Time Series investigation on CO2 concentration

Yutong Yuan 1004725103

April 10, 2022

Abstract

Climate change has always been a significant threat to humanity, and each time global warming intensifies, it cannot be ignored. The main reason of global warming is the concentration of CO2. In this article, you will realize the rapid acceleration of CO2 concentration over time. We will present the CO2 concentration data from 1960 to 2021 collected by Mauna Loa Observatory in Hawaii, with the emphasis on the seasonal change of co2 concentration level from 2015 to 2021. Additionally, the time series forecast presented in this article indicates that the CO2 concentration will continue to rise in a seasonal trend. Arima method is being used to find models and estimations.We also build up the 95% prediction interval and the 95% confidence interval to help us analyze the data. We conclude that human development activities are the main cause of the dramatic seasonal increase in carbon emissions. Since the industrial revolution, human-caused carbon dioxide emissions have grown year after year, and the greenhouse effect of the atmosphere has intensified, triggering climate change and posing a series of serious challenges.

Keywords: Time Series, SARIMA Model, CO2 concentration, global warming

Introduction

The dramatic increase in carbon emissions has continued to strengthen the greenhouse effect, causing the average global temperature to climb. In turn, global climate change is further causing more complex and dramatic climate change, triggering more frequent and more devastating natural disasters. At this time, It would be helpful if we could build a relatively accurate CO2 concentration model to predict future quality of atmosphere based on time series analysis. Therefore, human can take the action to prevent those possible disasters. To build the model, The dataset we used includes the Daily average CO2 concentrations (ppm) derived from flask air samples in Mauna Loa Observatory, Hawaii, for 2240 observations from 1960 to present. The dataset can be accessed from https://scrippsco2.ucsd.edu/data/atmospheric_co2/mlo.html Based on the recent science publication made by Mauna Loa (2022), Forecast of the annual rise in atmospheric carbon dioxide concentration measured at Mauna Loa, Hawaii for 2022. Their scientific forecast result shows that the annual average CO2 concentration at Hawaii will be 2.14 ± 0.52 parts per million (ppm) higher in 2022 than in 2021. This conclusion is consist with our prediction result of seasonal ARIMA model, the CO2 concentration will keep the upward trending after 2021 in short-term.

Statistical Methods

Figure 1.1 shows the pattern of weekly CO2 concentration from 1960 to 2021, and Figure 1.2 shows the weekly CO2 concentration from 2015 to 2021. As we can see, there exists a significant upward trend with seasonal pattern of CO2 concentration in both recent six decades and six years. In this report, we will focus on the co2 concentration in recent 6 years, which has 311 weekly observations. Firstly, data transformation is needed for better investigation. Since it is obvious that the data is not stationary yet. By differentiating the logged data, we removed the trends and get a time series. Observed that there are

persistence occurring, we are going to use a 52 order difference on the differentiated log data, since there are 52 weekly data collected in a year. From the plot we've plotted, we can observe that it is stationary with constant means and variance around zero.

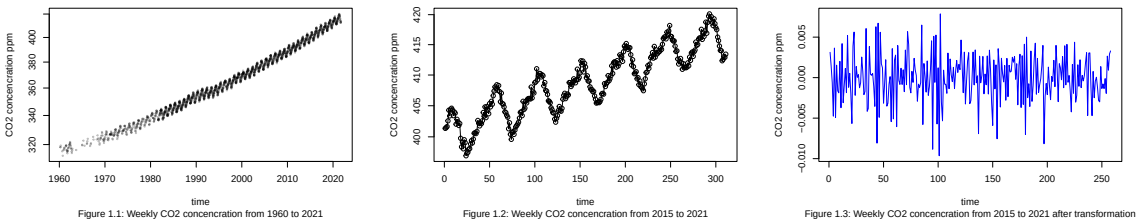

Figure 1: Original Data and Transformed Data

I'm going to use SARIMA model to forecast. We need to use the ACF and PACF plots to find our AR parameter and MA parameter.

Figure 2 shows the ACF and PACF for the transformed series.

