

Comparing Global CO_2 Emission Models from 1997 and Today

By CAROLYN DUNLAP
AYDA NAYEB NAZAR
QIAN QIAO
HECTOR RINCON*

In 1997, our team analyzed CO_2 concentration average monthly trends from 1959 to 1997. Our work culminated in a report that discussed multiple linear and ARIMA models of these concentration data and attempted to forecast CO_2 trends into the future, as far out as 2100. Given that we now have 25 more years of data, and at a highly weekly resolution, this report aims to revisit and evaluate our 1997 models. Using weekly average CO_2 concentration levels taking from the Mauna Loa observatory, we also analyze our data to fit new models and generate updated forecasts for potential CO_2 concentrations as far out as 100 years into the future.
Keywords: Replication, Modern Science

I. Introduction

In this report, we are continuing to evaluate atmospheric CO_2 trends (commonly referred to as the Keeling Curve), generating models and forecasts to help understand how we might expect atmospheric CO_2 levels to change in the future. Specifically, we are re-evaluating our best performing linear and ARIMA models from our 1997 report on CO_2 concentrations, using more recently generating weekly average CO_2 data up through the end of 2022. This report continues to use the CO_2 concentration data collected by the Mauna Loa observatory, with the key exception that we have weekly average CO_2 concentration data, compared to the monthly average concentration data used in 1997. Depending on the purposes of our analyses, we will either keep the weekly granularity or regenerate monthly averages as necessary. In addition to evaluating our 1997 models and forecasts, we will also use more recent data to generate new models and forecasts for CO_2 concentrations.

II. Measurement and Data

The data that we are using from this report was accessed via the United States' National Oceanic and Atmospheric Administration website, where they house continually updated weekly CO_2 concentration data collected from the Mauna

* : , : , : , : , : . The authors would like to thank their instructors from MIDS 271.

Loa observatory. This data was collected in a similar manner to the data in 1997 with a few key exceptions.

The first key exception is that in 2019, the observatory installed a new CO_2 analyzer that measures CO_2 concentration (defined as the number of CO_2 molecules out of a random million atmospheric molecules, or parts per million (ppm)) via cavity-ring down spectroscopy versus the previously used infrared absorption. We are presuming that this change in analysis technique does not dramatically alter the concentration of CO_2 . Another exception is that this dataset has a gap in December 1975: for the purposes of these analyses, we decided to fill all gaps using the previous CO_2 concentration gathered (e.g. we decided to use the CO_2 concentration taken on November 11, 1975 as the CO_2 values for all of the missing December 1975 datapoints).

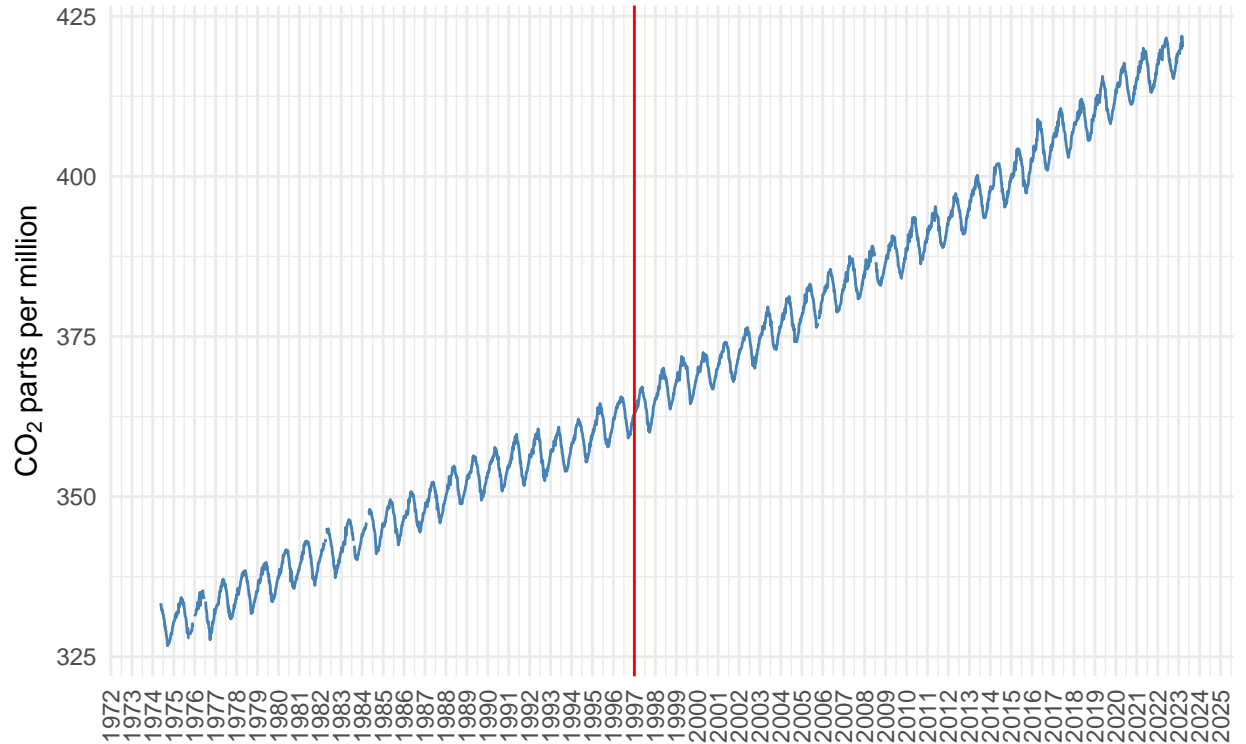
A final note about this data set is that in November 29, 2022 Mauna Loa erupted and thus all subsequent data points were collected at the Maunakea observatories, approximately 21 miles north of the Mauna Loa observatory. For ease and data consistency, we have decided to only use data through the end of 2022 in our models and forecasts (up to but not including the recent Maunakea observations).

