

Phenological trends and climate change in Minnesota

After you complete this activity, you should be able to:

- formulate an answerable question from a given data set
- use basic Excel skills to locate and sort desired data
- analyze phenological data to determine trends
- run a regression analysis on a phenophase and temperature
- determine if, and if so how, climate change has affected your chosen species

Seasonal events, for example flowering, the return of migrating birds, and fruiting, happen at particular times of the year. Some of these events happen in relation to climate, while others are dependent on other factors. As the climate changes, the timing of some events may change in some species. This exercise will help you evaluate **how climate change has already affected species in Minnesota**.

The American elm (*Ulmus americana*) is a deciduous species and its range includes most of the eastern United States and a bit of southeast Canada. This elm is a hardy species and can grow to a considerable size. It grows well in urban areas so it is common in many cities. Unfortunately, the species is also susceptible to Dutch elm disease and the tree populations can suffer from massive die offs. All these characteristics have led naturalists to be interested in this species, so we have lots of data about its phenophases.

Minnesota is in the northern part of the elm's range. As the climate changes we may wonder how it is going to affect the elm. Potential questions we may ask include: Does climate change increase the occurrence of Dutch elm disease? Does it increase the rate of transpiration (water evaporation from the leaf)? Does it affect flowering date? Does it affect when leaves change their colors in the fall? We can use the MN Phenological Network's database to determine which of these questions we can pursue and what the answers might be.

Activity A – Determine changes in flowering date for American elm in Ramsey Co, MN

1. **Open** the provided data sheet ("MNPN master data sheet 2018"). If you completed the Pre homework, sorted your data, and saved the data sheet, skip down to #4.
2. Take a **look** at the data and see if you can determine how it is initially presented (e.g. what is listed earlier in the spreadsheet compared to later.). Additionally, **note** how many different entries (rows) are in the spreadsheet. It is not possible to simply scroll through and find the information we want, so we will have to sort the data into an organization that will be more useful for our purposes.
3. Select *all* the data ("Control/Command" + "A"). If you are using Excel, choose the "**Sort and Filter**" tab in the "Data" tool bar. Click on the "Custom sort" dropdown option. If you are using Google Sheets, choose the "Data" tab, then "Sort Range," and "Data has header." How might we want to group the data so we can find information about American elm (*Ulmus americana*)?

Sort first by Lifeform, then "Add Level" or "Add another sort column" and sort by Group, then by Species (scientific name), then by Species (common name; there may be multiple organisms with the

same Latin species name, but different genus), then by Event. **Note:** as you add filter levels, be sure to add them sequentially (to the bottom) or it will drastically change how the data is sorted.

4. We are looking for data about the American elm (*Ulmus americana*). Do you want to search through 54000 rows of data to find it? No! We can **search** for the elm entry in the data base by finding elm entries (“Control/Command” + “F” in Excel, or “Edit” then “Find and Replace” in G Sheets) and entering “American elm.”

5. Now that that data is more organized, what are some phenological events (phenophases) recorded for the elm?

6. Based on the data, can we answer *all* of the questions we posed about elms and climate change (see page 1, “Potential questions we may ask”)? If not, what data are we missing? What data can we use to answer which question(s)?

7. After a long Minnesota winter, people may be ready to think about spring and flowers, so let us concentrate on the phenological event of flowering date for American elm. We can now ask the question, **how does climate change affect the flowering date of American elm?**

It may still be tough to see a trend in the data as is it presented, so we can create a figure to help us visualize these data. Before we do that, we want to think about what location the flowering data came from. Do we want to use data from all over the state or just one county? For the elm we will just use Ramsey Co. records because we have a lot of data from that small geographical location.

8. **Copy** (“Control/Command” + “C”) the elm **year** and **flowering day** (just these two columns) from Ramsey Co (not the Itasca Co data) from 1941-1991 and **paste** (“Control/Command” + “V”) them into a new spreadsheet. Tabs at the bottom of the window open a new worksheet in the same file. Alternatively, you can open a new window and save it as a new file.

9. What do you think the flower dates might look like over the 50 year span? **Sketch** out your predicted trend in Figure 1 below. Be sure to label your figure’s axes.
(To draw in a word program, look for an “insert” option and then select a draw/line/shape tool.)



Figure 1. Prediction of elm flowering date from 1941 -1991 in Ramsey Co, MN.



Figure 2. Actual elm flowering date from 1941 -1991 in Ramsey Co, MN.

10. Now **check** how close your prediction is to the real data. Highlight the data in your two columns and click on the “Insert” tab. To show the data, you will want to choose a scatterplot (in Excel select “scatter” in the ribbon; in G Sheets choose “Graph” then double click on the graph, select “set up” tab, then “scatter chart” in chart type). If you are using Excel, click on the gridlines of the chart and then hit delete to help focus on the data. In G Sheets, double click on the graph to open the Chart editor. Select the “Customize” tab, select “Gridlines and Ticks,” and then toggle between vertical and horizontal dropdown to remove the gridlines. What kind of pattern do you see in the data?

