Ball State Undergraduate Mathematics Exchange
`http://www.bsu.edu/libraries/beneficencepress/mathexchange`
Vol. 11, No. 1 (Fall 2017)
Pages 14 – 27

A Statistical Analysis of Atmospheric CO_2 Levels at Mauna Loa

Lauren A. Krzaczek and Philip A. Yates

Lauren A. Krzaczek graduated from Saint Michaels College with a degree in Mathematics in 2017. She now works as a Credit Analyst at Yankee Farm Credit, ACA in Vermont.

Philip A. Yates is an Assistant Professor of Mathematical Sciences at DePaul University, formerly an Associate Professor of Mathematics at Saint Michaels College and an Assistant Professor of Mathematics & Statistics at Cal Poly Pomona. His research interests include statistical applications in hydrology, environmental sciences, sports, political science, and statistics education.

Abstract This honors capstone project was an analysis of atmospheric carbon dioxide levels measured at the Mauna Loa Observatory in Hawaii. The Scripps CO_2 Program through the Scripps Institution of Oceanography collects data constantly and makes the data available to the public in the form of daily, weekly, and monthly observations. Due to the seasonal variation in the monthly observations, traditional modeling methods learned at the undergraduate mathematics and statistics level, like least-squares regression, are not appropriate. An exponential smoother, like the Holt-Winters model, can be used to fit these types of data as well as be used to forecast future observations. A brief discussion about atmospheric carbon dioxide levels and the potential impact on the climate follows the analysis.

Introduction

In late September 2016, a milestone that displeased a lot of climate scientists was reached [15]. The carbon dioxide levels at Mauna Loa, Hawaii increased above the 400 parts per million threshold. This threshold was first crossed as a global average in

March 2015 [16]. Why is this figure important? Scientists have warned that crossing this threshold could result in more global warming and the disasters associated with it, like sea-level rise and ocean acidification [14].

In 1957, geochemist Charles David Keeling won funding for the International Geophysical Year which helped him to design and build a carbon dioxide monitoring station at Mauna Loa [4]. The first measurements from Hawaii started in March 1958. Through these years, the plotting of the CO_2 parts per million over time has been called a Keeling curve. See Figure 1 [10].

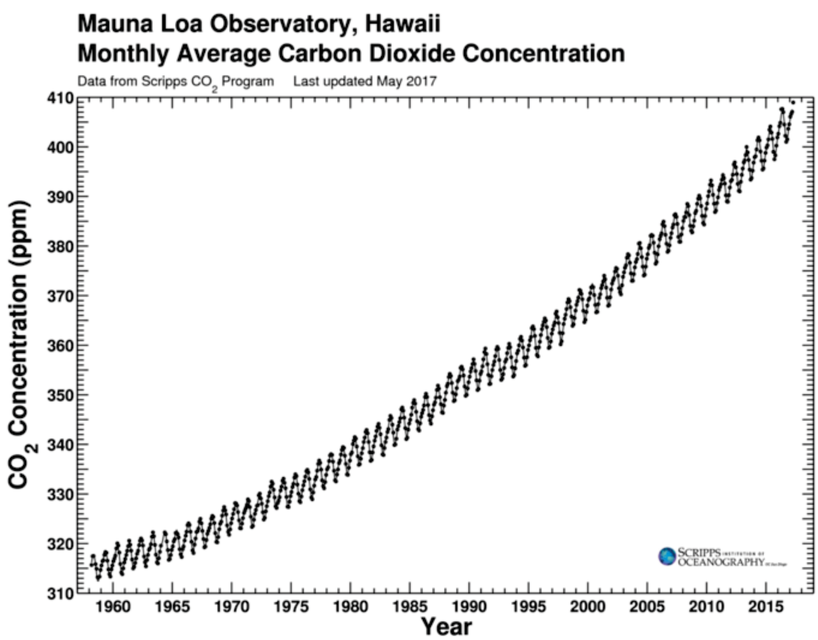

Figure 1: Keeling Curve: Mauna Loa Observatory, Hawaii

Due to the fact that the response variable, carbon dioxide levels measured in parts per million, is measured with respect to time, the given month and year the observation was measured, time series analysis techniques would be the best way to fit a model to the data and then use that model to forecast carbon dioxide levels for future increments of time. The rest of this paper will further describe the data collected from the Mauna Loa monitoring site and the reasoning behind the selection of the model used in the analysis. The fitted model and the forecasted carbon dioxide levels will be discussed in the forecasting results section. Finally, any final thoughts will be discussed in the summary and conclusions section.