A. Historical Trends in Atmospheric Carbon

To begin, we conducted the same analysis on this dataset as was done in 1997, beginning with a map of the time series data, shown below in Figure 1, with a vertical red line at 1997 to reflect what new data we have compared to the previous report.

Figure 1: Weekly Mean CO_2 at Mauna Loa 1974–2023

A continuation of the Keeling Curve (red line denotes 1997)



From this time series, we observe a steady continuation of the Keeling curve, with a continual increase in CO_2 concentration (ppm) as well as a notable seasonal component. The curve of the overall increasing CO_2 concentration, compared to data prior to 1997, does appear to be increasing at a slightly higher rate than what was previously observed.

To more carefully compare the distribution and autocorrelation aspects of this times series, we conducted further analysis, as shown below in Figure 2.

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

```
## Warning: Removed 18 rows containing non-finite values ('stat_bin()').
```

Figure 2a

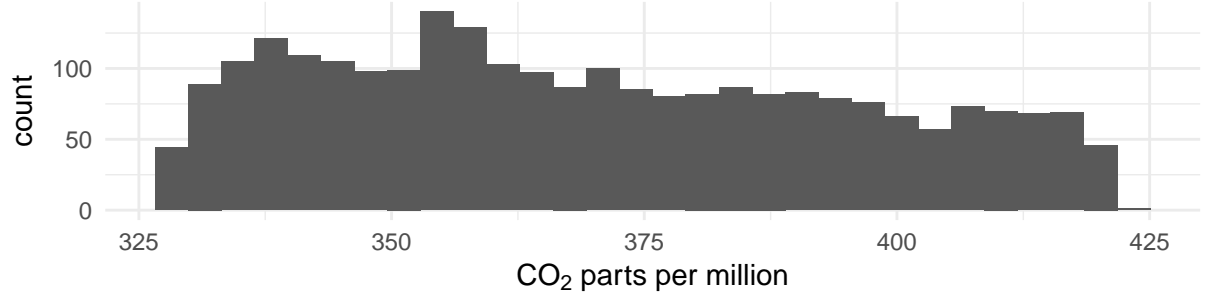
Histogram of Weekly Mean CO_2 

Figure 2b

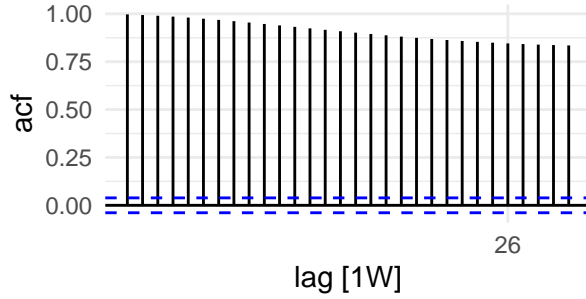
ACF of Weekly Mean CO_2 

Figure 2c

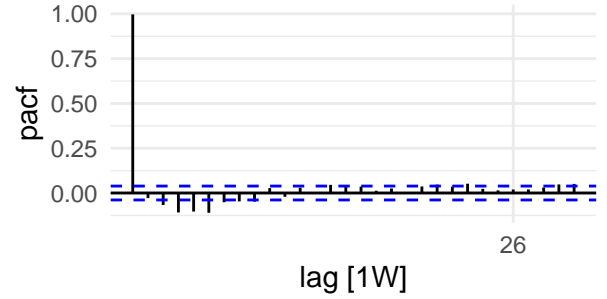
PACF of Weekly Mean CO_2 

Figure 2 reveals that the present dataset recapitulates the behavior observed in the 1997 dataset. Similar to what was seen in 1997, there is a strong autocorrelation in CO_2 concentration with very mild decay, maintaining its strong correlation even after a lag of 30 weeks (Fig 2b). The partial autocorrelation, PACF plot, reveals a first correlation of almost 1 and some cyclical correlations that may correspond to the seasonality of the data (Fig 2c). Taken together, this analysis suggests that there is a seasonal and potentially autoregressive component to the data.

