Ecological Niche Modeling- The Evolution of a Species

This unit was designed to give students the chance to accurately depict the current and future distributions of a species through the use of ecological niche modeling software. Students will observe how a specific population of their choice could evolve over the next 35 years. Additionally, students will actively participate in a citizen science project of their choosing in order to potentially increase the accuracy of future scientific predictions.

Ecological Niche Modeling- The Evolution of a Species

Center for Precollegiate Education and Training
UNIVERSITY of FLORIDA

Author: Jonathan Benskin – Boca Raton Community High School

Please direct inquiries to the Center for Precollegiate Education and Training at cpet@cpet.ufl.edu.

©2014 University of Florida

Center for Precollegiate Education and Training

PO Box 112010 • Yon Hall, Room 331

Gainesville, FL 32611

Phone 352.392-2310 • Fax 352.392-2344

Contents

Author's Note

It has become a yearly tradition that I spend part of my summer at the University of Florida working with staff and professors in the development of new high school curricula. Over the last couple of years the focus of my curricula has been molecular biology, but this year has ushered in a change of direction. After spending a week learning about evolution from some of the top professors in the nation I have decided to venture into the realm of ecology and population change. While listening to a presentation from the Soltis Lab, inspiration slowly began to sneak in. Part of their presentation was the use of software to model the current and future predicted distribution of a species. I found myself enjoying the simplicity of the software that made powerful predictive models of what the future of Florida might look like.

We all know that the climate is changing, but very few times do we think about how this change will affect other species directly around us. This module was created to show students (and teachers) how the distribution of a species could change within their lifetimes. Making predictions using this software was exciting, but while writing this module I often found myself frustrated with the lack of species data available. Because of this, I decided to incorporate "citizen science" into the last portion of the module in hopes of raising scientific awareness and empowering students to have a voice in the scientific community. These students will be creating policies in the future and it is my hope that they begin taking action now to preserve the beautiful state that I call "home".

I would like to thank Julie Bokor and the Center of Precollegiate Education and Training at the University of Florida for giving me the opportunity to write this module and creating an environment for teachers that fosters investigation and growth. Additionally, I would like to thank Dr. Charlotte Germain-Aubrey and the Soltis Lab at the University of Florida for the inspiration to write this module, the technical assistance, and the files used in the included activities.

Introduction

The climate is changing. Although it is vital that students understand how and why this is occurring, the focus of this module is to look at how this change will impact different species found in the state of Florida. The software used in this module is capable of predicting the current and future distribution of a specific species allowing students to visually see how climate change impacts the ecology around us. The depictions that are created offer a powerful look into the near future of our state that our students will live to see.

Although the software provides a powerful tool in the classroom and to researchers, it is very simple to use. All of the needed files can be found in one .zip file (see page 18) and no computer installation is required. I recommend running through the procedure (for “Lesson 1” and “Lesson 2”) prior to using it with students in order to familiarize yourself with its usage. Screenshots in the “Procedure” sections are provided for many of the steps and should offer some assistance. Although this software has other advanced features that you are welcome to explore, they will not be needed for this module.

As an actual high school biology teacher I do recognize that some students just really won’t care about how the distribution of a species could change in the future. Despite this fact, I want all students to take some sort of action that benefits the available body of knowledge. Citizen science projects offer a diverse range of topics that a student can get involved with and hopefully become passionate about. Introducing them to some of these ventures and having them brainstorm new project ideas could lead to some very creative solutions to some very complex problems.

Tips about this Curriculum

Lesson Plan Format: All lessons in this curriculum unit are formatted in the same manner. In each lesson you will find the following components:

KEY QUESTION(S): Identifies key questions the lesson will explore.
OVERALL TIME ESTIMATE: Indicates total amount of time needed for the lesson, including advanced preparation.
LEARNING STYLES: Visual, auditory, and/or kinesthetic.
VOCABULARY: Lists key vocabulary terms used and defined in the lesson. Also collected in master vocabulary list.
LESSON SUMMARY: Provides a 1-2 sentence summary of what the lesson will cover and how this content will be covered. Also collected in one list.
STUDENT LEARNING OBJECTIVES: Focuses on what students will know, feel, or be able to do at the conclusion of the lesson.
STANDARDS: Specific state benchmarks addressed in the lesson. Also collected in one list.
MATERIALS: Items needed to complete the lesson. Number required for different types of grouping formats (Per class, Per group of 3-4 students, Per pair, Per student) is also indicated.
BACKGROUND INFORMATION: Provides accurate, up-to-date information from reliable sources about the lesson topic.
ADVANCE PREPARATION: This section explains what needs to be done to get ready for the lesson.
PROCEDURE WITH TIME ESTIMATES: The procedure details the steps of implementation with suggested time estimates. The times will likely vary depending on the class.
ASSESSMENT SUGGESTIONS: Formative assessment suggestions have been given. Additionally, there is a brief summative assessment (pre/post-test) that can be given. Teachers should feel free to create additional formative and summative assessment pieces.
EXTENSIONS: (ACTIVITIES/LITERATURE) There are many activities and reading sources available to augment and enhance the curriculum. They have been included. If you find additional ones that should be added, please let us know.
RESOURCES/REFERENCES: This curriculum is based heavily on primary sources. As resources and references have been used in a lesson, their complete citation is included as well as a web link if available. All references and resources are also collected in one list.
STUDENT PAGES: Worksheets and handouts to be copied and distributed to the students.
TEACHER PAGES: Versions of the student pages with answers or the activity materials for preparation.

Collaborative Learning: The lessons in this curriculum have been developed to include many collaborative learning opportunities. Rather than presenting information in lecture format and teacher driven, the activities involve the students in a more engaged manner. For classrooms not accustomed to using collaborative learning strategies, have patience. It can be difficult to communicate instructions, particularly for students who are visual learners. For these students, use of visual clues such as flowcharts and graphics can help them understand how they are to move to different groups.

Groups: Some of these lessons are carried out in groups. While it isn't necessary for students to remain in the same groups the entire unit, if they work well together, it may foster students to think deeper as they are comfortable with their teammates and willing to ask questions of each other.

Inquiry-based: The lessons in the curriculum invite students to be engaged and ask questions. They work through background information in a guided fashion, but are challenged to think beyond what they have read or done. The teacher serves as the facilitator in these activities, not the deliverer of information.

Technology: Lessons have been written to be mindful of varying availability of technology in schools and homes. Some of the lessons would be very well suited to online environments and if your students are able, you might wish to engage in some of the technology modifications.

Content: Often we teach in a manner that is very content heavy. With high-stakes testing the norm, students are pushed to memorize and regurgitate numerous isolated facts. There is so much content that must be covered in a biology class, for example, that often it is difficult to synthesize those discrete facts into a compelling context or a story. This unit provides that opportunity: to take concepts learned such as muscles have a lot of glycogen or DNA codes for RNA, and put them in the context of disease. The lessons aren't designed to teach students what lysosomes do or transcription is, but rather *why* these ideas are important and *how* they can be used by researchers.

Implementation notes: This curriculum should be modified and adapted to suit the needs of the teacher and students. To help make implementation easier in this first draft, notes have been included in lessons as needed.

Extensions: Possible/recommended extension activities that can be completed in addition to the written curriculum.

Science Subject: Biology – AP, AICE, IB

Grade and ability: 10-12 grade Advanced Placement/AICE/IB Biology.

Science concepts: The overarching concepts within a unit.

Lesson Summaries

LESSON ONE: "THE FUNDAMENTALLY ECOLOGICAL NICHE"
Students will be introduced to Ecological Niche Modeling (ENM) through the use of "Maximum Entropy" habitat modeling. This software uses multiple layers of environmental data to estimate current and future niches/distributions of specific organisms. Preselected species and data information will be used to walk students through the basic usage of this software.

LESSON TWO: "HOW TO GET NICHE, FAST!"
Students will choose a specific Florida species in order to investigate distribution changes through the next thirty-five years. Data collection through the use of online databases will be used to compile coordinate information that can then be used in the ENM software to make these predictions.

LESSON THREE: "SCIENTIFIC CITIZEN THROUGH CITIZEN SCIENCE"
Students will be introduced to "citizen science" programs in order to see how their contributions can be used by researchers. Additionally, students will brainstorm possibly new citizen science programs.

Lesson Sequencing Guide

All lessons are based on a 55 minute class session:

	DAY 1	DAY 2	DAY 3
WEEK 1	**Lesson 1**	**Lesson 2**	**Lesson 3**

Vocabulary

ENM software- Ecological Niche Modeling software. Software that uses environmental information and georeferenced species data to model the possible distribution of an organism.

Eltonian niche- Defines a "niche" as the habitat in where the organism lives along with the specific adaptations that organism has in order to live and reproduce within that habitat.

Fundamental niche- Any location in the environment in which the organisms can survive.

Grinnellian niche- Defines "niche" as a "community-centric" concept which must take factors such as competition and predation into account.