Data & Model

The data for this analysis were collected by [10] from the Mauna Loa CO_2 monitoring site. Monthly observations started in March 1958 and ended with the most recent recorded month, April 2017. The variables include the monthly CO_2 concentrations in micromoles, CO_2 per mole (ppm), reported on the 2008A SIO manometric mole fraction scale; the month and year the measurement was recorded; and the month and year converted to a numeric value adjusted to the 15th of each month, i.e., March 1958 is 1958.203, March 1959 is 1959.203, etc.

Typically in most undergraduate applied statistical methods courses when faced with data like these, a simple linear regression model can be fit to the data. The simple linear regression model is most commonly defined as

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where the Y_i's are the values of the response variable, the X_i's are the values of the explanatory variable, β_0 and β_1 are the true values of the model's intercept and slope, and the ϵ_i's are error terms that are independent and identically distributed from a normal distribution with mean zero and variance σ^2. Estimates for β_0 and β_1 can be found using the method of least squares.

The least squares regression line is

$$\hat{Y}_i = -2686.912 + 1.529 X_i,$$

where $\hat{Y}_i$ is the estimated CO_2 concentration and X_i is the numeric value of the date. An interpretation of the estimated slope coefficient is that each year the CO_2 concentration is predicted or expected to increase by 1.529 ppm with a standard error of 0.00874 ppm. The least squares regression line and the CO_2 data can be found in Figure 2. This figure shows the original time series, the estimated trend component, the estimated seasonal component, and the estimated irregular component, also named the random component in the figure.

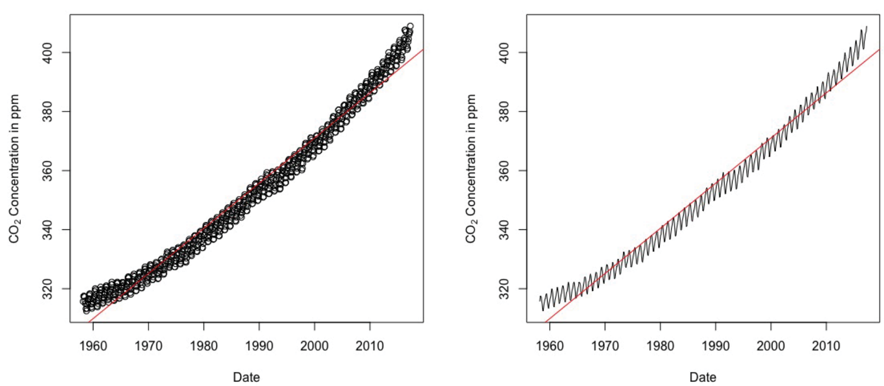

Figure 2: Least Squares Regression Line with Scatterplot & Lineplot of CO_2 Concentrations

To test the assumptions of the error terms in the model, residual plots, both the residuals versus the explanatory variable and the normal probability plot, are used. For the simple linear regression model fit to the CO_2 concentrations, the residuals plots are in Figure 3. For the graph on the left, if the error terms are independent and have constant variance σ^2, there should be constant spread above and below zero as well as no discernible pattern to the points. For the graph on the right, if the error terms are normally distributed, the observed residuals should be fairly close in value to the expected residuals from a normal distribution. This would mean that the points should fall close to the $y = x$ line. With both plots, what should be expected is not what is actually happening. This means that the simple linear regression model is not an appropriate model for these data.

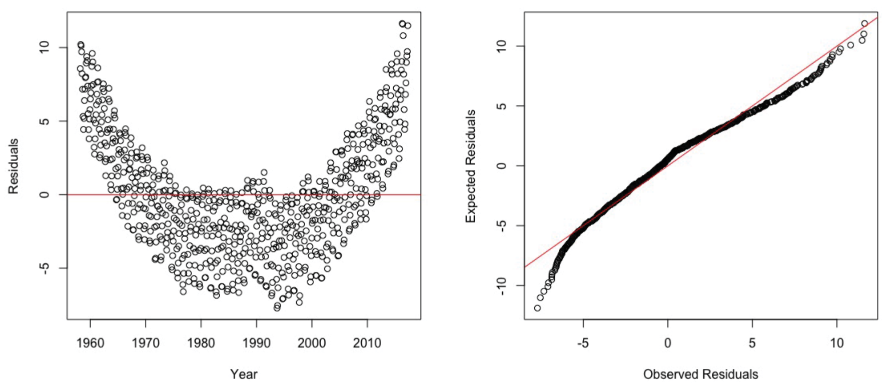

Figure 3: Residual Plots

Anytime observations are measured with respect to some unit of time, time series analysis should be used. [2] lays out a roadmap on how to analyze these type of data. The first thing that should be done is to decompose the time series into its constituent components, meaning its trend component, its irregular component, and if it has one, its seasonal component. Referring back to Figures 1 and 2, since the seasonal and random fluctuations seem to be roughly constant in size over time, an additive time series model seems appropriate for these data. An additive model is basically of the form

