

# Modeling Suitable Habitat for a Species of Conservation Concern: An Introduction to Spatial Analysis with QGIS

## Exercise

Peter J. Galante,<sup>i</sup> Katherine Hade,<sup>i</sup> Stefanie J. Siller,<sup>i</sup> Nadav Gazit,<sup>i</sup> and  
Suzanne K. Macey<sup>i</sup>

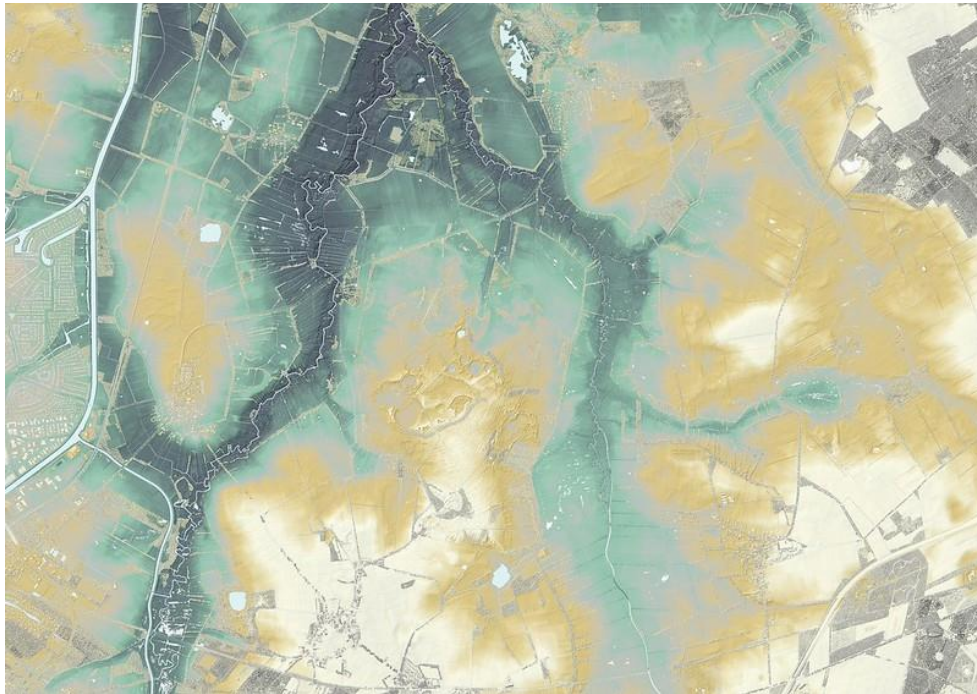


Image credit: Harry Bronkhorst/Flickr (CC BY-NC-SA 2.0)

## ABSTRACT

Spatial analysis has become a central practice in the field of conservation, allowing scientists to model and explore geographic questions on biodiversity and ecological systems. GIS (Geographic Information System) is an important integrative tool for mapping, analyzing, and creating data for spatial analyses. In this exercise, students use QGIS, a free open-source GIS program, to model suitable habitat for a cryptic mammal species. The exercise guides students through the process of: 1) organizing, cleaning, and clipping vector and raster data within QGIS; 2) analyzing climate, habitat, and additional geographic data along with species occurrence data; and 3) developing a map of suitable projected habitat for the species of interest. Students then apply their analyses to critically consider the implications for surveying and conservation action.

<sup>i</sup> Center for Biodiversity and Conservation, American Museum of Natural History, New York, USA

Reproduction of this material is authorized by the recipient institution for non-profit/non-commercial educational use and distribution to students enrolled in course work at the institution. Distribution may be made by photocopying or via the institution's intranet restricted to enrolled students. Recipient agrees not to make commercial use, such as, without limitation, in publications distributed by a commercial publisher, without the prior express written consent of AMNH.

All reproduction or distribution must provide full citation of the original work and provide a copyright notice as follows:

"Copyright 2020 by the author of the material, with license for use granted to the Center for Biodiversity and Conservation of the American Museum of Natural History. All rights reserved."

Any opinions, findings and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the American Museum of Natural History.

# Modeling Suitable Habitat for a Species of Conservation Concern: An Introduction to Spatial Analysis with QGIS

## Exercise

### LEARNING OBJECTIVES

After completing this exercise, students will be able to:

1. Use QGIS as a tool for overlay modeling and visualizing climate, habitat, geographic, and species occurrence data.
2. Apply spatial analysis results to make survey recommendations for a cryptic species of conservation concern.
3. Consider and discuss the potential uses and limitations of the data and the type of spatial analysis used in this exercise for making conservation decisions.

*“The ultimate goal is to learn how to solve problems spatially.”*

— Harder, C., and C. Brown. 2017. *The ArcGIS Book: 10 Big Ideas About Applying the Science of Where*. Esri Press, Redlands, CA, USA.

### INTRODUCTION

You’ve been asked to help a small non-profit conservation organization prioritize where they should survey for Treelepi\* in New York State. Treelepi (plural for Treeleop) are medium-sized mammals that are rare and cryptic—meaning they are not easily seen. Partially because Treelepi are elusive, we know very little about them, but what we do understand from past observations and studies is that Treelepi:

- have dense fur and prefer a specific temperature range,
- prefer natural areas (non-agricultural or built lands) that are forested and/or have plentiful woody vegetation, and
- prefer remote areas (Treelepi are shy and solitary animals and avoid major roads).

\*Treelepi are an imaginary mammal species for the purposes of this exercise. Although this exercise focuses on this imaginary species, the skills used in this exercise are real examples of how one might use spatial analysis for conservation purposes.

The New York State-based conservation organization you are consulting for would like to understand whether there should be more efforts to conserve this species at the state-level. Because they don’t have a firm understanding of how many Treelepi are in the state, they’d like to start a survey and monitoring project. Where should they start looking for Treelepi? You will be learning skills in spatial analysis to help answer this question and prioritize areas to survey in New York State.

## WHAT IS SPATIAL ANALYSIS?

*“Spatial analysis is a process in which you model problems geographically, derive results by computer processing, and then explore and examine those results. This type of analysis has proven to be highly effective for evaluating the geographic suitability of certain locations for specific purposes, estimating and predicting outcomes, interpreting and understanding change, detecting important patterns hidden in your information, and much more.”*

— Harder, C. and C. Brown. 2017. The ArcGIS Book: 10 Big Ideas About Applying the Science of Where. Esri Press, Redlands, CA, USA.

Spatial analysis offers a wide variety of tools, used in fields such as urban planning and geology. In the field of conservation, researchers use spatial analysis to help determine biodiversity hotspots and delineate protected areas, examine the effects of climate change on species or ecosystems, model the potential effects of the construction of dams, and more (Box 1).

### **Box 1: Spatial Analysis and the IUCN Red List**

Using spatial analysis to understand a species' current or potential geographic range is crucial for conservation action. The International Union for Conservation of Nature (IUCN) Red List uses two main measures of geographic criterion in evaluating the conservation status of a taxon: extent of occurrence and area of occupancy. Extent of occurrence refers to the smallest possible continuous boundary around all known or projected sites where the specific taxon occurs. This measure is useful for evaluating threats that may be present across a taxon's geographical distribution. Area of occupancy, on the other hand, is defined as the area within the extent of occurrence that is actually occupied by the taxon. This measure represents the amount of suitable habitat within the extent of occurrence. Figure 1 illustrates the differences between these two measurements. By estimating these spatial measures, conservationists can determine a taxon's conservation category and their extinction risk.

Image credit: IUCN. 2001. IUCN Red List Categories and Criteria: Version 3.1. Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK. Available from [www.iucn.org/content/iucn-red-list-categories-and-criteria-version-31](http://www.iucn.org/content/iucn-red-list-categories-and-criteria-version-31).

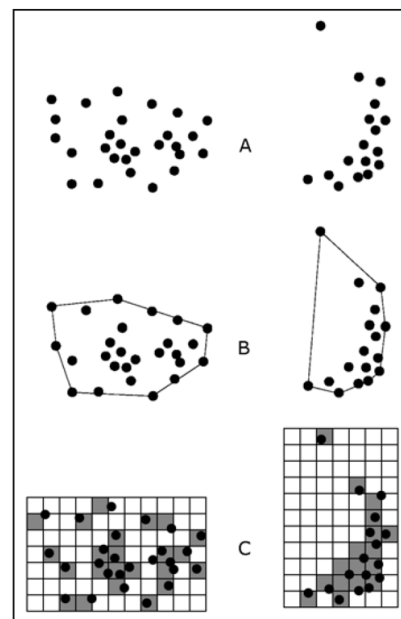


Figure 1. Two examples of the distinction between extent of occurrence and area of occupancy. (A) is the spatial distribution of known, inferred, or projected sites of present occurrence. (B) shows one possible boundary to the extent of occurrence, which is the measured area within this boundary. (C) shows one measure of area of occupancy which can be achieved by the sum of the occupied grid squares.

In this exercise, we will be overlaying descriptive habitat layers in a type of spatial analysis that will aid in determining suitable areas to survey for a rare species (the Treelop) based on known environmental tolerances such as temperature, or available habitat parameters. Throughout the exercise, take time to reflect and respond to the questions and prompts provided before moving on; these will help you think critically through the process and also help for writing your final summary and recommendations.

**Question 1:**

**Based on your understanding of the Treelop, what kind of geographical/climate data might be useful in determining suitable habitats? From a management/logistical perspective, what other variables should we consider?**

## **PART 1: DOWNLOADING QGIS AND ORGANIZING YOUR DATA**

### 1.1 DOWNLOADING QGIS

QGIS is an open-source Geographic Information System (GIS) that offers professional level tools for mapping, analysis, and creation of spatial data.

QGIS offers updated releases and bug fixes often. For the purpose of this exercise, we recommend that you download the latest Long Term Release (LTR) version, rather than the latest update. While the most up to date bug fixes and updates are available daily, these versions have gone through less scrutiny.

QGIS is available for Mac OS, Windows, and Linux distributions. We have provided screenshots and detailed steps using Mac OS and version 3.4 of QGIS. While these might differ a bit from what you will be using to complete this exercise, you should still be able to use the same functions. We only highlight the specific options for each step, and do not detail every option in the dialogue boxes unless necessary.

To download QGIS go to [qgis.org/en/site/forusers/download.html](http://qgis.org/en/site/forusers/download.html). You may also be interested in learning more about the background of some of the concepts of GIS that go beyond the scope of this exercise; QGIS provides a helpful documentation called “Gentle Introduction to GIS:” [docs.qgis.org/3.4/en/docs/gentle\\_gis\\_introduction/](http://docs.qgis.org/3.4/en/docs/gentle_gis_introduction/). Of particular relevance are the Introducing GIS, Coordinate Reference Systems, and Vector Spatial Analysis sections. Another great resource is the QGIS tutorial found at: [www.qgistutorials.com/en/index.html](http://www.qgistutorials.com/en/index.html).

## 1.2 SETTING UP A DIRECTORY

When using a GIS, being organized is imperative. Being unorganized will cost you time and effort!

It is best to manage your data in a single directory or a series of directories, meaning a single folder or a group of nested folders. When you are naming files or folders, be sure to name them carefully, descriptively, and in a standard way so they are easily understood in the future or by other researchers. Additionally, don't use spaces in file/folder names (for example, "file name.shp") when naming, because some GIS software programs have trouble reading certain characters in file/folder names; using underscores "\_" avoids this issue ("file\_name.shp"). Similarly, avoid starting a name with a number or special character (\*!@#, etc.) (e.g., "1\_file\_name.shp"). To avoid issues of this nature, we have suggested file names.

