442 hw2 Q1

Summary

We are interested in whether the fall of Berlin Wall and COVID-19 pandemic has impact on CO2 concentration. By using GAM model with smoothing term, we found that the speed of rising in CO2 concentration had slowed down after the two events occurred(but the CO2 concentration is still increasing). As time pass by even further, the CO2 concentration began to fall. The trend of random effects seems to be shallower after the two events(trend of increase in random effects seems to be slowed down). In addition, the estimated trend of increase in CO2 concentration becomes shallower after the two events. Thus, we conclude that the two events appears to have impact on CO2 concentration.

Introduction

The concentration of Carbon Dioxide is closely associated with climate change and Earth's future. In order to find out whether human activities are related to the change in CO2 concentration, an analysis is conducted based on the dataset provided by Scripps CO2 Program. In particular, we are interested in whether the fall of Berlin wall on November 9, 1989 and the COVID-19 pandemic staring February 2020 affected the global CO2 concentration.

Model and Method

Based on Figure 1 and Figure 2, the Carbon Dioxide concentration shows oscillating(seasonal) pattern alone time thus GLM might not fit the data. We might need random effects(random slope) since the slope of CO2 concentration changes frequently alone time. Thus, GAM might be appropriate to use since it allows us to add up few trigonometric functions to model the seasonal pattern. In addition, we also assume that the weekly average concentration of Carbon Dioxide follows Normal distribution since it seems approximately Normal(Shown in Figure 3). Thus, the model is given by:

$$E(Y_i) = \lambda_i = X_i\beta + U_{t_i}$$

$$Y_i \sim N(\lambda_i, \sigma_y^2)$$

Y_i represents our response variable which is weekly average CO2 concentration and we assumed it has normal distribution with mean of λ_i and variance of σ_y^2. $E(Y_i) = \lambda_i$ represents the expected weekly averaged CO2 concentration.

Since the slope of concentration of CO2 vs time changes frequently(based on Figure 2), we should use random slope(second order random walk) – U_{t_i} for smoothing purpose for random effects overtime. In addition we assume U_{t_i} are iid Normally distributed with following distribution:

$$[U_1...U_t]^T \sim RW2(0, \sigma_u^2)$$

X_i denotes the covariates defined by 4 trigonometric functions since there are oscillating patterns for CO2 concentration overtime, where β are the corresponding parameters. The four trigonometric functions are give by:

$$cos12 = cos(2 * \pi * timeYears)$$

$$sin12 = sin(2 * \pi * timeYears)$$
$$cos6 = cos(2 * 2 * \pi * timeYears)$$
$$sin6 = sin(2 * 2 * \pi * timeYears)$$

where timeYears is the numerical representation of time variable with 365.25 days as one standard deviation.

Since we do not know much about the data and do not have much experience about the prior, we set PC prior for σ_y^2 and σ_u^2, where σ_y has median of 1(the variation of CO2 concentration does not look severe from Figure 2) and σ_u has median of 0.001(the rate of change to slope of CO2 concentration alone time is small from Figure 2):

$$Prob(\sigma_y > 1) = 0.5$$
$$Prob(\sigma_u > 0.001) = 0.5$$

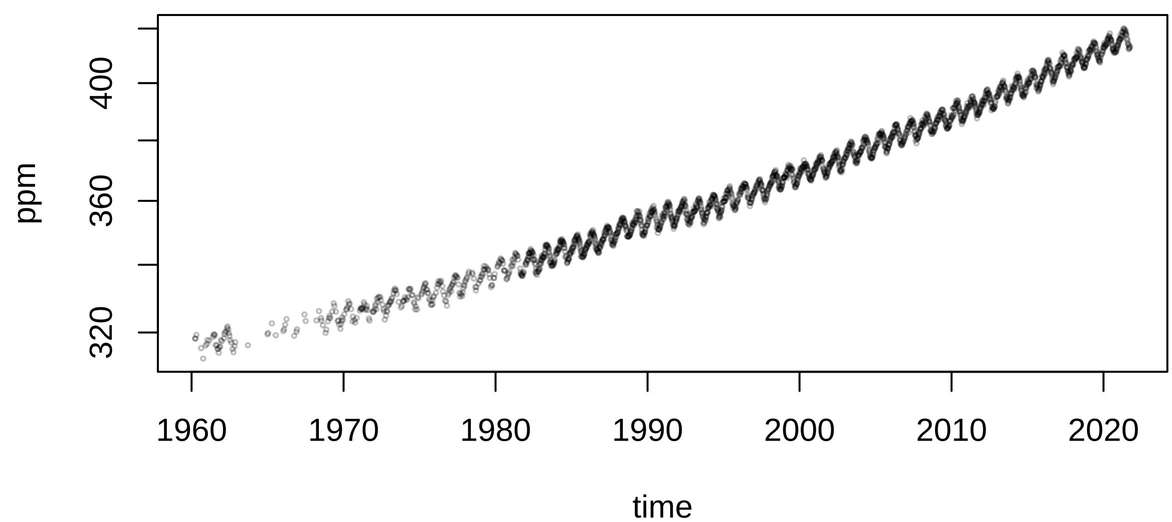


Figure 1: CO2 Concentration From Haiwaii Observatory over years

Results

To show the trend of Carbon Dioxide concentration after the events of Berlin wall's fall and COVID-19 pandemic, two lines corresponding the the approximate date(1989-11-10 and 2020-02-20) of the events have been drawn in Figure 4, Figure 5 and Figure 6. The horizontal dashed line in Figure 5 correspond to the value of 0 for CO2 concentration's derivative.

