

# W271 Lab 2

Investigating the Keeling Curve and forecasting CO2 levels in Earth's atmosphere

Denny Lehman, Mingxi Liu, Aruna Bisht, Deepika Maddali

## Abstract

The Keeling curve is a graph that represents the concentration of carbon dioxide (CO<sub>2</sub>) in Earth's atmosphere since 1958. In this paper, the Keeling curve was analyzed from two perspectives, one from a researcher in 1998 and one from today (2023). From the perspective of 1998, EDA was performed from 1959 to 1997 and a linear time trend model was fit. After analysis of the assumptions, a cubic polynomial model was selected. Using the Box-Jenkins method, a SARIMA model was constructed and forecast into 2100. To contrast those models, we present the same analysis from the perspective of 2023. A modern data pipeline was constructed for CO<sub>2</sub> data and the linear model and SARIMA model were compared to the actual CO<sub>2</sub> levels. Finally, new models were fit on the 2023 a forecasted into the future with the goal of reviewing model predictions in a future analysis.

## Contents

<b>1</b>	<b>Report from the Point of View of 1997</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.2	Data . . . . .	2
1.3	Linear model . . . . .	3
1.4	ARIMA times series model . . . . .	5
1.5	Atmospheric CO <sub>2</sub> growth Forecast . . . . .	7
<b>2</b>	<b>Report from the Point of View of the Present</b>	<b>8</b>
2.1	Introduction . . . . .	8
2.2	Data . . . . .	8
2.3	Compare linear model forecasts against realized CO <sub>2</sub> . . . . .	9
2.4	Compare ARIMA models forecasts against realized CO <sub>2</sub> . . . . .	9
2.5	Evaluate the performance of 1997 linear and ARIMA models . . . . .	10
2.6	Train our best model on present data . . . . .	10
2.7	How bad could it get? . . . . .	12
<b>3</b>	<b>Appendix</b>	<b>13</b>
3.1	Model Details of the cubic time-trend model in Section 1 . . . . .	13
3.2	Modern data pipeline code . . . . .	14
3.3	Plots mentioned in the text . . . . .	15

# 1 Report from the Point of View of 1997

## 1.1 Introduction

We all know the debate about global warming and its connection to human activities. But to study this topic in a scientific way, we need reliable data. The Keeling Curve is a milestone in this aspect. It shows the ongoing increase in atmospheric carbon dioxide (CO<sub>2</sub>) concentrations over time. It is named after Charles David Keeling, the scientist who initiated and maintained the measurements. Keeling began monitoring atmospheric CO<sub>2</sub> levels in 1958 at the Mauna Loa Observatory in Hawaii. He chose this location because it is remote and far from major sources of pollution, providing an ideal site to measure baseline CO<sub>2</sub> concentrations. The Keeling Curve graphically represents the seasonal variations in atmospheric CO<sub>2</sub> concentrations, as well as the long-term increasing trend. Keeling believes the seasonal pattern is a result of the Earth's vegetation absorbing CO<sub>2</sub> during the growing season and releasing it during the dormant period, while the trend is primarily driven by human activities, particularly the burning of fossil fuels such as coal, oil, and natural gas, which release large amounts of CO<sub>2</sub> into the atmosphere. The Keeling Curve is an important tool for scientists, policymakers, and the general public to understand the impact of human activities on the Earth's climate. It serves as a stark reminder of the need to reduce greenhouse gas emissions and address the causes and consequences of climate change.

Our research is based on the data from the Keeling Curve above. We first build a model based on data from 1959 to 1997 and make long-term predictions to the present. Then we combine the actual data with our prediction and discuss the implication of this comparison.

## 1.2 Data

The data measures the monthly average atmospheric CO<sub>2</sub> concentration from 1959 to 1997, expressed in parts per million (ppm). It was initially collected by an infrared gas analyzer installed at Mauna Loa in Hawaii, which was one of the four analyzers installed by Keeling to evaluate whether there was a persistent increase in CO<sub>2</sub> concentration.

Fig.1 shows a clear long-term upward trend, which is confirmed by Fig.2 where the growth rate for each year is above zero. Fig.2 also suggests the average growth rate after 1970 is higher than that before 1970, although there's no evidence of accelerating growth. The ACF plots in Fig.3 and Fig.4 suggest the original data is non-stationary but its first difference is stationary. More formally, the KPSS tests below confirm the observations above.

Table 1: KPSS test of original and 1st difference

	kpss_stat	kpss_pvalue
original	7.8173	0.01
1st_difference	0.0124	0.10

Another feature of the data is its robust seasonal pattern, with the peak in May and the bottom in October almost every year (see Fig.5). This seasonality can also be seen in Fig.4. Keeling believes it was the result of plant photosynthesis absorbing CO<sub>2</sub> from the atmosphere.

Fig.4 is the histogram of the remaining or irregular components after removing the trend and the

seasonal components from the data with STL<sup>1</sup>. It looks like a normal distribution without obvious outliers.

Fig.1 Atmospheric CO2 concentration monthly average, parts per million (ppm)