To begin, create a new, empty directory (folder) on your computer called "Treelop\_spatial\_analysis" to act as your working directory. In this directory, we will save all of the data provided in Part 1.3 below, as well as all of the new GIS layers we will create during the rest of the exercise.

## 1.3 DOWNLOAD THE DATA

Finding appropriate data is not without challenges, such as: finding accurate and reliable sources; accounting for variation in data quality or resolution; calibrating for biases and gaps in the available data; or, if needed, finding data that is freely available and is not behind a paywall.

For your convenience, we have provided data that you will use to make recommendations for surveying for Treelopi. These are some of the data that could be helpful in the process; later on, we'll ask you to reflect on what might be missing, or what next steps you would do after these initial spatial analyses.

Unzip "Treelop\_data.zip," as well as the compressed folders within, and put the Treelop\_data folder in your "Treelop\_spatial\_analysis" directory. Inside the "Treelop\_data" folder there are five types of data. See Table 1 for a list of the types of data including some metadata on the data and their sources. **Note:** if you are doing your own spatial analysis project, the listed data sources in Table 1 (e.g., [naturalearthdata.com](https://www.naturalearthdata.com), [worldclim.org](https://worldclim.org)) might be a good place to start your search for certain types of data. Additionally, these sites will normally provide a wealth of additional metadata, including how the data were collected, by whom, issues of data quality, as well as constraints and possible recommendations.

**Table 1.** Provided data to determine Treelopi survey locations.

Feature	Geographical extent	Scale/ Resolution*	Directory/ Filename for this exercise	Information/data	Source
Political Boundaries	United States of America	10 meter resolution	USA_states; USA.shp	Original name: ne_10m_admin_1_states_provinces	<a href="https://www.naturalearthdata.com/downloads/10m-cultural-vectors/10m-admin-1-states-provinces/">https://www.naturalearthdata.com/downloads/10m-cultural-vectors/10m-admin-1-states-provinces/</a>
Mean Annual Temperature	Northeastern United States of America	30 arc-seconds resolution	NY_regional_bio_1.tif	Original name: wc2.0_30s_tavg_01  Mean Annual Temperature in degrees Celsius. For this exercise, we've only included the New York regional data.	<a href="http://worldclim.org/version2">http://worldclim.org/version2</a>
Land cover	New York State + Parts of Surrounding States	1:250,000 scale	landcover; lulc250k.shp**	Collected by USGS and adapted by EPA. Cornell University Geospatial Information Repository (CUGIR) New York State land cover shapefile.	<a href="https://cugir.library.cornell.edu/catalog/cugir-008190">https://cugir.library.cornell.edu/catalog/cugir-008190</a>
Major Roads	Global	10 meter resolution	Roads; ne_10m_roads.shp	Original name: ne_10m_roads  Interstate, Federal, and State roads	<a href="https://www.naturalearthdata.com/downloads/10m-cultural-vectors/roads/">https://www.naturalearthdata.com/downloads/10m-cultural-vectors/roads/</a>
Treelopi Occurrences	New York State		Treelop_sightings.csv	These were generated for this exercise.	<a href="https://ncep.amnh.org">https://ncep.amnh.org</a>

\*<https://www.nrcan.gc.ca/maps-tools-publications/satellite-imagery-air-photos/remote-sensing-tutorials/satellites-sensors/spatial-resolution-pixel-size-and-scale/9407>

\*\*"lulc" is widely known to stand for "land use/ land cover" and the 250k represents the scale, which is "1:250,000."

### Reflection

Go back to your answer for Question 1. Compare these data to the variables you listed. What are the differences?



## PART 2: GETTING STARTED

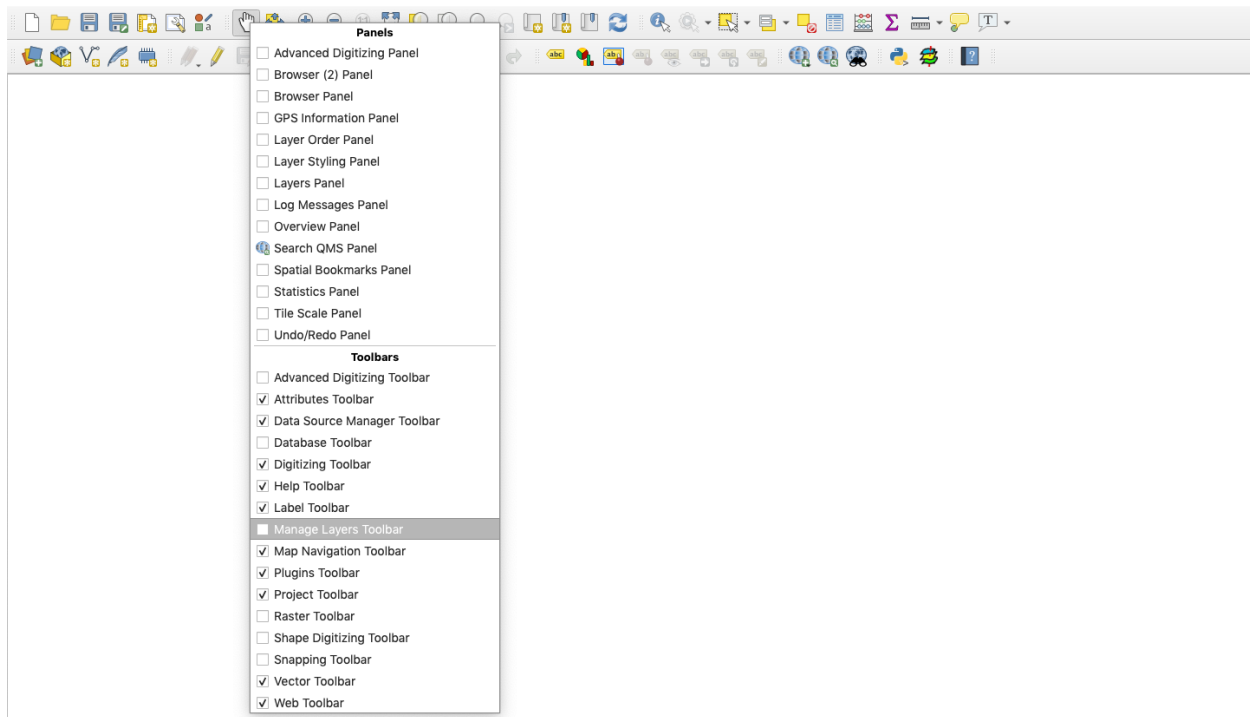
Below, we will be walking you through the steps of some aspects of spatial analysis using QGIS. In addition to explicit instructions, we provide screenshots from version 3.4 of QGIS on a Mac OS. We highly suggest you use QGIS version 3.4 to avoid compatibility issues with our instructions and the software. Additionally, unless otherwise noted, we suggest that you do not modify the default settings that will populate a dialogue box when you run a tool.

### 2.1 CREATE THE PROJECT

Open QGIS and save your new project as “Treelop\_spatial\_analysis.qgz” in the working directory (the .qgz extension is the default extension for a QGIS project).

### 2.2 SETTING UP YOUR WORKSPACE

To add some helpful toolbars, you can right click on the default toolbar that is on the top of your workspace and a menu will pop up. Select the “Manage Layers Toolbar” to add a new toolbar that will appear on the left side of your workspace.



### 2.3 LOAD THE DATA



Most aspects of spatial analysis can be represented through two types of data; vector<sup>1</sup> and raster<sup>2</sup>. The vector data we will use in this exercise are shapefiles (saved as file\_name.shp). These file types are standard in GIS and work across many GIS

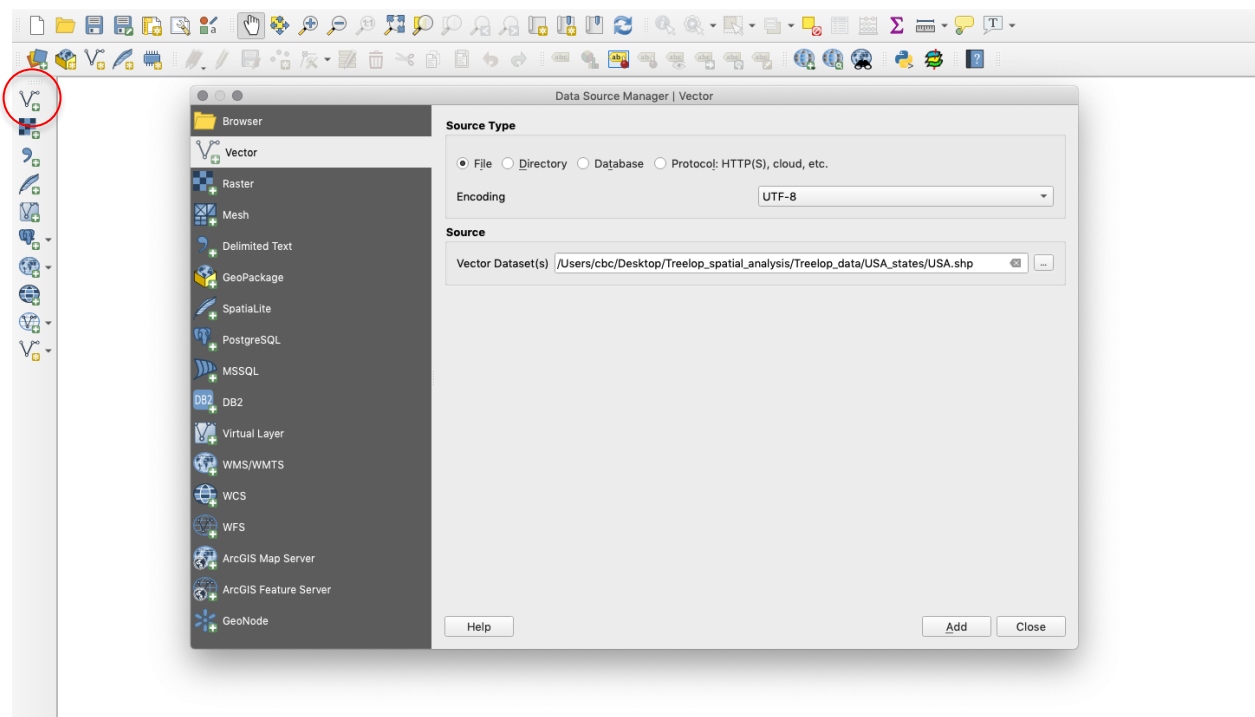


software (such as ArcGIS). Most vector data available online is in this file format. Vector data can represent point localities, lines, or polygons. Most raster data available online are in GeoTIFF format (file name.tif), an image file that has geographic referencing embedded within it. Rasters can represent continuous gradients such as elevation, temperature, or ocean salinity. All the files in this module (other than .csv files) can be “dragged and dropped” from a folder into the Layers panel in QGIS. Alternatively, below are step by step instructions to add layers.


