Global CO_2 Emissions in 2023 and beyond

By *

I. Introduction

In follow-up to our 1997 report we continue investigating the trend of rising CO_2 levels and whether or not it is likely caused by a larger trend or stochastic effects.

II. Mona Loa CO_2 Data

NASA provides daily measurement of $[CO_2]$ that we download for our analysis. In April of 2019 the Mauna Loa laboratory updated their equipment to measure CO_2 with a new technique called Cavity Ring-Down Spectroscopy (CRDS) in contrast to the prior infrared absorption technique. As such, all data from April 2019 onwards will contain measurements using the new method. Additionally, due to eruptions at the Mauna Loa site in 2022, data from December 2022 onwards are from a site at the Maunakea laboratory.

A. Recent Trends in Atmospheric Carbon

Conduct the same EDA on this data. Describe how the Keeling Curve evolved from 1997 to the present, noting where the series seems to be following similar trends to the series that you "evaluated in 1997" and where the series seems to be following different trends. This EDA can use the same, or very similar tools and views as you provided in your 1997 report.

Atmospheric carbon is plotted in Figure 1, and shows relentless increase in atmospheric CO_2 concentration even at the remote Hawaiian island.

III. Models and Forecasts

In this section we will evaluate performance of historic models and generate new models based on the most recent data.

A. 1997 Models vs realized data

Figure 1 demonstrates the recent trends in $[CO_2]$ along with the models fitted on the data prior to 1997. It is clear that ever accelerating trend continues, with

* : , .

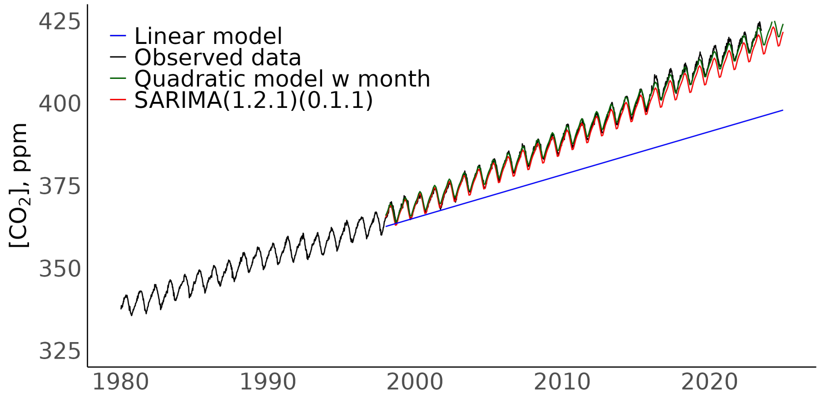


Figure 1. : Recent trend in $[CO_2]$ and models trained on pre-1997 data

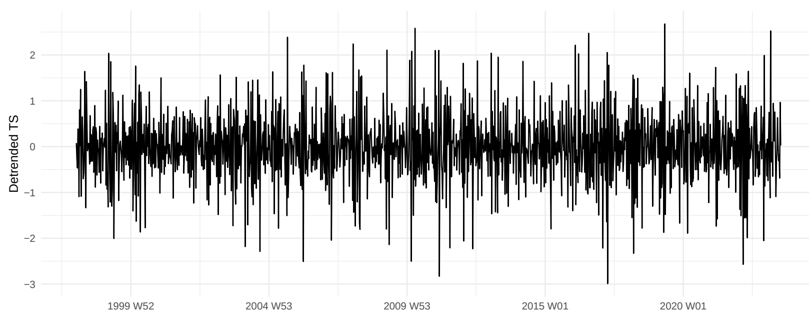
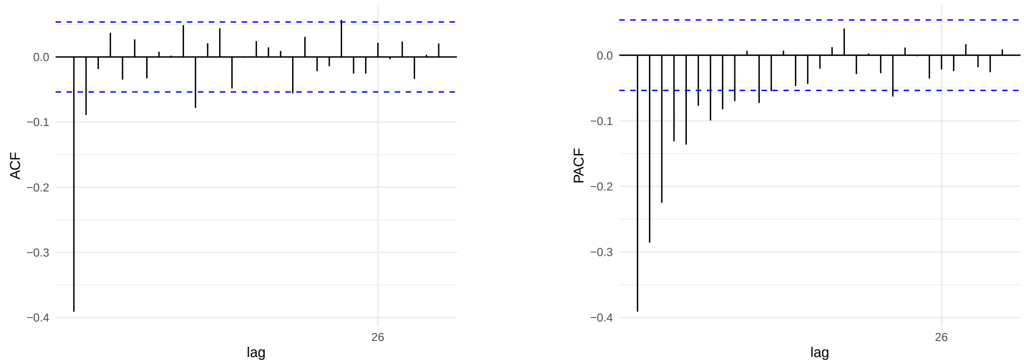