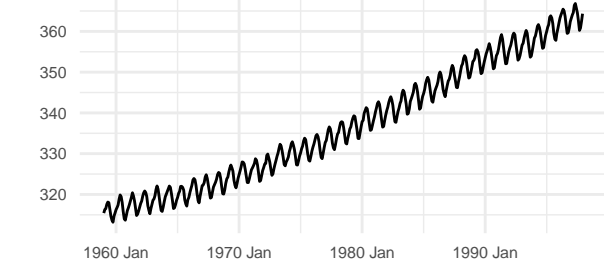


Fig.2 Annual growth rate of concentration, %

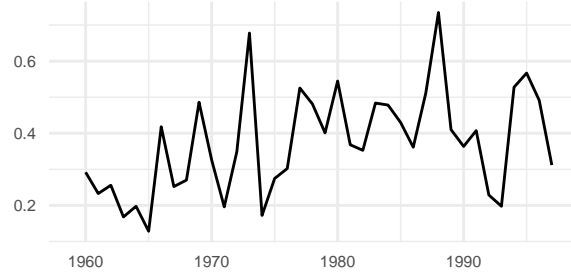


Fig.3 ACF of CO2 concentration

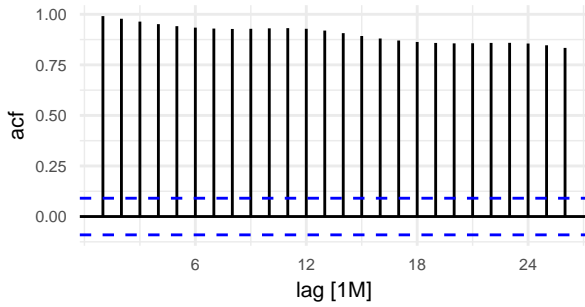


Fig.4 ACF of differenced CO2 concentration

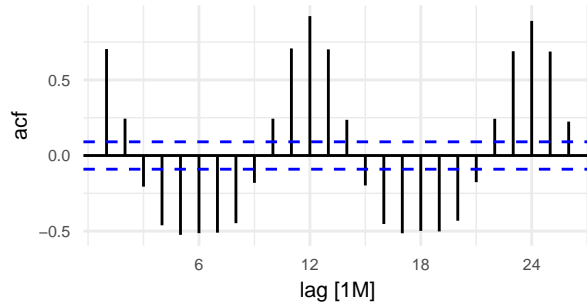


Fig.5 Seasonal plot of CO2 concentration

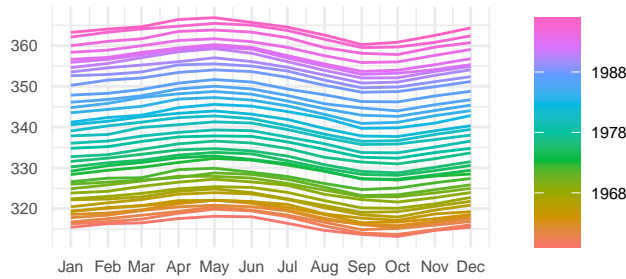
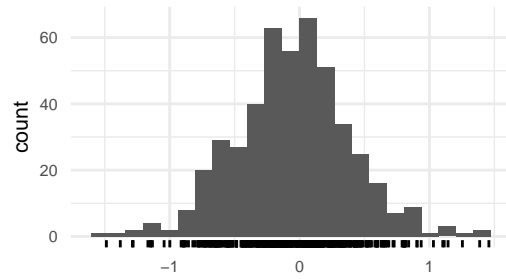


Fig.6 Histogram of irregular component by STL



### 1.3 Linear model

Before building the model, we need to consider whether the data needs a log transformation. Normally, a log transformation is required when the data shows exponential growth or the variance expands or shrinks over time. From Fig.1 and Fig.2 we can see the slope of the CO2 growth rate is stable, which suggests the growth is more linear than exponential. Also, Fig.5 shows the difference between the annual high and the annual low remained about the same over the years, suggesting the variance is nearly constant. Therefore, the log transformation is not necessary. We can first fit the original data with a linear time trend model as:

<sup>1</sup>Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1), 3–33.

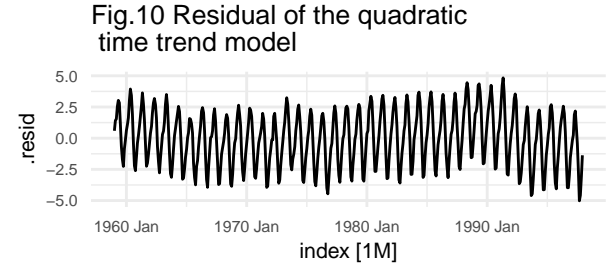
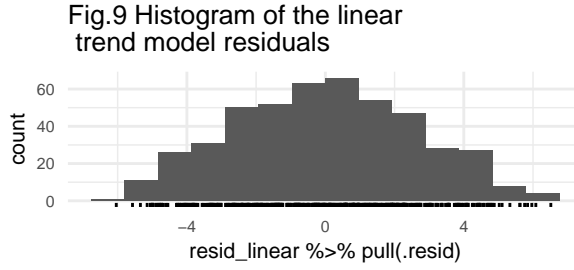
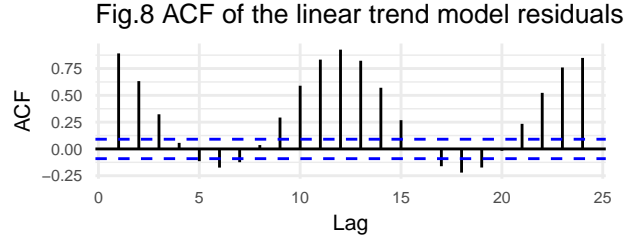
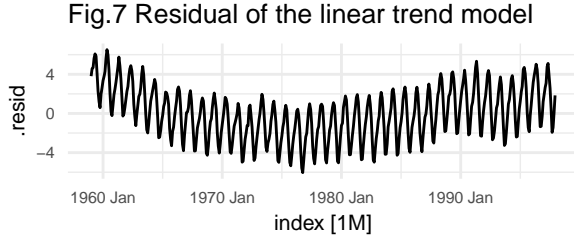
$$\text{CO}_2 = \beta_0 + \beta_1 t + \epsilon_t, \quad (1)$$

which gives the parameters as:

$$\text{CO}_2 = 311.5 + 0.11t + \epsilon_t \quad (2)$$

This linear trend model implies that the  $\text{CO}_2$  concentration increased by 0.11 ppm/month on average from 1959 to 1997. However, the residual plots in Fig.7 to Fig.10 suggest this simple linear trend model is not adequate in the following two aspects.