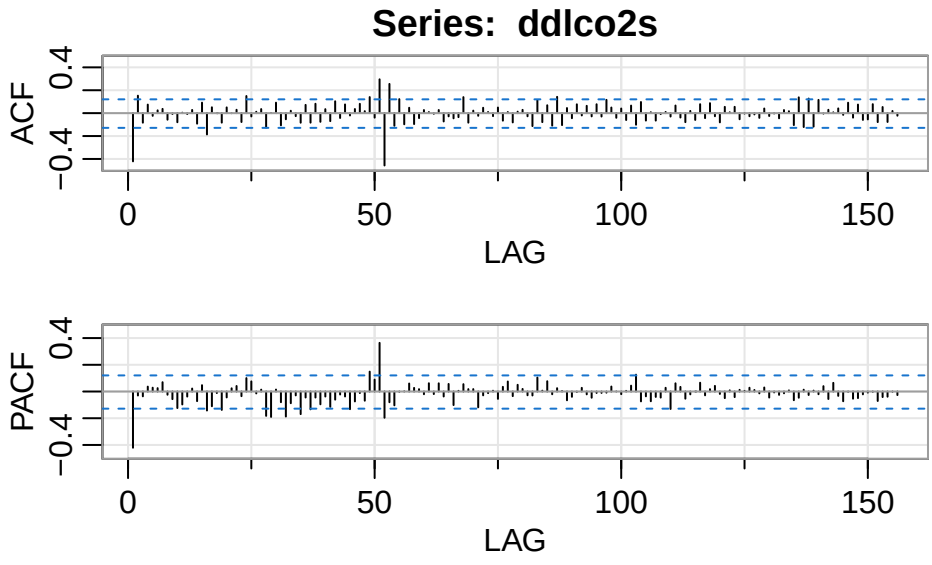

Figure 2: ACF & PACF of Transformed Data

We build Seasonal ARIMA models to fit the transformed series. Based on the Figure 2 plot of ACF and PACF, we can find that the ACF cut off at 1 or 0, PACF cut off at 1. Thus, we

set p=1 and q=1 or 0. We did non-seasonal difference once for the time series, thus d=1. We also did seasonal difference once, thus D=1. The ACF cut off at lag=1s, PACF cut off at lag=1s, therefore, we can set P=1, Q=1 with 1s equals to 52. We will try three models, Model1 $ARIMA(1,1,0)(1,1,1)_{52}$, Model2 $ARIMA(1,1,1)(1,1,1)_{52}$, and the last model is fitted if P=0, Model3 $ARIMA(1,1,1)(0,1,1)_{52}$.

Figure 3 shows the residual analysis for Model1 $ARIMA(1,1,0)(1,1,1)_{52}$.