Figure 2. : De-seasoned/-trended series for CO_2 concentration

no noticeable slow down or even reduced acceleration. Clearly, accuracy of predictions from the linear model does not justify even electricity spent on fitting this model. The trend is accelerating and linear model grossly underestimates the real concentration. Remarkably, simple quadratic model augmented with months as categories is extremely accurate, capturing both accelerating trend and seasonal variability 25 years into the future! Our best ARIMA model underestimates real observations only slightly and its performance is still unexpetedly good for predictions that far out. Worth noticing that both seasonal models perform extremely well up until approximately 2016, when the real trend seemed to have had an additional boost. At then time ARIMA model started falling behind at an increased rate. That might be a statistical fluke, or might indicate the change in the CO_2 emissions.



(a) Recent trend in CO_2 concentration

(b) Recent trend in CO_2 concentration

Figure 3. : Diagnostic plots for de-trended series

B. *Performance of 1997 linear and ARIMA models*

It appears that the first time $[CO_2]$ exceeded 420ppm was in Feb 2023 and the last time it was that low was in Mar 2023. It is unlikely it will return to this value any time soon. That correlates surprisingly well with out 25 year-old prediction that it will reach this threshold on 2023-04-01. As predicted, we are still far away from crossing 500 ppm threshold.

In this case, an RMSE of 1.9 means that our model, on average, makes an error of about 1.9ppm in its predictions.

We use Mincer-Zarnowitz regression to quantify predictive power of our model. To avoid extrapolation too far from the data, we shift both predicted and actual data by 365 ppm. Intercept of this regression is 0 and slope is 1.062. Given how close these parameters are to 0 and 1, we conclude that our ARIMA model has high predictive power. However, t-test rejects hypothesis that they are equal to 0 and 1 with p-values of 0.61479298 and $5.13297408 \times 10^{-79}$ respectively. This is likely due to the high number of data points, that makes small deviations highly significant.

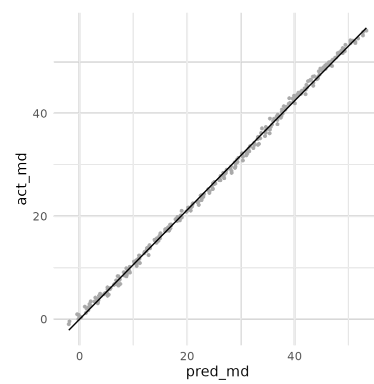


Figure 4. : Mincer-Zarnowitz regression for 1997 ARIMA model

C. Best models on present data

Seasonally adjust the weekly NOAA data, and split both seasonally-adjusted (SA) and non-seasonally-adjusted (NSA) series into training and test sets, using the last two years of observations as the test sets. For both SA and NSA series, fit ARIMA models using all appropriate steps. Measure and discuss how your models perform in-sample and (psuedo-) out-of-sample, comparing candidate models and explaining your choice. In addition, fit a polynomial time-trend model to the seasonally-adjusted series and compare its performance to that of your ARIMA model. In this section we will compare perfromance of various models in terms of their predictive power. For that we will save a small portion of our data, the most recent 2 years, as a test set. We will use the rest of the data (train set) to generate the models.

```
# Create a training set
co2_train <- co2_diff %>% filter((week > yearweek("1998-01-01")) &
  (week < yearweek("2021-06-01")))

# Cretae a test set
co2_test <- co2_diff %>% filter(week >= yearweek("2021-06-01"))
```

