# Title: Approximating Monthly Data of the Keeling Curve of Atmospheric CO<sub>2</sub> Concentration Using Exponential and Sinusoidal Functions

by Thomas Arend, 2023

### Abstract

<The Keeling Curve is a well-known time series that depicts the monthly measurements of atmospheric carbon dioxide ( $CO_2$ ) concentration at the Mauna Loa Observatory (MLO) in Hawaii. This study aims to develop an approximation model for the monthly data of the Keeling Curve by combining exponential and sinusoidal functions to capture both the long-term trend and the annual and semiannual oscillations (SAO).

By fitting the proposed model to the monthly data, the mean difference between the Keeling Curve and the approximation is found to be nearly 0 ppm (-1.462855e-15), indicating a close agreement between the two. Moreover, the standard deviation of the difference is calculated to be 0.7314717 ppm, highlighting the accuracy of the approximation in capturing the variability of the Keeling Curve data.

The developed approximation model provides valuable insights into the behavior of atmospheric  $CO_2$  concentration, allowing for the identification of underlying patterns and trends. It serves as a useful tool for understanding the impact of human activities and natural processes on  $CO_2$  levels in the atmosphere.

Overall, the proposed model successfully approximates the monthly data of the Keeling Curve, providing a comprehensive representation of the long-term trend and the annual and semiannual oscillations in atmospheric CO<sub>2</sub> concentration.

# Introduction

The Keeling Curve, initiated by Dr. Charles David Keeling in 1958<sup>1</sup>, is a seminal dataset that has provided valuable insights into the increase of atmospheric CO<sub>2</sub> concentration over the past six decades. The curve exhibits a clear upward trend, mainly driven by human activities, superimposed with seasonal variations due to natural processes such as plant growth and decay.

C. D. Keeling, S. C. Piper, R. B. Bacastow, M. Wahlen, T. P. Whorf, M. Heimann, and H. A. Meijer, Exchanges of atmospheric CO<sub>2</sub> and CO<sub>2</sub> 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.

In this study we develop an approximation of the Keeling Curve. For our investigation, we'll use the monthly in-situ CO<sub>2</sub> data set that has the file name of »monthly\_in\_situ\_co2\_mlo.csv.« We downloaded this file from

https://scrippsco2.ucsd.edu/data/atmospheric\_co2/primary\_mlo\_co2\_record.html

Antonio Stark<sup>2</sup> modeled the Keeling Curve with Quadratic Base and Sinusoidal Seasonal Trends. Instead of the quadratic base we will use a exponential base.

Ern et al.  $(2021)^3$  investigated the SAO in four different reanalyses for the period 2002–2018 but did not look at CO<sub>2</sub> concentrations.

Jiang et al.  $(2011)^4$  investigated the CO<sub>2</sub> SOA in the middle troposphere and at the surface. Using in situ measurements, they found a SAO in the midtropospheric and surface CO<sub>2</sub>.

Buermann et al  $(2007)^5$  found that after increasing trends in the seasonal amplitude of CO<sub>2</sub> at the MLO from the early 1970s to the early 1990s an updated record from the MLO indicates that the amplitude of the CO<sub>2</sub> seasonal cycle has been declining since the early 1990s until 2005. For our approximation we will ignore these trends.

### Methods

To approximate the monthly data (until 04/2023) of the Keeling Curve, we propose a mathematical model that combines an exponential function to capture the long-term trend and sinusoidal functions to represent the seasonal variations. The exponential function reflects the increasing CO<sub>2</sub> concentration over time, while the first sinusoidal function accounts for the annual oscillations. The second sinusoidal function accounts for the semiannual oscillations, which were discovered during the modeling process. We will come to this later.

The general form of our model is given by:

 $CO_{2(2023)}(t) = A^{*} \exp(r^{*} t) + B_{1}^{*} \sin(2\pi^{*} (t + \phi_{1})) + B_{2}^{*} \sin(4\pi^{*} (t + \phi_{2})) + C + \varepsilon_{3}(t)$ 

where:

- (t) represents the estimated CO<sub>2</sub> concentration at time t (in decimal date),
- A represents the base level of CO<sub>2</sub> concentration,
- r is the exponential growth rate,
- B1 represents the amplitude of the first sinusoidal function

<sup>2</sup> Antonio Stark: Using Data Science to Understand Climate Change: Atmospheric CO2 Levels (Keeling Curve) — Model Fitting and Time Series Analysis — Published in Towards Data Science, Jun 10, 2020

<sup>3</sup> Ern, M., Diallo, M., Preusse, P., Mlynczak, M. G., Schwartz, M. J., Wu, Q., and Riese, M.: The semiannual oscillation (SAO) in the tropical middle atmosphere and its gravity wave driving in reanalyses and satellite observations, Atmos. Chem. Phys., 21, 13763–13795, https://doi.org/10.5194/acp-21-13763-2021, 2021.

