

# ATMOSPHERIC CARBON DIOXIDE MODELING AND THE CURVE FITTING TOOLBOX

## DETERMINING THE NONLINEAR MODEL UNDERLYING CARBON DIOXIDE LEVELS IN THE ATMOSPHERE

BY MARY ANN BRANCH  
AND RUSS MINKWITZ

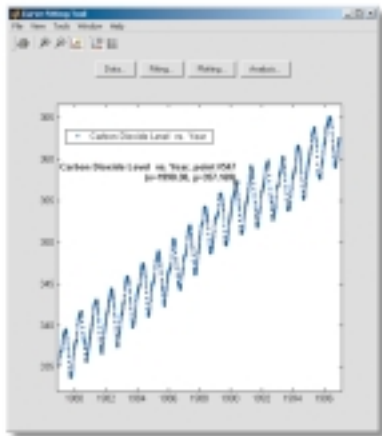
A common measurement for studying the effect of burning fossil fuels on weather patterns is the level of carbon dioxide (CO<sub>2</sub>) concentration in the atmosphere. Some scientists believe the upward trend in CO<sub>2</sub> levels could cause atmospheric temperatures to rise, polar ice caps to melt, and climates of different regions of the earth to change radically.

Since this data has cyclical phenomena, as well as an upward growth trend, nonlinear fitting techniques are appropriate, but can require added effort when compared to fitting a linear model. However, with the right tools and knowledge of your data, you can quickly determine the coefficients of a nonlinear model.

This article describes how to fit a nonlinear model to available CO<sub>2</sub> data using the newly released Curve Fitting Toolbox.

### Modeling the Data

In this example, we use CO<sub>2</sub> data, in parts per million, collected by the U.S. National Oceanic and Atmospheric Administration (NOAA) from 1979 to 1996 at the Mauna Loa Observatory in Hawaii<sup>1</sup>. The captured data is shown in the following illustration.



The Curve Fitting Tool found in the Curve Fitting Toolbox gives an immediate visual display of the data. By using the Data Tips feature we can interrogate specific points.

Cyclical phenomena in nature can be modeled by a sinusoid and its higher harmonics, that is,

$$y(t) = \sum_{k=1}^n A_k \sin\left(\frac{2\pi k}{T}(t + \theta_k)\right)$$

$A_k$ ,  $T/k$  and  $\theta_k$  are the amplitude, period, and phase, respectively, of the  $k$ -th harmonic. The sinusoidal terms capture seasonal trends, but the persistent upward trend requires an exponential term. A suggested model for this data, determined by examining records of total world fossil fuel production from the mid 1860s, is

$$y(t) = B + d \exp \alpha t + \sum_{k=1}^n A_k \sin\left(\frac{2\pi k}{T}(t + \theta_k)\right)$$

where  $t$  (time) is in years,  $\alpha$  is believed to be approximately 0.0444/year, and the other parameters are unknown (including the number of sinusoids needed)<sup>2</sup>.

To build the model, we go to the Curve Fitting Toolbox's Fitting GUI and create a custom equation with one sinusoid term.

$$y(t) = B + d \exp \alpha t + A_1 \sin\left(\frac{2\pi}{T}(t + \theta_1)\right)$$

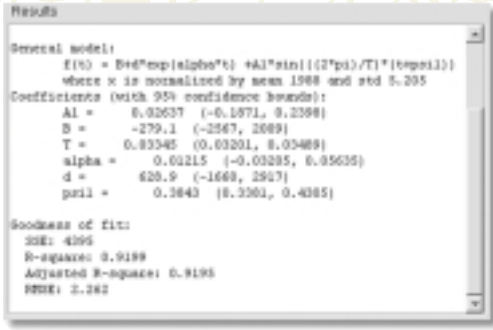
Because the time values are large, we can select the option to center and scale the time variable (remove the mean and divide by the standard deviation) to improve the conditioning of the numerical calculations.

With the Custom Equation panel, we enter start values of  $\alpha = .0444$ ,  $T = 0.192$  (the period of one year after centering and scaling), and random guesses for the rest of the parameters (below). When the fit completes and the result is plotted we see that the cyclical pattern has not been matched (left).