Hutchinsonian niche- Defines "niche" as an "abstract mapping of population dynamics onto an environmental space, the axes of which are abiotic and biotic factors that influence birth and death rates".

Macroevolution- Change to the frequency of an allele within a population over time

Maximum Entropy Species Distribution Modeling Software- Software that "learns" as it processes data in order to make an accurate prediction.

Niches- Any location in the environment in which the organisms can survive. A fairly broad term that can have multiple meanings (See above definitions or "Lesson 1").

Realized niche- The location in which an organism is actually found.

Next Generation Sunshine State Standards – Science

Benchmark	Lesson		
	1	2	3
SC.912.L.15.1 Explain how the scientific theory of evolution is supported by the fossil record, comparative anatomy, comparative embryology, biogeography, molecular biology, and observed evolutionary change.	X	X	
SC.912.L.15.14 Discuss mechanisms of evolutionary change other than natural selection such as genetic drift and gene flow.	X		
SC.912.L.15.3 Describe how biological diversity is increased by the origin of new species and how it is decreased by the natural process of extinction.		X	
SC.912.L.17.13 Discuss the need for adequate monitoring of environmental parameters when making policy decisions.		X	
SC.912.L.17.15 Discuss the effects of technology on environmental quality.	X	X	X
SC.912.L.17.4 Describe changes in ecosystems resulting from seasonal variations, climate change and succession.	X	X	
SC.912.L.17.8 Recognize the consequences of the losses of biodiversity due to catastrophic events, climate changes, human activity, and the introduction of invasive, non-native species.	X	X	
SC.912.N.1.1 Define a problem based on a specific body of knowledge, for example: biology, chemistry, physics, and earth/space science, and do the following: 1. Pose questions about the natural world 2. Conduct systematic observations, (Write procedures that are clear and replicable. Identify observables and examine relationships between test (independent) variable and outcome (dependent) variable. Employ appropriate methods for accurate and consistent observations conduct and record measurements at appropriate levels of precision. Follow safety guidelines). 3. Review what is known in light of empirical evidence, (Examine whether available empirical evidence can be interpreted in terms of existing knowledge and models, and if not, modify or develop new models). 4. Plan investigations, (Design and evaluate a scientific investigation). 5. Use tools to gather, analyze, and interpret data (this includes the use of measurement in metric and other systems, and also the generation and interpretation of graphical representations of data, including data tables and graphs), (Collect data or evidence in an organized way. Properly use instruments, equipment, and materials (e.g., scales, probeware, meter sticks, microscopes, computers) including set-up, calibration, technique, maintenance, and storage). 6. Pose answers, explanations, or descriptions of events, 7. Generate explanations that explicate or describe natural phenomena (inferences), 8. Use appropriate evidence and reasoning to justify these explanations to others, 9. Communicate results of scientific investigations	X	X	X

Benchmark	**Lesson**		
	1	**2**	**3**
SC.912.N.1.3 Recognize that the strength or usefulness of a scientific claim is evaluated through scientific argumentation, which depends on critical and logical thinking, and the active consideration of alternative scientific explanations to explain the data presented.	**X**	**X**	**X**
SC.912.N.1.4 Identify sources of information and assess their reliability according to the strict standards of scientific investigation.		**X**	**X**
SC.912.N.3.5 Describe the function of models in science, and identify the wide range of models used in science.	**X**	**X**	
SC.912.N.4.1 Explain how scientific knowledge and reasoning provide an empirically-based perspective to inform society's decision making.		**X**	**X**

Advanced Placement Biology Learning Outcomes & Science Practices

Outcomes & Practices	Lesson		
	1	2	3
LO 1.1 The student is able to convert a data set from a table of numbers that reflects a change in the genetic makeup of a population over time and to apply mathematical methods and conceptual understandings to investigate the cause(s) and effect(s) of this change.	X	X	
LO 1.2 The student is able to evaluate evidence provided by data to qualitatively and quantitatively investigate the role of natural selection in evolution.	X	X	
LO 1.3 The student is able to apply mathematical methods to data from real or simulated populations to predict what will happen to the population in the future.	X	X	
LO1.4 The student is able to evaluate data-based evidence that describes evolutionary changes in the genetic makeup of a population over time.	X	X	
LO 1.5 The student is able to connect evolutionary changes in a population over time to a change in the environment.	X	X	
LO 1.9 The student is able to evaluate evidence provided by data from many different scientific disciplines that support biological evolution.	X	X	
LO 1.12 The student is able to connect scientific evidence from many scientific disciplines to support the modern concept of evolution.	X	X	
LO 1.13 The student is able to construct and/or justify mathematical models, diagrams or simulations that represent processes of biological evolution.	X	X	
LO 1.22 The student is able to use data from real or simulated population(s), based on graphs or models of types of selection, to predict what will happen to the population in the future.	X	X	
LO 1.25 The student is able to describe a model that represents evolution within a population.	X	X	
LO 1.26 The student is able to evaluate given data sets that illustrate evolution as an ongoing process.	X	X	
Science Practice 1: The student can use representations and models to communicate scientific phenomena and solve scientific questions.	X	X	
Science Practice 4: The student can plan and implement data collection strategies appropriate to a particular scientific question.		X	X
Science Practice 5: The student can perform data analysis and evaluation of evidence.		X	X
Science Practice 6: The student can work with scientific explanations and theories.	X	X	X