$$\text{Data} = \text{Trend-cycle} + \text{Seasonal} + \text{Irregular},$$

[17]. The estimated seasonal components are in Table 1. The largest seasonal factor is May (3.064) and the smallest seasonal factor is October (−3.309). This seems to indicate that CO_2 concentrations are at their highest in May and at their lowest in October.

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
0.026	0.674	1.377	2.599	3.064	2.351	0.751	−1.393	−3.167	−3.309	−2.073	−0.902

Table 1: Estimated Seasonal Effects: CO_2 Concentrations at Mauna Loa

Figure 4 shows the estimated trend, seasonal, and irregular components of the monthly carbon dioxide concentrations at Mauna Loa.

Since there are seasonal components to the time series that can be described using an additive model, the time series can be seasonally adjusted by estimating the

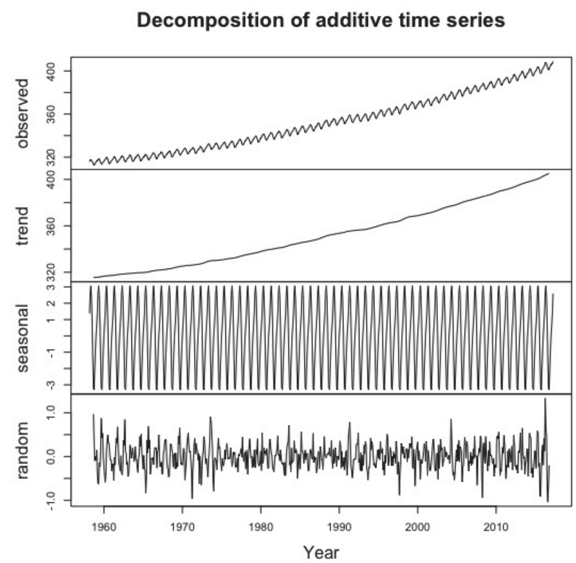

Figure 4: Decomposition of Monthly CO_2 Concentration Time Series

seasonal component and subtracting this component from the original time series. Figure 5 shows the adjusted time series. This series now just contains the monthly CO_2 concentration's trend component and irregular component.

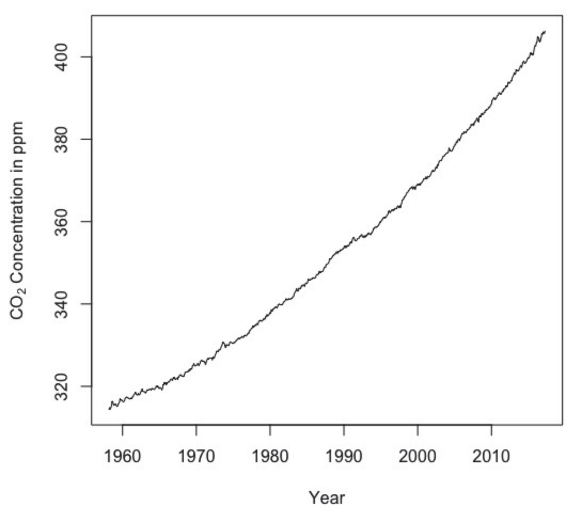

Figure 5: Seasonally Adjusted Monthly CO_2 Concentration Time Series

Forecasting Results

Often of the greatest interest when it comes to modeling, whether it is a simple linear regression model or an additive model for a time series, is making predictions or forecasts for future observations. When time series have both seasonal and trend components, Holt-Winters exponential smoothing can be used to make short-term forecasts [5, 18]. To make a forecast h units of time in the "future" and for a time series with period of seasonality m, the Holt-Winters additive method for exponential smoothing is [8]