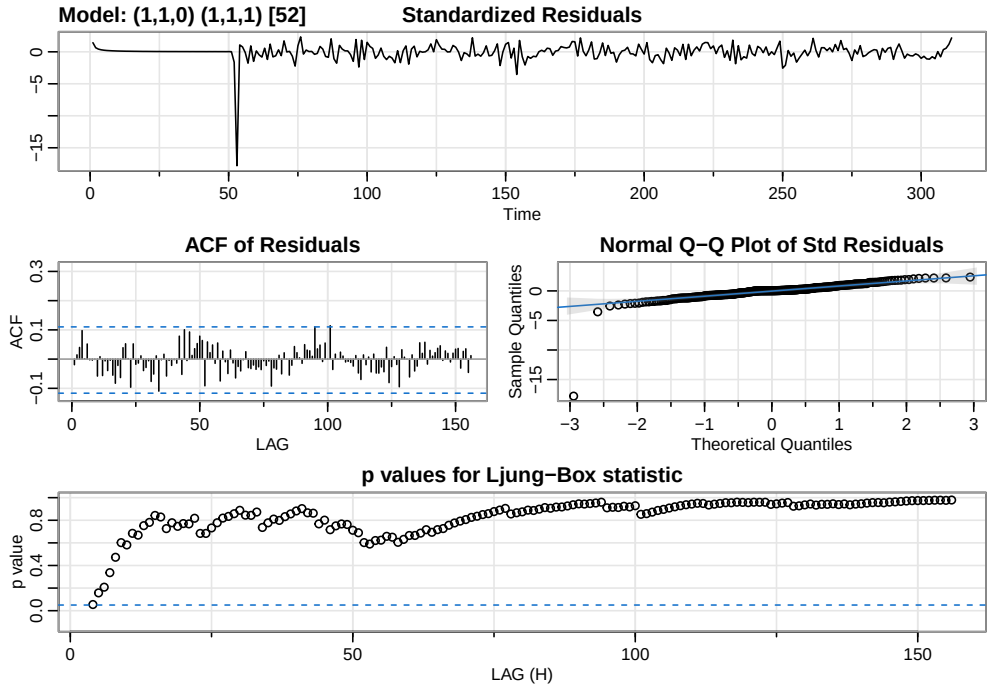

Figure 3: Residual Analysis for Model1

Figure 4 shows the residual analysis for Model2 $ARIMA(1,1,1)(1,1,1)_{52}$.

