Modeling Atmospheric Carbon Dioxide over the United States

Author
W. E. Long
tlong253@me.com

Advisor
J. Ernstberger
jernstberger@lagrange.edu
LaGrange College

Abstract

In this undergraduate research project, we implement a published, component-based model of the total carbon dioxide concentration in the atmosphere above the United States. To do so we use a simple ordinary differential equation to quantify the behavior of the carbon cycle (as described in published NOAA findings), compute curves of best fit to approximate CO_2 contributions due to emissions from fossil fuels and forest fires, and employ Henry's law to estimate the overall oceanic CO_2 absorption effect. Finally, via residual comparison, model parameters are estimated to determine a model approximation.

Introduction

Carbon dioxide is a commonly discussed greenhouse gas. In the following sections, we seek to understand the sources of carbon dioxide in the atmosphere. We examine a compartmental model of the total atmospheric carbon dioxide concentration in comparison to collected atmospheric CO_2 data. We explore each individual component of the model in the following sections. We have located previously recorded and published data for each component, and we have determined a numerical solution and model each component separately and in the aggregate using a variety of techniques, including differential equations and least squares approximations. In each of the following sections, we present our methodology for modeling the component as well as benefits and weaknesses in each particular method. We then present a final composite model representing the total of the four components discussed, which indicates an increasing atmospheric concentration.

Compartmental Model

To model the concentration of carbon dioxide in the atmosphere, we will use a compartmental model discussed in the Peters, Jacobson, and Sweeney paper [8]. The model used is

$$F(x,y,t) = \lambda_r F_{bio}(x,y,t) + \lambda_r F_{oce}(x,y,t) + F_{ff}(x,y,t) + F_{fire}(x,y,t).$$

Here, F_{bio} represents the concentration attributable to the natural carbon cycle, F_{oce} denotes the reduction in carbon dioxide through oceanic absorption, F_{ff} represents the addition of carbon dioxide through fossil fuel burning, and F_{fire} denotes the release of carbon dioxide from forest fires. The λ_r value is a scalar that is unique to each specific location. The presence of λ_r could give the model a global scope. The F_{ff} and F_{fire} components will increase the total amount of CO_2 in the atmosphere, where the F_{oce} will decrease the total concentration of CO_2 [8]. We implement the model to evaluate the CO_2 over the United States. Accordingly, we have treated the λ_r scalar as one.

Although more complex CO_2 models exist, the Peters, Jacobson, and Sweeney paper [8] is showing its validity through ongoing usage. The compartmental nature of this model has proven to be its greatest advantage. It allowed us to examine each component individually and then implement it in a manner congruent with data sources. Unfortunately, the relative simplicity of this model is also its greatest disadvantage. The model does not account for other sources of CO_2 separately, such as volcanism (volcanic eruption), which could cause us to oversimplify our model for each component.

Each of the following sections will describe the methods used to model the behavior from a particular component.
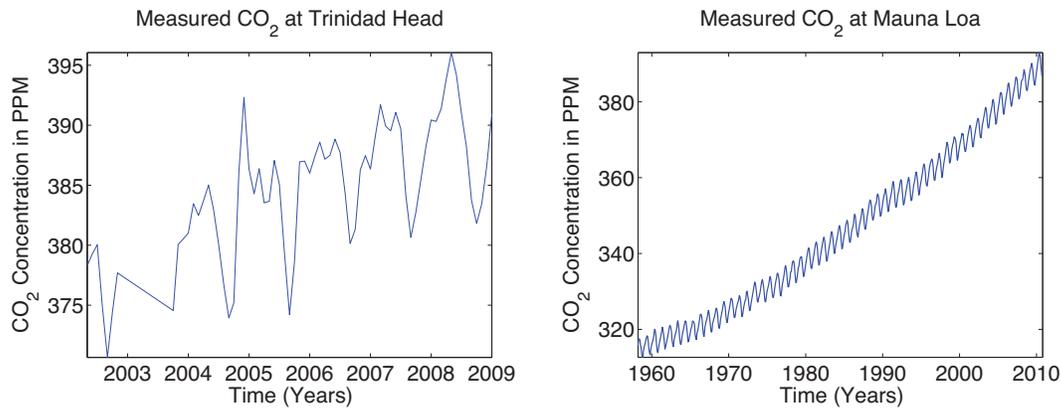
Figure 1: **Left:** Carbon cycle data collected at Trinidad Head [9]. **Right:** Carbon cycle data collected at Mauna Loa [6].