We first focus on seasolly adjusted data. As a first step we will perform a grid search of the model space, looking for the ARIMA model with the lowest BIC.

```
## Series: deseasoned
## Model: ARIMA(1,1,1) w/ drift
##
## Coefficients:
##          ar1      ma1  constant
##       0.1814  -0.8447   -0.0001
## s.e.  0.0356   0.0192    0.0028
##
## sigma^2 estimated as 0.388:  log likelihood=-1152.51
## AIC=2313.02   AICc=2313.05   BIC=2333.44

##
## Call:
## arima(x = co2_train$ppm, order = c(1, 1, 2), seasonal = list(order = c(0, 1,
##     1), period = 52))
##
## Coefficients:
##          ar1      ma1      ma2     sma1
##       0.08614  -0.67500  -0.09227  -0.81697
## s.e.  0.21123   0.21105   0.14537   0.02148
##
```

```
## sigma^2 estimated as 0.23105:  log likelihood = -830.66,  aic = 1671.33
```

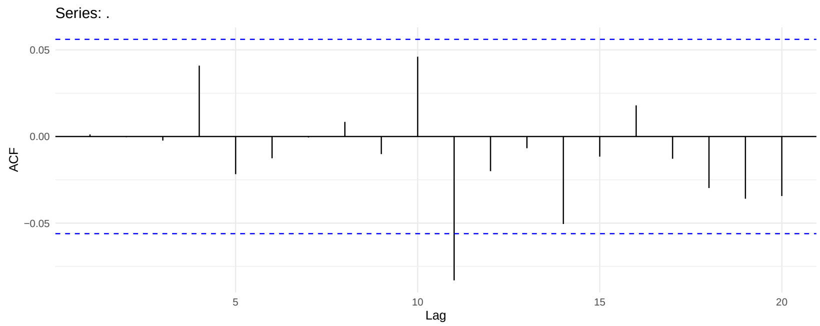


Figure 5. : ACF plot for SARIMA(1.1.1)(0.1.1)[52] residuals

Note: Residuals appear to be white noise

D. *Forecasts: 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?

```
## Don't know how to automatically pick scale for object of type <ts>. Defaulting
## to continuous.
```

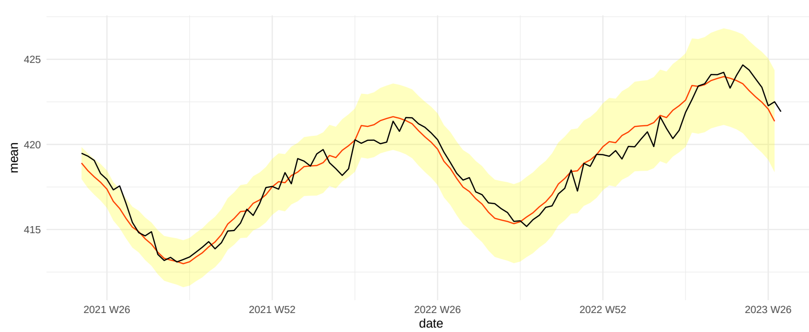


Figure 6. : Forecast for future CO2 concentration

Note: Residuals appear to be white noise

IV. Conclusions

What to conclude is unclear.

Appendix: Model Robustness

While the most plausible model that we estimate is reported in the main, "Modeling" section, in this appendix to the article we examine alternative models. Here, our intent is to provide a skeptic that does not accept our assessment of this model as an ARIMA of order (1,2,3) an understanding of model forecasts under alternative scenarios.