III. Models and Forecasts

Through our exploratory data analysis, we found that our weekly data from 1974-2022 appears to behave very similarly to the data that was analyzed in 1997. Stemming from this analysis, we move to specifically evaluating the best linear and ARIMA model from the 1997 report against this newer dataset.

A. Comparing 1997 Linear Models

To begin, we compared the best linear model from the 1997 report. In that report, we assessed four different models and determined that the quadratic model with a seasonal adjustment component most accurately described the data up to 1997. To compare this model, we forecasted this model 25 years in the future to encompass 2023, and graphed it against our current dataset, shown below in Figure 3.

Figure 3: Quadratic model forecast vs realised values 1997–2022

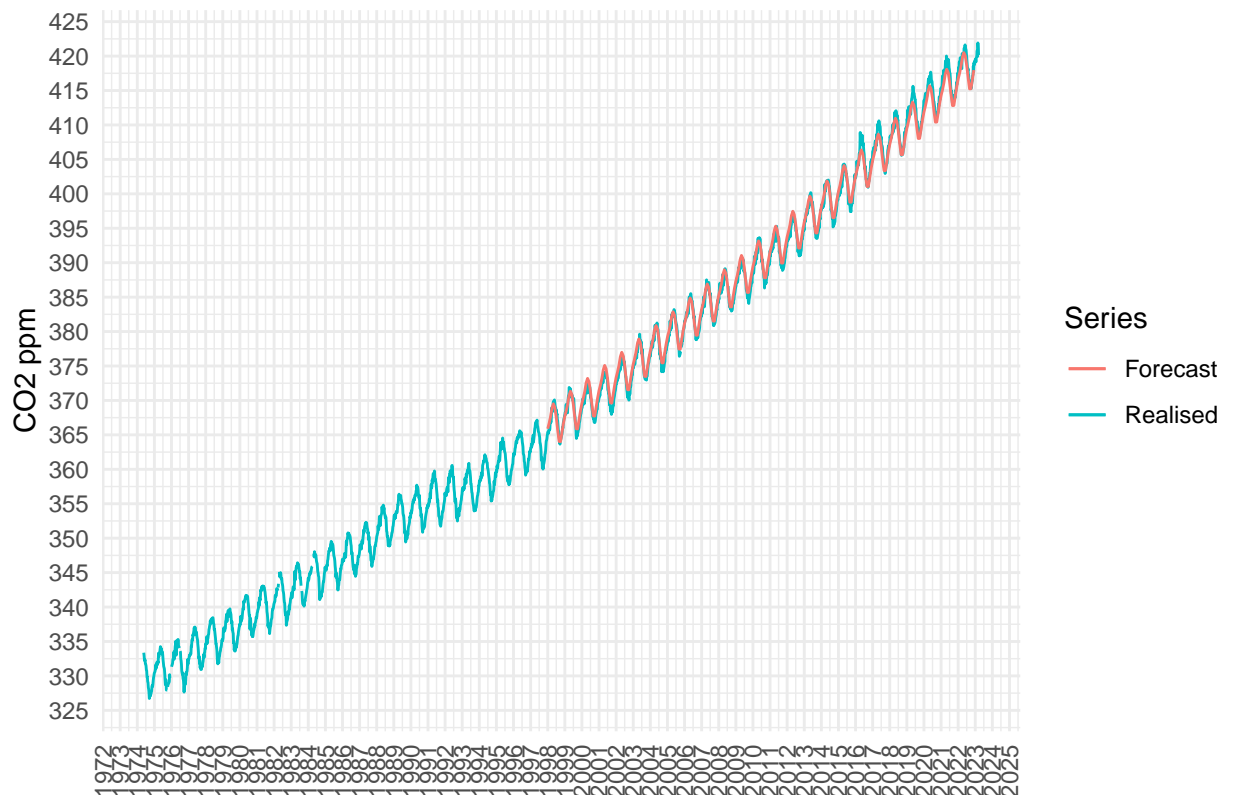
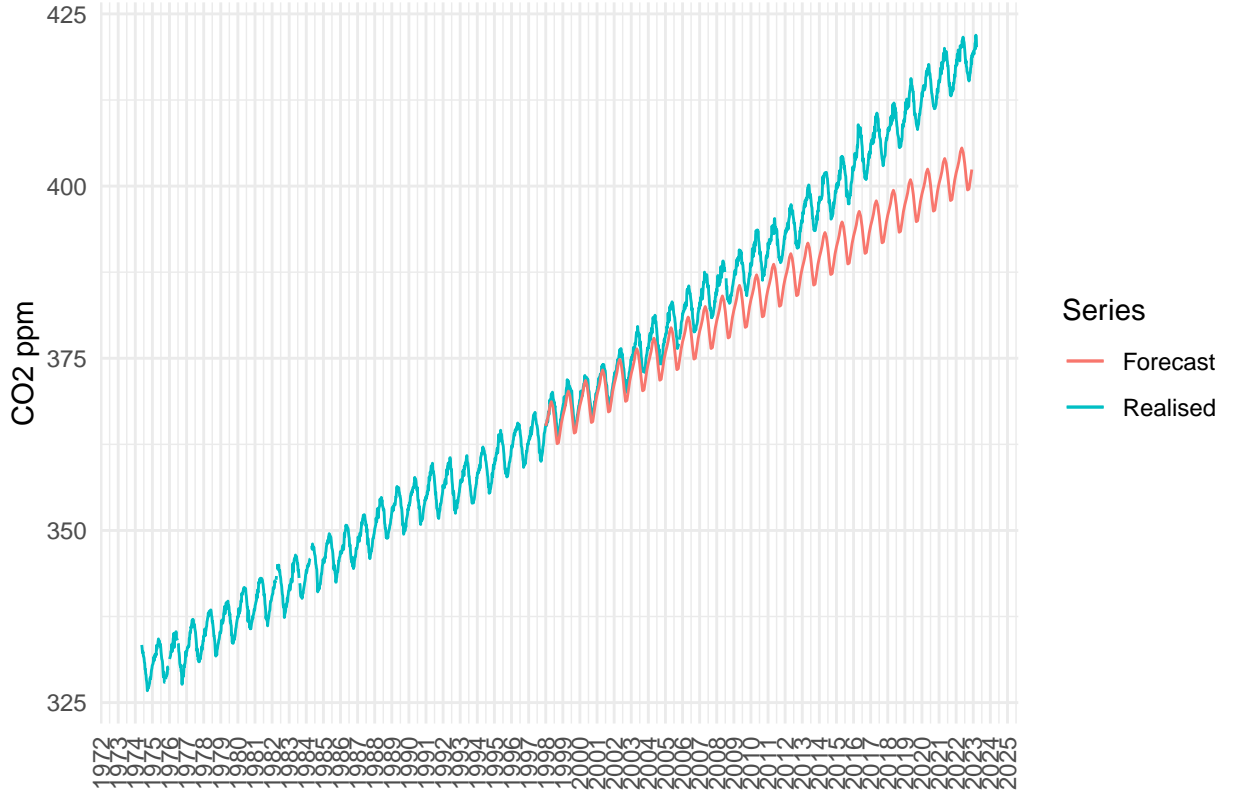


