

Question 1. How do the fall of the Berlin wall and the outbreak of COVID-19 influence the emission of CO₂

Summary

We investigate the effect of two human activities on the CO₂ Concentration. Both the fall of Berlin Wall in November 1989 and the global lockdown due to the outbreak of COVID-19 starting from February 2020 have slow down the CO₂ emission rate in the air. It enlighten us that it's important to develop green technology as the CO₂ Concentration are rising rapidly in recent decades since 1960.

Introduction

CO₂ plays a key role in global warming. In order to create a safe and healthy environment for us human beings, it's essential for researchers to monitor and control the concentration of CO₂ in the air. In this report, we would like to investigate whether CO₂ concentration is affected by human activities, in particular, the fall of the Berlin wall in November 1989 and the outbreak of COVID-19 starting from February 2020. Data was provided by Mauna Loa Observatory in Hawaii and edited by the Scripps CO₂, and it reported the daily average CO₂ concentration (ppm) based on flask samples from March 1960 to October 2020.

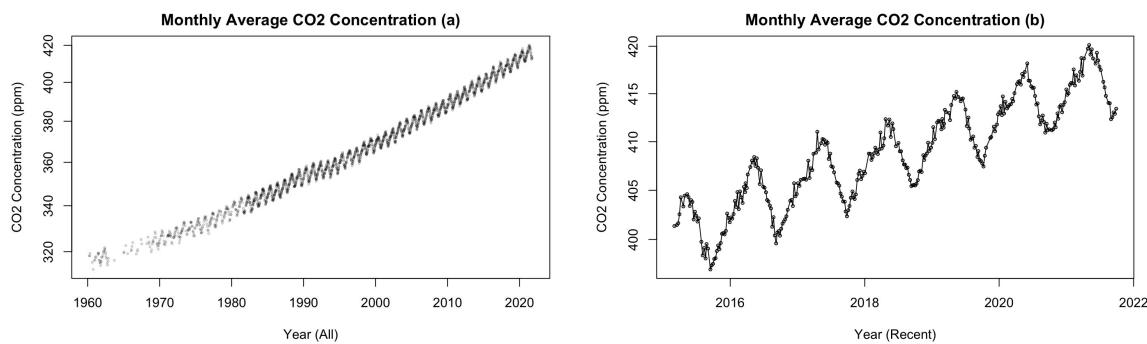


Figure 1: CO₂ at Mauna Loa Observatory, Hawaii

Methods

From the Figure 1(a), it can be seen that data is non-linearly increasing with seasonal pattern and the slope is different for each point. Then we implement a generalized additive model with random walk 2 in order to increase smoothness and assume the data follows normal distribution. The model is written as below:

$$Y_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = X_i \beta + U_{t_i}$$

where Y_i is the real daily average CO₂ concentration that follows normal distribution, μ_i is the expected daily average CO₂ concentration for i^{th} day, X_i is the covariates of all sinusoidal seasonal fixed effects and β is their parameters. U_{t_i} is the second-order random walk, since the slope either increases or decreases by 0.001 times daily. There are four sinusoidal functions for the fixed seasonal effects with the frequency of 12 month and 6 month, they are written as follow:

$$X_{i1} = \sin 12 = \sin(2\pi t_i / 365.25)$$

$$X_{i2} = \cos 12 = \cos(2\pi t_i / 365.25)$$

$$X_{i3} = \sin 6 = \sin(2\pi t_i / 182.625)$$

$$X_{i4} = \cos 6 = \cos(2\pi t_i / 182.625)$$

where t_i is the number of days since the originate date and divide by 365.25, the number of days it takes for the earth to go around the sun. The second-order random walk U_{t_i} is applied with a penalized term that:

$$\text{Prob}(\sigma > 0.001) = 0.5$$

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

Results

Figure 2 shows the CO₂ concentration from 1960 to 2020 produced by the fitted model based on the actual data, also the prediction within the confidence interval for the future CO₂ concentration. We could see that the CO₂ concentration steadily increases overall, despite a slight decrease around the year 1991 and 2020. The seasonal pattern can also be found on the graph and further details will be explain in the next paragraph.