$$\hat{Y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+},$$

where

$$\begin{aligned} \ell_t &= \alpha\left(Y_t - s_{t-m}\right) + \left(1-\alpha\right)\left(\ell_{t-1} + b_{t-1}\right), \\ b_t &= \beta\left(\ell_t - \ell_{t-1}\right) + \left(1-\beta\right)b_{t-1}, \\ s_t &= \gamma\left(Y_t - \ell_{t-1} - b_{t-1}\right) + \left(1-\gamma\right)s_{t-m}, \end{aligned}$$

and $h_m^+ = \lfloor (h-1) \bmod m + 1 \rfloor$. The parameters $0 \le \alpha \le 1$, $0 \le \beta \le 1$, and $0 \le \gamma \le 1-\alpha$ are estimated by optimizing the likelihood function for the data while searching over a restricted parameter space. This ensures the resulting model is forecastable. However, what do all of these smoothing equations mean? The estimated values of α, β, and γ for the CO_2 concentrations at Mauna Loa are 0.556, 0.0076, and 0.1152, respectively. The value of the estimate of α means that the estimate of the CO_2 concentrations, the level ℓ_t, at a given time point is based on recent observations and some observations from the distant past. The value of the estimate of β means that the estimate of the slope of the trend, b_t, is rarely updated over the course of the time series. The value of the estimate of γ means that the estimates of the seasonal components are more based off of observations in the more distant past instead of being based upon recent observations. Figure 6 shows how the forecasted model performs compared to the actual time series. The forecasted CO_2 concentrations are in red and the actual CO_2 concentrations are in black.

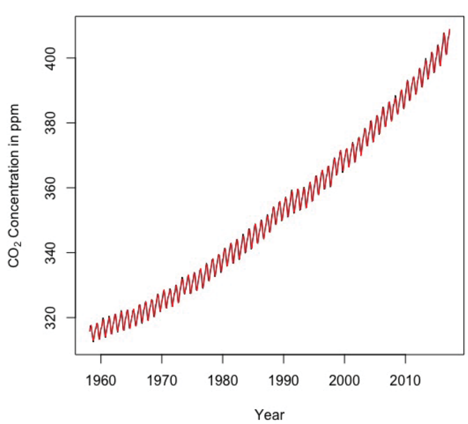

Figure 6: Time Series (Black) and Forecasted Time Series (Red) of CO_2 Concentrations

Suppose it was of interest to forecast the monthly CO_2 concentrations from May 2017 to December 2020, a period of 44 months, i.e., $h = 44$. Figure 7 shows the forecasts for these months (the solid red line) along with the forecasted time series for the previous twelve months. The 80% prediction intervals for these forecasts are the dashed red lines and the 95% prediction intervals for these forecasts are the dashed blue lines. The forecasted values can be found in Table 2 for 2020. Carbon dioxide concentrations were at their highest in May and their lowest in September.

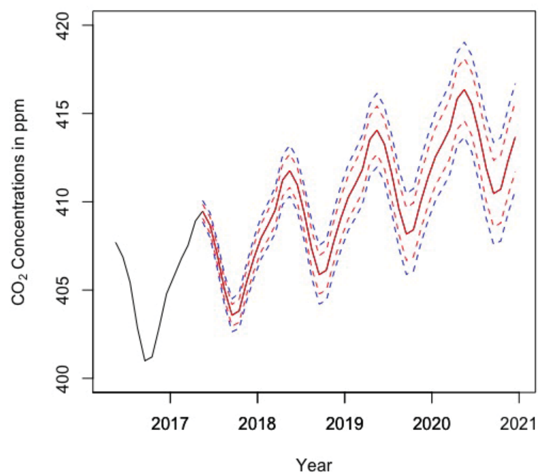

Figure 7: Forecast (solid red), 80% Prediction Intervals (dashed red), and 95% Prediction Intervals (dashed blue): Forecasts May 2017 to December 2020

Month	Point Estimate	80% Lower PI	80% Upper PI	95% Lower PI	95% Upper PI
January	412.551	410.930	414.172	410.072	415.030
February	413.335	411.682	414.988	410.807	415.863
March	414.129	412.444	415.814	411.552	416.706
April	415.828	414.102	417.554	413.189	418.468
May	416.357	414.599	418.115	413.669	419.045
June	415.566	413.777	417.355	412.830	418.303
July	413.929	412.108	415.750	411.144	416.714
August	411.903	410.050	413.756	409.070	414.736
September	410.465	408.580	412.349	407.583	413.347
October	410.712	408.796	412.628	407.782	413.643
November	412.326	410.378	414.274	409.347	415.305
December	413.668	411.689	415.648	410.641	416.696