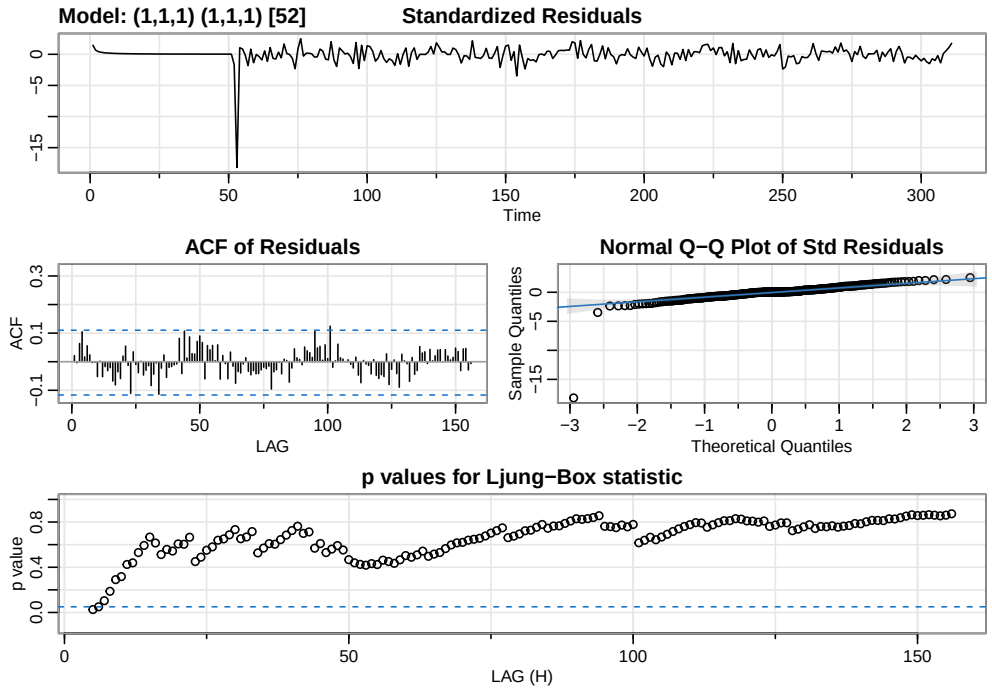

Figure 4: Residual Analysis for Model2

Figure 5 shows the residual analysis for Model3 $ARIMA(1,1,1)(0,1,1)_{52}$

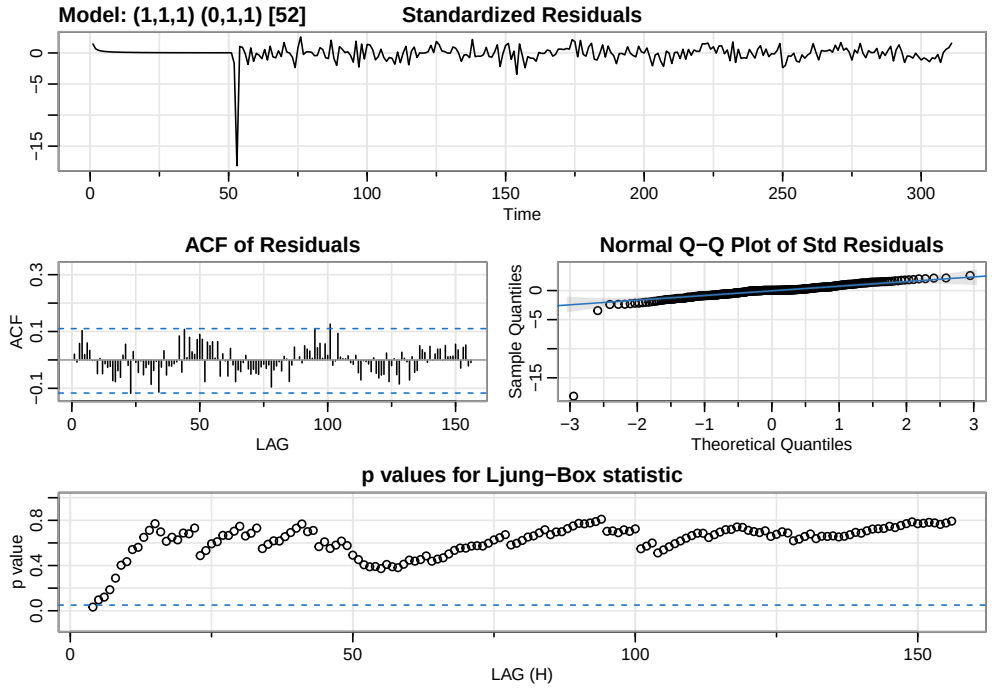

Figure 5: Residual Analysis for Model2

For all of three models, the standardized residuals after 1s, are similar to white noise with a no pattern plot, constant variance and zero mean. The ACF of Residuals are all almost within the limit, except for a singe point in model 1 and 2 are slightly over. It shows that all the ACF residuals are small enough to be ignored, which means that those residuals are all uncorrelated with each other. This conclusion agrees with the normal Q-Q plot that most points are on the line with very few outliers beside, shows no departure from the normal distribution assumption. Ljung-Box statistic p-value indicates that in our models, almost all the lags have a p-value above the significant level, larger than 0.05, which means that we do not reject the null hypothesis that residuals are independent. The further model analysis will based on AIC, AICc and BIC results.

Table 1 shows the AIC, AICc and BIC for three models

Table 1: AIC, AICc, and BIC for Three Models

Model	AIC	AICc	BIC
Model1	-9.130256	-9.129890	-9.075171
Model2	-9.145236	-9.144623	-9.076380
Model3	-9.144870	-9.144503	-9.089785

Table 1 shows the AIC, AICc and BIC for three models we built, we can tell that model 2 $ARIMA(1,1,1)(1,1,1)_{52}$ have the smallest value of AIC and AICc, with relative small BIC. Considering AIC, AICc, and BIC and their residual analysis results, we choose model 2 $ARIMA(1,1,1)(1,1,1)_{52}$as our prediction model as following format:

$$(1-\phi_1 B)(1-\Phi_1 B^{52})\nabla_{52}\nabla log x_t = (1+\theta_1 B)(1+\Theta_1 B^{52})w_t$$

where x_t indicates time series, w_t~$N(0,\sigma^2)$.

Results

By fitting the model selected in previous section, we got parameter estimations and corresponding p-values in Table 2.

Table 2: Estimated Coefficient and P-value

	Estimate	SE	t.value	p.value
ar1	-0.4252	0.1489	-2.8558	0.0046
ma1	0.1342	0.1603	0.8372	0.4033
sar1	-0.2278	0.1528	-1.4904	0.1374
sma1	-0.3329	0.1643	-2.0257	0.0438

Table 2 shows the estimated coefficients in our model, we can see that all coefficients, including non-seasonal AR(1), non-seasonal MA(1), and seasonal AR(1) seasonal MA(1), are significant with a p-value less than 0.05. Therefore, the final fitted model for $ARIMA(1,1,1)(1,1,1)_{52}$ is:

$$(1+0.4252B)(1+0.2278B^{52})\nabla_{52}\nabla log\hat{x_t} = (1+0.1342B)(1-0.3329B^{52})w_t$$

Therefore, since p = P = 1, this means that the number of auto-regressive term is 1. Also, d = D = 1 since we have to do one differentiating to make the series stationary.We also set q =1 and seasonal Q = 1, this means the number of moving average terms separately in the prediction equation, and s = 52 indicates seasonal length of 52 weeks in the data.

Prediction

After analyzing the fitted model, we can use it to predict the future CO2 concentration in Hawaii after 2021.

