

UNST 232 Mentor Section Assignment 5

Historical Climate Data

1 introduction

Informally, we can define climate as the typical weather experienced in a particular region. More rigorously, it is the statistical description, including mean and variation around that mean, of relevant variables over a specified period of time. The standard period is 30 years, defined by the World Meteorological Organization. The years 1951 to 1980 are used to define the mean in most analyses you will see that deal with global warming and climate change. Some common climate variables are temperature, precipitation, humidity, and wind measured at Earth's surface.

If you follow the popular dialog about global warming, you have probably heard statements about the significance, or lack of significance, of any given year being the hottest on record, not not as hot as the prior year, and so on. Sometimes temperature records from individual locations are used to make whatever particular point a speaker wishes to make. Given what we have discussed in lecture about the nature of climate and variability in the climate system, does this make much sense?

In this exercise, we will examine temperature and precipitation at two locations in the United States and as well as global mean annual temperature. You will be asked to use data arranged for you in an Excel workbook to perform some simple calculations and make graphs of those quantities over time and to make interpretations in terms of regional drought.

At the end of the document you will find numbered questions that are intended to assist you in studying these data sets. You should write answers to the questions using complete sentences and short paragraphs. Where the question asks for a plot from your data analysis, print the graph, number it, and refer to that number in your answer. You will have three mentor sessions to work on this project and the completed set of answers is due in the lecture section on 23 February.

2 the data

Climate records of one sort or another have probably been kept as long as people have been farming, perhaps longer. Such information forms the knowledge base upon which decisions about planting and harvesting can be made. “Modern,” instrumental records have been kept since the middle of the 17th century, although very few are continuous before the 18th century. By the start of the 19th century, reliable and continuous records were being kept in Europe and East Asia. More widely available and reliable records pick up the mid 19th century (North, 2006, Chapter 3). In North America, reliable climate records begin in the late 19th century.

When you open the Excel workbook, you will see four tabbed worksheet pages. The first page, labelled *combined data sets* is the one you will use in this project. The remaining three contain the original data from which the first page is composed, as indicated by the tab labels. Column headings at the top of each sheet indicate the data type and measurement units.

It is important in this analysis to remember the difference between climate and weather, as well as the spatial variability in climate that we have discussed in lecture. It is often useful to compare the trends you discover at a particular site to regional or global trends.

2.1 station data

Mean monthly and seasonal temperature records for McMinnville, Oregon (45.2 N latitude, 123.1 W longitude) and Wichita, Kansas (37.6 N latitude, 97.4 W longitude), have been assembled for you into an Excel workbook, using the archives maintained by NASA’s Goddard Institute for Space Studies. You can download the workbook at the class website. Both of these long-standing records are collected at regional airports. You can read more about the stations at <http://weather.gladstonefamily.net/site/KMMV> and <http://weather.gladstonefamily.net/site/KICT>.

Seasonal values were calculated: D-J-F (December-January-February) for the northern hemisphere winter season; M-A-M (March-April-May) for spring; J-J-A (June-July-August) for summer; and S-O-N (September-October-November) for fall. You will notice that some of the spreadsheet cells are blank. This indicates that no data is available for that observation interval. Record keeping can be variable for a number of reasons. Maybe the thermometer broke and there were insufficient funds to replace it or perhaps the pages on which the data were written have been lost.

2.2 global scale surface temperature analysis

The NASA Institute for Space Studies makes all the data they use in their global temperature anomaly analyses available online in a variety of formats. The homepage for everything they produce is <http://data.giss.nasa.gov/gistemp/>. Many details about their long-standing temperature analysis program can be found at the website. The data we will use are global, northern hemisphere, and tropical mean temperature anomalies computed using both land surface and ocean surface temperature measurements. The combination of land and ocean surface observations produces a slightly different anomaly than would land surface alone due to the large heat capacity of water. The anomaly is calculated relative to a base period from 1951 to 1980.