First, the mean of the residual forms a “U” shape over time, suggesting a quadratic or higher-order polynomial time trend model may be more appropriate. For instance, the residual from a quadratic time trend model shows a more constant mean over time, as shown in Fig.10.



In addition, the ACF plot in Fig.6 indicates strong seasonal patterns exist in the residuals, suggesting we should consider seasonal factors in the model. One solution is to include 11 dummy variables in the model to indicate the 12 months.

Based on the two points above, we compare the 2 candidates: a quadratic time trend model and a cubic one, shown below:

$$\text{Quadratic time trend: } \text{CO}_2 = \alpha + \beta_0 t + \beta_1 t^2 + \sum_{i=1}^{11} \gamma_i \text{Month}_{it} + \epsilon_t \quad (3)$$

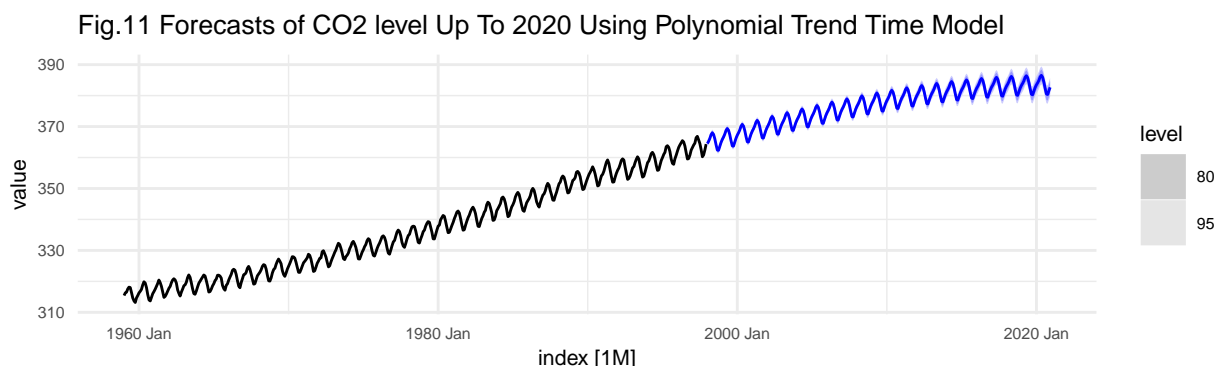
$$\text{Cubic time trend: } \text{CO}_2 = \alpha + \beta_0 t + \beta_1 t^2 + \beta_2 t^3 + \sum_{i=1}^{11} \gamma_i \text{Month}_{it} + \epsilon_t \quad (4)$$

We use the data before 1991 as the training set and the rest as the validation set (similar to an 80-20 split). Our final choice of the model depends on the combination of 2 guidelines: 1) the information criterion (AICc, BIC) from the model fitting process and 2) the root mean square error (RMSE) of predictions on the validation set, which are listed in Table.1. Both information criterion (AICc,

BIC) and RMSE favor the cubic model. Therefore, the cubic time trend model becomes our final choice. Its details are in the Appendix. We plot the forecast of this model until 2020 in Fig.9. One thing to note is that because the coefficient of the cubic term is negative, the predicted values will eventually begin to decrease when predicting the far future. In fact, we can see from Fig.11 that the predicted values have almost topped. This may be inappropriate extrapolation behavior. In that case, we should confine our predicting interval to the near term.

Table 2: Information Criterion of model fitting and RMSE of validation

.model	AIC	AICc	BIC	RMSE
cubic	-659.6147	-658.1324	-596.4044	2.112194
quadratic	-639.1525	-637.8481	-579.8928	2.796572



## 1.4 ARIMA times series model

We will use the Box Jenkins process to find the best ARIMA model via the following steps:

- Determine the appropriate model from EDA
- Find the best parameters
- Examine the residuals using diagnostic plots and statistical tests

The EDA revealed that the time series of CO2 had both autoregressive and seasonal components. Considering the ACF plot's low slow decay of autocorrelation, we expect differencing to be a key part of any time series model. In addition, we predict that the model will require seasonal components to model the 12 month cycle of seasonal variations. Therefore, we expect a seasonal ARIMA model (SARIMA) with differencing and seasonality terms to be best.

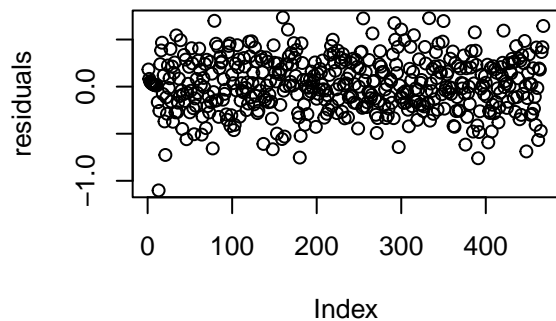
In this section, we fit the best SARIMA model and analyze the results. We choose BIC as our information criteria for model selection. Simplicity is a desirable property in data science models to help explain the relationship between variables. We choose BIC as our information criteria because it penalizes complex models more than AIC or AICc and therefore selects more simple models with fewer parameters as the best ones. Lower BIC scores are better.

```
## Series: value
## Model: ARIMA(0,1,1)(1,1,2)[12]
##
```

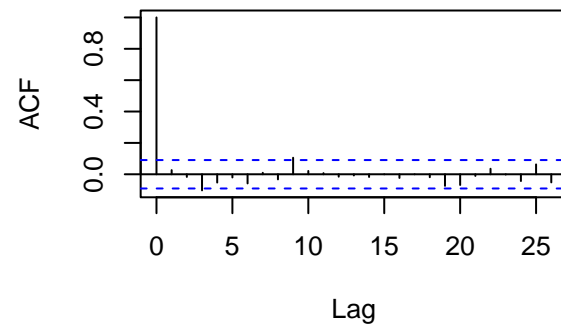
```
## Coefficients:
##          ma1      sar1      sma1      sma2
##      -0.3482 -0.4986 -0.3155 -0.4641
## s.e.   0.0499  0.5282  0.5165  0.4367
##
## sigma^2 estimated as 0.08603:  log likelihood=-85.59
## AIC=181.18   AICc=181.32   BIC=201.78
```

After searching over seasonal and non-seasonal P, D, and Q variables, the best model was an ARIMA(0,1,1)(1,1,2)[12] model with BIC score of 201.78. Next, we evaluate the model via diagnostic plots and statistical tests, concluding the Box Jenkins process.