11. One way to help visualize the data is to add a trendline (a line that represents the *overall* trend of the data) to your figure. **Add** a trendline to the graph (in Excel: find the layout or “add chart element” tools under the “chart tools” and “design” tab; in G Sheets double click on the graph, select “customize” tab, then “series” in the menu, then check the “trendline” box). Add a simple linear trendline.

Now what kind of pattern do you see in the data and what does the pattern show? How would you quantify the trend (i.e. use approximate values to describe the trend)? Based on the data, over the 50 years, has flowering happen earlier, later, or show no change in timing? How close was your prediction to the actual data?

Sketch the actual data trend above (Fig. 2); if appropriate, you can also copy and paste your figure from your spreadsheet. Be sure to label your figure’s axes.

12. As we look at the flowering date, what might be causing the change? **Identify** two ideas/hypotheses about what might be causing a change in this phenophase.

13. To investigate a relationship between climate and flowering date of American elm, we need climate data for Ramsey County from 1941-1991. Elm flowering date ranges from mid-March through mid-May so average temperature for March would be an appropriate period. For climate, it is good to gather data from at least a month. What do you think these data might look like? **Sketch** out your predicted trend in Figure 3 below. Be sure to label your figure’s axes.



Figure 3. Prediction of average March temperature in Ramsey Co, MN from 1941-1991.



Figure 4. Actual average March temperature in Ramsey Co, MN from 1941-1991.

14. **Find** the climate data for the period of 1941-1991. Go to the DNR Minnesota Climate Trends website (<https://arcgis.dnr.state.mn.us/ewr/climatetrends/>) for historical observations from specific Minnesota locations. **Enter** Counties as the geographical unit, select Ramsey County, select the climate variable of average temperature, for the month of March, and the appropriate date range. **Click** plot data. The data will appear in a figure and table below the selection area. In the table, highlight the temperature and year data from Ramsey Co. from 1941-1991. **Copy** and **paste** the data into empty columns of your spreadsheet with the elm data.

15. Now check how close your prediction about temperature matches the actual data. **Graph** the temperature data across time (1941-1991) using a scatterplot. To help focus on the data, delete gridlines. **Add** a trendline. How close was your prediction? **Sketch** the trend of the DNR temperature data in Figure 4; you can also copy and paste your figure from your spreadsheet. Be sure to label your figure's axes. If your prediction differs from the actual data, why might that be the case?

Preparations for Activity B and C

You may need to prepare your software for Activity B and C. If you are using Excel on a newer computer and your version does not have the "Data analysis" option under the data tab, you will need to **add** it. On a PC, in the file tab, choose Options; On a Mac OS choose Tools. Then select "Add-ins" from the menu on the left. Select "Analysis ToolPak" then click Go. Select Analysis tool pack then OK. The Data Analysis tab will pop up under the Data tab (it may request that you restart Excel to load the tool pack). Note: Older computers may not be able to add/run the analysis tool so ask your instructor for the regression statistics results.

In G Sheets, select the "Add-on" tab at the top of the sheet and select "Get add-ons." Search for "XLMiner Analysis ToolPak" in the search box and install that in your google sheet. Follow the prompts and once complete select the "Add-on" tab at the top of the sheet and verify "Linear Regression" is an option. **Note:** If you are not able to complete this step in G Sheets due to a verification backlog with Google, ask your instructor for the regression statistics results.

Activity B – Determine significance of changes in flowering date

Over the 50-year span, we see a trend for earlier flowering dates for American elm and warmer March temperatures, but we need to ask, does flowering date *depend* on temperature? Is there a significant dependent relationship between flowering and temperature in this species? It may help us to first look at the data.

16. Due to how spreadsheets handle data, you will need to have your independent data (temperature) in a column on the left side (i.e. the X axis) and your dependent data (flowering date) on the right side (i.e. the Y axis). If needed, copy and paste your data into new columns. Highlight these two columns of data and **create** (insert) a scatterplot. To help focus on the data, delete gridlines. **Add** a simple linear trendline to the graph. **Sketch** the data and trendline in Figure 5 and label your axes. What kind of pattern do you see in the data?



Figure 5. Elm flowering date and average March temperature from 1941 -1991 in Ramsey Co, MN.

(add trend and stats here)

17. A regression analysis helps determine if there is a statistically significant dependent relationship between two variables, in our case if flowering date depends on temperature. The regression analysis also measures the strength of this relationship between variables.

Run a regression analysis of the flowering date versus temperature. In Excel, click on the “Data” tab and then the “Data Analysis” option. Select regression in the pop up menu. Click on the “Y Range” box and highlight your flowering date, these are your dependent data. Click on the “X Range” box and highlight your temperature data, these are your independent data. Choose the Output Range button, click on the “Output Range” box and click in an open cell of your data sheet.

In G Sheets, you will need to make one more adjustment before you run your statistical analysis. Highlight your date data and select “Format” and then “Number” and “Number” again (G sheets will not recognize your date in the format from the phenology data set). Once your date is converted, click on the “Add-ons” tab at the top and select “XLMiner Analysis ToolPak,” then “start.” Scroll down and select “Linear regression.” Click on the empty “Input Y Range” box and highlight your flowering date, these are your dependent data. Click on the empty “Input X Range” box and highlight your temperature data, these are your independent data. Choose the Output Range box and click in an open cell of your data sheet.