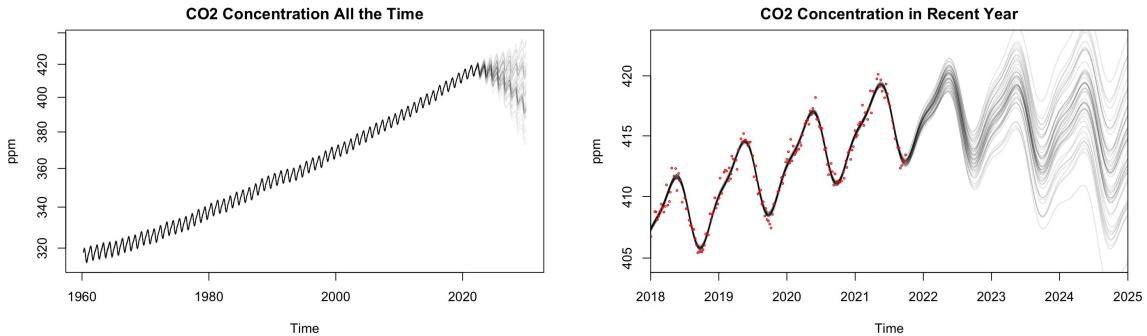


Figure 2: CO₂ Concentration Along the Time

The two vertical lines drawn on the Figure 3(a) indicates two important time period, the fall of Berlin wall in November 1989 and the lockdown during the COVID-19 pandemic from February, 2020. The derivative of CO₂ concentration at these two points are still positive, while the magnitude of these two points are lower smaller than before. Despite CO₂ concentration still increases in both scenarios, the speed is much slower than before. Then, the derivative gradually decreases below 0 in 2-3 year after the fall of Berlin wall in November 1989 according to Figure 3(a), which is a decrease in the CO₂ concentration. It has proven our hypothesis on the fall of Berlin Wall that it precedes a dramatic fall of industrial production in the Soviet Union and Eastern Europe which decreases the CO₂ production rate and finally leads to the decrease of CO₂ concentration in the air. While due to rapid global development in the later 2-3 decades, the derivative of CO₂ concentration increases rapidly. Similarly, as a result of the globe lockdown of COVID-19 and quarantine policy, the speed of CO₂ production decreases starting from February, 2020. The shut down of factories and the less greenhouse gas emission from car contributes to this phenomenon. Figure 3(b) shows the derivative of CO₂ concentration in the recent year, which shows the trend clearly in the recent year. According to our prediction in Figure 3(a), the derivative of CO₂ concentration might go below 0 in the future.

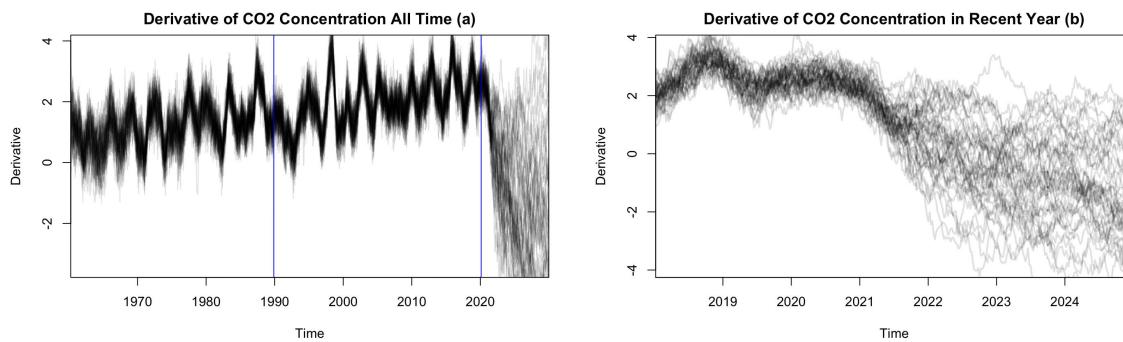


Figure 3: The Derivative of CO₂ Concentration Along the Time

Furthermore, Figure 4 shows the change of random effect over time where the effect increases. It also shows the predicted confidence interval in the future. It's a supplement to show that human activities of developing economics has increases the CO₂ emission.