Table 2: Forecasted CO_2 Concentrations, 2020

There remains the question on whether or not this model can be improved. This can be checked by looking at the residuals from the Holt-Winters forecasts. A correlogram can be constructed to investigate the autocorrelations of the sample forecast errors. See Figure 8. Lag 1 is just touching the significance bounds and lag 3 is slightly above the significance bounds. A Box-Ljung test [11] can be used to test the null hypothesis that there is independence in a given time series. In this situation, the question is whether or not it is safe to assume that the forecast error terms are independent over time. The p-value for this test is 0.05233. That means there is suggestive but inconclusive evidence that the forecast error terms are not independent over time; however, if one was comparing this p-value to the most commonly used significance level of 0.05, the decision would be to fail to reject the null hypothesis, i.e., there is not enough evidence to say that the forecast error terms are not independent. Figure 9 is a plot of the sample forecast errors with respect to the time of the forecast and a histogram of the sample forecast errors with a normal probability density function. For the first plot, the sample forecast errors all seem to have roughly the same spread and no discernible pattern. For the histogram, it seems safe to assume that the error terms are normally distributed. This means that the assumptions used to construct the prediction intervals, independent and normally distributed error terms with a common variance, are more than likely valid.

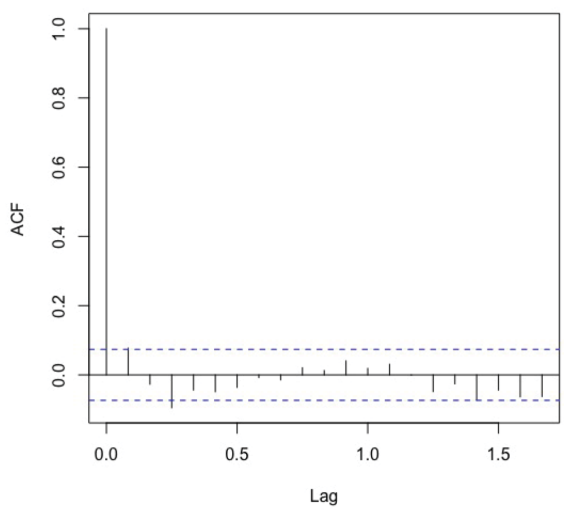

Figure 8: Correlogram of Autocorrelation of Sample Forecast Errors

Summary & Conclusions

The safe atmospheric carbon dioxide concentration level is 350 ppm [3]. At this level, Hansen et al. (2013) argue that the global temperature would stabilize at 1°C above pre-industrial levels. Unfortunately this threshold was first crossed at the Mauna Loa site in April 1987 and it has been above this threshold every month since November 1989. In fact, as stated in the introduction, the levels at Mauna Loa have surpassed the 400 ppm threshold and based on the forecasts, there are no signs that they will dip below that level.

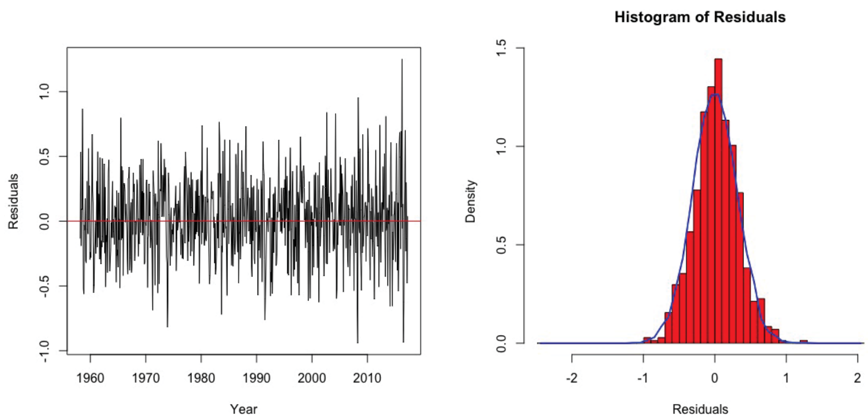

Figure 9: Sample Forecast Errors versus Time & Histogram of Sample Forecast Errors