1. Load political boundaries as a vector layer to your map.

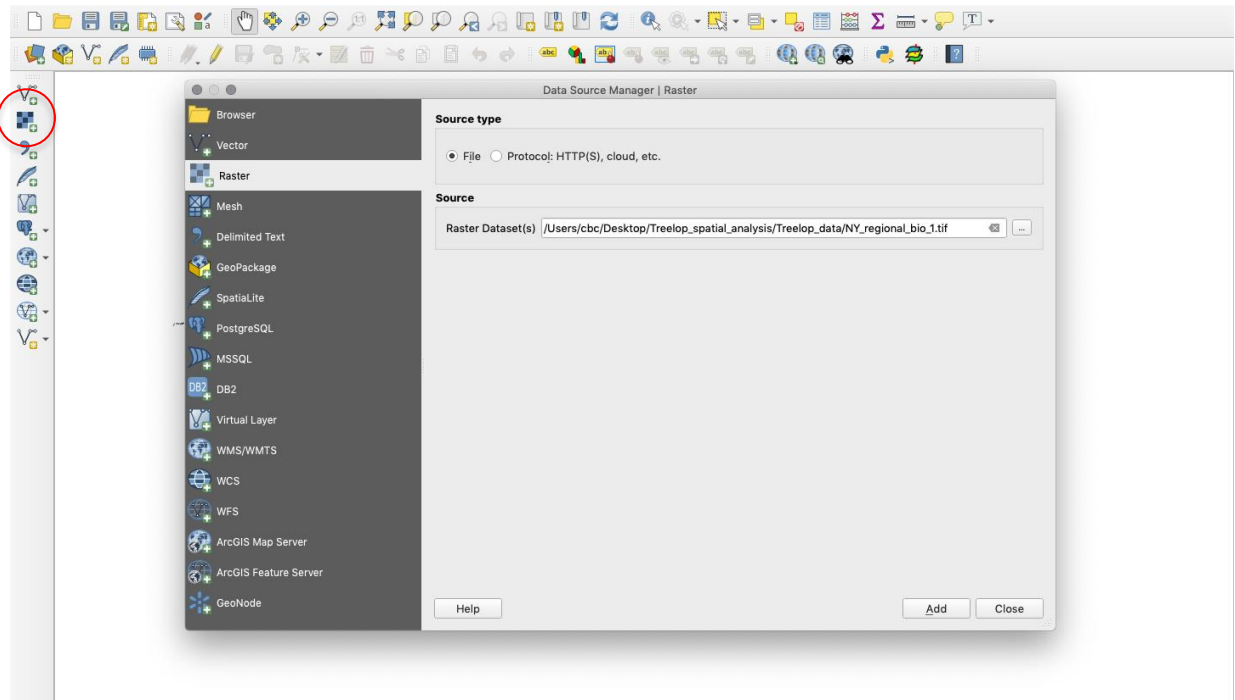
Shapefiles are a vector format data storage file originally developed by ESRI (Environmental Systems Research Institute) for storing geometry<sup>3</sup> and location information. These files are packaged with and linked to other files that are necessary to represent the geometry in GIS. For example, the file with the “.dbf” extension is the database file and contains the spreadsheet of information associated with each attribute. The “.shx” file is the index file, and the “.shp” file contains the shapefile geometry. These three files are necessary and need to be stored in the same directory for GIS to read the files, but only the file with the “.shp” extension needs to be loaded.

Navigate to the “Add Vector Layer” button (“”). Alternatively, you can go to the Layer panel > Add Layer > Add Vector Layer. In the “Vector Dataset(s)” box under “Source,” click browse (the small ellipses button “”) and navigate to your working directory and then to the “USA\_states” folder within “Treelop\_data” directory and select “USA.shp”. Click “Add” and then “Close” in the “Data Source Manager” dialog box.



2. Load mean annual temperature as a raster layer to your map.

Navigate to the “Add Raster Layer” button (“”). Alternatively, you can go to the Layer panel > Add Layer > Add Raster Layer. In the “Raster Dataset(s)” box under “Source,” navigate to the “Treelop\_data” folder and select “NY\_regional\_bio\_1.tif”.



3. Load land cover as a layer to your map.

Add “lulc250k.shp” through one of the methods above. Consider that this file is a “.shp” file. Is it a raster or a vector?


If you click on the Information Tab in the Layer Properties dialog for this layer (double click on the layer name in the Layers panel), you might notice that the CRS (Coordinate Reference System) of this layer is different from the previously added layers. We will explain more about CRS later in the lesson, but these details are important to note.

4. Load major roads as a layer to your map.

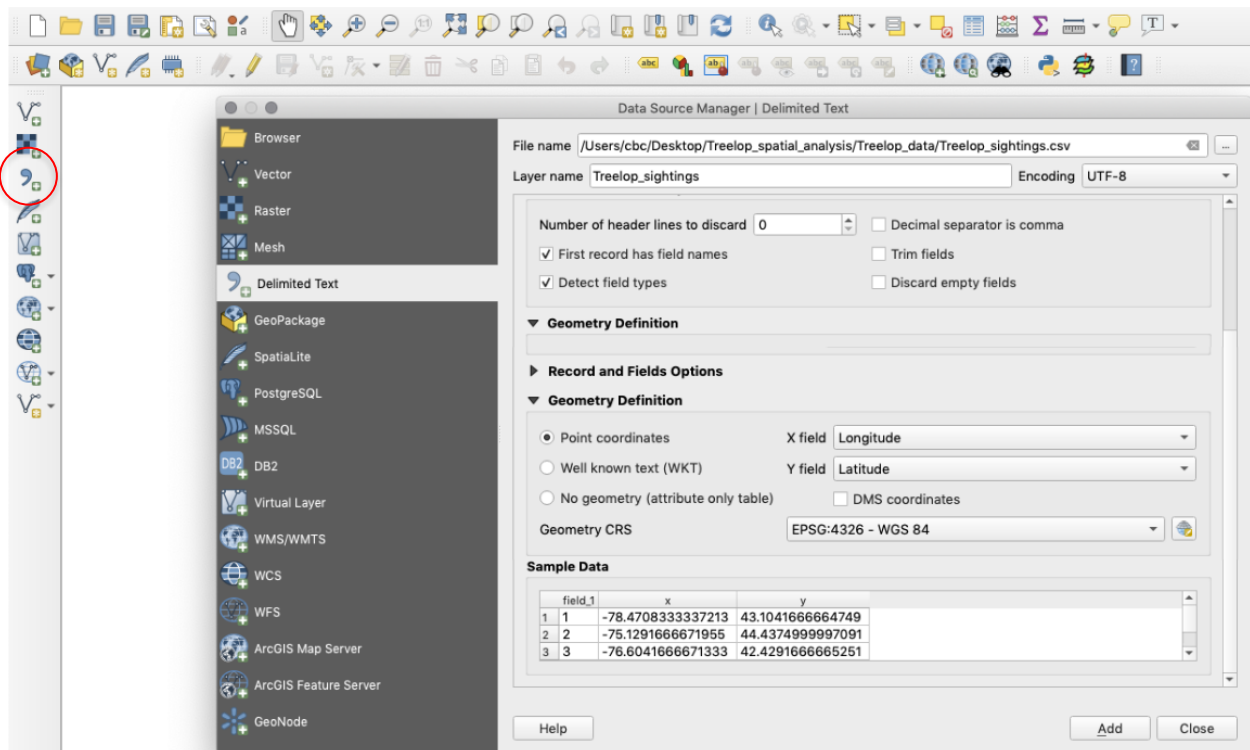
Add “ne\_10m\_roads.shp” through one of the methods above, or try dragging it from your working directory and dropping into QGIS.


5. Load Treelop occurrences as a layer to your map.

**Reminder:** “.csv” files cannot be dragged and dropped.


Navigate to the “Add Delimited Text Layer” button (“”). Alternatively, you can go to the Layer panel > Add Layer > Add Delimited Text Layer. In the “Data Source Manager” dialog box, under “File name,” navigate to “Treelop\_sightings.csv.” **Note:** This file is a comma-separated values file and was created in a format where periods represent decimal separators (instead of commas).

Under “Geometry Definition,” be sure to assign “Longitude” as the X field and “Latitude” as the Y field. (Just as on a typical graph with an X and Y axis, QGIS reads occurrence data based on these two axes: Longitude corresponds to the X axis, and Latitude corresponds with the Y axis.) Make sure that the Geometry CRS is set to “EPSG 4326 – WGS 84.” Click “Add” and then “Close” in the “Data Source Manager” dialog box.



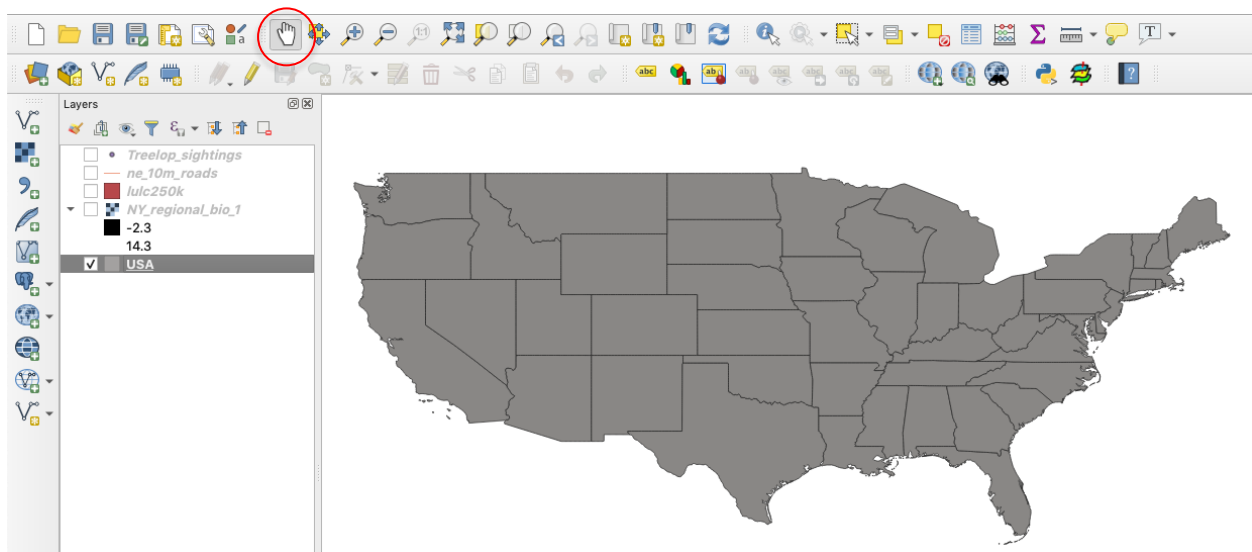
SAVE your work. Click on the “Save Project” icon (“”) in the upper left of the toolbar, go to Project > Save, or press Command-S on a Mac.

## PART 3: CLEANING UP THE DATA

Explore the data a bit. Using the “Pan Map” tool (“”), click and drag to move around the map. If using a mouse with a scrolling wheel, scrolling up and down will zoom in and out on your map. You’ll notice that some of our data are global, national, or regional. Our task, however, is to highlight potential areas to survey in New York State only, and retaining large areas rich with information requires more computer memory and could slow down processing time. To focus our area of analysis and save processing time, we

should crop the data down to just New York State. You'll need to have your Layers panel active, which lists all of your loaded data (layers) on the left side like a table of contents of your map. If the Layers panel is not active, right click anywhere on the top default toolbar and select "Layers Panel" as you did for adding the "Manage Layers Toolbar."