Figure 5 is the derivative of CO2 concentration alone time. The two points where two vertical line intercept with the plot has positive derivative value which means at the start of the two events, the CO2 concentration was still increasing. After the start of two events, we can tell that the derivative value all became smaller compared to before but was still positive. This means that the speed of increase in CO2 concentrations had slowed down after the two events. The reason might be: even though the two events undermined industrial production and the economy, the society need time to react and there are other things that were still emitting

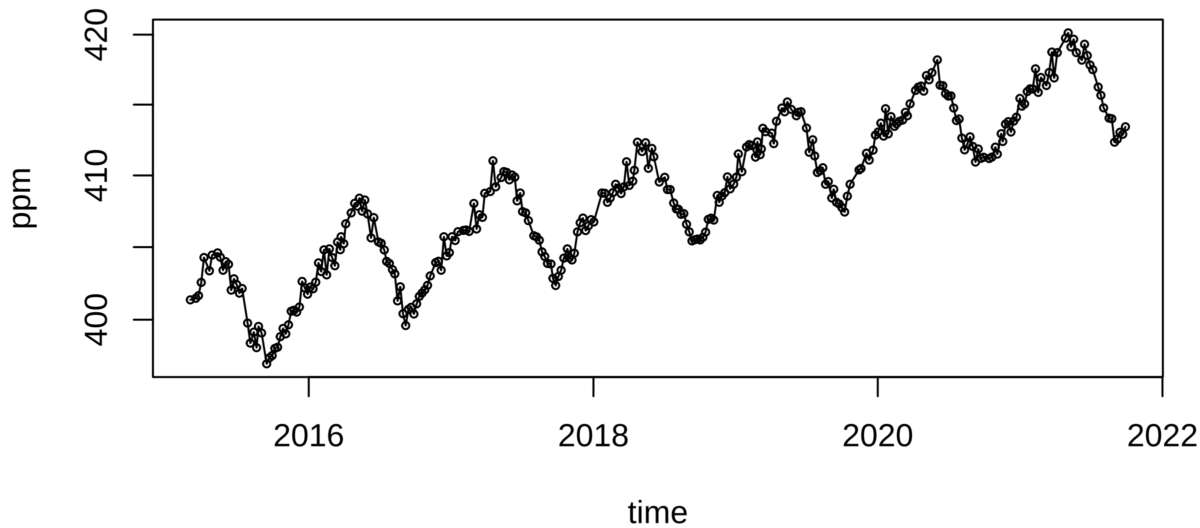


Figure 2: CO2 Concentration From Haiwaii Observatory recently

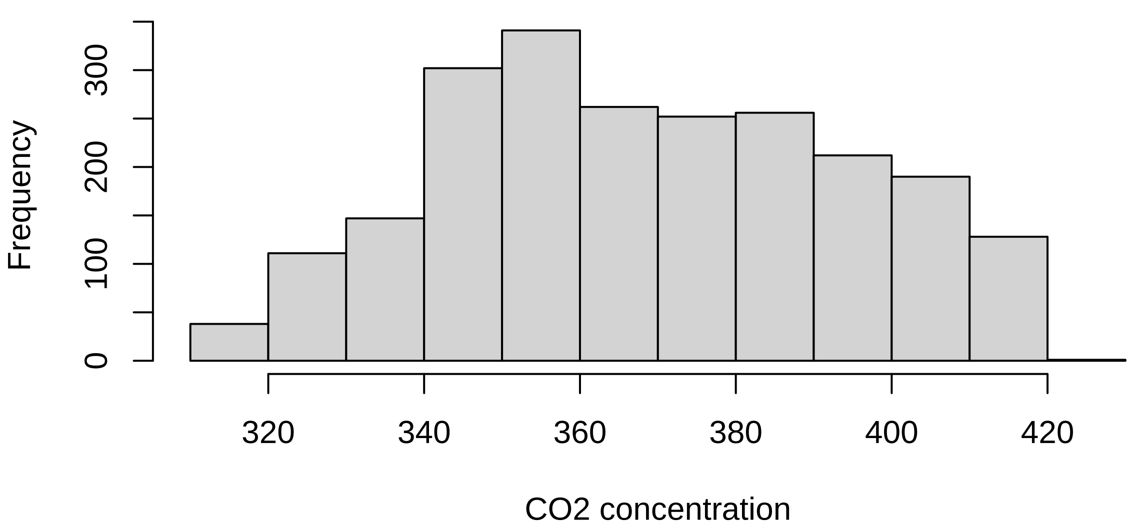


Figure 3: Histogram of concentration

CO2(i.e. cars). Thus, the concentration of CO2 is still increasing but at a decreasing speed. As time passed by even further after the two events, the derivative even became negative which means the concentration of CO2 was decreasing. The reason might be: the negative effects of these two events had made the economy even worse and the large-scale shut down of factories finally decrease the CO2 concentration. Thus, the two events seems to impact the CO2 concentration.

Figure 4 shows the smoothed random effect overtime to CO2 concentration(smoothed trend) where the dashed line shows the confidence interval for random effects. We can tell that the as time passed, the random effects becomes larger. About one year after the fall of Berlin Wall, the trend of increase in random effects seems to be slowed down(flatter). In addition, after start of COVID-19 pandemic, the random effects even decreased. Also, from Figure 6, which is the estimated trend of CO2 concentration alone time. We can tell that after the two events, the estimated trend of increase in CO2 concentration seems to be shallower(the speed of increase of CO2 concentration seems to be slowed down). This is consistent with the above analysis and further enhance the statement that the fall of Berlin Wall and COVID-19 pandemic appears to impact the CO2 concentration.