3 trend estimation

Trend estimation is a statistical method often used as a first step in studying the relationship between two variables. In our case, we are interested in the relationship between time (years in our observation record) and air temperature. Trend analysis tells us *if* a relationship exists but it doesn't tell us *why* it exists. For that, we need to know something about the physical processes involved in the system we are studying (in our case, Earth's climate).

Trend analysis is performed on data sets for which we expect a meaningful relationship because we understand at least a little bit about the system they represent. For our analysis to be useful though, we must be prepared for the case when no relationship exists, that is, the data are random. A random quantity may change in an unpredictable way every time it is measured. There should be no discernible relationship, no trend, in a series of measurements of a random variable. That may not be true for a small sample size and so statisticians have formal mathematical rules for determining the sample size required for meaningful analyses and the confidence that may be placed on the result.

In this exercise you are asked to examine the relationship between mean annual temperature and time in the temperature time series data. The Excel Trend Line tool computes this by finding the best fit of a line through the data. That tool allows you to fit a straight line (a line with constant slope) through the data or something more complicated. Will use the Linear Regression tool to produce a line with constant slope. The *best fit* is one that minimizes the differences between the actual data points and the trend line for all measurements in the time series.

The equation of a line made of of points (x, y) is

$$y = mx + b \tag{1}$$

where m is the slope of the line and b tells you where the line intersects the y -axis. The slope is the change in y over some interval of x . The goal in linear regression is to find a single equation of a line that best represents all of the data in our series. For all of the x in the data set, we want to find the combination of slope m and y -intercept b that gives the best prediction of the real values of y . This is accomplished using a mathematical formula that minimizes the differences between the real and predicted y 's for all of the x 's. The analysis yields a measure of the goodness of the fit often called the *R-squared* value.

In our example, the x 's are years and the y 's are temperature anomalies. The slope m is thus the rate of temperature change over the entire interval from the first to the last year in the record. The R-squared value tells us something about how much variation year-to-year there is around the linear trend. As you will discover by inspecting the data plots you produce in this project, a straight line is considerably more simple than the real time series. However, the linear best fit does provide a useful summary of the data.

4 Questions

Note: Figures you produce should have labels for all three axes, a title, a legend, and should be easy to read.

1. Which are the warmest and coolest years on record in McMinnville, Oregon and Wichita, Kansas?
2. It seems that every year we hear and read news reports about the prior year having been the warmest on record or comparisons between the current decade and some other, earlier, decade. In and of themselves, these pronouncements are difficult to interpret because they lack context. We don't know how individual years fit into a longer context or if what is happening right now in our region mirrors the the global trend. This is important because, as we have discussed in lecture, both interannual and spatial variability are important elements of Earth's climate.

In this question, you are asked to compare mean decadal temperatures in McMinnville and Wichita. The easiest way to do this is to use Excel's AVERAGE function. Calculate decadal mean winter and decadal mean annual temperatures for the two weather station sites as follows:

Begin by calculating the summer (J-J-A) mean of the years 1891 to 1900 for the McMinnville record:

```
=AVERAGE(D14:D23)
```

Cell G23 would be a good place to store the result of this calculation. Next, calculate the annual mean for the same years and location:

```
=AVERAGE(F14:F23)
```

Cell H23 would be a good place to store the result of this calculation.

Now, copy and past the two formulas in cells G23 and H23 to repeat the calculations for the next decade. The formula copied from G23 would be pasted into G35 and the formula copied from H23 would be pasted into H33. Do you see the pattern? It will be easy to proceed from here to compute the decade means for the entire McMinnville data set and to repeat the operation for the Wichita data set. Please do so and then answer the following questions:

- (a) Are the most recent two decades (use the 2000's and the 1990's) the warmest on record in McMinnville, OR and Wichita, KS? What data did you compare to arrive at this answer?
 - (b) Do the warmest decades also have the warmest summers? Explain how you arrived at your answer. One way would be to compare each decadal mean with the appropriate means for the whole data set.
 - (c) If you had been asked to make a future prediction based on the comparison between the 1930's and 1970's, what would you have predicted? *It's a good thing this isn't how climatologists make predictions!*
3. The analysis in question 2 is simple but a bit awkward. A better approach is to consider temperature anomalies, as we have done in lecture. In this question, you are asked to compute annual mean temperature anomaly records for Wichita, KS and McMinnville, OR and to plot the anomaly series (see Figure 1 for an example). The steps provided here should guide you through the process.