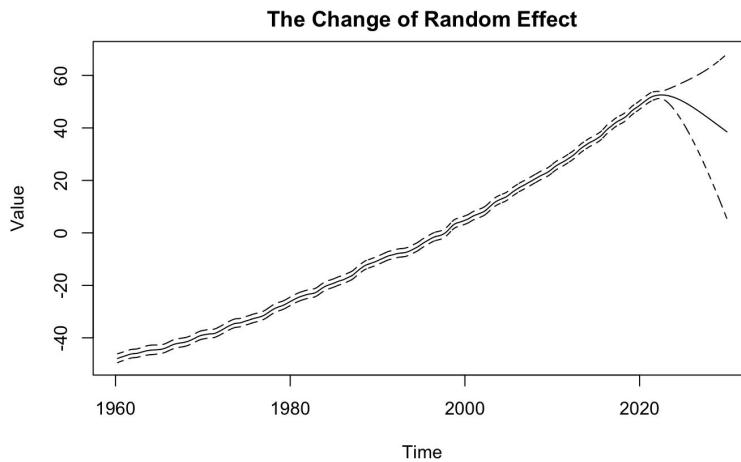


Figure 4: The Change of Random Effect Over Time

Question 2. The Effect of COVID-19 on Malignant neoplasms, Diseases of the heart, Accidents and chronic lower respiratory diseases

Summary

We investigate the hypotheses on the death counts on Malignant neoplasms, Diseases of the heart, Accidents and Chronic lower respiratory after the outbreak of COVID-19 based on data provided by Statistical Canada. To conclude, the hypotheses on the death counts for Malignant neoplasms, Heart Diseases and Respiratory Diseases is correct that the first two increases and the third one decreases after COVID-19. Surprisingly, there are excessive death for accidents after COVID-19 due to the shuttle of road safety facility.

Introduction

The COVID-19 pandemic has a huge impact on global morality rate. The objective of the report is to investigate people in Ontario from March to November 2020, we are interested in whether the morality rate of Malignant neoplasms and Diseases of the heart in increases as COVID-19 become a barrier for people to receive healthcare. Meanwhile, we suspect that the morality rate of Accidents and Chronic lower respiratory decreases due to quarantine, masks and social distance policy. Data records daily cause-specific mortality counts and is provided by Statistical Canada.

Methods

0.0.1 Data Visualization

Figure 5 displays death counts for Diseases of the heart, Malignant neoplasms, Accidents and Chronic lower respiratory disease from 2010 to 2020 separately. In Figure 5(a), there is a slight increase trend