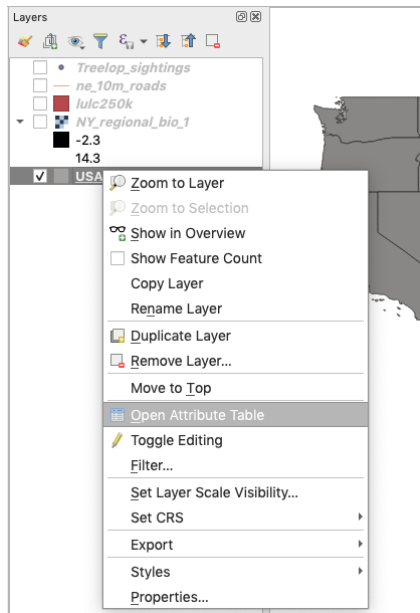
1. Begin with the states/provinces layer, "USA." In the Layers panel, hide all other layers by clicking on the check mark to the left of the layer to deactivate those layers.




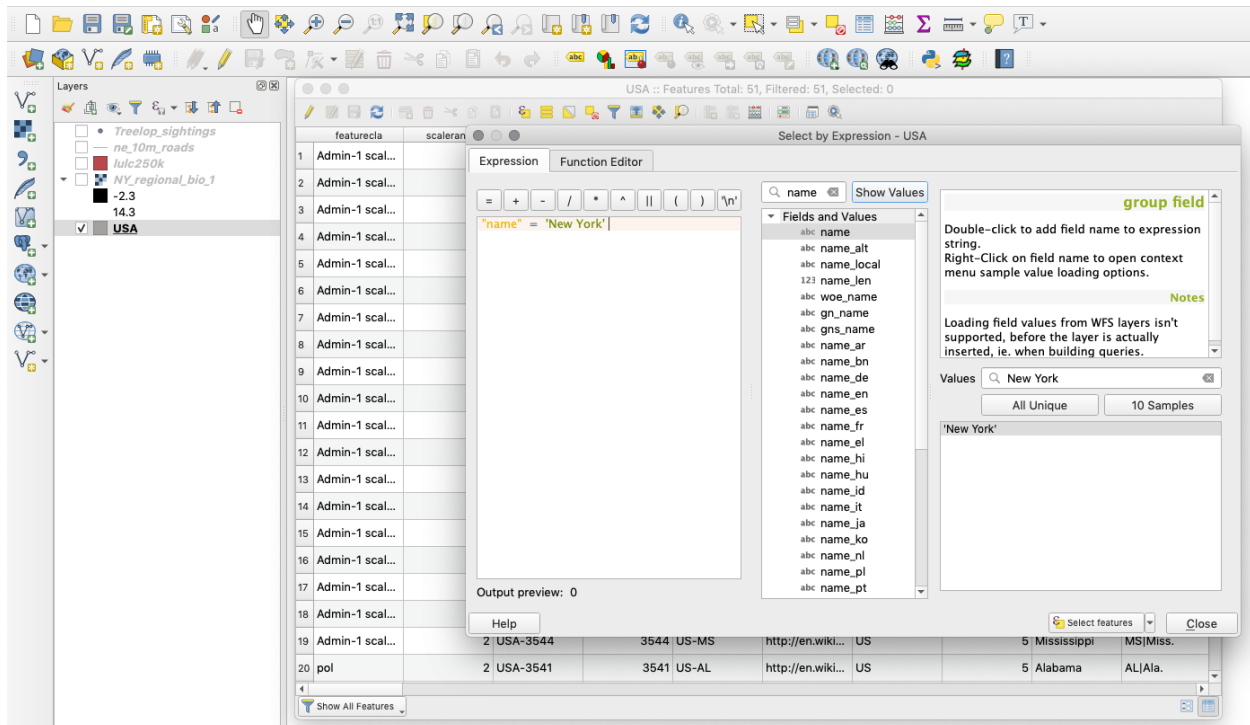
2. Using the "Zoom Full" button ("🗺️"), zoom to the extent of the shapefile "USA."  
**Note:** The Zoom Full function will pan out, or in, to include the extent of all layers currently selected in the Layer panel. Alternatively, you can right click on the layer of interest in the Layers panel and select "Zoom to Layer."



3. Create a new shapefile of New York:
  - a. An attribute table shows the data associated with geometry in the file. To open the Attribute Table, right-click on the "USA" layer in the Layers panel. Select "Open Attribute Table." Explore the Attribute Table. You will notice the wealth of information (attributes) available for each feature (each row) in the Attribute Table.



- b. Using an expression<sup>4</sup> string, select the features of the “USA” layer that correspond with the attribute “New York.” In the Attribute Table, click on “Select features using an expression” (“”). In the search bar, type “name” and double click on “name” in the window. It will appear in the Expression window. Click on the equals sign “=” above the Expression window. On the right side of the window is another search bar, “Values.” where you can search values within each field. Here, type in “New York,” then click “All Unique.” Double click on the field ‘New York’ and it should appear in the Expression window. Click “Select features.” On the map, New York State will be highlighted in yellow. Close the “Select by Expression” dialogue box.



- c. Back in the Attribute Table, in the lower left corner, change “Show All Features” to “Show Selected Features.” The row with data for New York State should be highlighted.
- d. Close the Attribute Table and inspect the map. New York should be highlighted. Right-click on the “USA” layer in the Layers panel and select “Zoom to Selection.” The map should jump to New York.
- e. To save the selection as its own layer, right-click on the layer in the Layers panel. Select “Export,” then “Save Selected Features As.” In “Format,” select “ESRI shapefile.” Navigate to the working directory by clicking on the ellipsis button next to the “File name” field. Name the file “NewYorkState” in the “Save As” field and hit “Save.” You will not need to change anything else in the dialogue box. Click “OK” and the new layer will be added to the map.

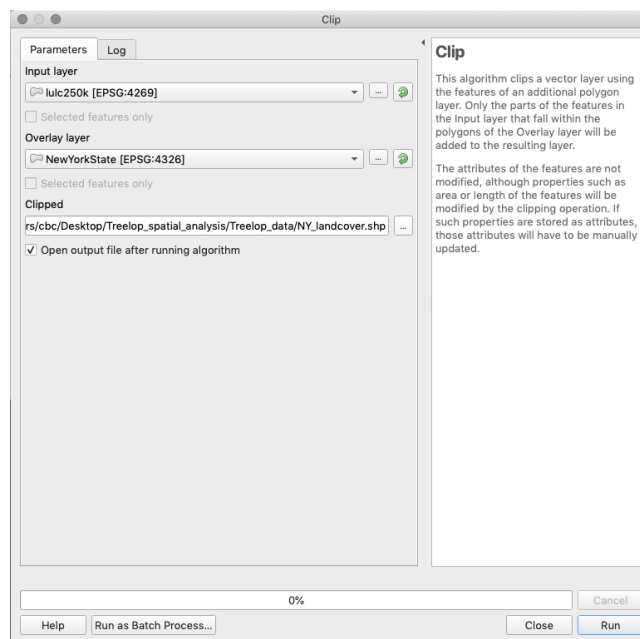
**Tip:** To avoid confusion with multiple similar layers, remove the “USA” layer by right-clicking on it and selecting “Remove Layer.”

**Note:** You should always use the ellipses button (“...”) to navigate to your working directory when creating a file and name the file in the “Browse” dialogue box to make sure your files are all located in one place.

SAVE the project

4. Next, we want to clip the land cover (“lulc250k”) and major roads (“ne\_10m\_roads”) layers to the extent of the New York State shapefile (we have already provided the “Treelop\_sightings” data for New York State sightings only). To choose New York State in the USA shape file, we only had to use the Attribute Table because “States” was one of many attributes in the file. The other files, however, do not have an attribute that would allow us to choose them by state and, therefore, we must use other tools in order to clip the data that we need.
- a. Clip the land cover layer: Go to Vector > Geoprocessing Tools > Clip. Since we are interested in clipping the landcover layer (“lulc250k”), select it as the “Input layer,” and the “NewYorkState” shapefile as the “Overlay layer.” Under “Clipped,” click on the ellipsis button, and select “Save to File.” Name the new file “NY\_landcover” and make sure that the file will be saved in your working directory (rather than the default output folder). Then make sure that you select “SHP files (\*.shp)” instead of the default “GPKG files (\*.gpkg)” in the dropdown menu at the bottom of the dialogue box (**GeoPackage files are in a standard, non-proprietary file format that is compatible QGIS but may not be readable by other GIS software**). Click “Save,” then “Run.” The clipped layer will automatically add to your Layers panel named as “Clipped.” Close the dialogue box and rename the “Clipped” layer to match the filename (“NY\_landcover”) by right clicking on the layer in the Layers panel and selecting “Rename Layer.”

**Tip:** if you add “.shp” to the filename it will automatically select shapefile as the output format.



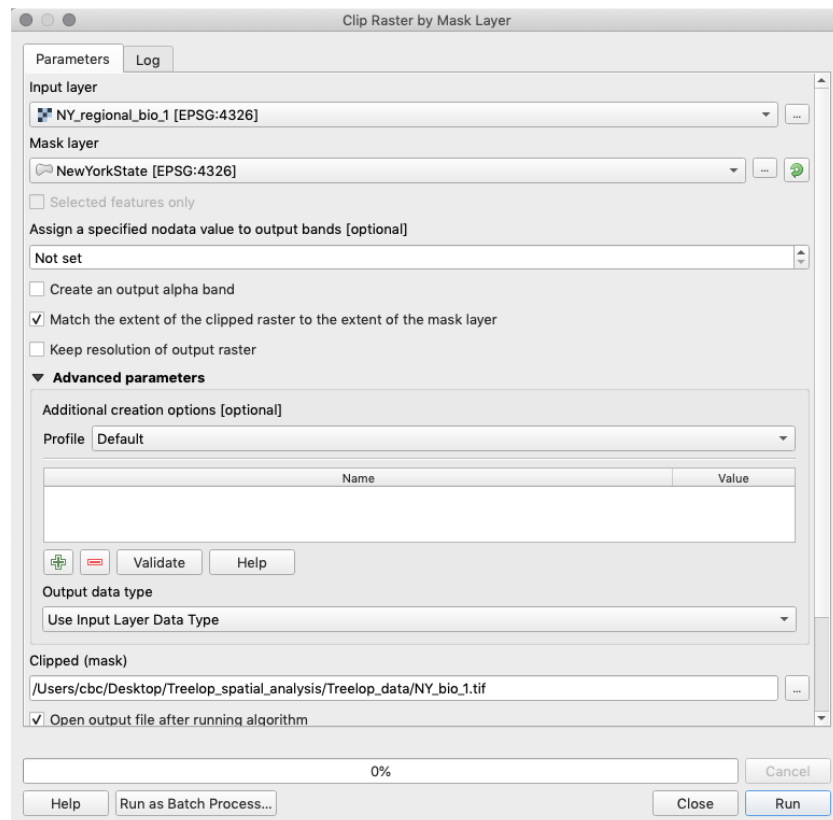


**Note:** If the “Vector” tab does not appear on the main toolbar, you will need to update all plugins. To update, go to Plugins > Manage and Install Plugins. Then click the button “Upgrade All.” The “Vector” tab should now be available.

- b. Repeat the clip procedure for the major roads layer (“ne\_10m\_roads”), naming it “NY\_roads”.
- c. After clipping the layers, select the check-boxes in the Layers panel to show only the newly clipped layers.

SAVE the project

5. Crop the Mean Annual Temperature raster to the New York State shapefile. All previous files were in vector format (see glossary). The temperature data is in raster format, so we need to use a different tool in order to clip the layer.
  - a. Go to Raster > Extraction > Clip Raster by Mask Layer. Set the “Input layer” to “NY\_regional\_bio\_1” and the “Mask layer” to “NewYorkState.” Click the ellipses next to “Clipped (mask)” and then “Save to File” and set the file directory. Name this new raster file “NY\_bio\_1.” Save in .tif format. Click “Run” (**Note:** you may need to scroll down the dialog box to view all options.)