The software will give you a good deal of information in the output from your test. **First**, look for the R squared (r^2) value. The r^2 value ranges between 0 and 1 and tells you how closely X and Y variables are related. If $r^2 = 0.52$, 52% of the variance of Y is “explained” by the X variable. In our phenology language, this means that 52% of the variation in flowering date is explained by average March temperature. What is your r^2 value? _____

Second*, look for the regression row in the ANOVA table and find the Significance F value. The Significance F value is the probability (P) that there is no relationship between the variables (the H_0 or null hypothesis). The lower the P -value the more confident you can be the statement of no functional relationship between the variables is false. It is standard in biology to use a P-value of 0.05, or 5%, as a cutoff for rejecting the statement that there is no relationship between variables. In plain English, if you have a P-value lower than 0.05, there is a really good chance that there *is* a relationship between your two variables (i.e. elm flowering date *does* depend on March temperature). If you have a P-value above 0.05, then it is most likely that there is no relationship between your two variables (i.e. flowering date *does not* depend on March temperature). What was your P-value? _____

Add into your figure caption a description of the trend you see in the data and the r^2 and p values.

*If you were unable to run the analysis in Excel or G Sheets, see the regression output information shared by your instructor.

18. When you look at the data in Figure 5 and the information from the regression analysis, what inferences can you make, or what can you **conclude** about the relationship between American elm flowering date and March temperature? How would you **quantify** the trend (e.g. For every “X°F”, flowering date changes “Y”)?

Activity C – Determine a phenophase (“event”) pattern for a Minnesota organism of your choice

19. Go back to the original data set from the MN Phenological Network. **Identify** a species you are curious about that has data for at least 20 years for a given event and location (you should have at least 20 data points over at least 20 years). What phenophases (events) are available for your species?

20. What phenological **question(s)** can you pose about your chosen species given this dataset?

21. Once you determined your species, **organize** the data by “event.” Keep in mind the location data and use the appropriate data. What county will your data come from? What time frame will you use?

Copy your data (phenophase dates and years) into a new spreadsheet. Pay attention to what years data are present. You may have gaps in the phenophase record and will need to adjust your temperature data to match the years available.

22. **Create** a graph of your data in Excel or G Sheets and **sketch** your data in Figure 6. Label your axes and add a figure caption.



Figure 6.



Figure 7.

23. What kind of pattern do you see in the data?

24. With the DNR MN Climate Trends website, **find** the climate (temperature) data for your *new* period appropriate for your *new* phenophase. **Copy** these data to your spreadsheet. **Adjust** or drop the temperature data for the years with no phenophase data (note in step 21 above). **Create** a figure with your climate data, and **sketch** your data in Figure 7. Remember to add axes labels and a figure caption.

25. **Create** a figure with your two variables (phenophase and climate data). **Sketch** your data in Figure 8. Remember to add axes labels and a figure caption. If appropriate, adjust your axis scale to fit your new data range (in Excel, right click the axis, select “Format Axis,” adjust Max and Min values; in G Sheets select the axis you want to adjust in the “Customize” tab of the Chart editor and add your Min and Max values).

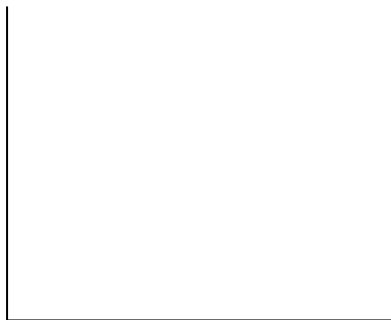


Figure 8.

26. **Run** a regression analysis of your chosen phenophase and climate data and measure the strength of the relationship between your variables. What is your r^2 value? _____

What was your P-value? _____

What inferences can you make, or what can you **conclude** about the relationship between your phenophase and your climate data (temperature)? How would you **quantify** the trend?

27. **Share** your Figure 8 with your classmates. Be sure to include the axes labels and the figure caption. These figures will be gathered and the class' data will be shared so you can answer question 28.

28. **Look** at all the different Figure 8s created by your classmates.

Which phenophases (if any) are now occurring earlier? How do the data help you reach this conclusion? Why might this/these phenophase(s) occur earlier?

Which phenophases (if any) are now occurring later? How do the data help you reach this conclusion? Why might this/these phenophase(s) occur later?

Which phenophases (if any) do not appear to have changed? How do the data help you reach this conclusion? Why might this/these phenophase(s) not be any different?

29. Was there anything that surprised you as you worked through this exercise?

Citations:

Minnesota Department of Natural Resources (DNR). 2020. Minnesota Climate Trends. Accessed [7.10.2020]. Available: <https://arcgis.dnr.state.mn.us/ewr/climatetrends/> .

Minnesota Phenology Network (MNPn). 2020. Datasets. Accessed: [7.10.2020]. Available: <http://mnpn.usanpn.org>.