MONT 100N – Modeling the Environment
Chapter 4 Project – The Mauna Loa Atmospheric CO2 Data Set.
October 18, 2017

The Mauna Loa Observatory (located at an elevation of about 3400 meters on the Mauna Loa volcanic mountain on the "big island" of Hawaii) is a research station maintained and staffed by the National Oceanic and Atmospheric Administration (NOAA), the major agency of the US government that studies weather and climate phenomena. NOAA maintains a web site, for instance, with the National Weather Service day-to-day forecasts and severe weather warnings that form the basis for most local weather reporting in the media. Among the data collected regularly at Mauna Loa are measurements of atmospheric concentrations of a host of trace gasses, including carbon dioxide, or CO2. The data set of those measurements goes back to 1958 and is one of the most complete records of the recent evolution of this aspect of the Earth's atmosphere. In this project we will apply the modeling techniques introduced in this chapter to try to understand what this data set is saying about changes in atmospheric CO2 over time.

Important Note

This is a well-known data set and you can find all sorts of discussions of various aspects of it on the web, if you look. I am going to ask all of you *not to look* until after you have worked through at least questions A through D below, though. The idea is for you to approach this entirely "fresh" and make your own observations and analysis and draw your own conclusions.

Getting Started

The data we will be looking at is contained in a (large) Excel spreadsheet called `MaunaLoaCO2Data.xls` that you will download from the course homepage. Begin by getting the spreadsheet and opening it in Excel. Note the layout:

- Column A gives the year the measurement was taken
- Column B gives the month (1 = January, through 12 = December)
- Column C gives a decimal equivalent of the middle of the month, so for instance January 1958 is given as 1958.04, since 1 month = 1/12 year = .08 year (roughly), and .04 year is about 15 days.
- Column D gives the average CO2 level observed at Mauna Loa that month in units of parts per million
- In Column D, if you look closely, you will see that a few of the entries near the start are -99.99. What do you suppose that means?
- In Column E, you will notice that most of the entries are the same as the corresponding entries in Column D, but the -99.99 entries have been replaced by other values. These are "interpolated" (estimated) values based on the trends from the nearby months. We will use Column E for all our values so that the -99.99's are not included.

Questions

The first thing you will notice if you look at the CO2 levels is that there is *a lot* of up-and-down variation. Is it completely random, though? And is there an underlying trend?

(A) To start to answer this question:

(1) Create scatter plots of the CO2 monthly averages for the calendar years 1965, 1975, 1985 (individually), versus the decimal year from Column C. This will require picking out the correct range of rows in Columns C and E for each of these years, and you may want to copy those values to other cells to create the scatter plots.

(2) Looking at these scatter plots, what do you notice about the way CO2 levels vary over these years? Describe what happens over the course of a typical year, and hypothesize a reason why the annual pattern works this way. Note: Mauna Loa is in the Northern Hemisphere and typical mixing patterns in the atmosphere mean that most of the air that passes over this location has come from other areas in the Northern Hemisphere. What happens through the months of May, June, July, August in the Northern Hemisphere, and how might that affect atmospheric CO2 levels?

(3) *Extra Credit*: How might you model the yearly variation of the CO2 readings? Suggest mathematical function(s) that might be useful and how you might apply them.

(B) Condensing the Data to a More Manageable Form. Our goal is to model how atmospheric CO2 levels have been changing over this period (but on the year-to-year level, not on the much more variable month-to-month level). This will be much more manageable if we identify some way to compute a "summary value" for each year to use as the representative CO2 level for that year.

(1) Identify (at least) three different ways that might be used to produce that sort of "summary value" and describe why they would be suitable.

(2) Choose one of your proposed ways to do this and give a reason for why you think that will be a reasonable way to "condense" the data for each year.

(3) Create new columns in your spreadsheet giving the number of years since 1959, and your summary CO2 value for the year. Since we don't have complete values for the years 1958 or 2017, just use the years 1959 to 2016 (58 years in all).

(C) "Let the modeling begin!"

(1) Using Excel, fit a linear model to your "condensed" data set and record your results. Give the equation of the regression line as a function of the years since 1959 (that is, $x = 0$ correspond to 1959, $x = 1$ corresponds to 1960, and so forth). Also give the R^2 value reported by Excel as a measure of goodness of the fit, and discuss the residuals for the linear model (in particular, is there a consistent pattern there)?

(2) What does your model predict concerning the CO2 level in 2020? (This is slightly outside the interval 1959 to 2016 of course, but not too far outside. So extrapolation from the linear model is at least a possibility!)

(D) Atmospheric CO2 levels are of concern, of course, because of the "greenhouse gas" properties of this compound–the way atmospheric CO2 can trap energy from reflected solar radiation and increase temperatures near the surface of the Earth. Some greenhouse effect is necessary for life on Earth, of course (our water-based form of life could not exist at the temperatures that would prevail with no greenhouse effect at all because all water would be frozen as ice). But have there been times in the past when CO2 concentrations were significantly higher than they are now? What were the Earth's climate and sea levels like then? (This may require some research–be sure to give the sources you used to compile your information.)

Assignment: Submit your edited Excel spreadsheet with the data and write up answers to the questions above in a separate document.