Unknowns	StartPoint	Lower	Upper
A1	0.895	-inf	inf
B	0.942	-inf	inf
T	0.192	-inf	inf
alpha	4.44e-02	-inf	inf
d	0.335	-inf	inf
phi	0.437	-inf	inf

Equation name: B + d\*exp(alpha\*t) + A1\*

$$y(t) = B + d \exp \alpha t + \sum_{k=1}^n A_k \sin \frac{2\pi k}{T} (t + \phi_k)$$



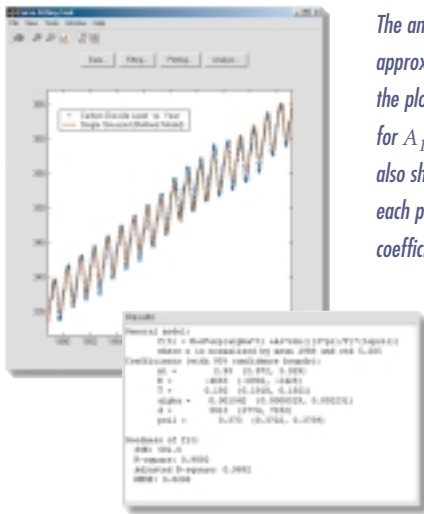
The results displayed indicate that the R-square value (a measure of goodness-of-fit) is not too far from one, implying the global trend is met, but the local trend is not. In this case, the amplitude of the sinusoid term is too small.

## Refining the Model

We can improve upon this model in several ways. First, we can determine the coefficients of the exponential term separately. We use the toolbox's built-in exponential model which automatically calculates optimal start points from the data,

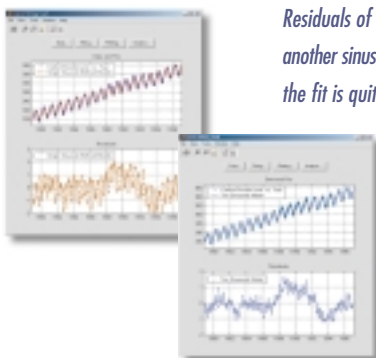
$$y(t) = d \exp \alpha t$$

Fitting this model results in coefficients of  $d = 349.6$  and  $\alpha = 0.02185$ . Using these values of  $d$  and  $\alpha$  as start values in the sinusoidal model, we get a new fit that matches the cyclical nature well.



The amplitude of this latest model is approximately three. We can verify on the plot that this is a reasonable value for  $A_1$  (left). The Results panel (below) also shows the confidence intervals for each parameter in parentheses after the coefficient value. Smaller confidence intervals imply more confidence in the parameter values. The confidence intervals are small for all the parameters except for  $B$  and  $d$ , this may indicate that  $B$  and  $d$  are highly correlated.

Now we can plot the residuals that still show a cyclical sinusoid pattern, so another sinusoid term is warranted.



Residuals of the single sinusoid model reveal yet another sinusoid term (left). With two harmonics, the fit is quite good (right). We can easily create equations with even more terms, examining the residuals of each resulting fit until we are satisfied.

## Extending the Analysis

When we have found a satisfactory model, we can send our fits to the MATLAB workspace for further analysis. We have several other observatory sites to examine (e.g., Samoa, the South Pole, and Barrow) and so can generate an M-file function that can recreate our fits. Then we can call the function in a loop to automatically fit other data sets to see how the resulting models compare across observatory sites.

This example touches on a few of the features that will help you fit models to data. Whether the models are linear or nonlinear, the Curve Fitting Toolbox enables you to quickly fit standard models from its library or custom models you create.

```

54
55 % --- Create fit "Single Sinusoid Model"
56 fo = fitoptions('method','NonlinearLeastSquares','Normalize','on');
57 st_ = [0.8952 0.9424 0.192 0.8444 0.3351 0.4374];
58 sec(fo,'Startpoint',st_);
59 ft_ = fittype('B+d*exp(alpha*t) + A1*sin((2*pi)/T)*(t+phi1)') ...
60 'dependent',{'Y'}, 'independent',{'t'}, ...
61 'coefficients',{'A1', 'B', 'T', 'alpha', 'd', 'phi1'};
62
63 % Fit this model using new data
64 cf = fit(time,co2,ft_ ,fo_);
  
```

M-file code to replicate the fits and plots created in the GUI can be generated from the GUI File menu. The M-file also reveals the toolbox commands that can be used at the command line to get the same results.

<sup>1</sup> GLOBALVIEW-CO2, "Cooperative Atmospheric Data Integration Project - Carbon Dioxide" (2001). [Available to the public at ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW]

<sup>2</sup> Kahner, D., Moler, C., Nash, S., Numerical Methods and Software, Prentice Hall (1989)

To learn more about the Curve Fitting Toolbox, visit: [www.mathworks.com/nncurvefitting](http://www.mathworks.com/nncurvefitting)