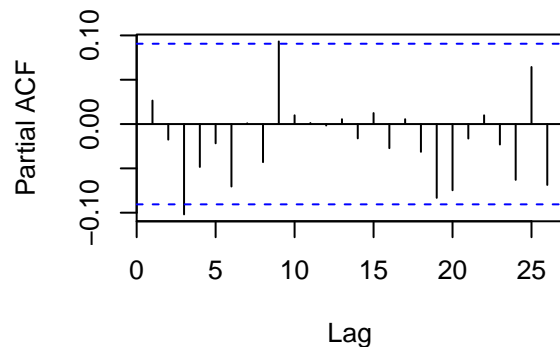
**Fig.12 Residuals of the SARIMA mode**



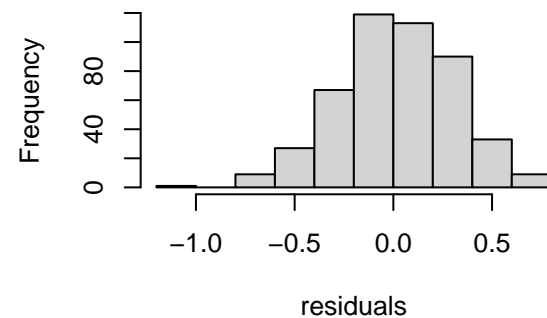
**Fig.13 ACF plot of residuals**



**Fig.14 PACF plot of residuals**



**Fig.15 histogram of residuals**



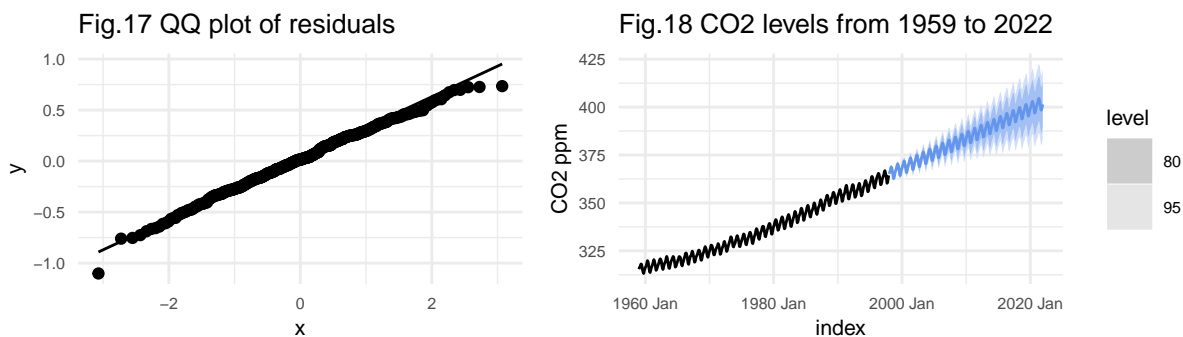
The residual plots (Fig 12-15) show that the SARIMA model was effective, with the residuals looking like stationary white noise (Fig 12). The time series of residuals appears to have a mean of 0 with constant variance, the ACF plot (Fig 13) shows no autocorrelation beyond the initial lag value. The PACF plot (Fig 14) appears to have a significant peak around the 3rd lag term, but this may be due to randomness, as it is barely passing the dashed blue line. The histogram (Fig 15) looks normally distributed at 0 with outliers creating a left tail.

We test the residuals for stationarity with the Augmented Dickey Fuller test (ADF). The ADF test has the null hypothesis that the data is non-stationary. With a p-value of 0.01, we reject the null hypothesis because there is enough evidence to say that the residuals are stationary.

The Box-Ljung test has the null hypothesis that the data presented is independently distributed. When presented with the residuals of the ARIMA model, the test had p-values of 0.566 and 0.144 for lag = 1 and lag = 10 respectively. For both of those lags, we fail to reject the null hypothesis and conclude that the data is independently distributed.

Finally, we visually inspect the histogram of the residuals (Fig.16) and the QQ plot (Fig.17) to see if the residuals appear normally distributed. The histogram has the Gaussian bell shaped curve with a few outliers. The QQ plot shows that the data matches up with the normal distribution's quantiles. With these plots, we can confidently say that the residuals are visually normally distributed.

To conclude, both diagnostic plots and statistical tests show that the residuals are stationary with mean 0, constant variance, and no autoregression or seasonality. We forecast our model to the year 2022 (Fig 18).



## 1.5 Atmospheric CO2 growth Forecast

We use our model to make predictions on future levels of CO2, specifically 420 and 500 ppm. We will investigate the earliest, best guess, and latest occurrence of these values. The earliest guess will be based on the first time the upper 95% confidence interval (CI) reaches the specified level and the latest guess will be the last time the value is within the lower 95% CI. The best guess will be the point estimate (mean) of the forecast.