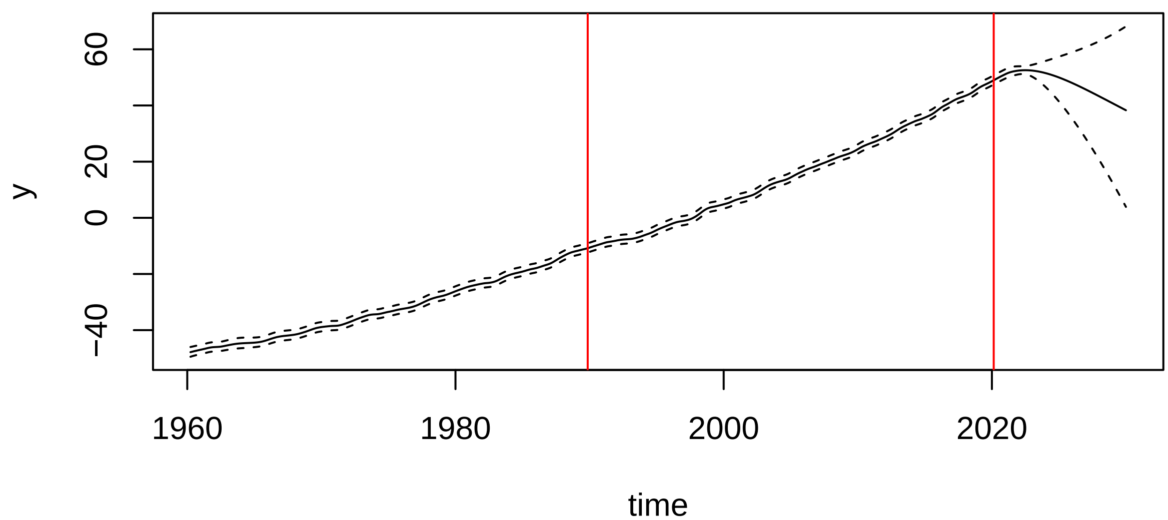


Figure 4: Change of random effects alone time

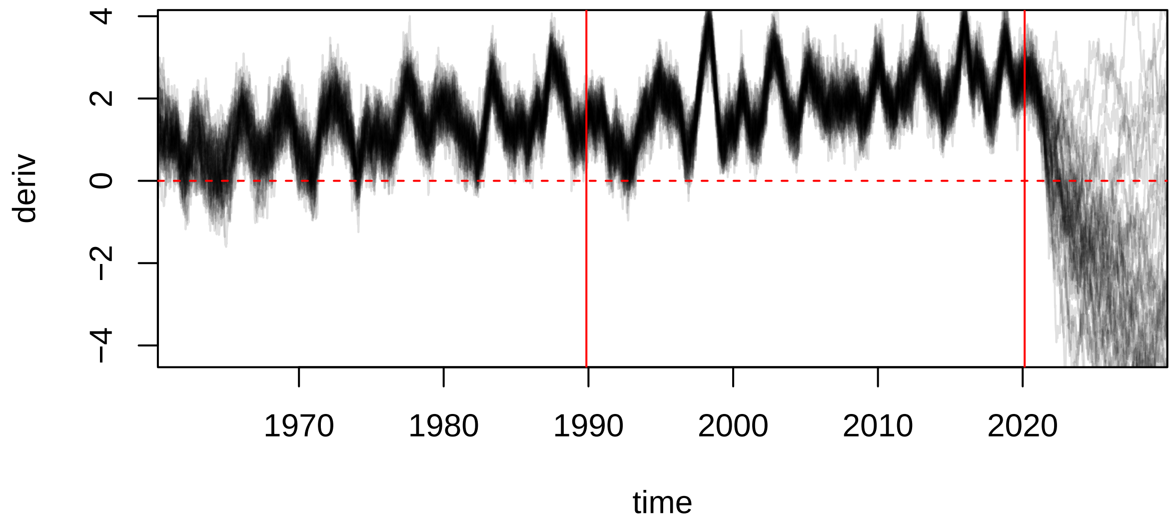


Figure 5: Dirivative of CO2 Concentration alone time

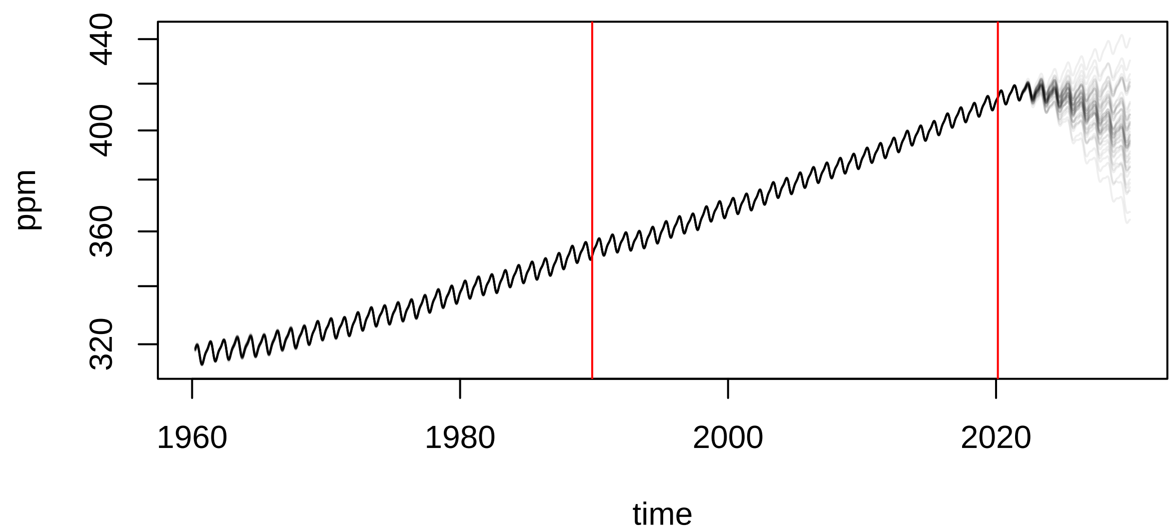


Figure 6: Estimated trend of CO2 concentration alone time

Question 2

(This question is being written up by Word instead of Rmarkdown for technical reason, the code is runned for 4 times, one for each type of diseases, and the only change is the name of diseases. Thus , the appendix shows only the code for “Accident”, for all other 3 types of diseases, just change the name)

Analysis for whether COVID-19 pandemic lead to excess death for four types of mortality.

Summary

By using the data provided by Statistics Canada, we analyze the excess (or deficit of) mortality related to heart attacks, neoplasms, respiratory diseases and accidents in Ontario from March to November to 2020-- the COVID-19 pandemic period. By using GAM model, we found out after the pandemic, the actual data shows an increasing trend of death numbers for heart attacks, neoplasms, and accidents whereas there shows a decreasing trend for respiratory diseases. For heart attacks, there is deficit death in March 2020 and starting April, there is excess death related to heart attacks. For neoplasms, from March 2020 to the start of July 2020, there is no strong evidence of excess death whereas after July, particularly late July and mid-August and the start of October, there is evidence of excess death related to neoplasms. For accidents, except the start of March, and the start of April, all other periods show excess death related to accidents. For respiratory diseases, the deficit of death exhibits for most of the period after May 2020, whereas the third week of March and the whole of April show some excess death.

Introduction

Statistics Canada is providing the cause-specific mortality counts for province Ontario. Among all of the ways that mortality happens, four types of mortality are in our specific interest: Malignant neoplasms(cancers), Diseases of the heart(heart attacks), Accidents(unintentional injuries), and chronic lower respiratory diseases. By our assumption, the first two types of mortality have increased during COVID-19 lockdown since there was limited access to healthcare for individuals. The last two types of mortality are assumed to be decreased since the social distance and masks will protect the individual from respiratory diseases during the pandemic. We intend to find the excess mortality for each type and whether they play a role in morality in Ontario from March to November to 2020—the COVID-19 pandemic period. We are also interested in whether there are particular months or weeks that have excess death.