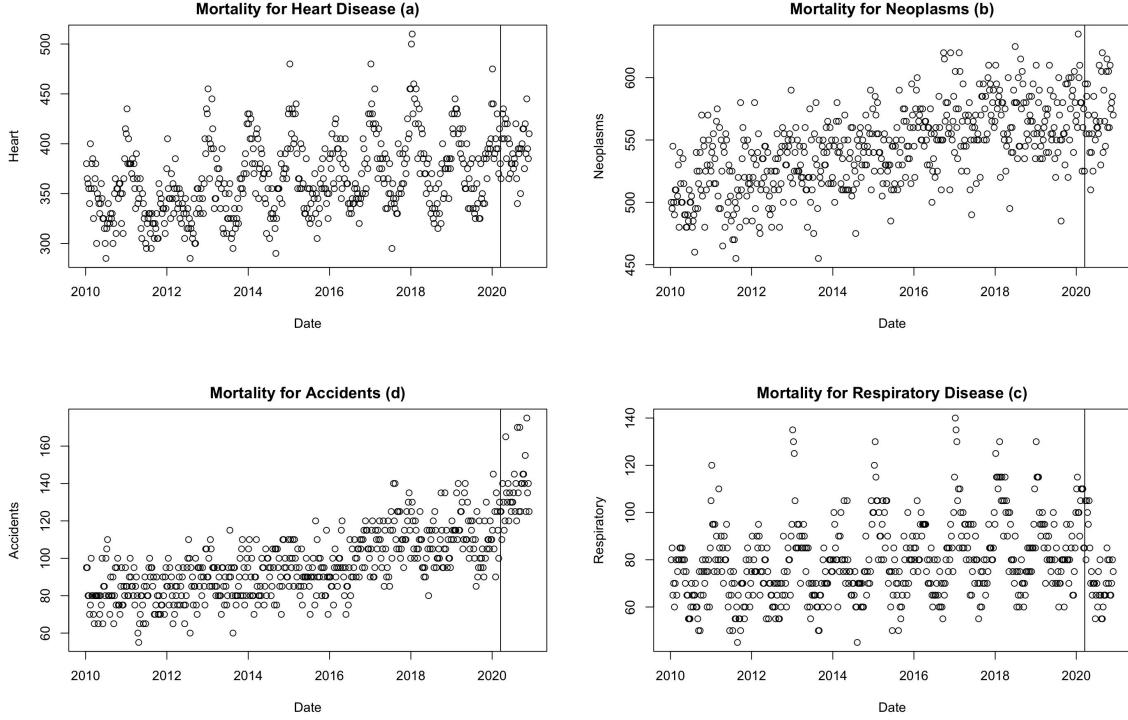


Figure 5: Death Counts in Ontario from 2010 to 2020

for heart disease starting from March 17, 2020. The similar trend can be found for Malignant neoplasms and Accidents where death counts increase. While the death counts for respiratory disease decreases starting from March 17, 2020. We can also tell that the data has seasonal effect and we will do further investigation in our model construction.

0.0.2 Model

The dataset consists of morality counts over a decade in Ontario, which is positive and discrete. Therefore, we can make the assumption that the death counts for people in Ontario follows Poisson distribution. In addition, there is seasonal pattern in the data, so we can fit a non-parametric model:

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

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

where Y_i is the real daily death counts in Ontario, λ_i is average expected death daily death counts. X_i is the covariates of all sinusoidal seasonal fixed effects and β is their parameters. U_{t_i} is the second-order random walk for the smoothing purpose as the slope changes by 0.001 times daily, and V_i is the overdispersion term. There are four sinusoidal functions for the fixed seasonal effects with the

frequency of 12 month and 6 month, they are written as follow:

$$X_{i1} = \cos 12 = \cos(2\pi t_i/365.25)$$

$$X_{i2} = \cos 6 = \sin(2\pi t_i/182.625)$$

$$X_{i3} = \sin 12 = \sin(2\pi t_i/365.25)$$

$$X_{i4} = \sin 6 = \sin(2\pi t_i/182.625)$$

where t_i is the number of days since the orginate date and divide by 365.25, the number of days it takes for the earth to go around the sun. The second-order random walk U_{t_i} is applied with a penalized term and the overdispersion term V_i follows Normal distribution with penalized prior:

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

$$\text{Prob}(\sigma_v > \log(1.2)) = 0.5$$

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

$$\text{Prob}(\sigma > 0.001) = 0.5$$

Results

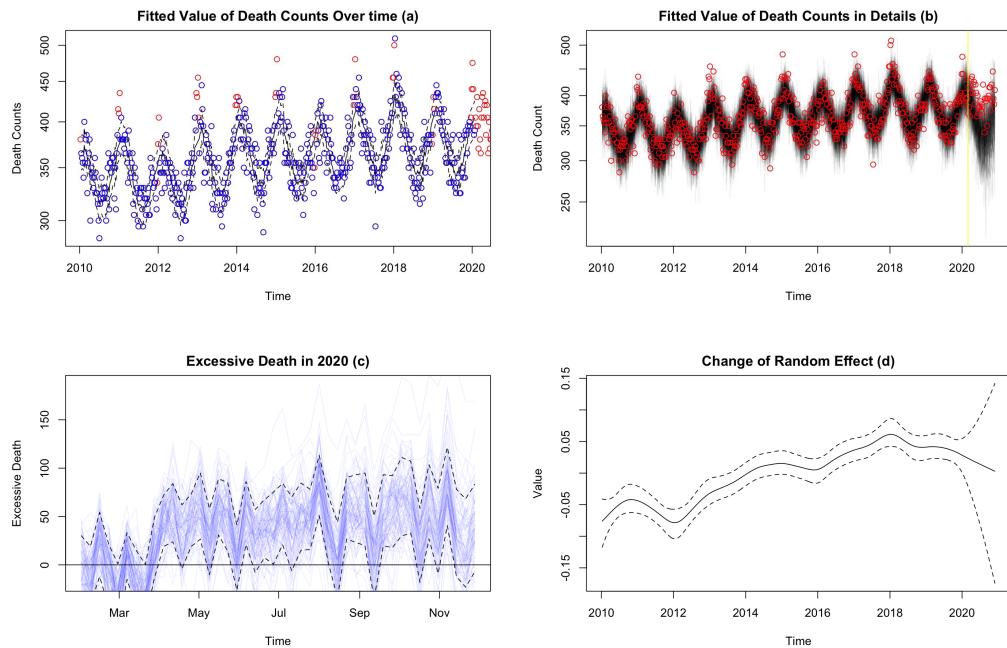


Figure 6: Death Counts of Heart Disease in Ontario from 2010 to 2020

Figure 6 makes the comparison of the deaths counts of Diseases of the heart, in Ontario before COVID-19 and after COVID-19. From Figure 6(a), the red dots represent the death counts after COVID-19 and for each Christmas. The death rate in Christmas is higher than other time, it might due to the lack of health care during the break. The vertical line in Figure 6(b) indicates the time March 2020 where the COVID-19 began. It's clear to see that the death of those diseases has a seasonal effect and a decreasing trend after the outbreak of COVID-19, and the red points are the actual value while the black lines are predicted value. It shows that our prediction is appropriate as the line and points overlap. Figure 6(c) shows excessive death of Heart Disease, the dashed lines represent the 80% credible interval for excess death. Most of the 80% credible interval is above 0, which means there is excessive death counts for heart attack after COVID-19. Figure 6(d) shows the change of random effect over time, the confidence interval for the predicted future change, and the slope at each point is changing and shows an increasing trend in general.

