

Ecology Data Exercise: Phenology

Name:

Learning objectives: Manipulate data to make and interpret scatterplot graphs and regressions of the phenology of bumblebees in relation to climate.

Data Source: All data for this module were collected by citizen scientists and are available from the National Phenology Network website.

Background Knowledge of Phenology

Part A: National Phenology Network. Go to the National Phenology Website and read the page [Why phenology?](#) Also, go to the [Volunteer Scientist](#) page to understand more about the process by which phenology data are collected.

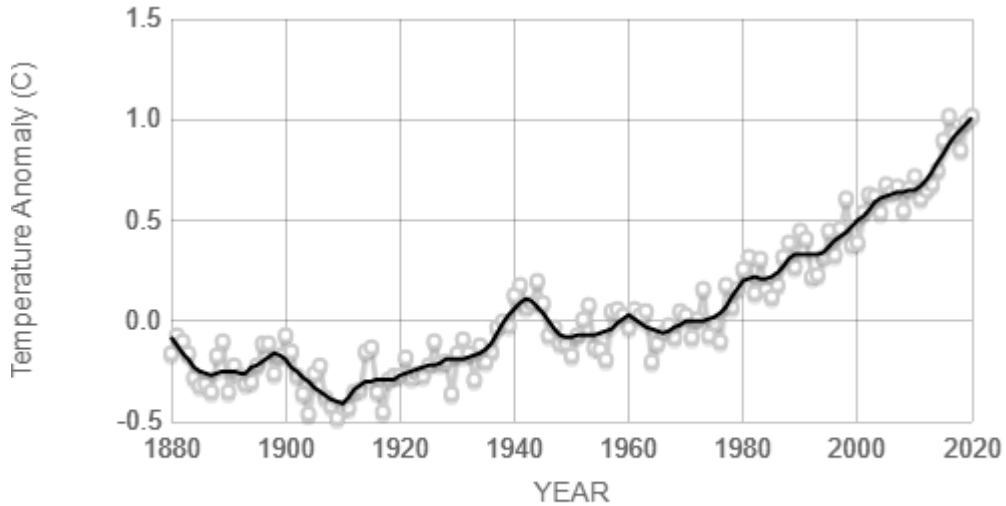
- 1) How would you define phenology?
- 2) Why is phenology vital for ecosystem function?
- 3) What is phenology necessary for human society?

Currently, spring has started on campus. As you walk around you see various buds forming and flowers developing. However, this process will continue for plant species differently throughout growing season (April – October).

- 4) What environmental cues are responsible for this phenological behavior?
- 5) Using your answer for #4, explain how patterns in phenology will differ by latitude.

Data Activity: How has bumble bee emergence changed over time?

PART B: One way to predict how a species will respond to climate change in the future is to examine at how it has responded to climate change in the past. The National Phenology Network provides recent data about the day of the year when the first sighting of adult bumblebees occurred at a site in the years from 2010 to 2019. Six of these years are among the top 10 warmest years on record (see image below).



Source: climate.nasa.gov

- 6) In the space below, sketch a graph of the relationship between phenology (first bee sighting) and year from 1980 to 2020. For phenology, assume the first active sighting for bees in 1980 was day 100 (around the second week of April). Below your figure, justify your prediction.

Now open 'phenology_data_student.xls'. The first worksheet contains the metadata file for the dataset. The metadata contains complete description of each variable. The second worksheet, 'phenophase only', contains bumblebee phenology data for all available data in the US and a subset of data for only Minnesota. Each dataset contains a column with the year of record for a site and a column with the day of the year when bumblebees were first recorded for the site.

Make a figure/s that visualizes the relationship between bumblebee activity and year for both datasets with a trend line and the R^2 statistic. If you place both datasets on one figure, graph the dataset with all sites first and then add a new series with data from Minnesota (because data points overlap). Copy-Paste your completed figure into the space below.

7) What does the R^2 value tell us about the variation of bumble bee activity and year for both datasets?

8) It should not be a surprise that different regions experience different trends in climate variables across years. Why might phenology in a region that includes Minnesota be affected differently than other regions (e.g. southwest or east coast)? How might you adjust your decisions for Q#6?

PART C: We examined change over time, but we can't be sure that changing climate is responsible for any patterns we observed. We can use environmental data recorded for each site to identify how much variation in bumblebee emergence phenology is explained by temperature or other climate-related variables.

- 9) Examine the variables described in the metadata worksheet carefully. Ecologists often calculate averages of climate variables over specific time periods within a year to express variability among seasons. Which of the climate variables in the dataset do you think will be the best predictor of emergence phenology (DOY)? Justify your reasoning below.

Now, evaluate the relationship between your chosen climate predictor and bee phenology both graphically and statistically. First, notice that the phenology column (Phenophase_Description) contains two different variables. Before beginning your analysis, sort the entire dataset by 'Phenophase_Description' to organize those different variables ('Sort & Filter' on the 'Home' Tab).

Data Cleaning. You are working with a real dataset. Often automated equipment for measuring climate generates missing data (i.e. power outage). Some of the columns have missing data that are recorded as '-9999' in the data sheet. This number is usually a standard output from a machine to let the user know the data is faulty. In Excel, we can replace those values with blank values. If not, our graphing efforts will be very wrong.

- Highlight the column of variables you would like to work with (or all columns).
- On the 'Home' tab click on 'Find and Select'. This is usually on the far right side.
- Click on 'Replace'.
- In the new window, enter -9999 into the value for 'Find What'.
- Leave 'Replace with' blank. We are replaced that specific number with nothing!
- Click 'Replace All'. It should tell you how many values were changed.
- Look to make sure it worked!
- Remove rows with empty data, if utilizing regression analysis.

- 10) Create scatterplots with a trend line and the R^2 statistic for both Flower visitation and Active adults (see metadata for descriptions) against your chosen climate variable. Analyze the relationship with regression analysis. Copy/Paste your figures into the space below and summarize your findings in sentence form.

11) This dataset has a broad coverage of multiple sites (states) and years. Additionally, the characteristics of each site (e.g. elevation) are different. Combined, this amount of variation may make it difficult to detect smaller patterns. Pick one of these additional sources of variation and explore how it impacts the relationship between the climate variable you have chosen and one of the phenology variables (Active adults or Flower visitation). You have several options in which to organize, visualize, test and report new findings. In the space below, Copy/Paste at least one new figure that visualizes your variables in a new way. For example, you could group variables into ranges of heights, specific time frames or geographic region. Analyze the relationship with regression analysis and summarize your new findings below the figure.

12) Based on your exploration of the data, how important do you believe temperature is for bumblebee emergence? Explain your reasoning.

13) What other kinds of data would you test if they were available?