Method and Model

Since we are counting for the weekly number of deaths in a given time interval, which is intuitively discrete and has positive value. Thus, it is reasonable to assumed that the expected weekly number of people die follows Poisson distribution. In addition, by looking at Figure 1-4, weekly mortality exhibits some seasonal pattern. Thus, a semi-parametric model is being fitted to model the data while maintaining explainability. The model is given by:

$$log(\lambda_i) = log\big(E(Y_i)\big) = X_i\beta + U_{t_i} + V_i$$

$$Y_i \sim \text{Poisson}(\lambda_i)$$

Y_i is our response—weekly number of people died which follows Poisson distribution by assumption and $\lambda_i = E(Y_i)$ represents the expected number of people died in week i. $log(\lambda_i)$ is the corresponding log mortality. X_i is the covariates defined by trigonometric functions and β are the corresponding parameters. U_{t_i} represents the second-order random walk—also referred to random slopes for smoothing purpose(i.e. it is the random effect smoothed with time). V_i represents the random effect of time(non-smooth variation). In addition, X_i contains four trigonometric functions for seasonal pattern(sin12, sin6, cos12, cos6) where:

$$cos12 = cos(2 * \pi * dataInt/365.25)$$

$$sin12 = sin(2 * \pi * dataInt/365.25)$$

$$cos6 = cos(2 * 2 * \pi * dataInt/365.25)$$

$$sin6 = sin(2 * 2 * \pi * dataInt/365.25)$$

“dataInt” is the numeric representation of time point. We use 365.25 refers to one year and use it as one standard deviation. We also assume the parametric part V_i follows i.i.d Normal distribution with mean of 0. Since we do not have much experience of the prior, we set pc.prec prior for σ_v such that the median is 1 (we can tell that difference of number of death between any two weeks are not that large):

$$V_i \sim N(0, \sigma_{\mathrm{v}}^2)$$

$$Prob(\sigma_v > 1) = 0.5$$

We also assume that the non-parametric part U_{t_i} follows Normal distribution with mean of 0 where the second order increment $\Delta^2 u_i = u_i - 2u_{i-1} + u_{i-2}$:

$$U_{t+1} | U_k, k < t \sim N(-2U_t + U_{t-1}, \sigma_{\mathrm{u}}^2)$$

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

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

In general, we have:

$$[U_1 \dots U_t]^T \sim RW2(0, \sigma_u^2)$$

Since we do not know much about the prior, we set pc.prec prior such that the median for σ_{u} is 0.00001(the rate of change between two points seems to be small so we assume 0.00001):

$$Prob(\sigma_{\mathrm{u}} > 0.00001) = 0.5$$

Figure 1 : The heart disease death number vs. Time

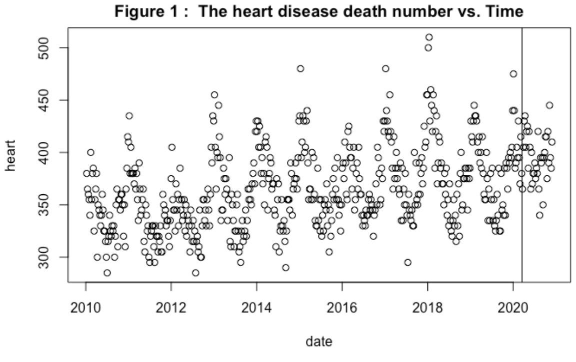
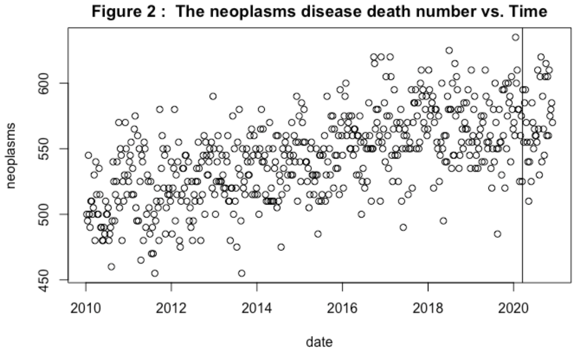


Figure 2 : The neoplasms disease death number vs. Time



Results

The vertical line shown in Figure 1-4 corresponds to the time of 2020/03/17 which could be interpreted as the start of the COVID-19 pandemic. Each of the four figures shows the trend of each type of mortality death number alone time. We can tell that after the start of the pandemic, the trend of death number seems to be increased for heart attack, neoplasms and accidents whereas the trend of death number seems to be decreased for respiratory disease. The reason for the increase in death numbers for heart attacks and neoplasms might be: the COVID-19 made people less access to health care. However, the laws enacted by the government that require people to wear masks and keep social distancing might decrease the probability of respiratory diseases and thus the mortality for respiratory diseases has decreased. One possible reason for the increased death number of accidents might be: the COVID-19 lockdown made safety facilities less accessible to people and increased the death number for accidents.

Some illustration for the Figures below: (X correspond to the first letter of 4 types of mortalities)

For Figure (X).1, the red dots represent the weekly number of death during Christmas holiday before COVID-19 and the weekly number of death during Covid lockdown period. The blue dots represent the weekly number of deaths related to neoplasms diseases before the pandemic and except Christmas holiday.
For Figure (X).4 shows the excess death related to heart attacks from March to November in 2020. The excess death means the actual death number minus the predicted death number. The dashed line represents the 80% credible interval for excess death.

Heart:

Figure H.1: The heart disease death number vs Time

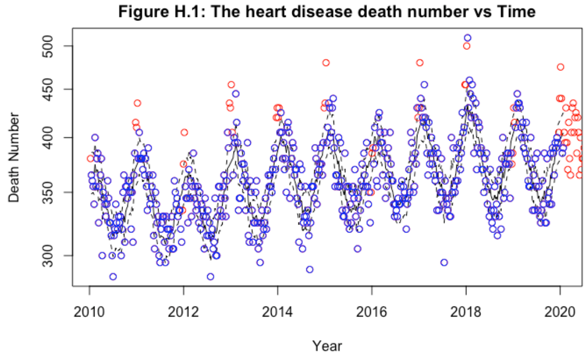


Figure H.2: Random effect of heart disease vs Time

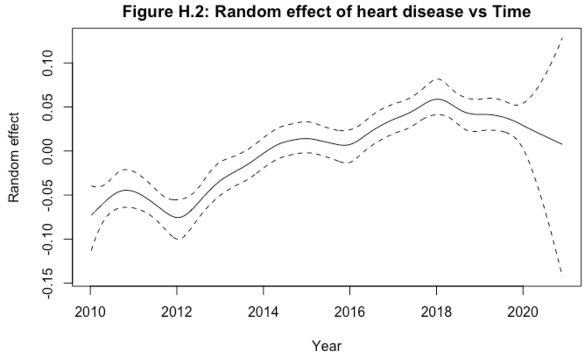


Figure H.3: Actual and predicted number of death of heart disease

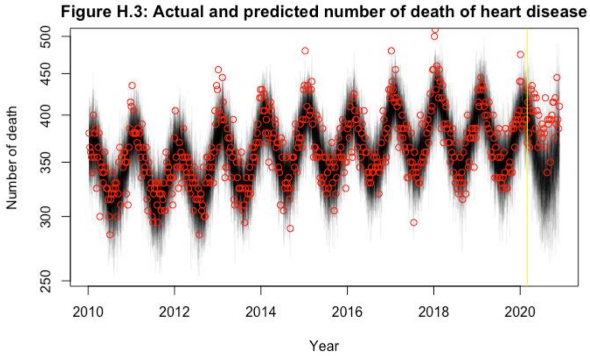
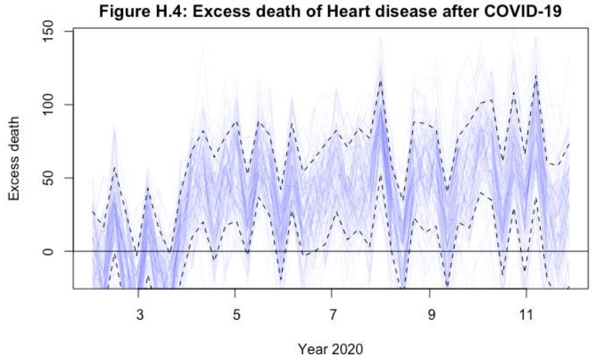


Figure H.4: Excess death of Heart disease after COVID-19



In Figure H.1 and Figure 1, we can see that the death number of heart attacks shows a slightly increasing trend throughout the years. However, the death number during Christmas is larger than other periods on average. The reason might be: the health system is disrupted by vacations and thus people with heart diseases are more exposed to heart attacks related mortality.

From Figure H.2 we can see that the predicted random effect is in an increasing trend although there are small fluctuations. However, after 2020, the trend seems to be shallower. Figure H.3 shows the comparison between the actual number of death from heart attacks and the predicted values. We can tell that pattern of the actual number of death and predicted values are pretty much the same before the pandemic. However, after the pandemic, the actual number of deaths from heart attacks is larger than predicted.

Referred to Figure H.4, for the whole month of March, the interval contains 0 and there is little evidence of excess death. Particularly, in late March, the interval is strictly negative which means the heart attack mortality is getting better and fewer people died from heart attacks. However, starting April, most of the intervals are positive and strictly exclude 0 which means the actual death number is larger than the predicted death number. Thus, after the outbreak of COVID-19, there was excess death related to heart attacks. The reason might be: the COVID-19 lockdown made people less access to health care and thus the mortality related to heart attacks increased.

Neoplasms:

Figure N.1: The neoplasms disease death number vs Time

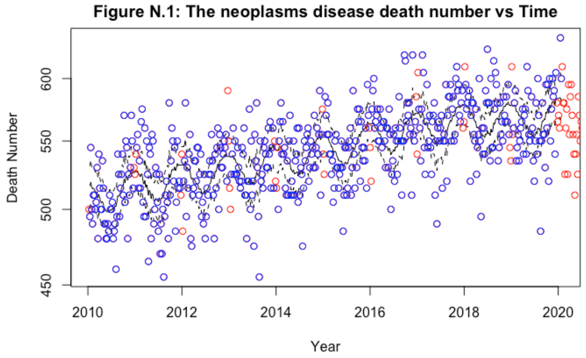


Figure N.2: Random effect of neoplasms disease vs Time

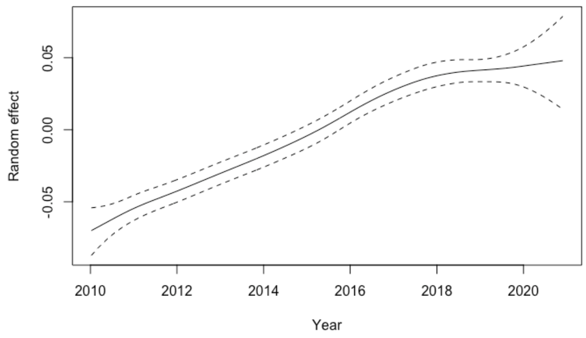


Figure N.3: Actual and predicted number of death of neoplasms disease

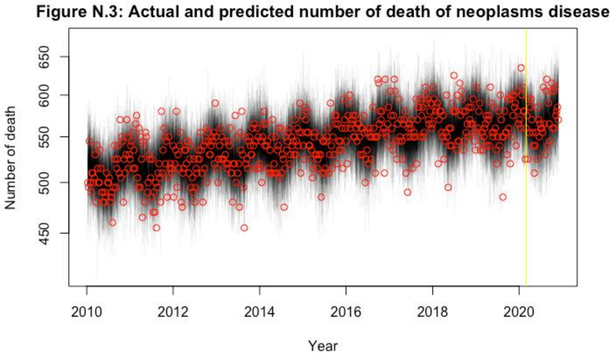
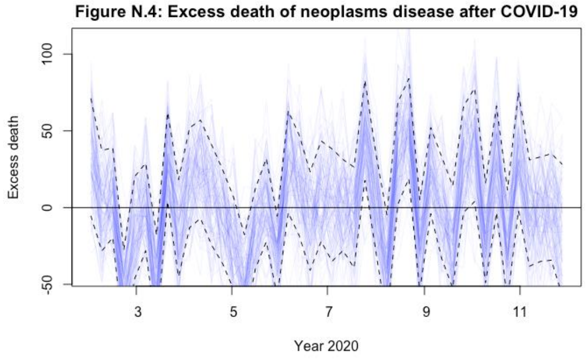


Figure N.4: Excess death of neoplasms disease after COVID-19



From Figure 2 and Figure N.1, we can see that there is an increasing trend of death number due to neoplasms from 2010 to 2020, no matter it's during Christmas or not. The death number continues to increase during COVID-period.