<sup>4</sup> Jiang, X., M. T. Chahine, Q. Li, M. Liang, E. T. Olsen, L. L. Chen, J. Wang, and Y. L. Yung (2012), CO2 semiannual oscillation in the middle troposphere and at the surface, Global Biogeochem. Cycles, 26, GB3006, doi:10.1029/2011GB004118.

<sup>5</sup> Buermann W, Lintner BR, Koven CD, Angert A, Pinzon JE, Tucker CJ, Fung IY. The changing carbon cycle at Mauna Loa Observatory. Proc Natl Acad Sci U S A. 2007 Mar 13;104(11):4249-54. doi: 10.1073/pnas.0611224104. Epub 2007 Mar 5. PMID: 17360510; PMCID: PMC1810339.

- B2 represents the amplitude of the second sinusoidal function
- φ1 represents the phase shift of the first sinusoidal function,
- $\phi$ 2 represents the phase shift of the second sinusoidal function,
- $C = C_1 + C_2 + C_3$  represents the baseline offset, and
- $\epsilon_i(t)$  represents the remaining error.

To determine the model parameters (A, r,  $B_1$ ,  $B_2$ ,  $\phi_1$ ,  $\phi_2$  and C), we employ a curve fitting technique, such as nonlinear least squares regression, to fit the model to the monthly data of the Keeling Curve. We will call this model Model 2023 because we used data from 1958 until April 2023.

In a first step we determine the model parameters for the exponential part of function.

 $CO_2(t) = A^* exp(r^* t) + C_1 + \varepsilon_1(t)$ 

After building the difference between the values and the exponential approximation  $\epsilon_1(t)$  we determine the model parameters for the first sinusoidal function.

 $\epsilon_1(t) = B_1 * \sin(2\pi * (t + \phi_1)) + C_2 + \epsilon_2(t)$ 

Calculating the error and studying  $\epsilon_2(t)$  we found that the error can be further approximated with semiannual sinusoidal function.

Finally we approximate  $\varepsilon_2(t) = B_2 * \sin(2\pi * (t + \phi_2)) + C_3 + \varepsilon_3(t)$ .

### Results

By fitting the proposed model to the monthly data of the Keeling Curve, we obtain estimates for the model parameters. The exponential component captures the long-term trend, while the sinusoidal component captures the seasonal variations. The goodness-of-fit of the model is assessed using statistical measures such as the coefficient of determination (R<sup>2</sup>) and root mean square error (RMSE).

```
CO<sub>2</sub>(t) = 1.31242863505765e-12 * exp( 0.0160453592326617 * t )
```

- + 2.85842937143706 \* sin( 2 \* pi \* (t 0.0575))
- 0.809084912332627 \* sin( 2 \* pi \* (t 0.0767))
- + 255.517201337755 + ε<sub>3</sub>(t)

We found  $R^2 = 0.9934622$  for the exponential part,  $R^2 = 0.8256447$  for the first and  $R^2 = 0.3814435$  for the second sinusoidal part of the approximation function.

The mean difference between the Keeling Curve and the exponential approximation ( $\epsilon_1(t)$ ) is found to be 0.03429826. The standard deviation of the difference is calculated to be 2.22735962.

Adding 1<sup>st</sup> sinusoidal part for the annual variations reduces the mean of the difference to nearly 0 (1.646100e-14) and the standard deviation is reduced to 0.9300533.

When adding the  $2^{nd}$  sinusoidal part for the semiannual variations the mean difference ( $\epsilon_3(t)$ ) is found to be nearly 0 (-1.624126e-14), indicating a close agreement between the two. Moreover, the standard deviation of the difference is reduced to 0.7314717 indicating that the approximation effectively captures the variability in the Keeling Curve data.

Four diagrams will be included to illustrate the results:

### **Keeling Curve**

Atmospheric CO2 concentrations (ppm)



© 2023 by Thomas Arend Source: R. F. Keeling, S. J. Walker, S. C. Piper and A. F. Bollenbacher; Stand: 15 Mai 2023

#### **Diagram 1: Keeling Curve with Approximations**

### Error ε<sub>1</sub>(t)



© 2023 by Thomas Arend Source: R. F. Keeling, S. J. Walker, S. C. Piper and A. F. Bollenbacher; Stand: 15 Mai 2023

Diagram 2: Error  $\varepsilon_1(t)$  and  $1^{st}$  and  $1^{st}$  and  $2^{nd}$  Sinusoidal Approximation The visual inspection of the error function  $\varepsilon_1(t)$  shows a clear annual sinusoidal oscillation. Therefore adding the first sinusoidal part to our model should improve the accuracy of the approximation.