Carbon Cycle

Data selection was important to model the natural carbon cycle, as it represents the natural fluctuations in atmospheric CO_2 concentrations. Since the objective is to model the carbon dioxide concentration over the United States, the first source examined was the NOAA observatory at Trinidad Head, California [9]. Since this location is within the continental United States, it initially appeared to be the ideal source for implementation of the model. The data from this source is plotted in Figure 1 (left).

Unfortunately, the data was not recorded at regular intervals, and the overall nature of the carbon cycle is not depicted well. Note that no recordings were made between January and August 2003 at the Trinidad Head site either [9]. There should be a natural cycle due to the growth and death of plants when seasons change, but this behavior is not apparent in this data. We would expect the concentration to be lower during the summer months and higher during the winter months. However, the higher and lower concentrations appear to be outliers, rather than a natural effect. There is also the potential that the CO_2 in this location would be higher than typical over the United States, due to the proximity to large cities. Thus, the effect of the fossil fuel component of the model could indicate a higher CO_2 emission than typical in the United States. Additionally, the data collection began in late 2002, which would limit the time period from which behavior could be interpreted, and an extended period would provide greater opportunity to interpret the trend [9]. Consequently, another data source needed to be located.

The source found and used in the model is the NOAA observatory at Mauna Loa, Hawaii. The carbon dioxide concentration has been measured monthly at this site since the late 1950s. At both sites, the data is recorded in the same manner, as well. A flask is used to collect surface air at the observatory, and the contents are then analyzed for their components [6]. Alongside the apparent regularity of data, due to its location, Mauna Loa is among the least altered by carbon dioxide emissions from fossil fuels. This leads us to believe the carbon cycle represented more closely resembles a natural pattern of the northern hemisphere [6]. It should be noted that Mauna Loa would be subject to an additional source of CO_2. The constant volcanic activity in Hawaii releases additional CO_2 into the atmosphere, which could contribute to an upward trend. The data is plotted in Figure 1 (right).

Here, the nature of the data and its trend can be seen more clearly. A natural yearly cycle can be seen in this data, with the higher and lower concentrations corresponding to the predicted seasons. There appears to be an upward trend to the data, as well, which we will approximate through other model components [6]. According to conventional wisdom, the primary source for the upward trend should be fossil fuels burning. Despite the potential increase in CO_2 concentration from volcanism, we believe that the Mauna
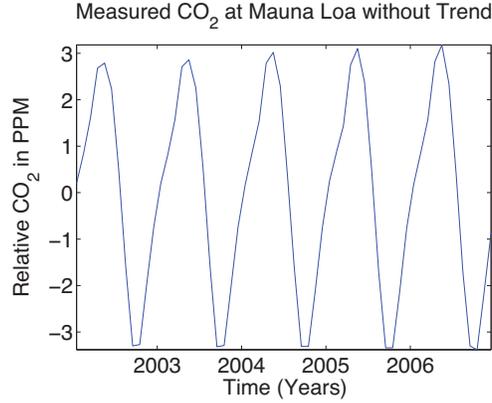
Figure 2: Observed carbon cycle at Mauna Loa with average trend removed.

Loa observatory provides us with the best representation of the carbon cycle for our model. We will revise the carbon cycle component of our model based upon this data.

NOAA provides data representative of the overall trend in the data. The average growth rate for the year is calculated as the difference between the average concentrations from November through February from those of the prior year [6]. To evaluate the natural carbon cycle without the effects of the other components, we have subtracted that trend from the data and plotted the results in Figure 2. By removing the effects of fossil fuels burning, we are able to model solely the natural carbon cycle, without human activity. If we were to include the effect of fossil fuel combustion, we would overstate its effect in the final combined model. As can be seen in Figure 2, the carbon concentration data takes a shape similar to simple harmonic oscillation. The maximum values occur in the winter months when there are fewer plants to absorb carbon dioxide. The minimum values occur during the summer months when more plants are absorbing carbon dioxide. We can model the behavior of a harmonic oscillator using a differential equation.