From Figure N.2 we can see that the predicted random effect is in an increasing trend but after the COVID-19, the trend became shallower. Figure N.3 shows the comparison between the actual number of death from neoplasms and the predicted values. We can tell that pattern of the actual number of death from neoplasms is pretty much the same as the predicted value. Overall, there is an increasing trend for the number of death from neoplasms throughout 2010-2020 and even during the covid period.

From Figure N.4, we could tell that from March 2020 to the start of July 2020, the interval is distributed around 0 which does not provide strong evidence of excess death. However, after July 2020, particularly late July and mid-August and the start of October, the credible interval is strictly positive and excludes 0 which means in these months, there was excess death from neoplasms. The reason might be: at the start of the pandemic, the society is still not fully locked down and people with neoplasms still have access to health care. As time passed, the COVID-19 situation became worse and health care place more focus on COVID-19 related diseases, thus people with neoplasms may be exposed to higher probabilities of mortality.

Accidents:

Figure A.1: The Accidents death number vs Time

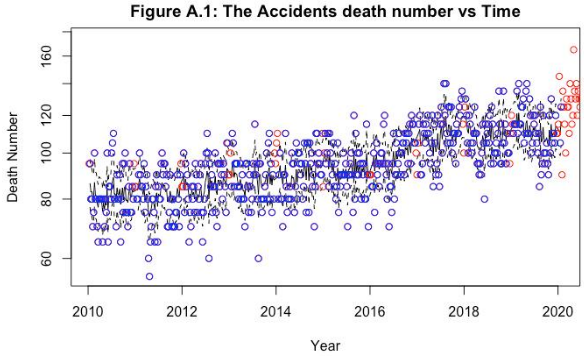


Figure A.2: Random effect of Accidents vs Time

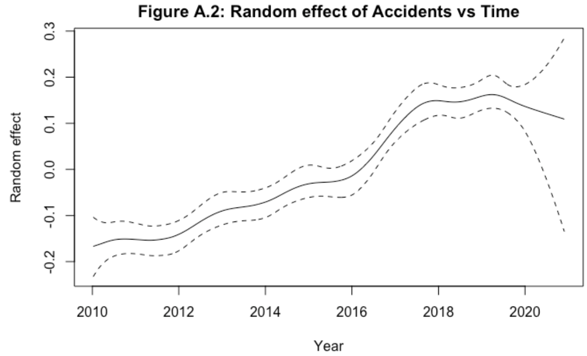


Figure A.3: Actual and predicted number of death of Accidents

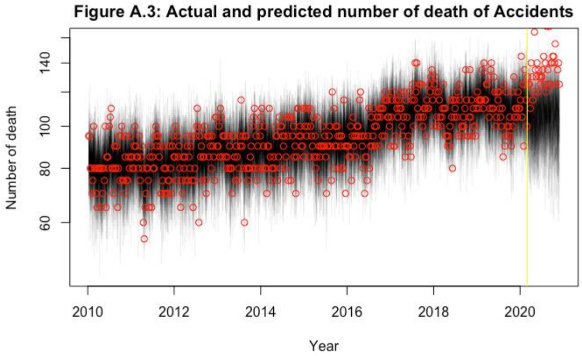
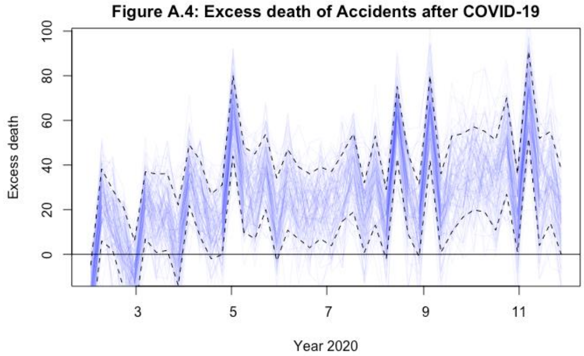


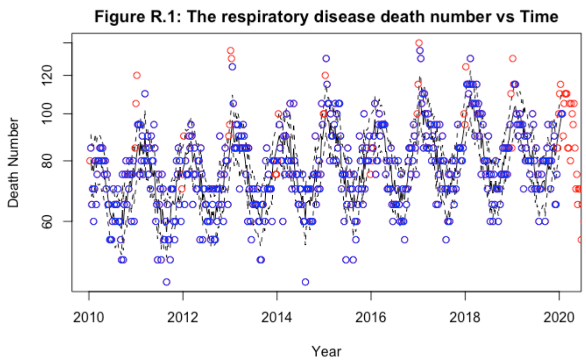
Figure A.4: Excess death of Accidents after COVID-19

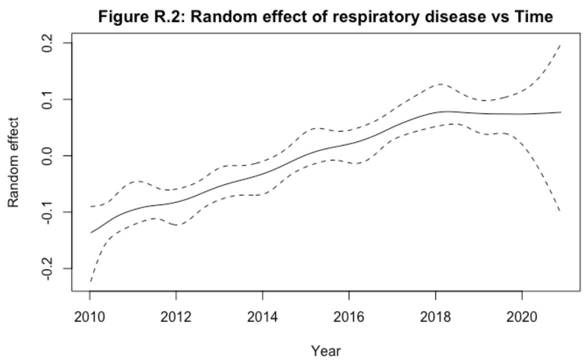


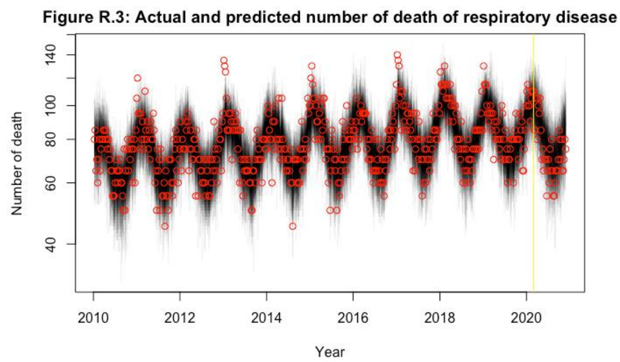
From Figure 4 and Figure A.1, we could tell that the death number related to accidents exhibits an increasing trend from 2010-2020, and the increasing pattern continues to exhibit after the COVID-19 pandemic. From Figure A.2, we can tell that the predicted random effect is in an increasing trend although there are small fluctuations. However, after the COVID-19, the trend became shallower. Figure A.3 shows the comparison between the actual number of death from accidents and the predicted values. We can tell that pattern of an actual number of death from accidents is pretty much the same as the predicted value before the pandemic. However, after the pandemic, the actual number of death seems to be larger than the predicted values.