Figure 6 shows the estimated next 10 weeks CO2 concentration's log value in red dots and

Table 3: Forecast weekly log CO2 concentration

Week	Prediction	Lower.bond.of.95..PI	Uower.bond.of.95..PI
1	6.024733	6.020079	6.029386
2	6.024113	6.018408	6.029817
3	6.022310	6.015414	6.029206
4	6.021031	6.013238	6.028824
5	6.022246	6.013604	6.030888
6	6.020211	6.010814	6.029608
7	6.020174	6.010072	6.030276
8	6.019064	6.008306	6.029823
9	6.019516	6.008137	6.030894
10	6.019397	6.007431	6.031363

corresponding prediction intervals in grey areas.

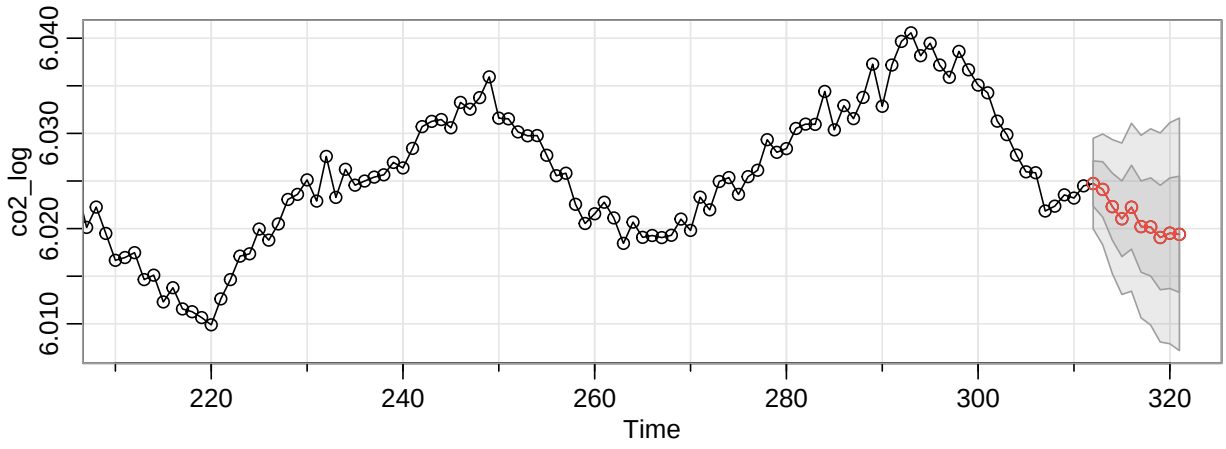

Figure 6: ten-week Forecast using Model 2

Table 3 shows the prediction values and the 95% confidence intervals for the ten prediction values.

From the forecasting results in Figure 6, we can see that the predicted 10 weeks log CO2 concentration is following the previous trend, but performing an upward pattern over the entire timeline. Since one cyclical season is 52 weeks, the prediction 10-week is only a small part in one season. This prediction indicates that, in the next 10 weeks, the log CO2

concentration will get lower, which is reasonable, since the time is July, the summer period always has the lower demand of heat and low emission if co2. However, based on two cyclical seasons of log CO2 concentration in figure 6, the prediction tells us that for the same week, the log CO2 concentration in prediction period will be higher than that in last year. Based on Table 3, we can see the range of confidence interval keep increasing as the prediction term increase, but the range of the confidence interval at 10 weeks later is still relatively narrow, which suggests an relatively accurate estimation. Figure 7 shows the periodogram of the weekly CO2 concentration in Mauna Loa Observatory, Hawaii.