Figure 3 reveals that the quadratic with seasonality model forecast fairly accurately predicts the actual data from 1997-2022. Up until about 2016, the 1997 model does seem to routinely over-predict the trough CO_2 values per year and after 2016, the 1997 quadratic with seasonality model routinely under-predicts the peak CO_2 values per year. This suggests that while the quadratic model fits the data fairly closely, a model that combines several localized trends might better capture the data than a single overarching model.

B. Comparing 1997 ARIMA Models

In addition to the linear model, the 1997 report also fit an ARIMA model to the Keeling Curve CO_2 data. We therefore wanted to additionally evaluate how well that ARIMA model, ARIMA(0,1,3)(0,1,1)[12], recapitulated the realized CO_2 concentrations from 1997-2022 (Figure 4 below).

Figure 4: ARIMA model forecast vs realised values 1997–2022



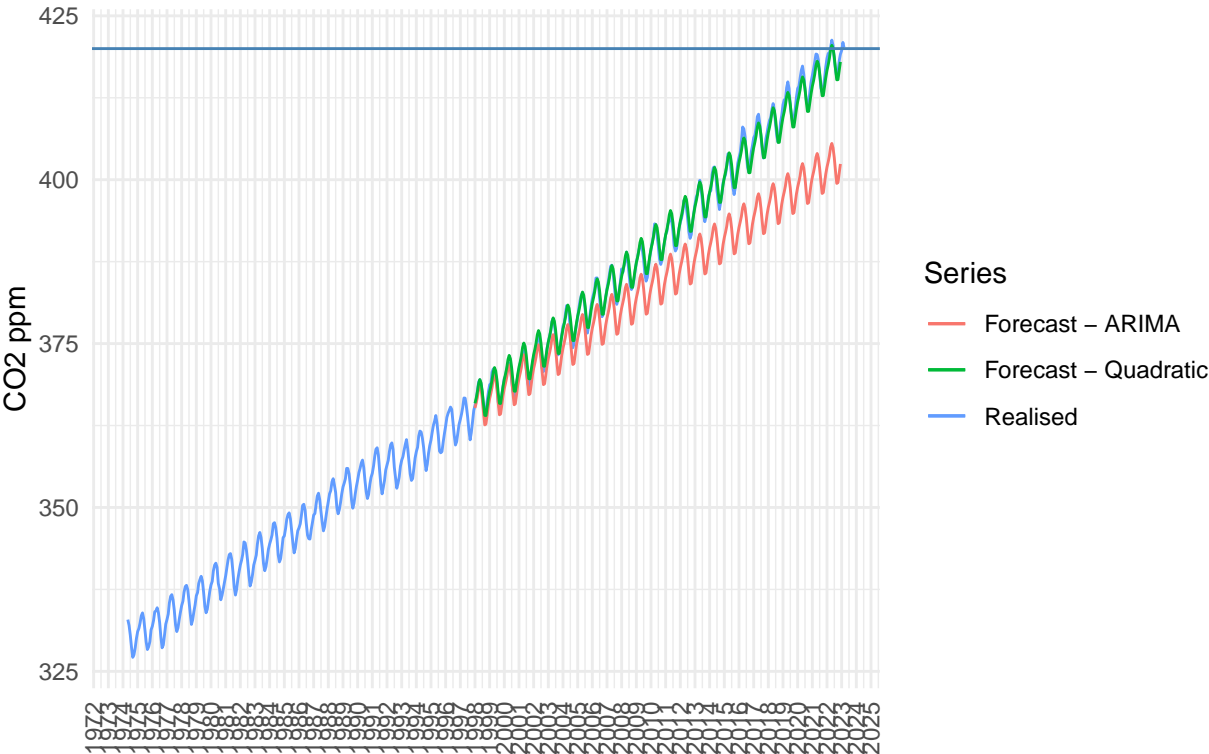
As evident from Figure 4, the 1997 ARIMA model poorly predicted the CO_2 concentrations from 1997-2022, especially compared to the quadratic with seasonality model also developed in 1997. From visual inspection, the 1997 ARIMA model predicted a linearly increasing Keeling Curve, rather than the quadratic process demonstrated in the previous model. Because the CO_2 concentration trends does appear to increase more each year post 1997, following a quadratic pattern more than a linear pattern, the ARIMA model consistently and more drastically underpredicted CO_2 concentrations at all points throughout the year post 1997.

C. Evaluation of 1997 Forecasts

To more formally compare the 1997 models, we evaluate the specific accuracy of several forecasts generated in the 1997 report to assess how closely they matched the current realized data. In 1997, we predicted the first and final times CO_2 concentration levels would reach 420 ppm and 500 ppm. We also evaluated predicted CO_2 concentrations in the year 2100. As of the date of writing this report (published March 2023), we have not yet crossed the 500 ppm threshold nor have data on CO_2 in the year 2100, however we have already reached 420 ppm.

In the 1997 report, 420 ppm predictions were generated using the ARIMA(0,1,3)(0,1,1)[12] model only, because at the time that model was assessed to best recapitulate the data. In the 1997 report, CO_2 concentrations was predicted to first cross the 420 ppm threshold in April 2032, however, the current data reveals that we actually first crossed the 420 ppm threshold in May 2021 according to the weekly data and April 2022 when the current data is aggregated monthly, a full 10-11 years prior to the 1997 ARIMA model estimate. As discussed above, the 1997 ARIMA model severely underestimated the growth rate of the Keeling Curve past 1997. In contrast and as can be shown below in Figure 5 (the blue horizontal line corresponds to the 420 ppm threshold), the quadratic model with seasonality developed in the 1997 report predicted a May 2022 estimate for monthly mean CO_2 concentration, which compared to the ARIMA model is a much more accurate forecast.

Figure 5
ARIMA model forecast, quadratic model forecast, and realised values 1997–2022



To more rigorously compare the accuracy of the two models, we assessed numerous accuracy measures for each 1997 model against the aggregated monthly average data used in the this report, comparing forecasted accuracy from 1997-2022. The results are summarized in Table 1 (below) and reveal that across each metric (mean error (ME), root mean squared error (RMSE), mean absolute error (MAE), mean percentage error (MPE), mean absolute percentage error (MAPE), mean absolute scaled error (MASE), root mean squared scaled error (RMSSE), and autocorrelation errors at lag 1 (ACF1)), the 1997 quadratic with seasonality model outperformed the 1997 ARIMA model.