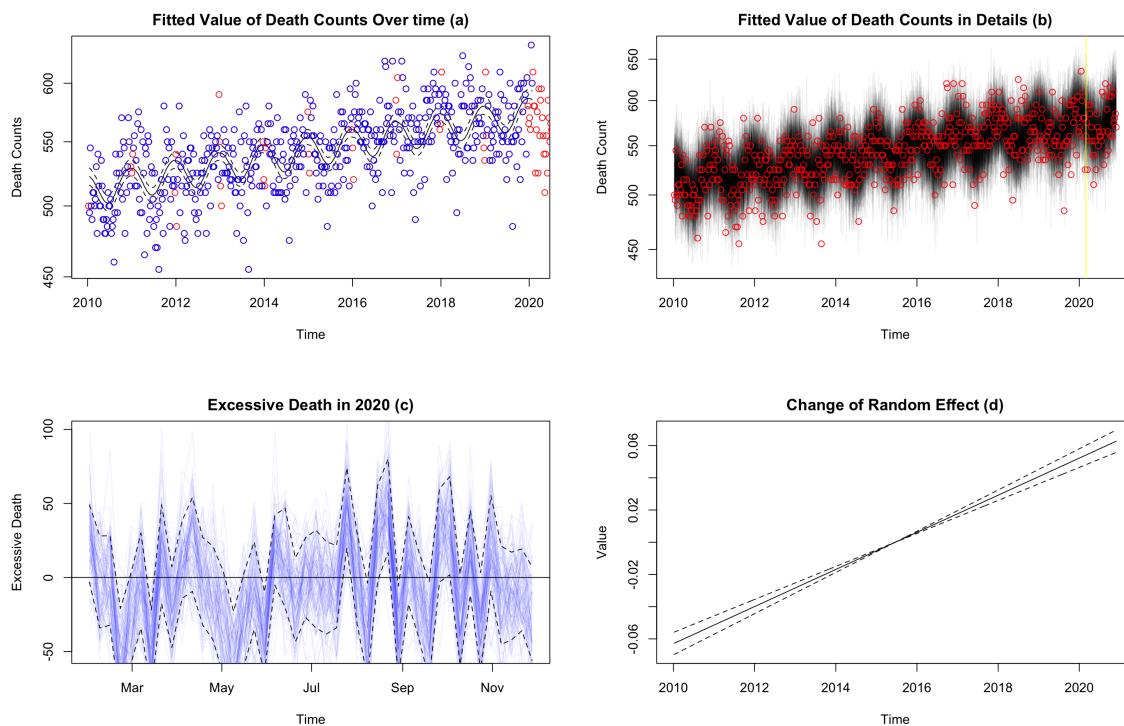


Figure 7: Death Counts of Malignant neoplasms in Ontario from 2010 to 2020

Similarly, Figure 7 describes the situation for Malignant neoplasms. The death counts for Christmas is approximately the same compared to the normal time. Also, the death counts after the COVID-19 fluctuates and follows the same trend as before. More details could be seen in Figure 7(c), the 80% credible interval for excess death counts gathers around 0 except a small increase around August, which means there isn't significant change for Malignant neoplasms after COVID-19. Figure 7(b) shows that our predicted values and actual values fit well as black line and red line overlaps each other despite some red points at the top our model fail to capture. The random effect has an increasing trend. Therefore, the death counts for Malignant neoplasms has a slight increase after COVID-19.