We will model this cycle using a second order differential equation of the following form

$$\frac{d^2 F_{bio}}{dt^2} + c_1 \frac{dF_{bio}}{dt} + c_2 F_{bio} = 0.$$

We will use a MATLAB differential equation solver to approximate a solution to the above equation. To expedite the process, this harmonic oscillation equation can be rewritten as the following system of first order differential equations:

$$\frac{dF_1}{dt} = F_2$$
$$\frac{dF_2}{dt} = -c_2 F_1 - c_1 F_2.$$

Here, $\frac{dF_1}{dt}$ represents the term $\frac{dF_{bio}}{dt}$ in the second order differential equation, and $\frac{dF_2}{dt}$ represents $\frac{d^2 F_{bio}}{dt^2}$.

The constant c_1 represents the damping term for the oscillation. Since seasonal plant growth could continue indefinitely, it is logical that the curve will not decay over time. Hence, we set $c_1 = 0$. Thus, the amplitude of the curve will remain constant in our model. To improve the model by optimal choice of c_2, we use the *fminsearch* method in MATLAB to minimize the least-squares residual functional. The function *fminsearch* is based on the well known Nelder-Meade process [4]. The optimization problem is then stated as follows

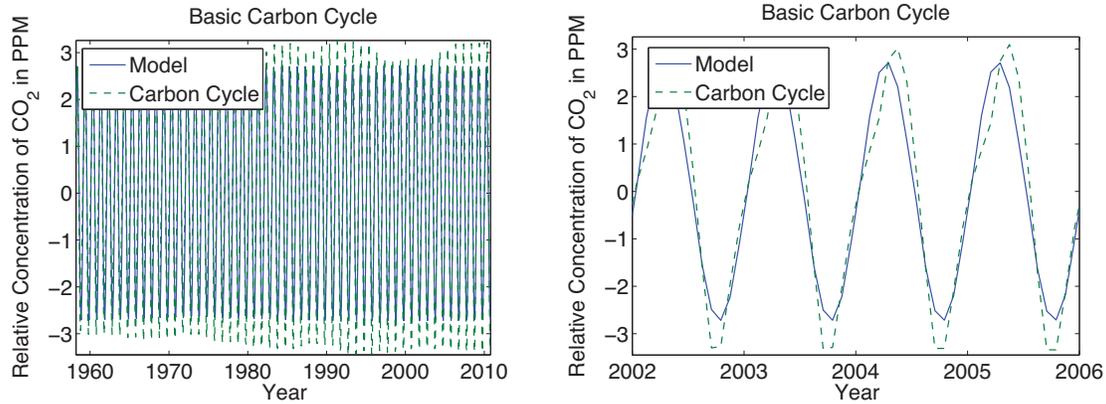$$\min_{c_2} \sum_{i=1}^{n} |F_{bio}(t_i, c_2) - \tilde{F}_{ml}(t_i, c_2)|^2,$$

Figure 3: **Left:** Model fit (-) to the observed carbon cycle data (–). **Right:** Model fit (-) to the observed carbon cycle data (–) from 2002-2006.

Approximation Order	Polynomial
Cubic	$y_3(t) = a_3t^3 + a_2t^2 + a_1t + a_0$
Quadratic	$y_2(t) = a_2t^2 + a_1t + a_0$
Linear	$y_1(t) = a_1t + a_0$

Table 1: Polynomials of order one, two and three used to approximate fossil fuel emissions.

where $\tilde{F}_{ml}(t)$ is the carbon dioxide data at time t_i and F_{bio} is the numerical solution to the differential equation computed via a standard MATLAB ordinary differential equation solver. The MATLAB ODE45 solver is based on the Runge-Kutta Method [5, 7]. The minimization of the residual provides us with the best estimate of the parameter c_2. Thus, we have the value of c_2 that will provide the least squared error in the model.