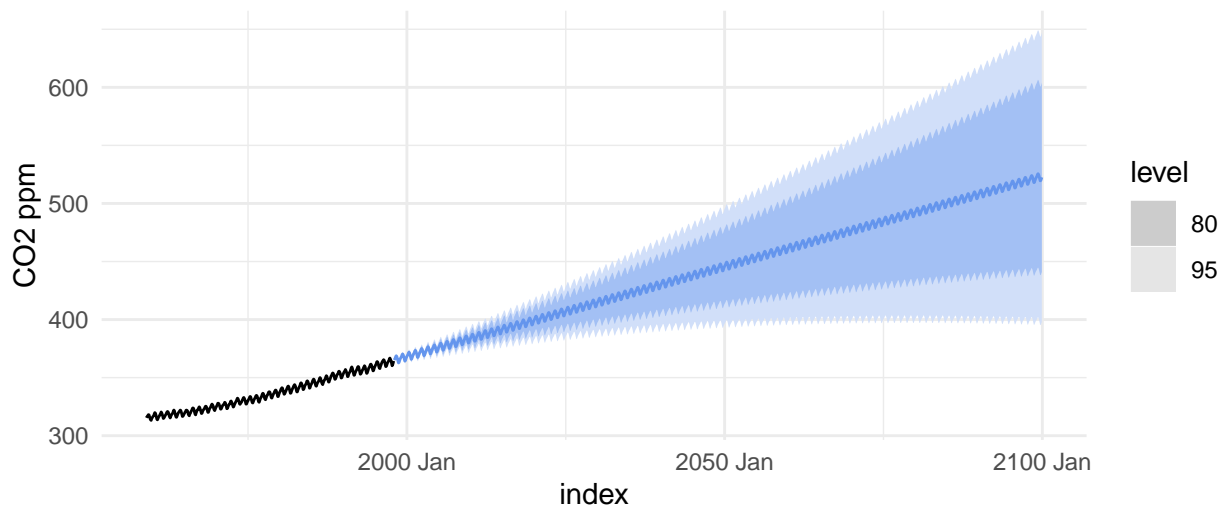
Table 3: Predicted occurrences of key CO2 levels

CO2.ppm.level	earliest.occurrence	point.estimate	final.occurrence
420 ppm	2022 Apr	2031 May	never
500 ppm	2054 Apr	2083 Apr	never

Based on our model, the first time we could potentially see CO2 at 420 ppm is 2022-04-01 because that is when the upper 95% confidence interval (CI) of our model first reaches 420 ppm. The model's lower 95% CI hovers around 420, so there is no predicted final time. Knowing what we know today in 2023, 419 ppm was reached on May 2021, which was *before* our model's earliest guess. CO2 levels have risen faster than our model anticipated. This is a precursor to the analysis provided later in this paper. The first time our model predicts the earth to reach 500 ppm CO2 on 2054-04-01, which is when the 95% CI reaches 500 ppm. The model's lower 95% CI never reaches 500, so there is no predicted final time.

Below is the prediction of our model to the year 2100. Confidence intervals are shown fanning outward. The error of the predictions compounds overtime which expands the confidence intervals into a funnel shape. The farther out in time from the recorded data points, the less accurate the prediction.

**Fig.19 CO2 levels from 1959 to 2100**



## 2 Report from the Point of View of the Present

### 2.1 Introduction

In our original 1997 paper, we made several predictions on the expected level atmospheric CO<sub>2</sub>. Currently, we will evaluate the accuracy of those predictions using time series analysis and extrapolate from present data to make predictions about the future.

### 2.2 Data

A modern data pipeline was constructed to load both weekly and monthly CO<sub>2</sub> data from January 1959 to June 2023. This will allow us to compare our forecasts in the previous section to the actual CO<sub>2</sub> levels. The code for the pipeline can be found in the appendix.



Fig.20 Atmospheric CO2 concentration monthly average, parts per million (ppm)

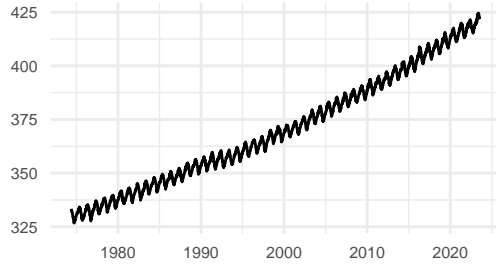


Fig.22 Seasonal plot of CO2 concentration

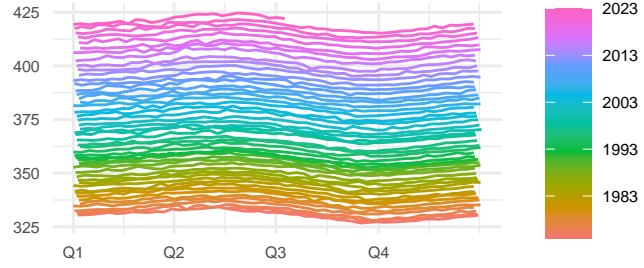


Fig.23 ACF of CO2 concentration

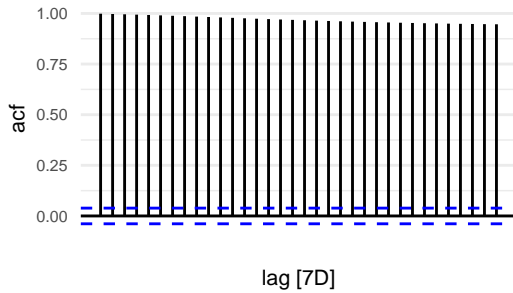
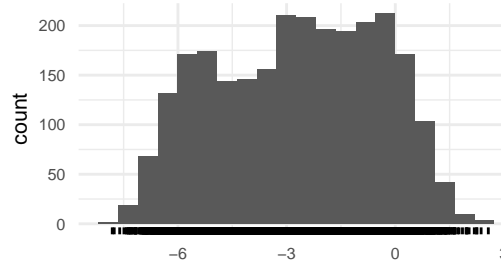