Advanced International Certificate of Education (AICE) Biology Learning Outcomes

Outcome	Lesson		
	1	2	3
K. a) define the terms niche and population and be able to recognize examples of each	X	X	
P. d) explain, with examples, how environmental factors can act as stabilizing or evolutionary forces of natural selection;	X	X	
P. h) use the knowledge gained in this section in new situations or to solve related problems.	X	X	

Background Information

General background information is given here. More information can be found within the individual lessons.

The usage of Ecological Niche Modeling (ENM) software has become more abundant over the last twenty years as both computer software and hardware have advanced. This software allows researchers (and now students) to use published environmental data to construct current and future predictions of a species distribution. Visualizing and teaching macroevolution can be tricky, but these predictions are showing exactly what macroevolutions is- change in allele frequency in a given area over time. Allowing students to use the included data to explore the use of this software will allow them to actually visualize macroevolution.

Students might believe that the predefined, published data used in “Lesson 1” was specifically chosen by this author to depict a rare process of change. Because of this, “Lesson 2” was created to demonstrate that the distribution of most organisms will change over the next 35 years. Students are given the freedom to choose a Florida organism to investigate and will quickly find that there are two possible outcomes of their investigation. Outcome 1 is that the predicted distribution of that organism will change over the next 35 years signifying macroevolution will take place (it is extremely unlikely that there would be no predicted change for a species in the next 35 years, but this author did not exhaustively model all possible Florida species). Outcome 2 is the distinct possibility that not enough data about that specific species exists to make an accurate prediction.

“Lesson 3” focuses on the possibility of the aforementioned “outcome 2”- the absence of data. In “Lesson 1”, students observed how the predicted distribution map drastically changed when less data were used. Although some location data do exist for most species, often not enough data is available to accurately predict how a distribution could change. Because of this, “Lesson 3” focuses on the digitizing of data through “citizen science”. The digitization of data has become a recent trend within the scientific community and will allow researchers to better predict distributions. Students will actively contribute to the body of available georeferenced data available to researchers and will then investigate other citizen science projects that they could contribute to. Upon completion of these activities, students will have real predictions of macroevolution that will occur around them and a better understanding of how their current efforts through citizen science can benefit themselves and others.

LESSON ONE: "THE FUNDAMENTALLY ECOLOGICAL NICHE"

KEY QUESTION(S): What is a "niche"? How do scientists map niches? How can we use software to estimate current and future species distribution?

KEY SCIENCE CONCEPTS: Biological niches, evolution, scientific predictions, environmental change, distribution modeling.

OVERALL TIME ESTIMATE: One day (55 minutes).

LEARNING STYLES: Visual, logical.

VOCABULARY:
Eltonian niche- Defines a "niche" as the habitat in where the organism lives along with the specific adaptations that organism has in order to live and reproduce within that habitat.[1]
Fundamental niche- Any location in the environment in which the organisms can survive.[4]
Grinnellian niche- Defines "niche" as a "community-centric" concept which must take factors such as competition and predation into account.[2]
Hutchinsonian niche- Defines "niche" as an "abstract mapping of population dynamics onto an environmental space, the axes of which are abiotic and biotic factors that influence birth and death rates".[5]
Maximum Entropy Species Distribution Modeling Software- Software that "learns" as it processes data in order to make an accurate prediction.
Realized niche- The location in which an organism is actually found.

LESSON SUMMARY: Students will be introduced to Ecological Niche Modeling (ENM) through the use of "Maximum Entropy" habitat modeling. This software uses multiple layers of environmental data to estimate current and future niches/distributions of specific organisms. Preselected species and data information will be used to walk students through the basic usage of this software.