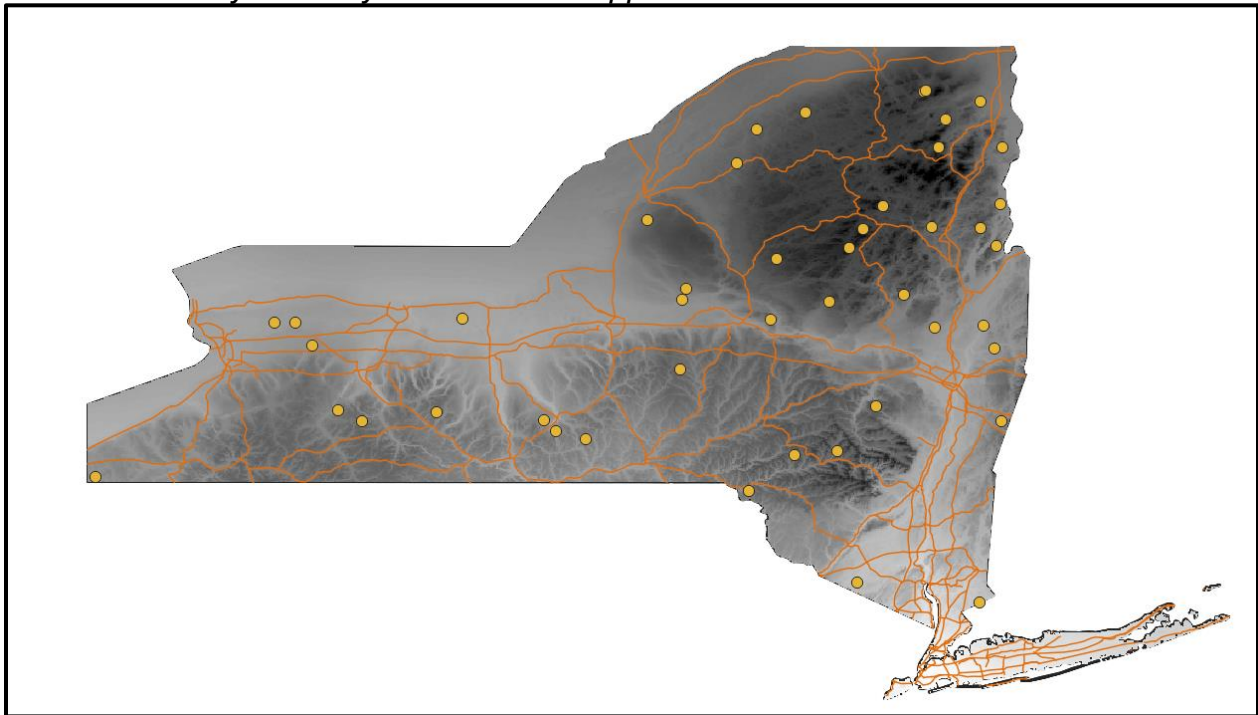
- b. Close dialogue box and rename the “Clipped (mask)” layer on the Layers panel to match the new raster file name (“NY\_bio\_1”).
- c. **Important:** the newly clipped raster will draw on your map over the previous NY clipped layer. You can reorder the layers in your Layers panel so they draw on the map in a different order (allowing you to see them all clearly). Simply drag and drop the newly clipped layers in the Layers panel to your preferred order.

**Tip:** To quickly find a tool by its name, you can activate the Processing Toolbox: go to View > Panels > Processing Toolbox Panel. Here you can type in a tool’s name and select.

SAVE the project

It should look something like this:

*This is with the NY\_bio1 layer on top of the landcover layer so you cannot see landcover, but if you were to uncheck the NY\_bio1 layer then you would be able to see the landcover layer. All layers should be clipped to the NY extent.*



## PART 4: ANALYZING THE DATA

### SPATIAL ANALYSIS

Now that we have layers that are specific to New York State, we can begin the analysis for our report. Until this point, we have been “cleaning up the data.” Now, we want to use spatial analysis to answer questions.

In this part of the exercise, we will walk you through the spatial analyses steps, but if you were doing this on your own, you’d want to consider the following: how can you use these data to give an assessment of potential survey locations? Given what we know about Treelopi, what are some processes you could use to analyze the data? How might the data interact and inform us about where we could find Treelopi?

### *Reflection*

Brainstorm and write down your thoughts on these prompts. If there are other questions you have, what are they?

Let's begin analyzing our data and create a visual output that can show us potential suitable habitats for Treelopi that match what we know about the species (refer to page 3 of this exercise).

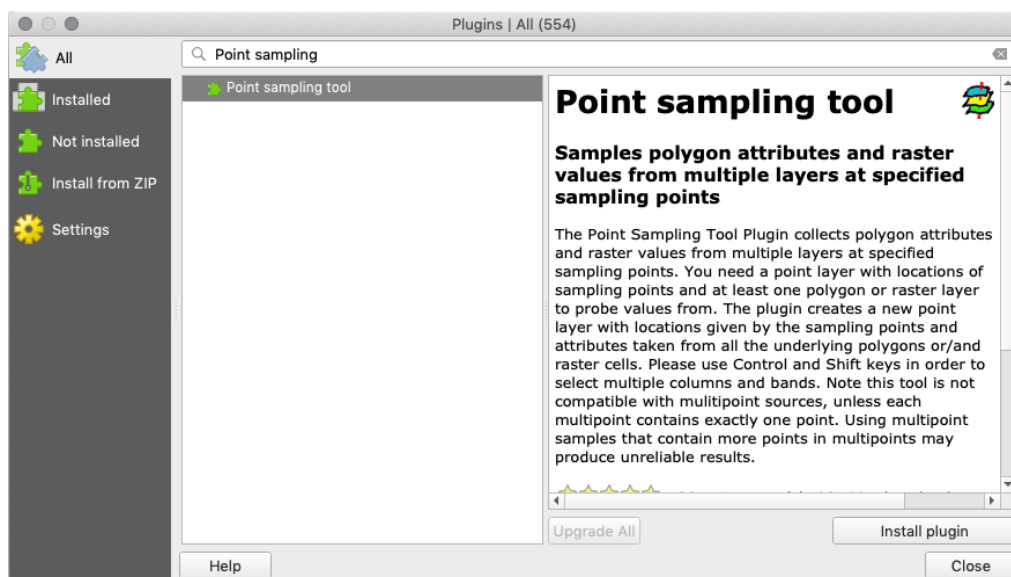
## STEPS:

- 4.1 Analyze climate data: Identify areas of New York State with suitable temperature for Treelopi.
- 4.2 Analyze habitat data: Identify areas of suitable habitat from land cover data.
- 4.3 Analyze road data: Determine areas of New York State that are not adjacent to major roads.
- 4.4 Identify the areas of overlap where there is suitable climate, habitat, and distance from roads.

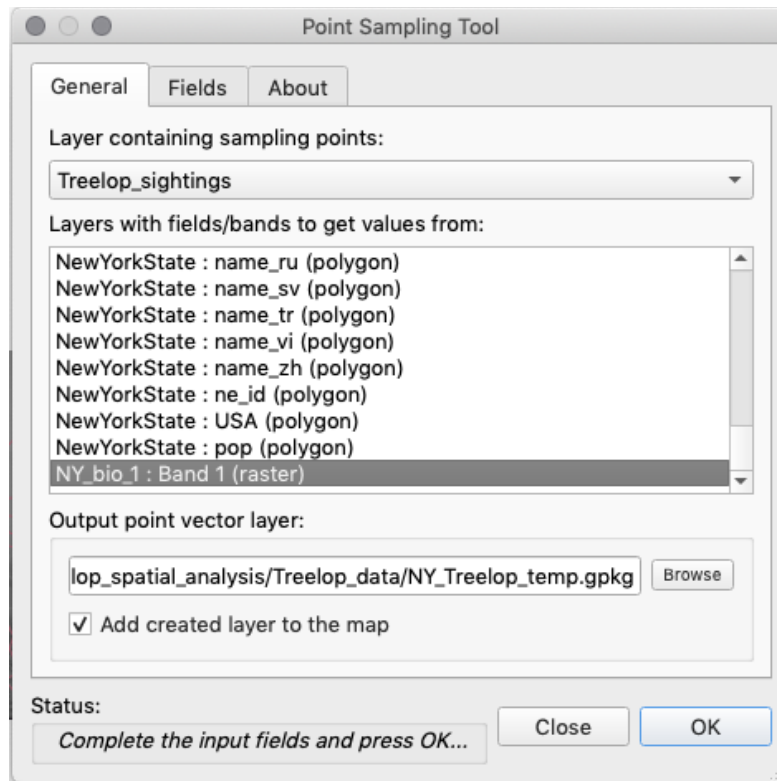
### 4.1 ANALYZE THE CLIMATE DATA

1. We want to determine the areas with suitable mean annual temperatures for Treelopi. To do so requires several steps. We have sightings of Treelopi across New York State, and we can make the assumption that the mean annual temperatures associated with those locations are suitable for the species. Therefore, the first step is to reduce the raster of mean annual temperature to show only those areas that are within the temperature range in which Treelopi have been recorded.

For this step, we'll need to install a plugin. Plugins are a way for QGIS developers to increase the utility of the software past the base QGIS download. Go to Plugins > Manage and Install Plugins. In the "All" tab on the left, search for the "Point sampling tool" then click "Install plugin." When the installation is over, close the dialogue box.



2. Next, we will extract the temperature values from “NY\_bio\_1” where Treelop has been seen.
  - a. Select the “Trellop\_sightings” layer in the Layers panel. Then, go to Plugins > Analyses > Point Sampling Tool. Select the “Trellop\_sightings” layer for “Layer containing sampling points” (this should be selected by default). For “Layers with fields/bands to get values from,” select “NY\_bio\_1.” Make sure the “Add created layer to the map” box is checked. Navigate to your chosen file directory by clicking “Browse” next to “Output point vector layer.” Save the file as “NY\_Trellop\_temp” with the extension of a GeoPackage file (.gpkg). Again, GeoPackage files are in a standard, non-proprietary file format that is compatible QGIS, but may not be readable by other GIS software. Here, we will use it only for this step. Click “OK.” Once the new layer has shown up on the Layers panel, click “Close” to close the dialogue box.



- b. Open the Attribute Table of “NY\_Trellop\_temp.” Examine the values of Mean Annual Temperature (“NY\_bio\_1” column) for each locality of a Treelop sighting. **Tip:** You can sort the values in ascending or descending order by clicking on the field (column) header.

NY\_Treelop\_temp.gpkg :: Features Total: 45, Filtered: 45, Selected: 0

	fid	NY_bio.1
1	30	2
2	32	2.79999995...
3	29	3.29999995...
4	18	4.59999990...
5	25	4.59999990...
6	38	4.69999980...
7	14	4.90000009...
8	19	5
9	15	5.199999809...
10	45	5.300000190...
11	7	5.40000009...
12	42	5.69999980...
13	4	5.69999980...
14	36	5.800000190...
15	39	5.90000009...
16	2	5.90000009...