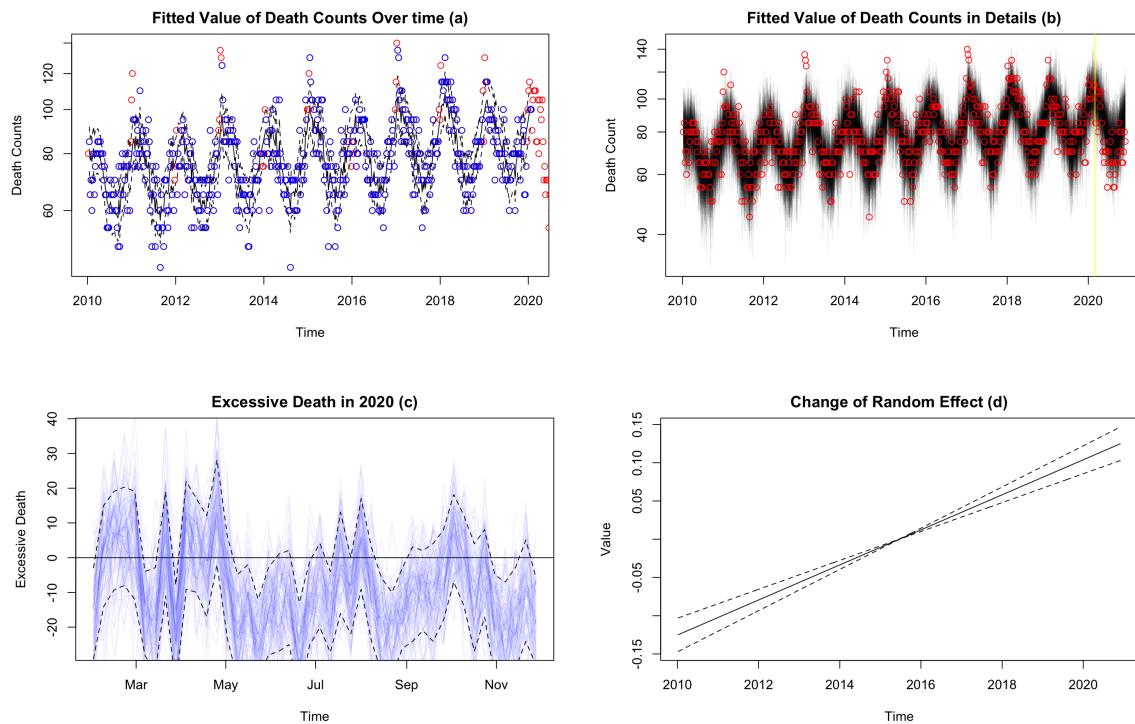


Figure 8: Death Counts of Respiratory Diseases in Ontario from 2010 to 2020

As for Chronic lower respiratory diseases, the death counts over Christmas is higher than usual. Since people like to gather and party during Christmas and it's more difficult to access to health care. While the death counts after COVID-19 shows a downward trend, according to Figure 8(c). Only the 80% credible interval for excess death in March to April is above 0, starting from May the excess death goes down below 0 which is deficit due to the policy of wearing masks and quarantine. While the data rises a bit in August as the government policy about quarantine move to step three of the roadmap to reopen. Our prediction fits well with the actual data according to Figure 8(b).

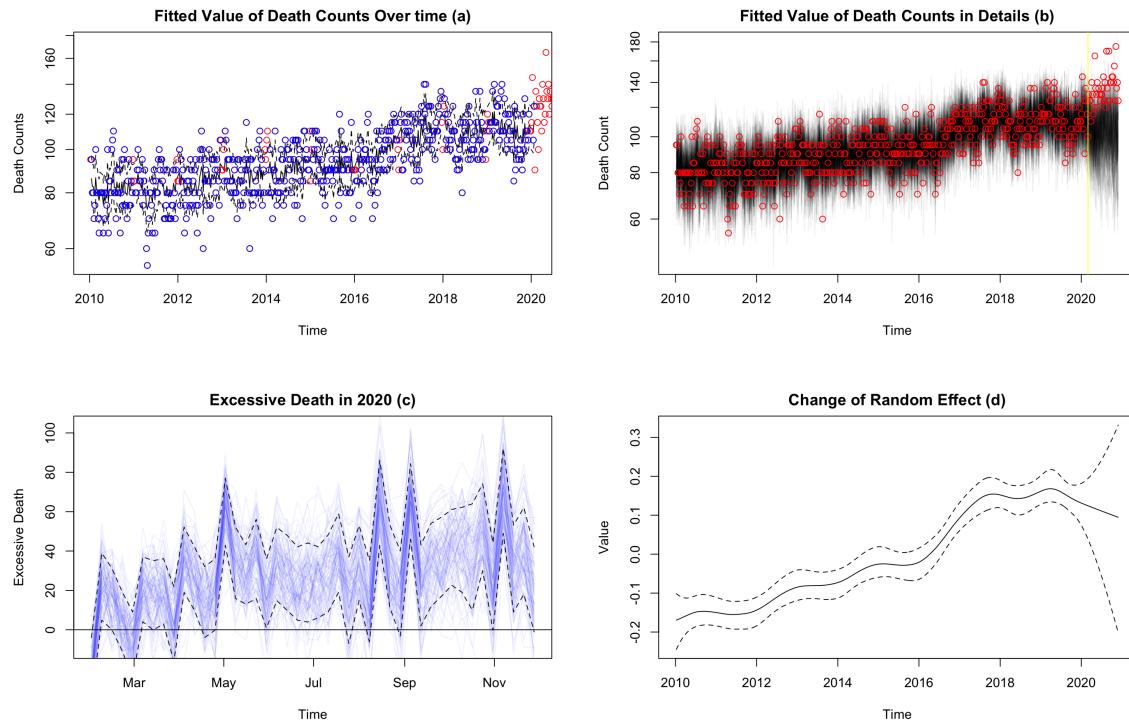


Figure 9: Death Counts of Accidents in Ontario from 2010 to 2020

According to Figure 9, there is a dramatic increase trend for the death counts of accidents after COVID-19. Death counts over Christmas is approximately the same as other normal time. It could be seen from Figure 9(c) that almost all 80% credible interval for excess death are above zero starting at March. Accountable reasons could be that it takes longer time for ambulance to arrive and people choose to drive with safety concerns after COVID-19. Figure 9(d) draws an upward trend for random effect.