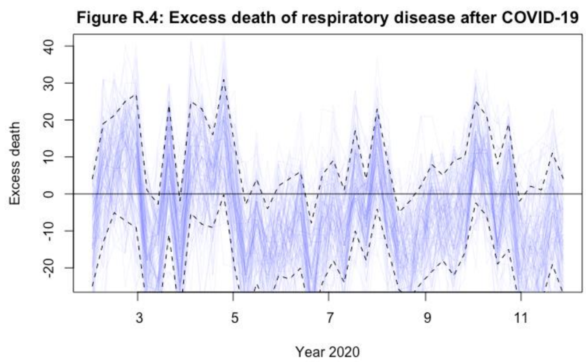
From Figure A.4, for the start of March, and the start of April, the interval contains 0 and there is no strong evidence of excess death. However, for all other periods, the intervals are positive and strictly exclude 0 which means the actual death number is larger than the predicted death number. Thus, after the break out of COVID-19, there was actually excess death related to accidents. The reason might be: the lockdown made safety-related infrastructure less accessible to individuals and people are more willing to go out by driving cars. Thus, accident-related excess death increased after COVID-19 pandemic.

Respiratory:









From Figure 3 and Figure R.1, we could tell that the death number has an increasing trend whereas, after the COVID-19 pandemic, there is a decreasing trend. Figure R.2 shows the predicted random effects of respiratory diseases. There is an increasing trend of it whereas the trend becomes shallower after the COVID-19 pandemic. Figure R.3 shows the comparison between the actual number of death from respiratory diseases and the predicted values. We can tell that pattern of the actual number of death from respiratory diseases is pretty much the same as the predicted value. Overall, there is an increasing trend for the number of death from neoplasms throughout 2010-2020 but and decreasing trend after the pandemic.

From Figure R.4, the interval for the third week of March and the whole of April has a large proportion that is above 0. This means that there might be some excess death from respiratory diseases. However, after May 2020, the interval becomes more “negative”. For most of the period after May, the proportion that the interval below 0 accounts for about 80%. This means that there might be a deficit of death(actual death number from respiratory diseases is less than predicted values). The reason might be: at the start of the pandemic, the health systems were not prepared to face the situation and everything is in a mess. Besides, COVID-19 itself is a type of respiratory disease and has a high death rate. Thus, in the third week of March and the whole of April, there is excess death from respiratory diseases. However, with the law that requires people to keep social distance and wear face masks, the spread of respiratory diseases could get decreased and thus there exhibit a deficit of death compared to before.

Appendix-Q1

```
#install.packages("INLA",repos=c(getOption("repos"),INLA="https://inla.r-inla-download.org/R/stable"),
library("INLA", quietly = TRUE)
#install.packages("Pmisc", repos = "http://r-forge.r-project.org")
library("Pmisc", quietly = TRUE)
#install.packages('lubridate')
library(lubridate)
library(R.utils, quietly = TRUE)
library(Biobase)
#install.packages("R.utils")
#if (!requireNamespace("BiocManager", quietly = TRUE))
#    install.packages("BiocManager")

#BiocManager::install("Biobase")
#install.packages('tinytex')
#tinytex::install_tinytex()
cUrl=paste0("http://scrippsco2.ucsd.edu/assets/data/atmospheric/","stations/flask_co2/daily/daily_flask

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"

co2s$date=as.Date(co2s$day)
co2s$time=strptime(paste(co2s$day, co2s$time),format ="%Y-%m-%d %H:%M",tz ="UTC")

# remove low-quality measurements
co2s=co2s[co2s$quality==0, ]



plot(co2s[co2s$date>as.Date("2015/3/1"),c("date","co2")],log ="y",type ="o",xlab ="time",ylab ="ppm",ce

co2s$dateWeek=as.Date(lubridate::floor_date(co2s$date,unit ="week"))
co2s$timeYears=as.numeric(co2s$date)/365.25
co2s$cos12=cos(2*pi*co2s$timeYears)
co2s$sin12=sin(2*pi*co2s$timeYears)
co2s$cos6=cos(2*2*pi*co2s$timeYears)
co2s$sin6=sin(2*2*pi*co2s$timeYears)
allDays=seq(from =min(co2s$dateWeek),to =as.Date("2030/1/1"),by ="7 days")

table(co2s$dateWeek%in%allDays)
co2s$dateWeekInt=as.integer(co2s$dateWeek)

library("INLA",verbose =FALSE)
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(dateWeekInt,model ="rw2",values =as.integer(allDays),prior ="pc.prec",param =c(0.001,0.
```

```r
            data =co2s, family ="gaussian",
            control.family =list(hyper =list(prec =list(prior ="pc.prec",param=c(1,0.5)))),control.inla

qCols=c("0.5quant","0.025quant","0.975quant")
1/sqrt(co2res$summary.hyperpar[, qCols])


sampleList = INLA::inla.posterior.sample(50, co2res)
sampleMat = do.call(cbind, Biobase::subListExtract(sampleList,
                                                   "latent"))
sampleMean = sampleMat[grep("dateWeekInt", rownames(sampleMat)),
]
sampleDeriv = apply(sampleMean, 2, diff) * (365.25/7)
forSinCos = 2 * pi * as.numeric(allDays)/365.25
forForecast = cbind(`(Intercept)` = 1, sin12 = sin(forSinCos),
                    cos12 = cos(forSinCos), sin6 = sin(2 * forSinCos),
                    cos6 = cos(2 * forSinCos))
forecastFixed = forForecast %*% sampleMat[paste0(colnames(forForecast),
                                                 ":1"), ]
forecast = forecastFixed + sampleMean
#1 CO2 Concentration From Haiwaii Observatory over years
plot(co2s$date, co2s$co2,log ="y",cex =0.3,col ="#00000040",xlab ="time",ylab ="ppm")
#2 CO2 Concentration From Haiwaii Observatory recently
plot(co2s[co2s$date>as.Date("2015/3/1"),c("date","co2")],log ="y",type ="o",xlab ="time",ylab ="ppm",ce
#3Histogram of concentration
hist(co2s$co2,xlab = "CO2 concentration", main="")
#4
matplot(allDays, co2res$summary.random$dateWeekInt[,qCols],
        type = "l", col = "black", lty = c(1, 2,2), xlab = "time", ylab = "y")
abline(v = as.numeric(as.Date("1989-11-10")) , col = "red")
abline(v = as.numeric(as.Date("2020-02-20")) , col = "red")
#5 derivative
matplot(allDays[-1], sampleDeriv, type = "l", lty = 1,
        xaxs = "i", col = "#00000020", xlab = "time", ylab = "deriv",
        ylim = quantile(sampleDeriv, c(0.025, 0.995)))
abline(v = as.numeric(as.Date("1989-11-10")) , col = "red")
abline(v = as.numeric(as.Date("2020-02-20")) , col = "red")
abline(h = 0, lty = "dashed", col = "red")
#6
matplot(allDays, forecast, type = "l", col = "#00000010",
lty = 1, log = "y", xlab = "time", ylab = "ppm")
forX = as.Date(c("2018/1/1", "2025/1/1"))
forX = seq(forX[1], forX[2], by = "1 year")
toPlot = which(allDays > min(forX) & allDays < max(forX))
abline(v = as.numeric(as.Date("1989-11-10")) , col = "red")
abline(v = as.numeric(as.Date("2020-02-20")) , col = "red")
```

