab = "time",ylab = "Number of Death")
points(x[x$age == "0-49 years old", c("time", "dead")], cex = 0.4,
col = "red")
title("Figure 16: Predicted Number of Deaths with Original data")

matplot(xForInlaTotal$time, res$summary.random$timeForInla[,
c("0.5quant", "0.975quant", "0.025quant")], type = "l",
lty = c(1, 2, 2), col = "black", ylim = c(-1.5, 2) *
0.1,ylab = "random slope", xlab = "time")
title("Figure 17: Time Effect")
sampleList = INLA::inla.posterior.sample(30, res, selection = list(Predictor = 0))
sampleIntensity = exp(do.call(cbind, Biobase::subListExtract(sampleList,
"latent")))
sampleDeaths = matrix(rpois(length(sampleIntensity),
sampleIntensity), nrow(sampleIntensity), ncol(sampleIntensity))
par(mfrow=c(2,1))
#matplot(xForInlaTotal$time, sampleDeaths, col = "#00000010",
#lwd = 2, lty = 1, type = "l", log = "y")
#points(x[x$age == "0-49 years old", c("time", "dead")], col = "red",
#cex = 0.5)
matplot(xForInlaTotal$time, sampleDeaths, col = "#00000010",
lwd = 2, lty = 1, type = "l", log = "y", xlim = as.Date(c("2019/6/1",
"2020/11/1")), ylim = c(0.03, 0.1) * 1000, ylab = "deaths", xlab = "", )
points(x[x$age == "0-49 years old", c("time", "dead")], col = "red",
cex = 0.5)
title("Figure 18: Sample Deaths VS. Real data")
xPostCovidTotal = xPostCovid[xPostCovid$age == "0-49 years old",
]
xPostCovidForecast = sampleDeaths[match(xPostCovidTotal$time,
xForInlaTotal$time), ]
excessDeaths = xPostCovidTotal$dead - xPostCovidForecast
matplot(xPostCovidTotal$time, xPostCovidForecast, type = "l",ylim = c(30, 100), col = "black",xlab = "2020",ylab = "deaths")
points(xPostCovidTotal[, c("time", "dead")], col = "red")
title("Figure 19: Post COVID-19 Forecast VS. Real data")
matplot(xPostCovidTotal$time, excessDeaths, type = "l",
lty = 1, col = "#00000030", ylab = "excess deaths",xlab = "time")
title("Figure 20:Excess Deaths")
excessDeathsSub = excessDeaths[xPostCovidTotal$time >
as.Date("2020/03/01") & xPostCovidTotal$time <
as.Date("2020/06/01"), ]
excessDeathsInPeriod = apply(excessDeathsSub, 2, sum)
#round(quantile(excessDeaths[nrow(excessDeaths), ]))
par(mfrow=c(2,1))
plot(ecdf(quantile(excessDeathsInPeriod)), main = "",ylab = "Quantiles",xlab = "Number of Deaths")
title("Figure 21: Quantile Plot from 2020-03 to 2020-06")
plot(ecdf(quantile(excessDeaths[nrow(excessDeaths), ])),ylab = "Quantiles", xlab = "Number of Deaths",main = "")
title("Quantile Plot for Recent Weeks")
```