Fig.24 Histogram of irregular component by STL



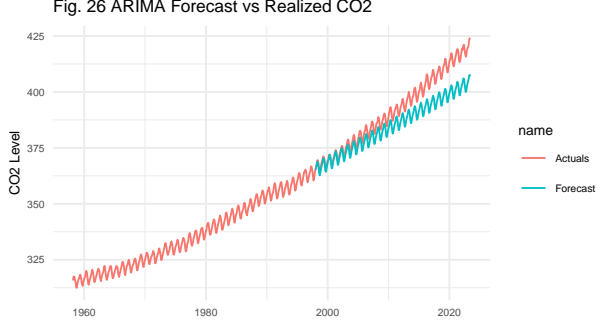
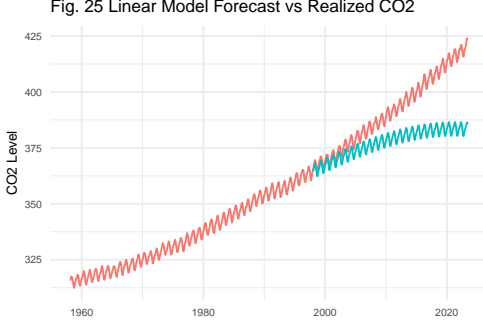
The time series plot in Fig.20 shows, before 1990, the CO2 levels increased steadily at certain rate. During the 1990's, the trend seemed to slightly flatten. But starting from the late 1990's, the trend became steeper again, suggesting the CO2 levels grew at a steady but higher speed (see Appendix for additional plots of this phenomenon). The ACF plot in Fig.23 confirms the data is not stationary. The seasonal pattern remained the same after 1997, as shown in Fig.22. However, the histogram in Fig.24 shows that the irregular component of the data is less likely to be normal than the data before 1997.

## 2.3 Compare linear model forecasts against realized CO2

Next we compare the cubic time trend linear model (the “linear model”) to the realized CO2 levels. We forecast the linear model onto the actual CO2 levels (Fig.25) and discover that the forecast may not have capture the realized trend of CO2 levels. The forecast appears to predict a stabilization in the CO2 levels around year 2020, whereas the actual CO2 level trend increased.

## 2.4 Compare ARIMA models forecasts against realized CO2

In Fig.26, we compare the ARIMA model from 1998 data to the realized CO2 levels. The ARIMA forecast is much closer to the realized CO2 levels than the linear model forecast. The only difference observed, is that the ARIMA model appears to have forecasted a linear trend, while the realized CO2 levels followed an almost exponential growth.



## 2.5 Evaluate the performance of 1997 linear and ARIMA models

Table 4: Accuracy of predictions from 1997

.model	ME	RMSE	MAE	MPE	MAPE	ACF1
Best Polynomial	14.32	17.91	14.32	3.55	3.55	0.99
Best ARIMA	6.78	8.32	6.78	1.68	1.68	0.99

We evaluate the accuracy for the best polynomial and ARIMA models built on the data till 1997. The RMSE of prediction from the best polynomial model reaches 17.91, and that of the best ARIMA model is 8.32.

## 2.6 Train our best model on present data

One interesting feature of Keeling and colleagues' research is that they were able to evaluate, and re-evaluate the data as new measurements were released. Now, we want to do the same. First, we build models on both the original non-seasonal-adjusted data (NSA) and the seasonal-adjusted data (SA) using the STL package. We split both NSA and SA series into training and test sets, using the last two years of observations as the test sets.

For the NSA data, since we know it has a trend and an annual seasonality, we can first difference it with period=52 (for weekly data). After that, both ACF and KPSS test suggest the seasonally differenced data is still non-stationary, therefore another difference is needed. The cut off at lag=2 in ACF and decay in PACF (see Fig.27) suggest both the non-seasonal and seasonal part are likely to be MA(2) process, therefore an  $ARIMA(0, 1, 2)(0, 1, 2)_{52}$  is our manually selected candidate. We compare it with the auto-selected model based on BIC, which is an  $ARIMA(0, 1, 1)(2, 1, 0)_{52}$ . Our manual model outperforms the auto-selected model on training set but slightly underperforms on test set, as shown in Table.5. The RMSE for both model are much higher on the test set than on the training set, which may suggest overfitting. We finally choose the auto-selected  $ARIMA(0, 1, 1)(2, 1, 0)_{52}$  since it has lower RMSE on test set.

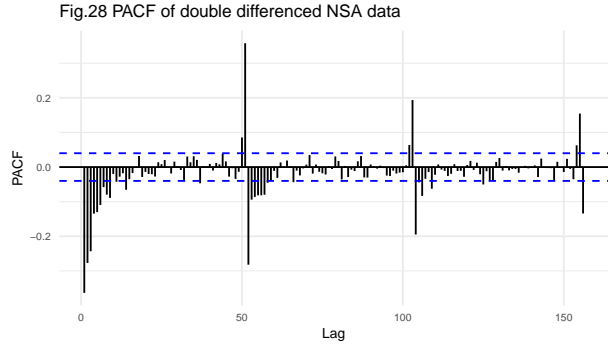
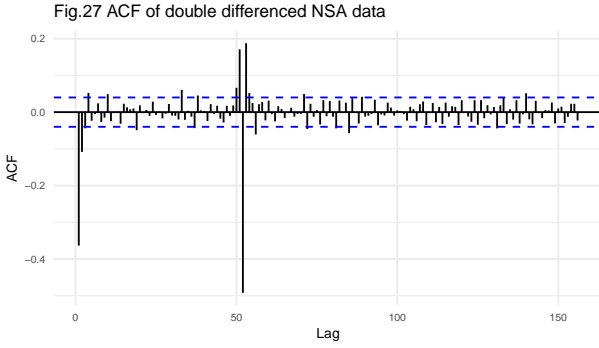
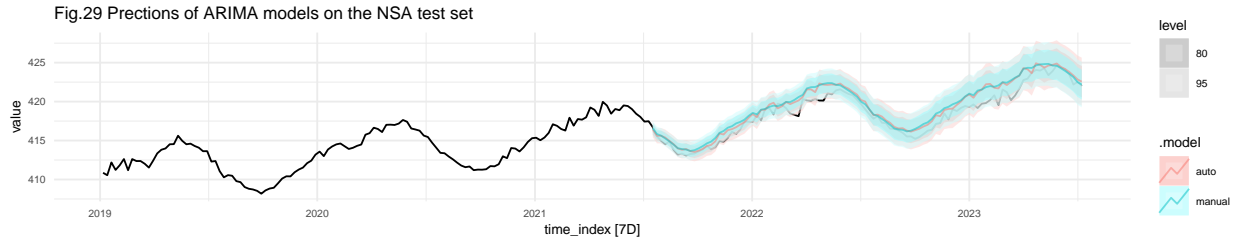