The optimized model approximation to the carbon cycle data is displayed in Figure 3. As can be seen in the figure, the model captures the frequency, amplitude, and phase shift of the oscillation. The model does not approximate the relative maxima and minima of each oscillation, though. The differential equation used provides for a constant amplitude in the sinusoidal curve. Thus, it will not describe as well those years with higher or lower CO_2 than typical. Since it does approximate the overall curve well, we can then implement this component in the final model.

Fossil Fuels

The data for carbon dioxide emissions from the burning of fossil fuels was compiled by the Carbon Dioxide Information Analysis Center (CDIAC)[1]. The mass of carbon dioxide emitted from 1800 to 2007 is given as an average emission for the year. This will force us to conclude the implementation of the model at 2007. We considered both polynomial and exponential models to estimate the trend in the CO_2 emissions from fossil fuel combustion. We tested the polynomial models of the forms found in Table 1 to determine which would most accurately model the data.

Since we have an overdetermined system of equations we will seek a least-squares solution to the system. To do so we build three matrices: the time data matrix $\mathbf{X}$, the vector $\vec{c}$ containing the unrealized polynomial coefficients a_i, and the collected emissions data $\vec{Y}$. For the quadratic fit, with $i = 1 \ldots n$ data measurements, $(t_i, y(t_i))$, the matrices would be defined as
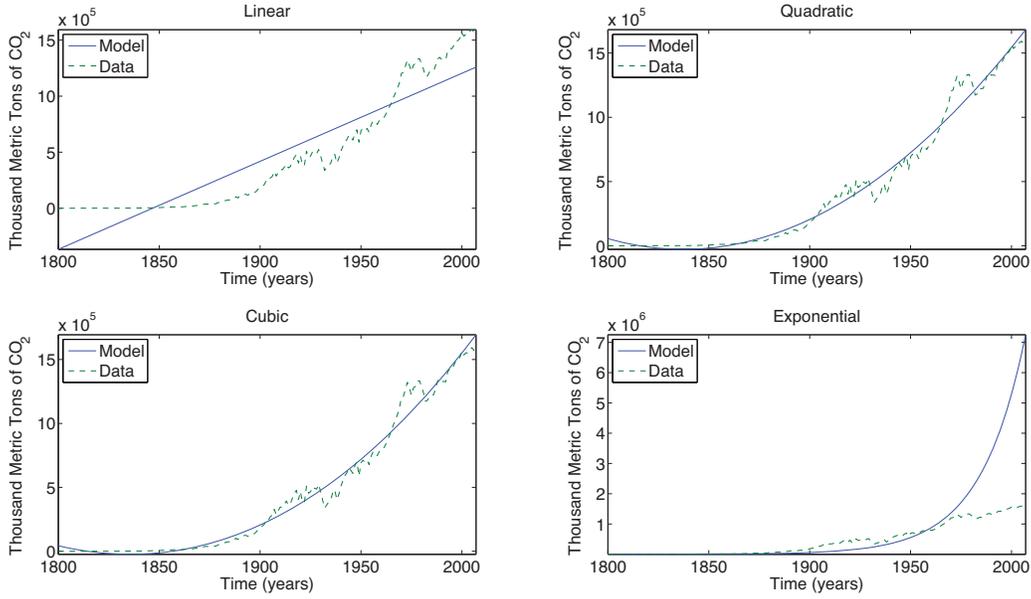
Figure 4: (Upper left to bottom right)-Linear, quadratic, cubic, and exponential fits (-) to fossil fuel data (–) [1].

$$\mathbf{X} = \left[\begin{array}{ccc} t_1^2 & t_1 & 1 \\ \vdots & \ddots & \vdots \\ t_n^2 & t_n & 1 \end{array} \right] \quad c = \left[\begin{array}{c} a_2 \\ a_1 \\ a_0 \end{array} \right] \quad Y = \left[\begin{array}{c} y(t_1) \\ \vdots \\ y(t_n) \end{array} \right]$$