Show All Features

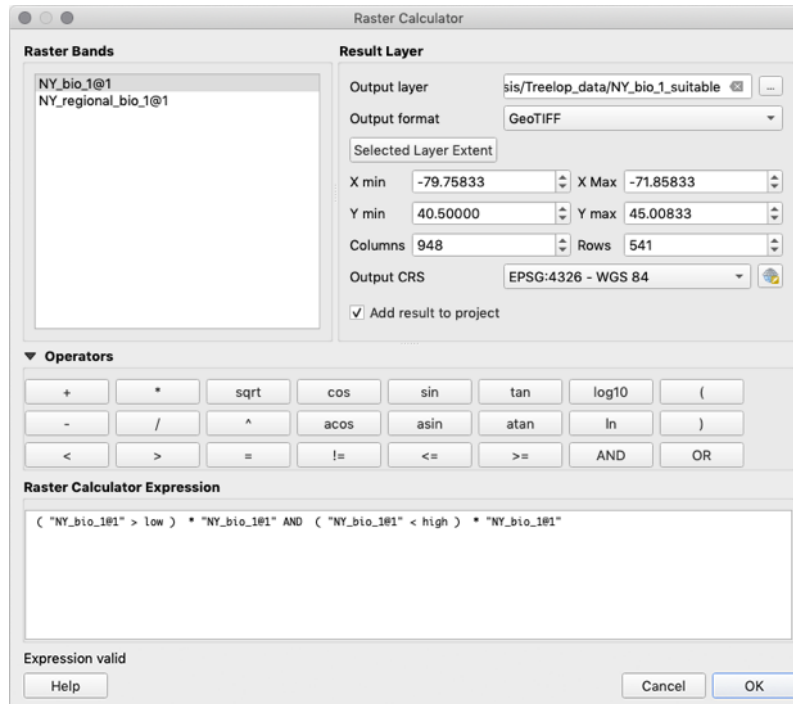
3. To determine what is the Treelop's preferred temperature range, we are only interested in the minimum and maximum mean annual temperatures observed for Treelopi. Sort the observed mean annual temperatures by clicking on the header in the Attribute Table. Click once to sort by the lowest mean annual temperature associated with a Treelop sighting and click the header again to see the highest mean annual temperature for a Treelop sighting. Our assumption is that areas outside of this range of temperature values are not suitable for Treelop.

**a. Question 2: Record the values. Round the minimum value down to the nearest full degree Celsius. Round the maximum value up to the nearest full degree Celsius. We will round down and up to ensure that all sighting points are included within the range while also erring toward a conservative estimate of suitable areas.**

- b. We assume that only areas within this range are suitable for Treelopi; areas where we should be focusing our efforts. We will remove what we assume to be the unsuitable areas from the map by setting values below the minimum and above the maximum temperatures where we have Treelop occurrences to "nodata" (meaning there is no data associated with areas outside this range of mean temperatures). The rest of the values (within the minimum and maximum range) will be set to "1" to denote that it is "suitable." This will leave all of the geographic areas with suitable temperature on the map, while removing areas not within the suitable temperature range.
- c. Using the Raster Calculator: Using a logic string and minimum and maximum values you observed in the Attribute Table from the step above, you will create a raster showing all areas of New York State that have

suitable (1) and unsuitable (0) temperatures (a binary raster mask). Open the Raster Calculator by going to Raster > Raster Calculator. Enter in the formula below in the “Raster Calculator Expression” field, with “low” being the recorded minimum mean annual temperature and “high” being the maximum. *DO NOT* type in symbols; use the operator buttons and the items under “Raster Bands” whenever possible.

(NY\_bio\_1@1 > low)\*NY\_bio\_1@1 AND (NY\_bio\_1@1 <  
high)\*NY\_bio\_1@1



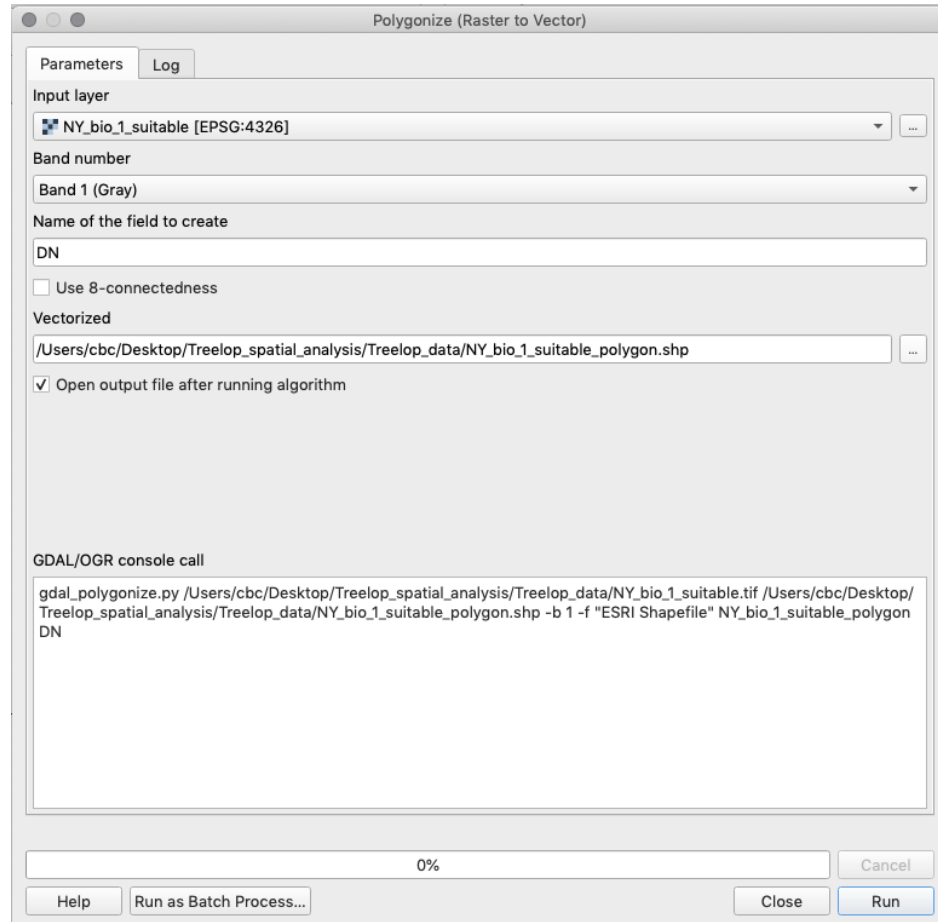
- d. Click the ellipsis button next to the “Output layer” field to select your working directory and name this file “NY\_bio\_1\_suitable” and click “Save.”
  - e. In the “Raster Calculator” dialogue box, make sure the “Output format” is in .tif format (“GeoTIFF”). **Note:** after filling out the expression, “Expression valid” should appear in bottom left corner of the calculator. If “Expression invalid” appears, you will not be able to perform the calculation. Click “OK” and view the areas of New York State that have unsuitable temperatures for Treelop (in black).
4. We now have, what we assume to be, the temperature-appropriate areas for Treelop in raster format. The rest of our data (i.e., land cover, roads), however, are in vector format. To be able to perform the future steps of our analysis, we will convert this binary raster—with suitable areas as the value 1 and unsuitable areas as the value 0—into a shapefile (specifically into a polygon<sup>5</sup>, which is a vector data type). **Note:** we could clip a raster by a vector layer and omit these



steps, but the computer processing time can be prohibitive, so we will do the conversion instead.

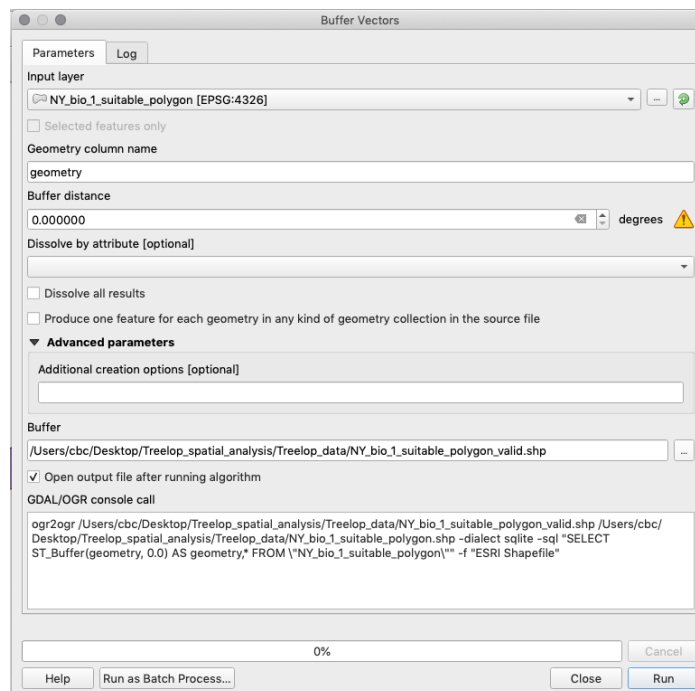
For this new polygon, we want to focus on only the suitable areas (denoted by the value 1), therefore, for simplicity, we will delete all the areas with a value of 0. After we convert the raster into a polygon, we will delete the geometries (rows in the attribute table) showing the value 0 (unsuitable area) and save the edited polygon.

- a. Go to Raster > Conversion > Polygonize (Raster to Vector). Input the “NY\_bio\_1\_suitable” layer for “Input layer.” Next to “Vectorized,” click the ellipsis button, “Save to File,” set file directory and name it “NY\_bio\_1\_suitable\_polygon.” Be sure to switch the file extension to a shapefile (.shp file) rather than a geopackage (.gpkg file). Click “Save” and then click “Run.”



- b. The layer will draw to your map and be automatically given the named “Vectorized” in the Layers panel. Rename it to match the file name (“NY\_bio\_1\_suitable\_polygon”).

- c. You will notice that this polygon does not look like the original raster of suitable areas. To fix this, delete polygons that are not equal to 1. Click on the polygon layer “NY\_bio\_1\_suitable\_polygon”, and view its Attribute Table. Click “Toggle editing mode” (“✎”). Click on the “DN” field (column) header to sort. Select all rows with a DN value not equal to 1. Click “Delete selected features” (“🗑”) to delete these entries. Click “Toggle editing mode” again, and save changes.
- d. It is not uncommon for GIS to have some issues when going through analyses (especially conversions). In this case, the software encounters a problem with the process of drawing the polygons that are on the edges and some of them are, therefore, considered “invalid.” There is a fix to this issue, which is forcing the software to redraw the polygons. To do this, we will make a new layer that takes the polygon and buffers the geometries by 0 (meaning no addition or subtraction). This will redraw the polygon with no invalid geometries.
  - i. Go to the Processing > Toolbox. In the search bar, type in “buffer” and under “Vector geometry,” click on “Buffer.” Select the correct “Input layer” (“NY\_bio\_1\_suitable\_polygon”), and set “Buffer distance” as 0. Save as a shapefile named “NY\_bio\_1\_suitable\_polygon\_valid” by clicking on the ellipses next to “Buffer.” Click “Run.”