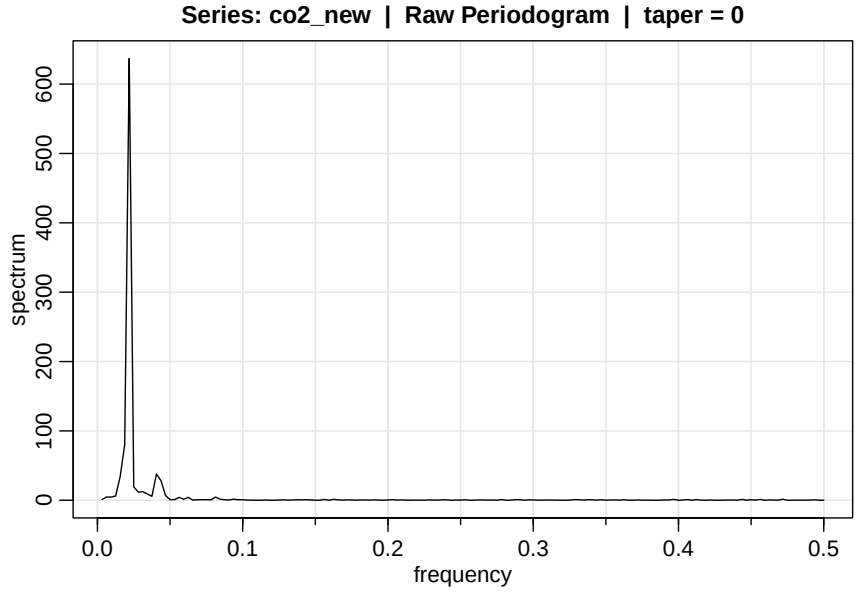

Figure 7: Periodogram of Weekly CO2 concentration

We can tell from Figure 7 that there is a peak at around 0.02, closing to the value of 1/52. This consistent with previous seasonal analysis with s=52. Since 1/s = 1/52 = 0.0192, from where on the graph we can tell it reaches a peak at around 0.02. This is close to our value and this supports that it is consistent with out previous seasonal analysis with s = 52. Table 4 shows the first three predominant spectrum with the 95% confidence intervals.

From table 4, we find out the 95% confidence intervals are relatively large. Hence we cannot conclude anything from this, the table cannot tell much information about significance.

Table 4: First Three Predominant Frequency of Weekly CO2 concentration

Order	frequency	period	spectrum	lower	upper
First	0.0219	45.7143	636.9402	12.392462	1805.6184
Second	0.0187	53.3333	79.8337	14.457859	2106.5528
Third	0.0406	24.6154	37.8756	6.672866	972.2564

Discussion

Based on our data analysis and prediction of the CO2 concentration over time, we can find that the CO2 concentration is always increasing with the seasonal pattern from 318.27ppm in 1960 to 413.43ppm in 2021. This may be caused by the fast-growing global industry with better technology and larger demand. Also, the pattern of CO2 concentration is fluctuating seasonally within one year, peak at winter, large drop in summer, which is reasonable since the larger heating demand in winter and lower demand in other season, also includes some other seasonal human activities differences.However, there still exist some limitation of our model. For the prediction of the next ten weeks, we made the co2 concentration forecasting for 2021 July, which is the time of COVID-19 with the lock-down and work-from-home policy in worldwide. But in this model, we did not consider how will the COVID pandemic influence the CO2 emission or global industry. Secondly, there may exist some considerations on the SARIMA coefficients setting. One may argue that our PACF does not cut off when lag is 1, or ACF does not cut off when lag is 1, because there are still some higher lags where ACF or PACF exceeds the limit. Meanwhile, whether the seasonal should be set as one year which is 52 weeks or 2 years which is 104 weeks may need some further study. To summarise, human society's growth and fast industrialization have resulted in a significant increase in carbon emissions since the industrial revolution. Simultaneously, in 2021, human technology will transition the energy system away from fossil fuels and toward renewable energy sources such as solar energy, therefore reducing CO2 emissions.

Bibliography

C. D. Keeling, S. C. Piper, R. B. Bacastow, M. Wahlen, T. P. Whorf, M. Heimann, and H. A. Meijer, Exchanges of atmospheric CO2 and 13CO2 with the terrestrial biosphere and oceans from 1978 to 2000. I. Global aspects, SIO Reference Series, No. 01-06, Scripps Institution of Oceanography, San Diego, 88 pages, 2001.

Mauna Loa. (n.d.). Mauna Loa Carbon Dioxide Forecast for 2022. Met Office. Retrieved April 17, 2022, from https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/long-range/forecasts/co2-forecast

Sampling station records. MLO Station Data | Scripps CO2 Program. (n.d.). Retrieved April 17, 2022, from https://scrippsco2.ucsd.edu/data/atmospheric_co2/mlo.html