The system created from the matrix $\mathbf{X}$ and the vectors $\vec{c}$ and $\vec{Y}$ is given as

$$\mathbf{X}\vec{c} = \vec{Y}$$

We computationally implement this operation via the $\vec{c} = \mathbf{X}\backslash\vec{Y}$ command in MATLAB. This provides us with coefficient approximates that yield the minimized least-squares residual of the model versus the data [5]. We employ this process for linear, quadratic, and cubic polynomial fits to the data. In essence, we used the data to find a solution to the system of equations. A nonlinear estimation technique (*fminsearch*) was used to determine parameters of an exponential model $y_{exp}(t) = Ce^{\alpha t}$ that could minimize the least-squares data versus model residual. The method is similar to that used to model the carbon cycle above. The fossil fuel data takes a similar form to that of an exponential curve, with continual growth since the early 1900s. Thus, we do consider an exponential fit to the data as well. The optimized results of the linear, quadratic, cubic, and exponential fits are plotted alongside the data in Figure 4.

Through a calculation of the sum of the squared residuals similar to the minimization seen above in the carbon cycle component, we determine that the cubic model provides the best fit to the data. Even visually, the model provides a good representation of the overall trend in the data. However, the cubic model creates the possibility of nonphysical model growth beyond the recorded data. It is then possible that, beyond 2007, the cubic model (as well as the quadratic and exponential models) would not provide an accurate representation of fossil fuel emissions, indicating an increase in carbon emissions higher than reasonably possible. Consequently, we employ the linear model. The linear model unfortunately does oversimplify the
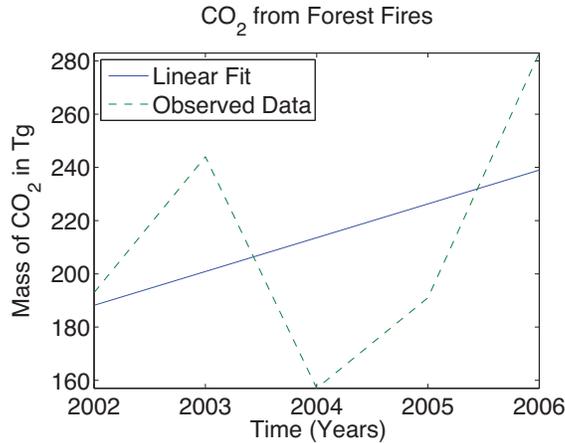
Figure 5: Data (–) and model (-) of forest fire CO2 emissions from 2002-2006 [10].

trend, for the entire data set. However, as indicated in the subsequent section, we are forced to limit our model to a smaller time period, and the linear fit does provide a better approximation for the shorter time period than it does for the 200 year period shown in Figure 4.

Forest Fires

There are limited data sets for carbon dioxide released from forest fires in the United States. Further, since forest fires can be a natural phenomenon or man-made, the data was recorded at irregular time intervals at best, making information approximation difficult. There is randomness to the occurrence of forest fires, both in frequency of occurrence and in intensity, each of which will influence the total mass of CO_2 released from a given fire. The size of trees and density of the forested area affected by the fire would also increase the unpredictability associate with forest fire emissions[10]. This means that a recoverable trend is unlikely to be present. The suboptimal data employed limits our ability to model the behavior accurately and we consequently choose a simple linear regression to the data displayed in Figure 5. We determine the linear fit to the data using the same method described in the fossil fuel section. Additionally, the forest fire data obtained was only recorded from 2002 through 2006, limiting our model to a five year time period.