There are a myriad reasons why these levels have been increasing since scientists began monitoring them. Be it manmade (i.e., burning of fossil fuels or the deforestation of the rain forests) or more biological in nature (i.e., ocean-atmosphere exchange, plant and animal respiration, plant decay and decomposition, or volcanic eruptions), the planet faces increasing levels of carbon dioxide concentrations in the atmosphere. The Paris Climate Pact adopted in December 2015 and eventually signed by 195 countries was an attempt to minimize the human impact on atmospheric carbon dioxide concentrations with the goal of holding the increase in the global average temperature to below 2°C above pre-industrial levels [1]. Recently, the United States has decided to opt out of this agreement, possibly putting the entire climate pact in jeopardy. Is all hope lost? A report from the Organisation for Economic Co-operation and Development [12] stated that if carbon dioxide concentrations are kept under the 450 ppm threshold, there is a 50% chance of stabilizing the average global temperature at an increase of 2°C above pre-industrial levels. A report from PricewaterhouseCoopers (PwC) [9] stated that if the global fossil fuel usage stays at its current levels, the 450 ppm threshold will be surpassed by 2034. There is still time to act, but the global leaders would have to do it sooner rather than later.

Bibliography

[1] Chan, S. (2015), "Key Points of the Paris Climate Pact," *The New York Times*, Available: https://www.nytimes.com/interactive/projects/cp/climate/2015-paris-climate-talks/key-points-of-the-final-paris-climate-draft, 12 December 2015.

[2] Cochran, A. (2017), *A Little Book of R For Time Series, Release 0.2*, Parasite Genomics Group, Wellcome Trust Sanger Institute, Cambridge, UK.

[3] Hansen, J., Kharecha, P., Sato, M., Masson-Delmotte, V., Ackerman, F., Beerling, D. J., Hearty, P. J., Hoegh-Guldberg, O., Hsu, S.-L., Parmesan, C., Rockstrom, J., Rohling, E. J., Sachs, J., Smith, P., Steffen, K., Susteren, L. V.,

von Schuckmann, K., and Zachos, J. C. (2013), "Assessing "Dangerous Climate Change": Required Reduction of Carbon Emissions to Protect Young People, Future Generations and Nature," *PloS ONE*, 8, 1–26.

[4] Harris, D. C. (2010), "Charles David Keeling and the Story of Atmospheric CO_2 Measurements," *Analytical Chemistry*, 82, 7865–7870.

[5] Holt, C. C. (1957), *Forecasting Trends and Seasonals by Exponentially Weighted Moving Averages*, Carnegie Institute of Technology, Office of Naval Research Memorandum, No. 52, Pittsburgh, PA.

[6] Hyndman, R. J. (2017), "`forecast`: Forecasting Functions for Time Series and Linear Models," R Package, version 8.0. Available: http://github.com/robjhyndman/forecast.

[7] Hyndman, R. J. and Khandakar, Y. (2008), "Automatic Time Series Forecasting: The `forecast` Package," *Journal of Statistical Software*, 26, 1–22.

[8] Hyndman, R. J., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008), *Forecasting with Exponential Smoothing: The State Space Approach*, Berlin: Springer-Verlag.

[9] Johnson, L., Grant, J., and Low, L. P. (2013), "Busting the Carbon Budget: Low Carbon Economy Index 2013," PricewaterhouseCoopers (PwC) LLP, UK, Available: http://www.pwc.co.uk/assets/pdf/low-carbon-economy-index-2013.pdf.

[10] Keeling, C. D., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M., and Meijer, H. A. (2005), "Atmospheric CO_2 and $^{13}CO_2$ Exchange with the Terrestrial Biosphere and Oceans from 1978 to 2000: Observations and Carbon Cycle Implications," in *A History of Atmospheric CO_2 and its Effects on Plants, Animals, and Ecosystems*, eds. Ehleringer, J. R., Cerling, T. E., and Dearing, M. D., New York: Springer Verlag, pp. 83–113.

[11] Ljung, G. M. and Box, G. E. P. (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.

[12] OECD (2012), *OECD Environmental Outlook to 2050*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264122246-en.

[13] R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

[14] Restuccia, A. (2015), "Global CO_2 Concentrations Surpass 400 ppm for First Time," *Politico*, Available: http://www.politico.com/story/2015/05/global-co2-concentrations-surpass-400-ppm-for-first-time-117689, 6 May 2015.

[15] Schwartz, J. (2016), "A Milestone for Carbon Dioxide in the Atmosphere," *The New York Times*, Available: https://www.nytimes.com/2016/10/04/science/atmospheric-carbon-dioxide-400-ppm.html, 3 October 2016.

[16] Vaughan, A. (2015), "Global Carbon Dioxide Levels Break 400ppm Mileston," *The Guardian*, Available: https://www.theguardian.com/environment/2015/may/06/global-carbon-dioxide-levels-break-400ppm-milestone, 6 May 2015.