Table 5: Comparison of manual and auto ARIMA models for NSA data

Model	Train AICc	Train BIC	Train RMSE	Test RMSE
manual	2921.32	2950.23	0.44	1.12
auto	3276.88	3300.01	0.47	1.03



For the SA data, we will build two types of models: an ARIMA model and a polynomial time-trend model. For the ARIMA model, seasonal difference is not needed, but KPSS test suggest the data is not stationary and requires a first difference. The ACF of the differenced data cuts off at lag = 2 and PACF seems to decay, suggesting an ARIMA(0,1,2) model. Again, we compare this manually selected model with the auto-selected model, which is an ARIMA(1,1,1). Table.6 suggests our manual model slightly outperforms the auto-selected model on both training and test set. The RMSEs are also higher on the test set than on the training set, although the difference is smaller than the models for NSA data. Our final choice for the SA data would be our manual model ARIMA(0,1,2).

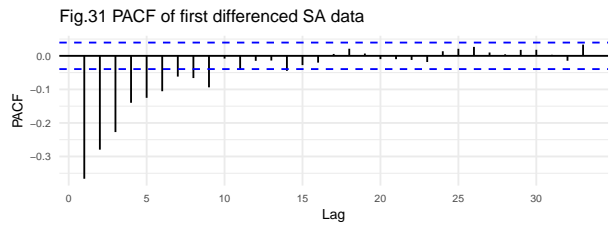
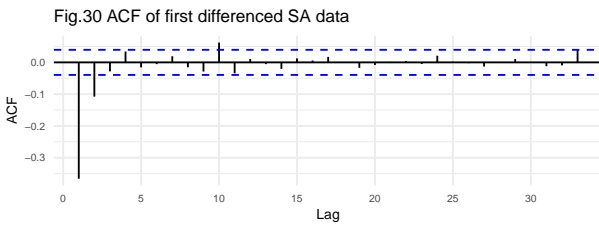


Table 6: Comparison of manual and auto ARIMA models for SA data

Model	Train AICc	Train BIC	Train RMSE	Test RMSE
manual	1988.55	2011.76	0.36	0.59
auto	1991.14	2014.36	0.36	0.60

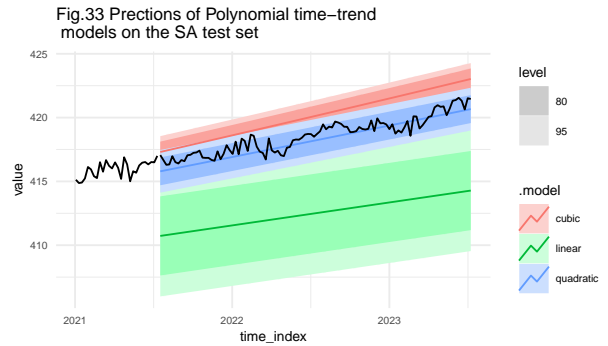
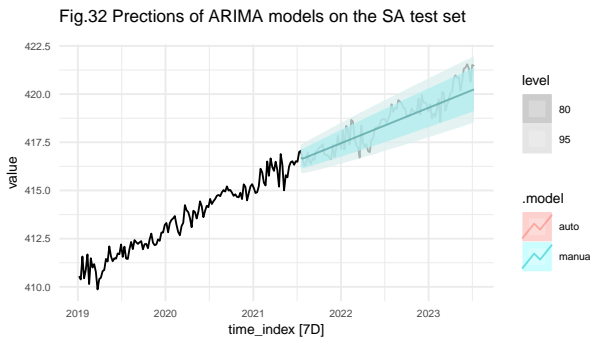
For the polynomial time-trend model, we compare models with degree from 1 to 3. The comparison on both training and test sets are in Table.7. The linear model performs the worst on both sets. The cubic model is the best on the training set but underperforms itself and the quadratic one on the test set, suggesting an overfitting. The quadratic model has the lowest RMSE on test set and its test RMSE is even lower than the training RMSE, suggesting a good modelling. Therefore, the quadratic model is our final choice for the polynomial time-trend model. However, the quadratic model underperforms the best ARIMA model on both training and test sets, as shown in Table.8.

Table 7: Comparison of polynomial models for SA data

Model	Train AICc	Train BIC	Train RMSE	Test RMSE
linear	4346.91	4364.33	2.42	6.11
quadratic	-786.28	-763.06	0.85	0.65
cubic	-2263.13	-2234.11	0.63	1.68

Table 8: Best ARIMA model vs. Best polynomial time-trend model

Model	Train AICc	Train BIC	Train RMSE	Test RMSE
Best ARIMA: (0,1,2)	1988.55	2011.76	0.36	0.59
Best Poly: Quadratic	-786.28	-763.06	0.85	0.65



## 2.7 How bad could it get?

Now in July 2023, we've already observed that the CO2 concentration exceeded 420 for the first time during the week of 2022-03-27, so it's no more a prediction. Based on the best SARIMA model

for the NSA data, we predict atmospheric CO2 levels in the year 2122. This also provides us the time when atmospheric CO2 is expected to be at 420 ppm and 500 ppm levels for the first and final times, as listed in Table.9. We need to point out that although the table suggest the CO2 level may never pass 420, this is because we use the lower bound of the 95% confidence interval as the prediction. As we predict further in to the future, the confidence interval expands rapidly with its lower bound decreasing back to below 420. But according to the actual data, the CO2 level hasn't dropped below 420 since 2023-02-26.