Appendix-Q2

```
knitr::opts_chunk$set(echo = TRUE)
#install.packages('tinytex')
#tinytex::install_tinytex()
#tinytex::install_tinytex()
#tinytex:::install_prebuilt()


#install.packages("INLA",repos=c(getOption("repos"),INLA="https://inla.r-inla-download.org/R/stable"),
library("INLA", quietly = TRUE)
#install.packages("Pmisc", repos = "http://r-forge.r-project.org")
library("Pmisc", quietly = TRUE)
#install.packages('lubridate')
library(lubridate)
library(R.utils, quietly = TRUE)
#if (!requireNamespace("BiocManager", quietly = TRUE))
#    install.packages("BiocManager")

#BiocManager::install("Biobase")
library(Biobase)
deadFile=Pmisc::downloadIfOld("https://www150.statcan.gc.ca/n1/tbl/csv/13100810-eng.zip")
(deadFileCsv =deadFile[which.max(file.info(deadFile)$size)])
x=read.csv(deadFileCsv)
x[1:2, ]
x$date=as.Date(as.character(x[[grep("DATE",names(x))]]))
x$province=gsub("[,].*","", x$GEO)
x=x[x$date<as.Date("2020/12/01")&x$province=="Ontario", ]

diseases_type <- c("heart","neoplasms","respiratory","Accidents")
for(i in seq(1,4)) {
  cur_disease = diseases_type[i]
  plot(x[grep(cur_disease, x$Cause),c("date","VALUE")],ylab =cur_disease)
  abline(v =as.Date("2020/03/17"))
  title(capture.output(cat("Figure", i, ": ","The", cur_disease, "disease death number vs. Time")))
  }

dateSeq=sort(unique(x$date))
table(diff(dateSeq))
dateSeqInt=as.integer(dateSeq)
x$dateInt=x$dateIid=as.integer(x$date)
x$cos12=cos(2*pi*x$dateInt/365.25)
x$sin12=sin(2*pi*x$dateInt/365.25)
x$sin6=sin(2*2*pi*x$dateInt/365.25)
x$cos6=cos(2*2*pi*x$dateInt/365.25)
x$dayOfYear=as.Date(gsub("^[[:digit:]]+","0000",x$date))
x$christmasBreak=(x$dayOfYear>=as.Date("0000/12/21"))|(x$dayOfYear<=as.Date("0000/01/12"))

#Here I do 4 times to generate 16 photos, for 4 diseases, and there are 4 photos for each diseases.
xSub=x[grepl("Accidents", x$Cause,ignore.case =TRUE)&x$province=="Ontario", ]
xPreCovid=xSub[xSub$date<as.Date("2020/02/01")&(!xSub$christmasBreak), ]
library("INLA")

##change pc.prec
```

```r
resHere=inla(VALUE~cos12+cos6+sin12+sin6+f(dateInt,model ="rw2",values =dateSeqInt,prior ="pc.prec",para
               f(dateIid,values =dateSeqInt,prior ="pc.prec",
                 param =c(1,0.5)),data =xPreCovid,family ="poisson",
             control.compute =list(config =TRUE),control.predictor =list(compute =TRUE))

#1
matplot(resHere$.args$data$date, resHere$summary.fitted[,paste0(c(0.025,0.975,0.5),"quant")],type ="l",l
points(xSub$date, xSub$VALUE,col ="red")
points(xPreCovid$date, xPreCovid$VALUE,col ="blue")
title("Figure A.1: The Accidents death number vs Time")



#2
matplot(dateSeq, resHere$summary.random$dateInt[,paste0(c(0.025,0.975,0.5),"quant")],type ="l",lty =c(2
toPredict=cbind(`(Intercept):1`=1,`cos12:1`=cos(2*pi*dateSeqInt/365.25),`sin12:1`=sin(2*pi*dateSeqInt/36
dateIntSeq=paste0("dateInt:",1:length(dateSeqInt))
dateIidSeq=paste0("dateIid:",1:length(dateSeqInt))
resSample=inla.posterior.sample(n =100, resHere)
resSampleFitted=lapply(resSample,function(xx) {toPredict%*%xx$latent[colnames(toPredict), ]+xx$latent[da
resSampleFitted=do.call(cbind, resSampleFitted)
resSampleLambda=exp(resSampleFitted)
resSampleCount=matrix(rpois(length(resSampleLambda),resSampleLambda),nrow(resSampleLambda),ncol(resSampl
title("Figure A.2: Random effect of Accidents vs Time")

#3
matplot(dateSeq, resSampleCount,col ="#00000010",type ="l",lty =1,log ="y",ylab="Number of death",xlab=
points(xSub[,c("date","VALUE")],col ="red")
abline(v =as.Date("2020/03/01"),col ="yellow")
is2020=dateSeq[dateSeq>=as.Date("2020/2/1")]
sample2020=resSampleCount[match(is2020, dateSeq),]
count2020=xSub[match(is2020, xSub$date),"VALUE"]
excess2020=count2020-sample2020
title("Figure A.3: Actual and predicted number of death of Accidents")
#4
matplot(is2020, excess2020,type ="l",lty =1,col ="#0000FF10",ylim = range(-10, quantile(excess2020,c(0.
title("Figure A.4: Excess death of Accidents after COVID-19")
matlines(is2020,t(apply(excess2020,1,quantile,prob=c(0.1,0.9))),col="black",lty=2)
abline(h=0)
quantile(apply(excess2020,1,sum))
sample2020
```