[17] Wei, W. W. S. (1990), *Time Series Analysis: Univariate and Multivariate Methods*, Redwood City, California: Addison-Wesley.

[18] Winters, P. R. (1960), "Forecasting Sales by Exponentially Weighted Moving Averages," *Management Science*, 6, 324–342.

Appendix: R Code

The following are the lines of R code [13] used in the analysis of the CO_2 concentration analysis. The `forecast` package [7, 6] was used in the Forecasting Results section of the paper. The function `plotForecastErrors` is from [2].

Simple Linear Regression

Fitting the simple linear regression model.

```
> lm1 <- lm(CO2~Date.Num)
> summary(lm1)

Call:
lm(formula = CO2 ~ Date.Num)

Residuals:
    Min      1Q  Median      3Q     Max
-7.7151 -2.8513 -0.4161  2.4514 11.6365

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.687e+03  1.738e+01  -154.6   <2e-16 ***
Date.Num     1.529e+00  8.744e-03   174.9   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.979 on 708 degrees of freedom
Multiple R-squared:  0.9774,Adjusted R-squared:  0.9773
F-statistic: 3.058e+04 on 1 and 708 DF,  p-value: < 2.2e-16
```

Plotting the time series as a scatterplot and a line plot (Figure 2).

```
> plot(Date.Num,CO2,xlab="Date",
+ ylab=expression(paste(CO[2]," Concentration in ppm")))
> abline(lm1,col="red")
> plot(Date.Num,CO2,xlab="Date",
+ ylab=expression(paste(CO[2]," Concentration in ppm")),type="l")
```

Plotting the residuals versus year and the normal probability plot of the residuals (Figure 3).

```
> abline(lm1,col="red")
> e <- residuals(lm1)
> plot(Date.Num,e,xlab="Year",ylab="Residuals")
> abline(h=0,col="red")
> MSE <- sum(e^2)/708
> expected <- NULL
> n <- length(e)
> for(i in 1:n){
+ expected[i] <- qnorm(i/(n+1),0,sqrt(MSE))
+ }
> plot(sort(e),expected,xlab="Observed Residuals",ylab="Expected Residuals")
> abline(0,1,col="red")
```

Time Series Analysis

Reading the data as a time series and decomposing the time series (Figure 4).

```
> CO2.ts <- ts(CO2,start=c(1958,3),end=c(2017,4),frequency=12)
> CO2.decompose <- decompose(CO2.ts)
> CO2.decompose$seasonal
             Jan         Feb         Mar         Apr         May
1959  0.02607684  0.67399350  1.37734120  2.59908689  3.06437425
1960  0.02607684  0.67399350  1.37734120  2.59908689  3.06437425
             Jun         Jul         Aug         Sep         Oct
1959  2.35130672  0.75138574 -1.39287431 -3.16683691 -3.30868719
1960  2.35130672  0.75138574 -1.39287431 -3.16683691 -3.30868719
             Nov         Dec
1959 -2.07269472 -0.90247202
1960 -2.07269472 -0.90247202
> plot(CO2.decompose,xlab="Year")
```

Plotting the seasonally adjusted time series. (Figure 5).

```
> CO2.adjust <- CO2.ts-CO2.decompose$seasonal
> plot(CO2.adjust,xlab="Year",ylab=expression(paste(CO[2]," Concentration in ppm")))
```

Fitting the Holt-Winters exponential smoother using the `ets` function in the `forecast` library.

```
> library(forecast)
> CO2forecasts <- ets(CO2.ts,model="AAA")
> CO2forecasts
ETS(A,A,A)

Call:
 ets(y = CO2.ts, model = "AAA")

  Smoothing parameters:
    alpha = 0.556
    beta  = 0.0076
    gamma = 0.1152

  Initial states:
    l = 314.5802
    b = 0.0738
    s=0.6799 0.0594 -0.8278 -1.8735 -3.0033 -2.7398
           -1.1856 0.6469 2.1334 2.6649 2.2746 1.1709

  sigma:  0.3096

     AIC     AICc      BIC
3030.371 3031.255 3107.981
```

Plot the time series and forecasted time series (Figure 6).

```
> plot(CO2.ts,xlab="Year",ylab=expression(paste(CO[2]," Concentration in ppm")))
> lines(Date.Num,fitted(CO2forecasts),col="red")
```