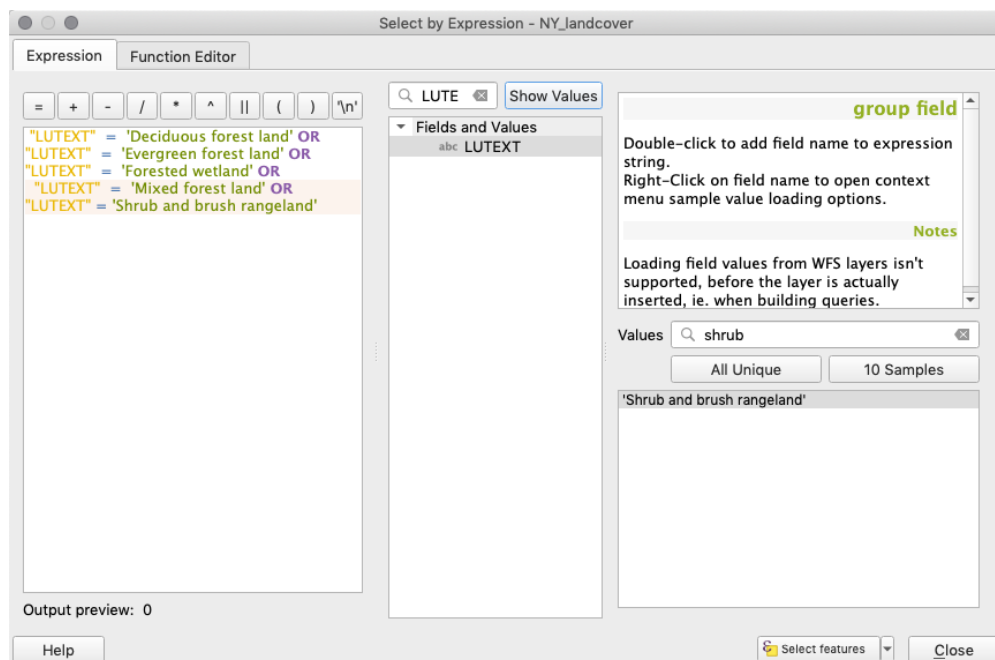
- ii. Close the dialogue and rename the new layer “Buffer” to match the file name (“NY\_bio\_1\_suitable\_polygon\_valid”).

SAVE the project

## 4.2 ANALYZE THE HABITAT DATA

In the previous steps, we created a shapefile (vector format) layer of all the areas in New York State that have suitable temperature ranges for Treelopi. We also know, from the few natural history studies of this species, that Treelopi also prefer certain kinds of habitats. In this next step, we extract only those areas from our land cover layer that are preferred and create a new shapefile. To do so, we will use logical operators and expressions to select attributes that match what we know about Treelopi. Logical operators such as “and,” “or,” “not,” or “equal to” can be used to select attributed that fulfill our needs.

1. Navigate to the Attribute Table of the land cover layer, “NY\_landcover.” Explore the Attribute Table, especially the column labeled “LUTEXT” (stands for “Land Use Text”). This column shows land cover type for each geometry (polygon feature) in New York State. Given what we know about Treelopi (e.g., prefer forested areas), which land cover types do you think we should include?
2. Click the “Select features using an expression” icon. Considering that Treelopi prefer natural areas (non-agricultural lands), are forest generalists, and are often associated with woody vegetation, write an expression using “OR” that satisfies the conditions where the field “LUTEXT” is either “Deciduous forest land,” “Evergreen forest land,” “Forested wetland,” “Mixed forest land,” or “Shrub and brush rangeland.” Remember to use the search bar in the middle panel of the dialogue box to select your column of interest (double-click), use the search bar and “All Unique” button to select your variable of interest for that field, and use the expression buttons above the “Expression” box while writing your expression. Click “Select features.” Click “Close” to close the dialogue box. Notice these features are now selected on the map of land cover.



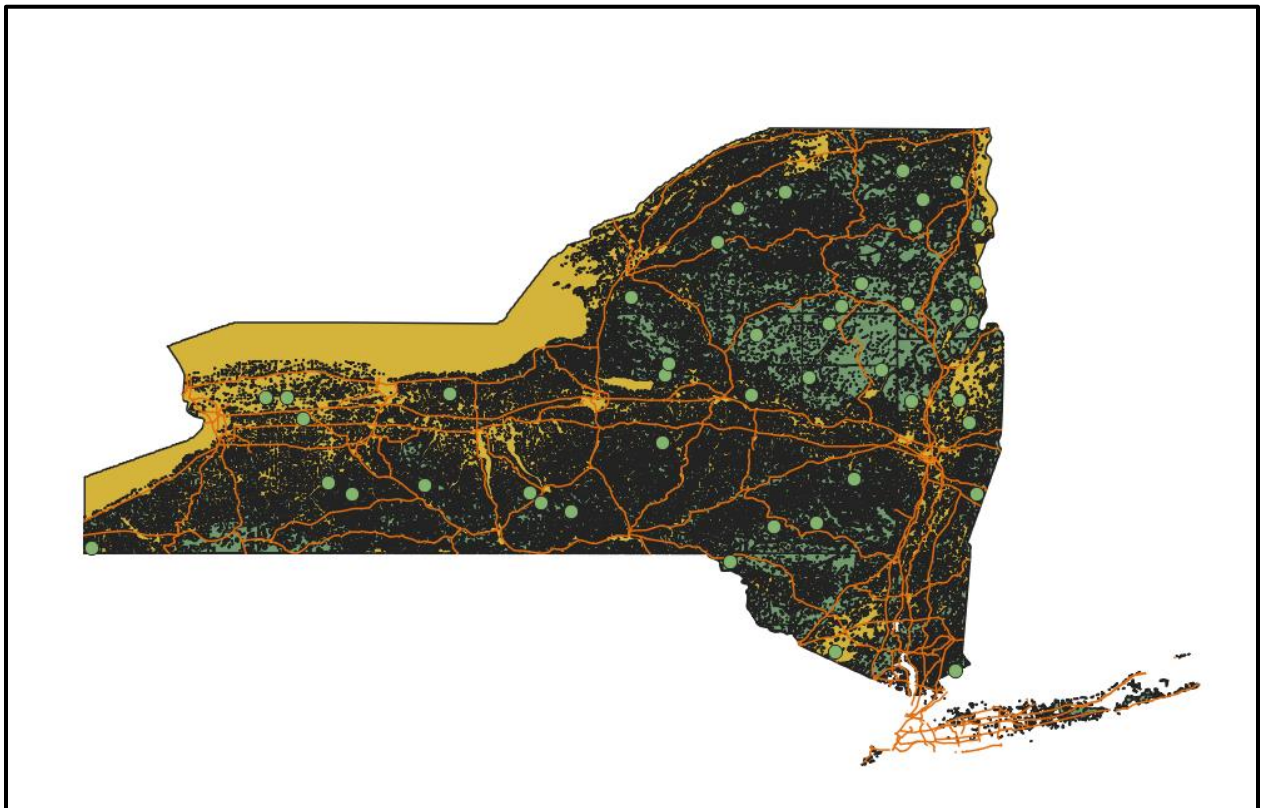
3. To create a new shapefile from this new selection, right-click on the “NY\_landcover” layer in the Layers panel and select “Export,” then “Save Selected Features As.” Export as a new shapefile named “Treelop\_preferable\_landcover” (be sure to save in your working directory).  
**Note:** Often in QGIS, processes have more than one solution. Exporting selected features in this way will result in the same outcome as deleting features that are NOT of interest (as in step 4.1.4).

### *Reflection*

What are some ways this important alternative might be useful to know?

SAVE the project

*This is what your map should look like now (assuming you have your stacked in the same way):*



### 4.3 ANALYZE THE ROAD DATA

We often use latitude and longitude, a 3-dimensional spherical coordinate system, to define the location of points on the Earth (which is an oblate spheroid), but in many cases we transform, or project, these coordinates into a 2-dimensional grid so it is easier to process. QGIS uses a Coordinate Reference System (CRS)<sup>6</sup> to define the coordinate system used for each layer. For cases when you display layers that have a different CRS than the one defined for your QGIS project, those layers will be “projected on the fly” to transform them so they overlay with the other layers in your project. For visualization, on the fly projections are quite helpful, but if you are going to do analyses—like measuring areas or distances between objects—it is better for you to choose an appropriate CRS that preserves spatial qualities such as shape, area, direction, and distance to meet the needs of your project.

In the following steps, you will convert the data into a CRS that preserves area because we will need to measure area in future steps. Because there are many Coordinate Reference Systems to choose from, we will tell you which CRS is best applicable to the data we have in the region we are studying. For more information about CRS, consider reviewing the Coordinate Reference Systems chapter of *A Gentle Introduction to GIS* ([docs.qgis.org/3.4/en/docs/gentle\\_gis\\_introduction/index.html](https://docs.qgis.org/3.4/en/docs/gentle_gis_introduction/index.html)).

1. First, we will need to convert your project from a spherical model (3-dimensional) to a planar model (2-dimensional). The Universal Transverse Mercator (UTM)<sup>7</sup> projection is a planar model measured in meters, and QGIS will automatically convert the units.
  - a. Go to Project > Properties. Navigate to the “CRS” tab in this window. In the “Filter” search field, type “EPSG:26918 ” and then under “Coordinate reference systems of the world,” select “NAD83 / UTM zone 18N EPSG:26918” and then click “OK.” **Note:** NAD83 / UTM zone 18N stands for the North American Datum 1983 and Universal Transverse Mercator zone 18 North. See the region it covers in the small map on the right part of the dialogue box.

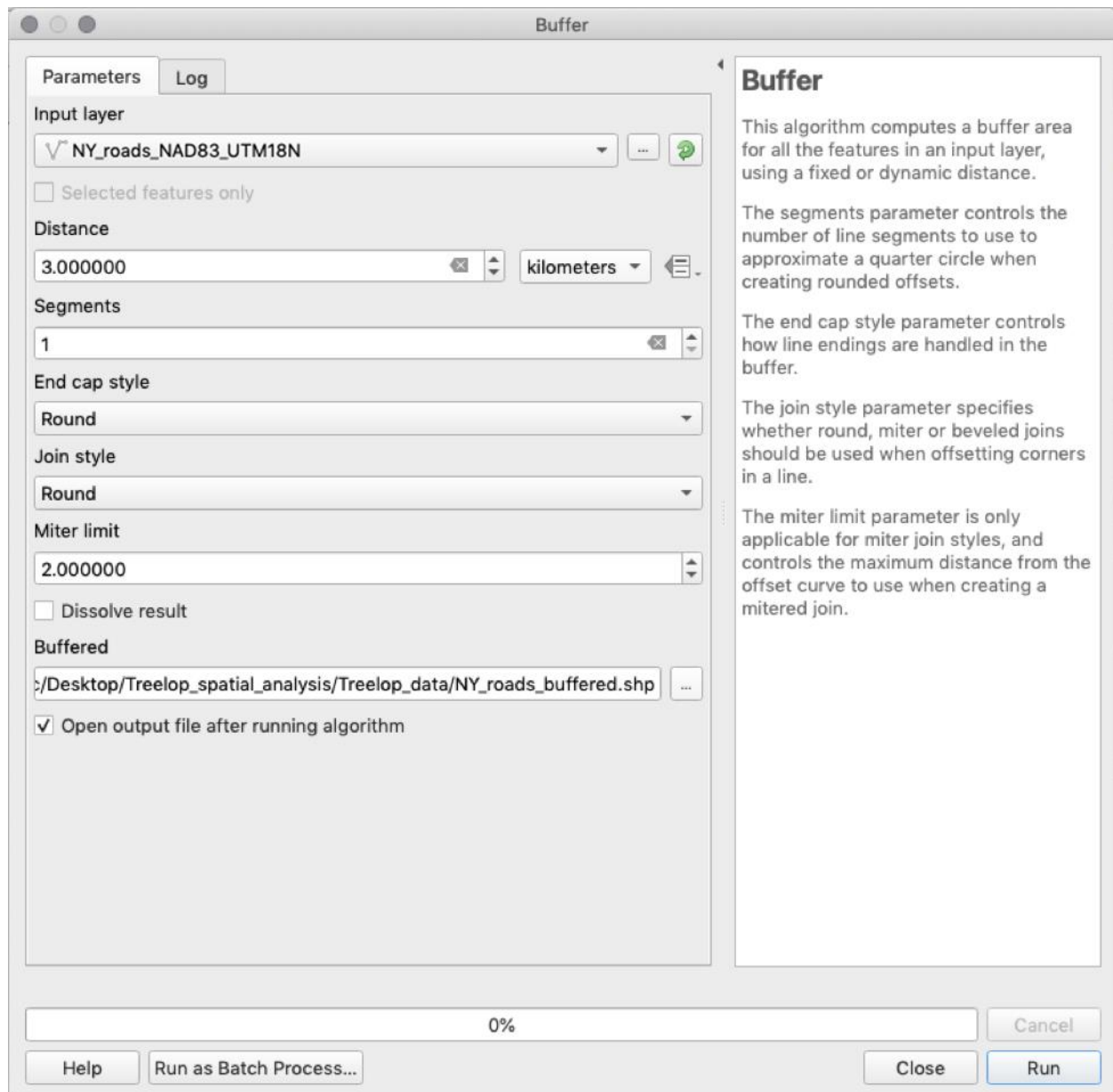
**Question 3: We know that a Treelop is a shy, reclusive creature. What do you think are some characteristics of roads that would make the area around them unsuitable?**