Table 1—: 1997 Model Forecasting Accuracy 1997-2022

.model	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
arima013011	6.974	8.496	6.97	1.737	1.737	4.727	5.382	0.985
TSLM(value ~ trend() + I(trend()^2) + season())	0.006	0.919	0.72	-	0.183	0.488	0.582	0.837

IV. Training New Models Based on Present Data

A. ARIMA Models

Since we have conducted a thorough analysis of the models developed in the 1997 report, we now wanted to devise updated models on this more granular and recent CO_2 concentration data. In order to do this, we seasonally adjusted the weekly data, and split both the seasonally adjusted (SA) and non-seasonally adjusted (NSA) data into training and test sets. Because we are working with time series data, our test sets were comprised of the most recent 2 years of weekly CO_2 concentrations (2020-2022). We fitted both SA and NSA training data sets to ARIMA models. Before fitting ARIMA models, we took the first difference of the data to remove the increasing trend and evaluated whether or not the weekly data was stationary post differencing (since stationarity is an assumption of ARIMA models).

```
##
## Phillips-Perron Unit Root Test
##
## data:  train.diff.sa$diff[-1]
## Dickey-Fuller = -108, Truncation lag parameter = 8, p-value = 0.01

##
## Augmented Dickey-Fuller Test
##
## data:  train.diff.sa$diff[-1]
## Dickey-Fuller = -19, Lag order = 13, p-value = 0.01
## alternative hypothesis: stationary
```

As shown above, using both the Phillips-Perron Unit Root Test and Augmented Dickey-Fuller Test we reject the null hypothesis that the first difference of the SA weekly CO_2 concentration data is non-stationary.

```
##
## Phillips-Perron Unit Root Test
##
## data:  train.diff.nsa$diff[-1]
## Dickey-Fuller = -53, Truncation lag parameter = 8, p-value = 0.01

##
## Augmented Dickey-Fuller Test
##
## data:  train.diff.nsa$diff[-1]
## Dickey-Fuller = -16, Lag order = 13, p-value = 0.01
## alternative hypothesis: stationary
```

Similarly, using both the Phillips-Perron Unit Root Test and Augmented Dickey-Fuller Test we also reject the null hypothesis that the first difference of the NSA weekly CO_2 concentration data is non-stationary. We therefore call the first difference of both datasets stationary processes and can proceed to fit ARIMA models.

```
## Series: ppm
## Model: ARIMA(1,0,1)(1,1,0)[52] w/ drift
##
## Coefficients:
##          ar1          ma1          sar1    constant
##          0.9838   -0.7132   -0.4967     0.0446
## s.e.    0.0044    0.0208    0.0179     0.0029
##
## sigma^2 estimated as 0.2571:  log likelihood=-1757
## AIC=3524   AICc=3524   BIC=3553

## Series: ppm
## Model: ARIMA(1,0,1)(1,1,0)[52] w/ drift
##
## Coefficients:
##          ar1          ma1          sar1    constant
##          0.9838   -0.7132   -0.4967     0.0446
## s.e.    0.0044    0.0208    0.0179     0.0029
##
## sigma^2 estimated as 0.2571:  log likelihood=-1757
## AIC=3524   AICc=3524   BIC=3553

## Series: season_adjust
## Model: ARIMA(0,1,2)(1,0,0)[52] w/ drift
##
## Coefficients:
##          ma1          ma2          sar1    constant
##          -0.6127   -0.1583   -0.2323     0.0429
## s.e.    0.0199    0.0195    0.0202     0.0017
##
## sigma^2 estimated as 0.1265:  log likelihood=-916
## AIC=1841   AICc=1841   BIC=1870

## Series: season_adjust
## Model: ARIMA(0,1,3)(2,0,0)[52] w/ drift
##
## Coefficients:
##          ma1          ma2          ma3          sar1          sar2    constant
##          -0.6033   -0.1430   -0.0095   -0.2882   -0.2660     0.0546
```

```
## s.e.    0.0204    0.0228    0.0196    0.0198    0.0201    0.0017
##
## sigma^2 estimated as 0.1174:  log likelihood=-848
## AIC=1711    AICc=1711    BIC=1751
```

For the NSA data, we fit two ARIMA models, ARIMA(1,0,1)(1,1,0)[52] based on the best (lowest) AIC (AIC = 3524) and ARIMA(1,0,2)(1,1,0)[52] based on the best (lowest) BIC (BIC = 3423) from all permutations of ARIMA models with up to 10 autoregressive processes, 2 unit roots differences, 10 moving average processes, and the equivalent in seasonality components. We used similar techniques to fit the SA data, and found that the same ARIMA model, ARIMA(0,1,2)(1,0,0)[52], had the lowest AIC (AIC = 1841) and BIC (BIC = 1870) for the SA data (model shown below).

For each of these models, we evaluated the model accuracy against the training (in-sample) and test (out-sample) data sets.

EVALUATING ACCURACY

To evaluate the ARIMA models, we first assessed each model (the best model via AIC or BIC for either the NSA or SA data sets) accuracy against the training data provided, which is described in Table 2 below.