- First, calculate the reference mean annual average temperature from 1951 to 1980. The Excel formula for the McMinnville, OR mean annual average is:

```
=AVERAGE(F11:F133)
```

- Next, calculate yearly anomalies by subtracting the reference mean from each individual year's annual average temperature. If you stored the reference mean in cell F135, the Excel formula to calculate the anomaly for the year 1888 would look like this:

=F11-\$F\$135

You can copy and paste this formula into cells for the remaining time series. The \$ symbols establish fixed references to column F and row 135. This ensures that you always subtract the value stored in that cell. If you stored that value in a different cell than T2, you will need to adjust the formula accordingly.

- The best way to keep track of the time ordering of your anomaly series is to store the results of the calculations in cells that stay in line with years of the original data. The formula =F11-\$F\$135 could be typed into cell I11, for example, and then copied into cells I12 to I133.
 - Repeat the procedure to calculate an anomaly time series for the Wichita, KS data set.
 - Use the anomaly time series you just calculated to make temperature anomaly plots for both McMinnville and Whichita. Use *Straight Marked Scatter* in the (X, Y) Scatter Plot Chart type, found in the Insert Chart tool.
 - Select the data series you want to plot. If you have followed the directions thus far, the McMinnville series are I11:I133. These are the ordinate data (“Y-series data”) on your plot.
 - Click the *Straight Marked Scatter* button to begin. Once you have a basic plot made, you will need to select the abscissa data (“X-series”) and make a few other changes to make the plot as easy to read as possible.
 - Define the *X-values* using the *Source Data...* option in the *Chart* pull-down menu. The data are the years corresponding to the mean annual temperature anomalies, A11:A133 for McMinnville. You can also name the series using this tool.
 - Plot attributes can be formatted by using the mouse to select the feature you want to change. Either double click the feature (for example, the vertical axis) to bring up the Format tool or click once and select the feature in the *Format* pull-down menu.
 - Use *Chart Options* to create a title and axis labels.
 - Take care in how you handle years for which the mean annual temperature is missing from the record.
- (a) Please print and turn in your anomaly plot for McMinnville, OR. You can include the figure in the main body of your text or print it separately.
- (b) Compare the temperature anomalies since the 1950’s in the two towns. Do you see the same trend in both locations? If so, what is it? If not, how do they differ?

4. Plot the global, northern hemisphere, and tropical mean temperature anomaly time series (columns V, AA, and AF) on one chart and use the chart to answer the questions below.
 - Start by creating a *Straight Marked Scatter* plot with the global time series, following the procedure you used to create the McMinnville plot.
 - Next, use the *Add Data* option in the *Chart* pull-down menu to add the northern hemisphere and tropical time series to the plot.
 - (a) Of the three, which record has the largest interannual (year-to-year) temperature range? Why do you think this is the case?
 - (b) Of the three, which record has the smallest interannual range? Why do you think this is the case?
5. Plot the 5-year running mean temperature anomalies for the global, northern hemisphere, and tropical time series (columns W, AB, and AG) and answer the questions below.

Two of the questions require you compute a linear trend through the data.