#### 2023-05-16

### Error ε<sub>2</sub>(t)



#### Difference between atmospheric CO2 concentrations (ppm) and 1st sinusoidal approximation, $\epsilon 2(t)$

Diagram 3: Error  $\varepsilon_2(t)$  after Exponential and 1<sup>st</sup> Sinusoidal Approximation Our visual inspection of the error function  $\varepsilon_2(t)$  finds a semiannual oscillation. This semiannual

oscillation explains the small deviations form the sinus function in the first error function.

Therefore we add a second sinusoidal part to our model to improve the approximation further.

#### Error ε₃(t)



Diagram 4: Error  $\varepsilon_3(t)$  after Exponential, 1<sup>st</sup> and 2<sup>nd</sup> Sinusoidal Approximation Our visual inspection of the error function  $\varepsilon_3(t)$  shows no obvious pattern. Therefore we stop at this point and do not add further parts to our model.

### Extrapolation

To test the quality of the general model we restrict the data and extrapolate the approximation into the future. We determine the parameter of our model only on the values until December 1999 and call it Model 1999.

When we restrict the time frame from 1958 to 1999 the approximation differs only marginal.

CO<sub>2(1999)</sub>(t) = 1.26738983189575e-12 \* exp(0.0160596707252981 \* t) + 2.80725394368403 \* sin(2 \* pi \* (t - 0.0617)) - 0.750902310503362 \* sin(2 \* pi \* (t - 0.0794)) + 256.03745907830

When we extrapolate this model we can »predict« the future CO<sub>2</sub> concentration and test the predictive power of this approach. For April 2023 we predict the CO<sub>2</sub> concentration to be 419.9 ppm. The real value was 422.73 ppm which is 2.83 ppm higher. Taking into account that approximation based on only 41 year and the prediction was more then 23 years in the »future« this is a very good guess.

Using both models to predict the values for December 2049 we calculate 506.5894 ppm with Model 1999 and 507.5627 ppm with Model 2023. These values are less than 1 ppm apart.



Diagram 5: Keiling Curve with Approximations based on values until December 1999 Diagram 5 visually shows that the extrapolation of the model fits extremely good with the real values.

# Discussion

The proposed model provides a reasonable approximation of the monthly data of the Keeling Curve by combining an exponential function for the long-term trend and two sinusoidal functions for the annual and semiannual oscillations. While the factors causing the annual oscillation can be attributed to natural processes such as plant growth and decay, the factors contributing to the second sinusoidal oscillation with the semiannual frequency are not fully understood. Further research is needed to investigate the underlying mechanisms responsible for this phenomenon.

# **Limitations and Future Research**

Although the proposed model captures the general behavior of the Keeling Curve, it is important to note that it may not account for all factors influencing CO<sub>2</sub> concentrations. The inclusion of additional sinusoidal components allows for a more accurate representation of the seasonal variations, but the underlying drivers of the semiannual oscillations remain a subject of ongoing scientific inquiry. Future research efforts should focus on unraveling the complex interactions between various climatic, biological, and atmospheric processes that contribute to the observed semiannual variations in CO<sub>2</sub> concentrations.

# Conclusion

This study presents a mathematical model that combines an exponential function and sinusoidal functions to approximate the monthly data of the Keeling Curve. The model captures both the long-term trend and the annual variations in atmospheric CO<sub>2</sub> concentration, providing valuable insights into the impact of human activities and natural processes. However, the factors contributing to the second sinusoidal variation with the semiannual oscillation require further investigation to enhance our understanding of the dynamics of CO<sub>2</sub> concentrations in the atmosphere.

When the approximation is extrapolated over a longer period using only data up to 1999, the predictive accuracy of the approach is impressive.

### **Limitations and Future Research**

Although the proposed model captures the general behavior of the Keeling Curve, it may not account for all factors influencing CO<sub>2</sub> concentrations. Other variables, such as climate patterns and anthropogenic influences, might introduce additional complexities. Future research could explore the inclusion of additional terms or alternative mathematical functions to improve the accuracy of the approximation model.

### Conclusion

This study presents a mathematical model that combines an exponential function and a sinusoidal function to approximate the monthly data of the Keeling Curve. The model captures both the long-term trend and the seasonal variations in atmospheric CO<sub>2</sub> concentration. The proposed approximation facilitates the analysis of CO<sub>2</sub> concentration patterns and provides a valuable tool for understanding the dynamics of climate change and its underlying drivers.

# Keywords

Keeling Curve,  $CO_2$  concentration, exponential function, sinusoidal function, approximation, trend, seasonal variations, climate change.