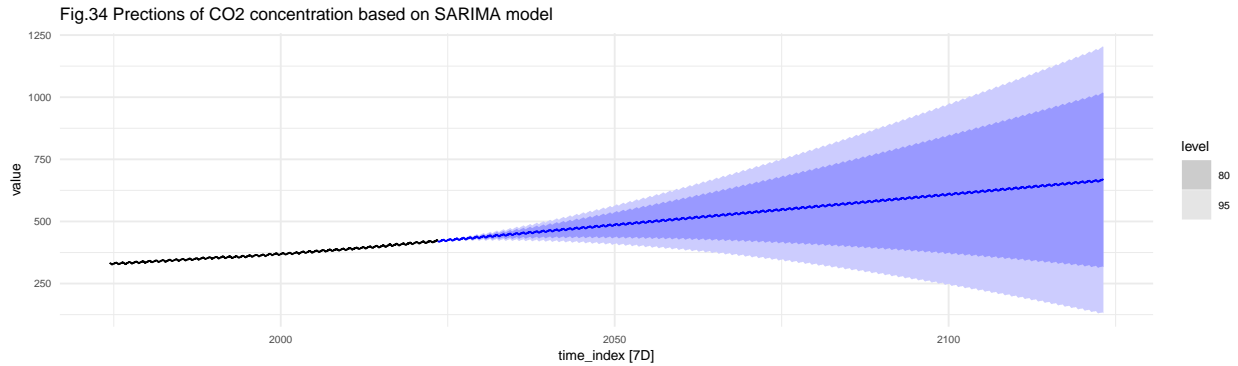


Table 9: Predicted occurances of key CO2 levels with present data

CO2.ppm.level	earliest.occurance	point_estimate	final.occurance
420 ppm	2022-03-27	2022-03-27	never
500 ppm	2040-05-06	2054-04-19	never

### 3 Appendix

#### 3.1 Model Details of the cubic time-trend model in Section 1

```
## Series: value
## Model: TSLM
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.172427 -0.262499 -0.003472  0.264990  1.476396
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   3.157e+02  1.094e-01 2884.373 < 2e-16 ***
## trend()       4.471e-02  1.922e-03  23.266 < 2e-16 ***
## I(trend()^2)  1.856e-04  1.159e-05  16.016 < 2e-16 ***
## I(trend()^3) -9.331e-08  1.979e-08  -4.715 3.44e-06 ***
## season()year2  6.565e-01  1.036e-01   6.334 6.94e-10 ***
## season()year3  1.383e+00  1.036e-01  13.342 < 2e-16 ***
## season()year4  2.495e+00  1.036e-01  24.076 < 2e-16 ***
```

```
## season()year5    2.986e+00  1.036e-01   28.809 < 2e-16 ***
## season()year6    2.366e+00  1.036e-01   22.829 < 2e-16 ***
## season()year7     9.035e-01  1.037e-01    8.717 < 2e-16 ***
## season()year8    -1.123e+00  1.037e-01  -10.832 < 2e-16 ***
## season()year9    -2.908e+00  1.037e-01  -28.050 < 2e-16 ***
## season()year10   -3.139e+00  1.037e-01  -30.273 < 2e-16 ***
## season()year11   -1.970e+00  1.037e-01  -19.001 < 2e-16 ***
## season()year12   -8.910e-01  1.037e-01   -8.592 2.44e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4145 on 369 degrees of freedom
## Multiple R-squared:  0.9988, Adjusted R-squared:  0.9988
## F-statistic: 2.194e+04 on 14 and 369 DF, p-value: < 2.22e-16
```

### 3.2 Modern data pipeline code

```
co2_present_raw=read.csv(
  "https://gml.noaa.gov/webdata/ccgg/trends/co2/co2_weekly_mlo.csv",skip=51)
co2_present <- co2_present_raw %>%
  mutate(time_index=lubridate::make_date(year,month,day)) %>%
  dplyr::select(time_index,average) %>%
  as_tsibble(index = time_index) %>%
  mutate(average =replace(average,average<=-999,NA)) %>%
  mutate(average = na.approx(average))
```

```
co2_present_raw=read.csv(
  "https://gml.noaa.gov/webdata/ccgg/trends/co2/co2_mm_mlo.csv",skip=56)

co2_present_month <- co2_present_raw %>%
  mutate(time_index=lubridate::make_date(year,month)) %>%
  dplyr::select(time_index,average) %>%
  as_tsibble(index = time_index) %>%
  mutate(average =replace(average,average<=-999,NA)) %>%
  mutate(average = na.approx(average))
```

```
glimpse(co2_present)
```

```
## Rows: 2,565
## Columns: 2
## $ time_index <date> 1974-05-19, 1974-05-26, 1974-06-02, 1974-06-09, 1974-06-16~
## $ average      <dbl> 333.37, 332.95, 332.35, 332.20, 332.37, 331.73, 331.69, 331~
```

```
glimpse(co2_present_month)
```

```
## Rows: 784
## Columns: 2
## $ time_index <date> 1958-03-01, 1958-04-01, 1958-05-01, 1958-06-01, 1958-07-01~
## $ average      <dbl> 315.70, 317.45, 317.51, 317.24, 315.86, 314.93, 313.20, 312~
```

### 3.3 Plots mentioned in the text