Table 2—: Evaluating ARIMA model accuracy for present data (Training Data)

model	IC	.modeltype	ME	RMSE	MAE	MAPE	MAP	MASE	RMSSCF	ACF1
ARIMA(1,0,1)(1,1,0)[52] (NSA data)	AIC	auto Training	0.008	0.501	0.382	0.002	0.104	0.208	0.249	0.127
ARIMA(0,1,2)(1,0,0)[52] (SA data)	AIC	auto Training	-	0.355	0.269	-	0.073	0.148	0.178	0.003
			0.001			0.001				
ARIMA(1,0,1)(1,1,0)[52] (NSA data)	BIC	auto Training	0.008	0.501	0.382	0.002	0.104	0.208	0.249	0.127
ARIMA(0,1,2)(1,0,0)[52] (SA data)	BIC	auto Training	-	0.342	0.259	-	0.070	0.141	0.170	0.000
			0.001			0.001				

Based on the in-sample accuracy comparisons, for non seasonally adjusted data, we chose the ARIMA(1,0,4)(1,1,0) model because it had slightly better (lower) root mean squared error and ACF1 compared to the other ARIMA model (ARIMA(1,0,2)(1,1,0)). For the seasonally adjusted data, we only had one ARIMA model, ARIMA(0,1,2)(1,0,0), and use that model for out-sample evaluation.

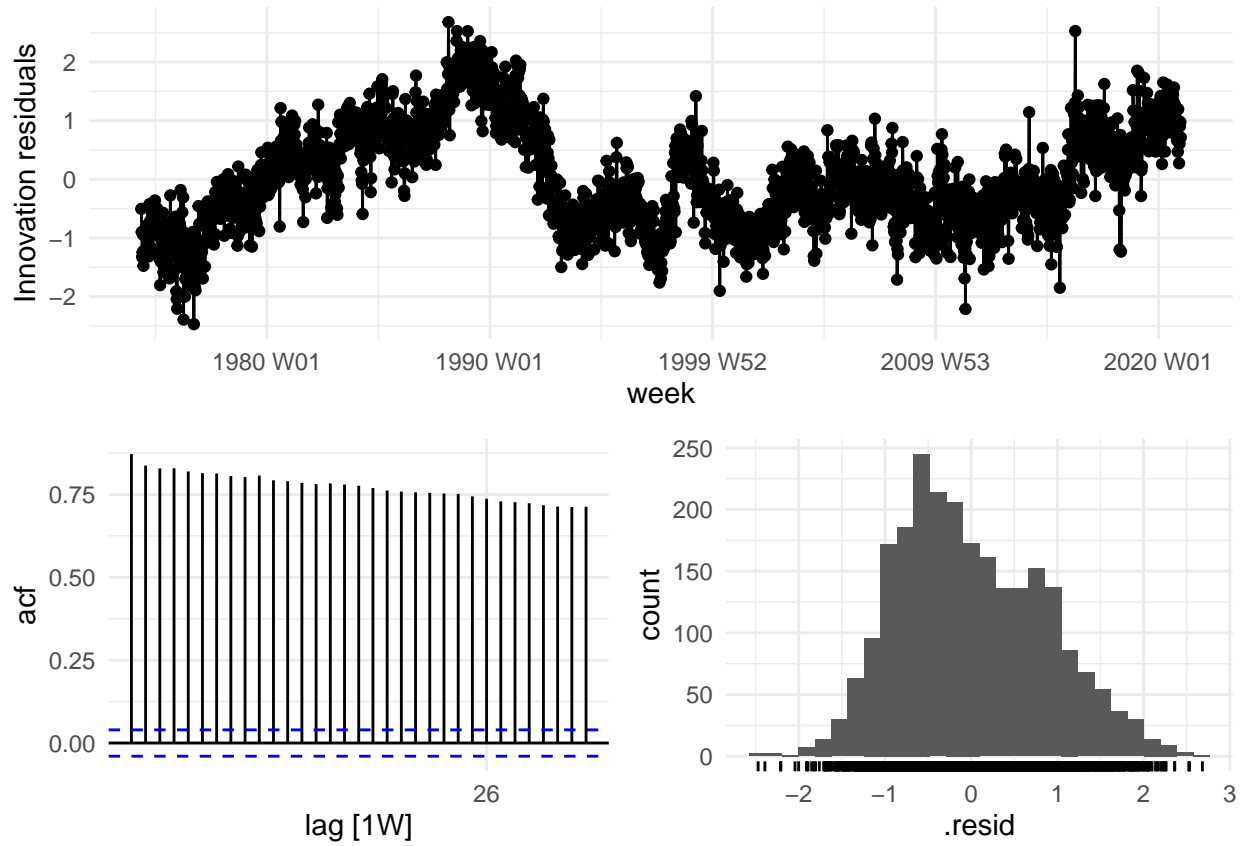
To evaluate our chosen ARIMA models further, we used each model (for NSA and SA data respectively) to forecast 2 years in the future (2020-2022) and compared the accuracy of each model against our test data sets (both for NSA and SA data). These results are described below in Table 3.