Oceanic Effect

The oceans act as a carbon sink, reducing the total amount of carbon dioxide in the atmosphere [8]. The absorption amount is proportional to the relative atmospheric concentration of carbon dioxide. The oceanic effect would then follow a similar curve to that of the natural carbon cycle. Thus, as the CO_2 concentration in the atmosphere increases, the total amount of CO_2 absorbed into the oceans will increase as well [8]. This component should be particularly important to the concentration observed at the Mauna Loa observatory, due to the proximity to the Pacific Ocean. We calculate the absorption using Henry's Law which states that $p = kc$ where p is the partial pressure of carbon dioxide in atmospheres, c is the concentration of the gas in the liquid in molarity, and k is a constant for the specific gas. Here $k = 29.41 \frac{L*atm}{mole}$ [2]. Thus, the absorption of CO_2 into the ocean is dependent upon the total concentration of CO_2 in the atmosphere. Then, the decrease in the total atmospheric carbon dioxide concentration will be relative to the increase in carbon dioxide from the fossil fuel and forest fire factors. We assume that there is no upper limit to the absorption within the range of CO_2 concentrations considered in our model. We will calculate the total reduction based on the sum of the other three components. We have not plotted the results of the oceanic effect individually,
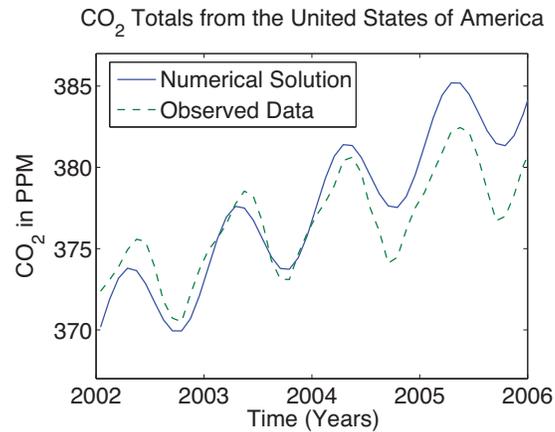
Figure 6: Optimized compartmental CO2 concentration model (-) in comparison to CO2 data (–).

as the effect is entirely dependent on the other components. Instead, this component will only be shown in the final model.

Model Results

Combining each component in Figure 6, we display the results of the aggregate model of the carbon dioxide concentration. The numerical solution plotted is the sum of the differential equation seen in Figure 3, the linear fit plotted in Figure 4, and the linear fit plotted in Figure 5 less the oceanic reduction given as a percentage of the total concentration from the other three components. The model is only compared to the data from the years 2002-2006 due to the limited available forest fire data. If more extensive forest fire data were available then the model timespan could be extended.

Conclusion

The model does quantify an increase in the total concentration of carbon dioxide in the atmosphere. Based on the relative size of each component, the largest addition to the natural carbon cycle comes from the burning of fossil fuels. The primary sources of error in the computation of the model may involve unit conversions–specifically converting parts per million to mass and molarity. Both calculations were based on crude estimates of the mass of the atmosphere and the volume of the ocean [3]. More accurate forest fire data and parameter estimations (in particular to λ_r) are required to obtain a better model characterization of the CO2 concentration over the United States.

Acknowledgment

References

[1] T. BODEN, G. MARLAND, & B. ANDRES, National CO_2 Emissions from Fossil-Fuel Burning, Cement Manufacture, and Gas Flaring: 1751-2008, cdiac.gov, (2011).

[2] BROWN, LEMAY & BURSTEN, Chemistry: The Central Science, New York, Pearson, (2007).

[3] "Carbon Dioxide Information Analysis Center - Conversion Tables," National Academy of Science, cdiac.gov, n.d.

[4] fminsearch mathworks.com Mathworks, n.d.

[5] F. GIORDANO ET AL, Mathematical Modeling, Belmont, CA, Cengage Learning, (2009).

[6] Mauna Loa Observatory, noaa.gov, NOAA, n.d.

[7] ode23, ode45, ode113, ode15s, ode23s, ode23t, ode23tb, mathworks.com, Mathworks, n.d.

[8] W. PETERS, A. JACOBSON, C. SWEENEY, ET AL, An Atmospheric Perspective on North American Carbon Dioxide Exchange: CarbonTracker, National Academy of Science, (2010), 18925-18930.

[9] Trinidad Head Observatory, noaa.gov, NOAA, n.d.

[10] C. WIEDINMYER & J. NEFF, Estimates of CO2 from fires in the United States: implications for carbon management, cbmjournal.com. Carbon Balance and Management Journal, (2007).