Numerical analysis involves the practical use of mathematical calculations. Much like the Babylonian approximation of \( \sqrt{2} \) (which turned out to have tremendous practical applications), modern numerical analysis does not seek exact answers, because exact answers are often impossible to obtain in practice. The current “arguments” revolving around “Climate Science” are a great example of this. The naysayers tend to focus on the existence of uncertainty as a reason to doubt everything. In fact, the constraints on the uncertainty are precisely why we should be paying great heed! Numerical analysis is concerned with obtaining approximate solutions while maintaining reasonable – as accurately as possible – bounds on errors.

Numerical analysis naturally finds applications throughout engineering and the physical sciences. The current century finds that life sciences, social sciences, and the arts have adopted various elements of numerical analysis and statistics as fundamental to their practice.

The overall goal of the field of numerical analysis is the design and quantitative study of techniques to give approximate but accurate solutions to hard problems, the variety of which is suggested by the following:

- Advanced numerical methods are essential in making highly accurate numerical weather prediction commonplace. (Consider the availability and specificity of the upcoming winter and seasonal forecasts compared to when you were in grade school.)

- Computing the trajectory of a spacecraft requires the accurate numerical solution of a system of ordinary differential equations. Recently, a rocket was not just launched but landed in vertical position. Others have been retrieved at sea on ships.

- Car companies can improve the crash safety and autonomous operations of their vehicles by using computer simulations. Such simulations essentially consist of solving partial differential equations numerically.

- Hedge funds (private investment funds) use tools from all fields of numerical analysis to attempt to calculate the value of stocks and derivatives more precisely than other market participants. (Google “Voodoo” economics sometime!)

- Airlines use sophisticated optimization algorithms to decide ticket prices, airplane and crew assignments and fuel needs. Historically, such algorithms were developed within the overlapping field of operations research.
Insurance companies use numerical programs for actuarial analysis.

Purpose

The purpose of this assignment is to provide you with some experience exploring and analyzing data numerically, i.e. without using an information visualization system. You will be provided a data set (that can be imported into Excel) about the ambient atmosphere. You should explore and analyze this data using Excel or any other graphing or visualization tools with which you are familiar (e.g. Mathematica, MATLAB, S-Plus). Your goal here is to perform an exploratory analysis of the data set, to better understand the data set and its characteristics, and to develop insights about the environmental data.

Report

The report that you turn in should consist of five things.

1. First, list (bullet list of items) five "insights", chunks of knowledge, or deeper questions that you either encountered or gained while exploring the data. An insight could be some understanding of the data and its characteristics that is not relatively obvious or intuitive. It is something that most people might not realize initially. Note that an insight or knowledge chunk simply may be a deeper question that arose in your mind while exploring the data and your cursory analysis may not have been sufficient to answer the question.

2. Plot the time series of the key environmental variable(s) supplied to you by the instructor. Based on your visual inspection, describe the temporal behaviors of each parameter and their apparent relationships to each other.

3. Perform and report on the outcomes of the following statistical techniques

   a. Determine the median and geometric means of each environmental parameter for all daytime hours, all nighttime hours, for the average day (24 hours) and over the entire observational period. Compute and report the average deviation and variance for each statistical measure.

   b. Pick out the two groups (e.g. daytime, nighttime, full day, full observational period, or another descriptive measure of the data set) whose means are closest to each other. Do a two-sample t-test on the two groups you picked, using the Excel function described on the t-test page. Report the P-value and propose a plausible explanation what it means.

   c. Do a two-sample t-test on the two groups in your data set with means that are the furthest apart. Report the P-value and propose a plausible explanation for what it means.

   d. Perform a Pearson Correlation for all pairs of time-dependent variables from the data set. Report the R-values and propose a plausible explanation for what it indicates.

4. Attempt to “fit” the data with a mathematical function (e.g. linear, 2nd-degree polynomial, 3rd-degree polynomial). Which functions worked the best and for which parameters? What do you think is the purpose of such an exercise? What advice would you give someone who needs to fit a similar data set?

5. Write two paragraphs about challenges or problems that you encountered in doing the analysis this way. Did anything limit or frustrate you? If nothing did, perhaps there was something that was more difficult than you thought it should be. Nothing is perfect, so you should be able to list some potential issues here.
Download Data for Analysis Here

3_AirConc_200507.xls