Plot the forecasts for future months (Figure 7).

```
> CO2forecasts2 <- forecast(CO2forecasts,h=44)
> plot(CO2forecasts2,xlab="Year",
+ ylab=expression(paste(CO[2]," Concentrations in ppm")),main="")
> CO2forecasts2
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
May 2017       409.4661 409.0693 409.8628 408.8593 410.0729
Jun 2017       408.6753 408.2184 409.1322 407.9765 409.3741
Jul 2017       407.0383 406.5269 407.5497 406.2562 407.8204
Aug 2017       405.0121 404.4503 405.5739 404.1529 405.8713
Sep 2017       403.5739 402.9647 404.1832 402.6422 404.5057
Oct 2017       403.8215 403.1672 404.4758 402.8208 404.8222
Nov 2017       405.4350 404.7375 406.1326 404.3682 406.5018
Dec 2017       406.7776 406.0383 407.5168 405.6470 407.9081
Jan 2018       407.9567 407.1771 408.7364 406.7644 409.1491
Feb 2018       408.7408 407.9218 409.5597 407.4883 409.9933
Mar 2018       409.5351 408.6777 410.3924 408.2238 410.8463
Apr 2018       411.2343 410.3251 412.1435 409.8438 412.6248
May 2018       411.7630 410.8175 412.7086 410.3169 413.2092
Jun 2018       410.9722 409.9908 411.9536 409.4713 412.4732
Jul 2018       409.3352 408.3185 410.3520 407.7802 410.8902
Aug 2018       407.3090 406.2573 408.3607 405.7006 408.9175
Sep 2018       405.8709 404.7846 406.9571 404.2096 407.5322
Oct 2018       406.1185 404.9979 407.2390 404.4048 407.8321
Nov 2018       407.7320 406.5775 408.8864 405.9664 409.4975
Dec 2018       409.0745 407.8864 410.2626 407.2574 410.8916
Jan 2019       410.2537 409.0321 411.4753 408.3854 412.1219
Feb 2019       411.0377 409.7829 412.2925 409.1186 412.9568
Mar 2019       411.8320 410.5441 413.1199 409.8624 413.8017
Apr 2019       413.5312 412.1996 414.8629 411.4946 415.5678
May 2019       414.0600 412.6958 415.4241 411.9737 416.1463
Jun 2019       413.2692 411.8726 414.6657 411.1334 415.4050
Jul 2019       411.6322 410.2034 413.0610 409.4470 413.8173
Aug 2019       409.6060 408.1450 411.0670 407.3716 411.8404
Sep 2019       408.1678 406.6747 409.6609 405.8843 410.4513
Oct 2019       408.4154 406.8902 409.9406 406.0828 410.7480
Nov 2019       410.0289 408.4718 411.5861 407.6474 412.4104
Dec 2019       411.3715 409.7823 412.9606 408.9411 413.8018
Jan 2020       412.5506 410.9296 414.1717 410.0715 415.0298
Feb 2020       413.3347 411.6818 414.9876 410.8067 415.8626
Mar 2020       414.1290 412.4442 415.8137 411.5523 416.7056
Apr 2020       415.8282 414.1022 417.5542 413.1885 418.4678
May 2020       416.3569 414.5993 418.1146 413.6689 419.0450
Jun 2020       415.5661 413.7768 417.3554 412.8296 418.3026
Jul 2020       413.9291 412.1081 415.7501 411.1441 416.7141
Aug 2020       411.9029 410.0503 413.7556 409.0695 414.7363
Sep 2020       410.4648 408.5804 412.3491 407.5829 413.3467
Oct 2020       410.7123 408.7962 412.6284 407.7819 413.6428
Nov 2020       412.3259 410.3780 414.2737 409.3469 415.3048
Dec 2020       413.6684 411.6888 415.6480 410.6408 416.6960
```

Plot the correlogram of the autocorrelations (Figure 8) and run the Box-Ljung test.

```
> acf(CO2forecasts2$residuals,lag.max=20,main="")
> Box.test(CO2forecasts2$residuals,lag=20,type="Ljung-Box")

Box-Ljung test

data:  CO2forecasts2$residuals
X-squared = 31.222, df = 20, p-value = 0.05233
```

Plot the sample forecast errors versus time and the histogram of the sample forecast errors (Figure 9).

```
> plot.ts(CO2forecasts2$residuals,xlab="Year",ylab="Residuals")
> abline(h=0,col="red")
> plotForecastErrors(CO2.forecasts$residuals)
```