- Select the data series of interest by clicking on it with your mouse.
 - Select *Add Trendline...* in the *Chart* pull-down menu.
 - In *Type* choose *Linear* and in the *Options* check *Display equation* and *Display R-squared value*.
 - Excel will probably place the resulting text box inside the plot axes, making it difficult to read. Move the text box if necessary.
- (a) Which of the three time series shows the largest change over the entire period of record?
 - (b) Why might the record you identified in part (a) have experienced larger total change over this time period than the others?
 - (c) What is the mean rate of global temperature change over the period of record? Make sure to include units in your answer.
 - (d) What is the rate of tropical temperature change over the period of record? Why is the R-squared value different for the tropical data set than it was for the global data set?
6. Compare the Oregon and Kansas mean annual temperature anomaly time series with the global anomaly time series.

- (a) Are the magnitudes of the interannual variations in the various time series different? Why do you suppose that is the case?
- (b) During the 1930's, the U.S. great plains region experienced a sequence of severe drought events (the “Dust Bowl” era) that caused the regional farm economy to collapse, a forcing that affected the entire country both economically and socially. Were those years anomalously warm everywhere? Explain how you arrived at your answer.

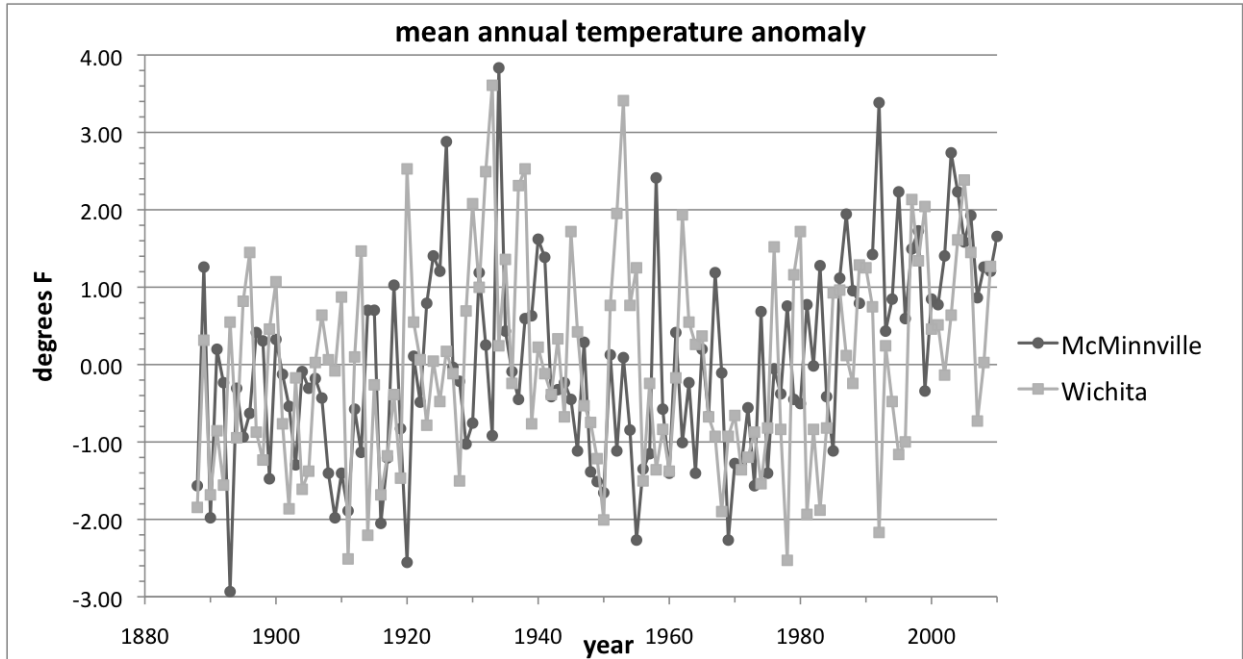


Figure 1: Example figure with anomalies for both weather stations.

5 references

North, G.R. et al, 2006, Surface Temperature Reconstructions for the Last 2,000 Years, G.R. North, chair: Committee on Surface Temperature Reconstructions for the Last 2,000 Years, National Research Council, The National Academies Press, Washington, DC.

NASA GISTEMP data archive, <http://data.giss.nasa.gov/gistemp>. Accessed 31 December 2010.