For this analysis, we have data for the locations of major roads and can use those data to narrow down our potential survey locations by omitting sub-optimal areas. We chose a 3 km distance from roads because other (fictional) studies on this species showed that the negative effects of roads extended into natural areas up to that distance. **Note:** Although Treelopi are not likely to be near busy roads, they *have* been observed near these roads (although rarely). If you were doing a similar analysis on your own question, you may similarly find references or expert opinion on what distance threshold to use.

2. In GIS, a common tool to create an area that surrounds a feature (e.g., road) by a certain distance is to create a buffer. To do this for our analysis, we will start by making a new polygon layer showing a 3 km buffer around the roads layer.

**Consider:** How is this step different from the previous ones (how is the data different)?

- a. First, because we will be doing analysis that requires measurement of distance in km, we will need to project the “NY\_roads” layer to the same CRS that we set to the project before. In the “Toolbox” (Processing > Toolbox), type in “reproject” and under “Vector general,” double-click on “Reproject layer.” Add “NY\_roads” as the “Input layer” and the same CRS above (“NAD83 / UTM zone 18N EPSG:26918”) for “Target CRS.” Under “Reprojected,” save this layer as a shapefile named “NY\_roads\_NAD83\_UTM18N” in your working directory. Hit “Run,” and then rename the layer in the Layers panel to match the file name.
- b. To buffer the new “NY\_roads\_NAD83\_UTM18N” layer, navigate to Vector > Geoprocessing Tools > Buffer. Read the right sidebar to learn more about the “buffer” function. Select the correct “Input layer,” and under “Distance,” select 3, and make sure the unit next to the distance text box is “kilometers.” Under “Segments,” select 1. Under “Buffered,” save this as a shapefile named “NY\_roads\_buffered” hit “Run.”



SAVE the project

#### 4.4. IDENTIFY THE AREAS WHERE THE LAYERS OVERLAP

1. Take some time to clean up your Layers panel. Drag the “NY\_roads\_buffered,” “Treelop\_preferable\_landcover,” and “NY\_bio\_1\_suitable\_polygon\_valid” to the top of the Layers panel.

We now have our discrete data layers, each adapted to reflect where Treelopi might appear or not based on what we know about their behavior and needs. Our next step in our analysis, then, would be to see how all of these data combined can help us create a map of potential hotspots where Treelopi might be found: areas of New York State that



have suitable temperature, adequate land cover types, and are far enough from roads. For each of the steps below, be sure to read the description of the tool, which will appear on the right-side panel within the dialogue box.

2. Calculate the intersection between the suitable land cover and suitable temperatures. Go to Vector > Geoprocessing Tools > Intersection. As the "Input layer," use the vector of suitable temperature ("NY\_bio\_1\_suitable\_polygon\_valid"), and set the "Overlay layer" as preferable land cover ("Treelop\_preferable\_landcover"). Under "Intersection," select your working directory and save as a shapefile named "LC\_temp\_intersection." Click "Save" and then click "Run."

Close the dialogue box and rename the "Intersection" layer to match the file name.

SAVE the project

3. Remove areas near roads using the "Difference" tool. Go to Vector > Geoprocessing Tools > Difference. Use the "LC\_temp\_intersection" layer as the "Input layer," and use the "NY\_roads\_buffered" as the "Overlay layer." Under "Difference," save as a shapefile named "LC\_temp\_roads\_difference" in your working directory. Remember to rename the layer in the Layers panel.

SAVE the project

If you look at the Attribute Table of this layer, you'll notice that all of the original landscape geometries remain. We are interested in the total suitable area in New York State, so we need to dissolve these geometries into one. Go to Vector > Geoprocessing Tools > Dissolve. Read the right sidebar to learn more about the "dissolve" function. Use the most recent layer ("LC\_temp\_roads\_difference") as the "Input layer." Under "Dissolved," save this layer as a shapefile "Treelop\_habitat\_NY" in the working directory. Click "Save" and then click "Run." This process will take a few minutes. Once complete, rename the layer and check the Attribute Table; there should only be one row. Please note that the attribute table for this output will show data from the first row only and the contents of the row do not represent the attributes of the layer anymore.

4. Your instructor may ask you to take a screenshot of the map of your "Dissolved" Layer (all other layers not selected) or to make a printable map (one with a title, scale bar, legend, compass arrow, data source, author, date, etc.). These steps are not included here but may be provided by your instructor.


SAVE the project

## PART 5: APPLY YOUR UNDERSTANDING OF THE RESULTS TO A CONSERVATION QUESTION

### 5.1 CONSIDER YOUR RESULTS AND THE DATA

You have completed a preliminary map from your analysis. Congratulations! Now, let's think through what this map means for your assignment to prioritize where the small non-profit conservation organization should survey for Treelopi in New York State.

If the organization was going to survey all of the land on this map, how many square kilometers would they have to survey?

To calculate the area of habitable land in New York State, open the Attribute Table of "Treelop\_habitat\_NY." Click on the "Toggle editing mode" to make the layer editable. Then, click "Open field calculator" (). Create a new field with "Output field name" as "area\_sqkm". The expression for calculating the area of every row in an attribute table is simple: under the Search text box, select Geometry-> \$area. However, this will calculate the area in map units. Since we are using the Universal Transverse Mercator projection, the units are currently in meters. Since we want kilometers, we need to multiply the area by 0.000001. This will show the total area (in km<sup>2</sup>) of habitable land in New York State that is suitable for Treelopi.

What is the total area of our estimated "optimal" habitable land for Treelopi?

Examine the results of your analysis (i.e., map and calculated area). What do you think some limitations would be for surveying for Treelopi?

Given these limitations, would you feel comfortable recommending a priority area to survey? Why or why not?

What do you think are some ways to further narrow down priority areas for surveys? Go back to your response to Question 1 on page 5. Were there data types listed there that might help you? What else would you like to know that could alter your priority areas for surveys?

Based on the assumption that all of the habitat on your map is suitable and your understanding that this species is considered rare, do you think that Treelopi are limited by suitable habitat? Hypothesize what else might be limiting their numbers.

A brand-new study is published on Treelopi. The results suggest that Treelopi need large contiguous habitat (at least 20 km<sup>2</sup>) for breeding. How can we use this new information to better understand their suitable habitats? Explain how you might further refine your map through spatial analysis to incorporate this information. (Do not list step by step instructions of how you would accomplish this with the software rather a summary of what those steps would accomplish.)

Have you considered the limitations of the provided data? Let's review the data listed in Table 1.

Political boundaries are mainly stable and because we are focused on New York State, this data doesn't appear to have a limitation but rather defines our scope.

Temperature data in this exercise is only provided as mean annual temperature. How might mean annual temperature mask variation in temperatures throughout the year? What other measures of temperature (or other climate variables) might be important? Why?

If you navigated to the source of the data ([worldclim.org](http://worldclim.org)), you'd see that the mean annual temperature is calculated from a 30-year period (from 1970–2000). What might have changed since then? (**Note:** averages for climate data are normally obtained by looking at a 30-year interval of years.)

Likewise, if you look at the metadata associated with the land use/land cover dataset, you'd notice that it is from 1990. How might things have changed in the past ~30 years? How could updated data source help narrow down your survey area? (**Note:** up to date data layers may or may not be available. Sometimes more recent data are only available after a fee is paid. Depending on your project

needs and budget, it might be worthwhile to spend money on more up to date datasets.)

Roads data is updated to 2018, but only contains Interstates, State, and Federal roads. If you thought that Treelopi were averse to all types of roads, you might want to find a data source that included county, town, and private roads.

Treelopi sightings data has no metadata associated with it. Unfortunately, it is common that some data sources will not provide helpful information beyond the spatial information (e.g., only GPS coordinates). What sort of data do you wish you knew about these sightings? How would it be helpful for supporting your analysis and recommendations?

## 5.2 REPORTING BACK TO THE CONSERVATION ORGANIZATION

Based on your work and responses above, write out a brief overview of what you did to get the preliminary map (don't list steps, rather summarize the final output). Explain your recommendation for where to do surveys, and/or what other steps should happen before a final survey recommendation is set.

## GLOSSARY

Glossary terms are derived from: QGIS Development Team. 2019. QGIS Geographic Information System. Open Source Geospatial Foundation Project [qgis.osgeo.org](http://qgis.osgeo.org). See “A Gentle Introduction to GIS” [docs.qgis.org/3.4/en/docs/gentle\\_gis\\_introduction/](http://docs.qgis.org/3.4/en/docs/gentle_gis_introduction/) for more helpful background information.

---

<sup>1</sup> *Vector*: a type of spatial data made up of vertices and paths which uses geometry to represent real-world features in GIS. Vector data are presented as points, lines, and polygons (areas).

<sup>2</sup> *Raster*: a type of spatial data made up of grids of pixels (cells), often used for data that is continuous across an area and cannot easily be divided into vector features, e.g., the backdrop behind vector data.

<sup>3</sup> *Geometries*: points, lines, and polygons that represent a spatial feature.

<sup>4</sup> *Expression*: a way to manipulate attribute value, geometry, and variables in QGIS to perform functions like changing the geometry style, the position of the label, the height of a layout item, etc.

<sup>5</sup> *Polygon*: a closed-shape vector object defined by X and Y coordinate pairs.

<sup>6</sup> *Coordinate Reference System (CRS)*: uses coordinates to define how a two-dimensional GIS map is related to real places on the three-dimensional earth. Deciding on which map projection and CRS to use depends on many factors, including the regional extent of the area you want to work in, the analysis you want to do, and the availability of data.

<sup>7</sup> *Universal Transverse Mercator (UTM)*: a global map projection which assigns coordinates to locations on the surface of the Earth. The UTM CRS divides the globe into 60 equal zones (numbered 1 to 60) starting at the international date line. Each 6 degrees wide in longitude. Each zone is further divided into zones north and south of the equator.