Table 3—: Present ARIMA Model Forecasting Accuracy
2021,2022

data	.model	ME	RMSE	MAE	MPE	MAPE	ACF1
Non-seasonally adjusted	auto	0.277	0.608	0.507	0.066	0.121	0.252
Seasonally adjusted	auto	0.858	1.015	0.885	0.205	0.212	0.439

B. Polynomial Model

```
## Series: season_adjust
## Model: TSLM
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.471 -0.644 -0.126  0.634  2.682
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.31e+02   5.19e-02   6376   <2e-16 ***
## trend()      2.23e-02   9.84e-05    226   <2e-16 ***
## I(trend()^2) 4.95e-06   3.91e-08    127   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.852 on 2431 degrees of freedom
## Multiple R-squared:  0.999,    Adjusted R-squared:  0.999
## F-statistic: 9.83e+05 on 2 and 2431 DF, p-value: <2e-16
```



EVALUATING ACCURACY

Table 4—: Polynomial model accuracy 2021, 2022

.model	.type	ME	RMSE	MAE	MPE	MAP	MASE	RMSS	ACF1
TSLM(season_adjust ~ trend() + I(trend()^2))	Test	0.68	0.853	0.723	0.163	0.173	NaN	NaN	0.432

V. How Bad Could It Get?

A. (3 points) Task Part 6b: How bad could it get?

With the non-seasonally adjusted data series, generate predictions for when atmospheric CO2 is expected to be at 420 ppm and 500 ppm levels for the first and final times (consider prediction intervals as well as point estimates in your answer). Generate a prediction for atmospheric CO2 levels in the year 2122. How confident are you that these will be accurate predictions?

Table 5—: Predictions on crossing the 420 and 500 CO2 ppm

target	week	.mean	ci.80.lower	ci.80.upper	ci.95.lower	ci.95.upper
420	2022 W13	421	419	422	419	422
420	2025 W40	421	418	423	417	424
500	2065 W09	500	492	508	488	512
500	2068 W35	500	492	508	488	513

We now show predictions for the year 2122:

Table 6—: Predictions for the year 2122

week	.mean	ci.80.lower	ci.80.upper	ci.95.lower	ci.95.upper
2122 W02	605	593	616	587	623
2122 W03	605	593	617	587	623
2122 W04	605	593	616	587	623
2122 W05	604	592	616	586	622
2122 W06	604	592	616	586	622
2122 W07	604	592	615	586	621
2122 W08	603	592	615	585	621
2122 W09	603	591	614	585	620
2122 W10	602	591	614	584	620
2122 W11	601	590	613	584	619
2122 W12	601	589	612	583	618
2122 W13	601	589	612	583	618
2122 W14	600	589	612	583	618
2122 W15	600	588	611	582	618
2122 W16	600	588	611	582	617
2122 W17	599	587	611	581	617
2122 W18	599	587	610	581	617
2122 W19	599	587	610	581	617
2122 W20	598	587	610	581	616
2122 W21	598	587	610	581	616
2122 W22	599	587	610	581	616
2122 W23	599	587	610	581	616
2122 W24	599	587	611	581	617
2122 W25	599	587	611	581	617
2122 W26	599	588	611	582	617

week	.mean	ci.80.lower	ci.80.upper	ci.95.lower	ci.95.upper
2122 W27	600	589	612	582	618
2122 W28	600	588	612	582	618
2122 W29	601	589	613	583	619
2122 W30	601	590	613	583	619
2122 W31	601	589	613	583	619
2122 W32	602	590	613	584	619
2122 W33	602	590	614	584	620
2122 W34	603	591	614	585	620
2122 W35	603	591	614	585	621
2122 W36	603	591	614	585	620
2122 W37	604	592	615	586	621
2122 W38	604	592	615	586	621
2122 W39	604	592	616	586	622
2122 W40	604	593	616	586	622
2122 W41	604	592	615	586	621
2122 W42	604	592	615	586	622
2122 W43	604	592	615	586	622
2122 W44	604	592	616	586	622
2122 W45	604	592	616	586	622
2122 W46	605	593	616	587	622
2122 W47	605	593	617	587	623
2122 W48	606	594	617	588	624
2122 W49	606	594	617	588	624
2122 W50	606	594	617	588	623
2122 W51	606	595	618	589	624
2122 W52	606	595	618	589	624
2122 W53	607	595	618	589	625

VI. Conclusions

What to conclude is unclear.