STUDENT LEARNING OBJECTIVES:
Students will be able to:
- Use modeling software to predict current population distributions
- Use modeling software to predict future population distributions and the evolution of that species
- Analyze data sets to determine if enough data is present to make accurate predictions
- Graphically depict how a niche can change over time
- Determine the most important factor(s) in the evolution of that population's distribution

STANDARDS
Next Generation Sunshine State Standards

SC.912.L.15.1	**SC.912.L.15.14**
SC.912.L.17.15	**SC.912.L.17.4**
SC.912.L.17.8	**SC.912.N.1.1**
SC.912.N.1.3	**SC.912.N.1.3**

Advanced Placement (AP) Biology Learning Outcomes(LO)/Science Practices (SO)*:*

LO 1.1	**LO 1.2**	**LO 1.3**
LO 1.4	**LO 1.5**	**LO1.9**
LO 1.12	**LO 1.13**	**LO 1.22**
LO 1.25	**LO 1.26**	

Science Practice 1, 6

Advanced International Certificate of Education (AICE) Biology Learning Outcomes:

K. a	**P.d**	**P. H**

MATERIALS:
Computers with Internet access and required .zip file.

BACKGROUND INFORMATION:

The term "biological niche" has been defined multiple different ways over the last century. Overall, there are three major definitions of this term that can be used. Students might be somewhat familiar with what a niche is, but there are minor distinguishing features between these three definitions. This background will distinguish between them and then define the term "niche" in context of this module.

The term "niche" was used in the context of ecology almost a century ago by Joseph Grinnell in his paper "The niche-relationships of the California Thrasher".[1] He ultimately describes that a niche is defined by the habitat in where the organism lives along with the specific adaptations that organism has in order to live and reproduce within that habitat.[1] This relatively simple definition focuses on the specific traits an organism has in order to make it successful and is now known as a "**Grinnellian niche**".[2] Soon after, Charles Elton slightly redefined a niche by incorporating their trophic position along with their traits.[3] This type of niche has been described as "community-centric" and is now known as an "**Eltonian niche**".[2] The third type of commonly recognized niche is the "**Hutchinsonian niche**" first proposed in 1957 by George Hutchinson.[4] This complex niche is described as an "abstract mapping of population dynamics onto an environmental space, the axes of which are abiotic and biotic factors that influence birth and death rates."[5] Hutchinson coined the term "**fundamental niche**" which can be described as any location in the environment in which the organisms CAN survive.[4] Additionally, the "**realized niche**" is the location in which the organism is actually found (and will be within the fundamental niche). In this module, we will be predicting the fundamental niches of specific species.

This module will use **Maximum Entropy Species Distribution Modeling software** in order to predict the fundamental niches of specific species. This software uses GPS data of where a specific species was found and environmental maps to predict the distribution of an organism. Although there are different programs that can be used to do this, the software chosen for this module is free and requires no installation (can be found inside the required .zip file included with this curriculum or at http://www.cs.princeton.edu/~schapire/maxent/). This software will use prepared data to make predictions in a user-friendly manner. In order to run this software, two different categories of data must be provided to it. The first category is species data which feature the exact locations of sightings or specimen collection. "Lesson 2" of this learning module will focus on how to obtain these data sets, but for "Lesson 1" they will be provided in the .zip file included with this module. The second category of data required for the software is prepared environmental layers. These different layers include environmental data such as rainfall, geography, and temperatures and can be for past, current, or predicted future conditions. All layers used in this learning module are prepared and included in the required .zip file. These prepared data layers were downloaded from http://www.worldclim.org and then modified by Dr. Charlotte Germain-Aubrey of the University of Florida to make usable for this module. Current and future (2050) environmental layers are provided.

ADVANCE PREPARATION:
Make sure that all students have access to the Internet and required .zip file found at http://www.cpet.ufl.edu/teachers/ssi/evolution/. Click on the file entitled "Ecological Niche Modeling- The Evolution of a Species" and save to desktop.

Required Laboratory Equipment:
Computers with Internet access

PROCEDURE AND DISCUSSION QUESTIONS WITH TIME ESTIMATES:

Background information, 20 minutes

Have a short discussion of what a biological niche is and how it is relevant to distributions. Use the information found in the background for this discussion; modify according to the level of the class. Students will then use the website http://scienceasaverb.wordpress.com/2010/10/29/introduction-to-ecological-niche-modeling-environmental-niche-modeling-species-distribution-modeling-part-two-what-is-niche-modeling/ to answer questions in order to get a fundamental understanding of what niche modeling is and how it is used.

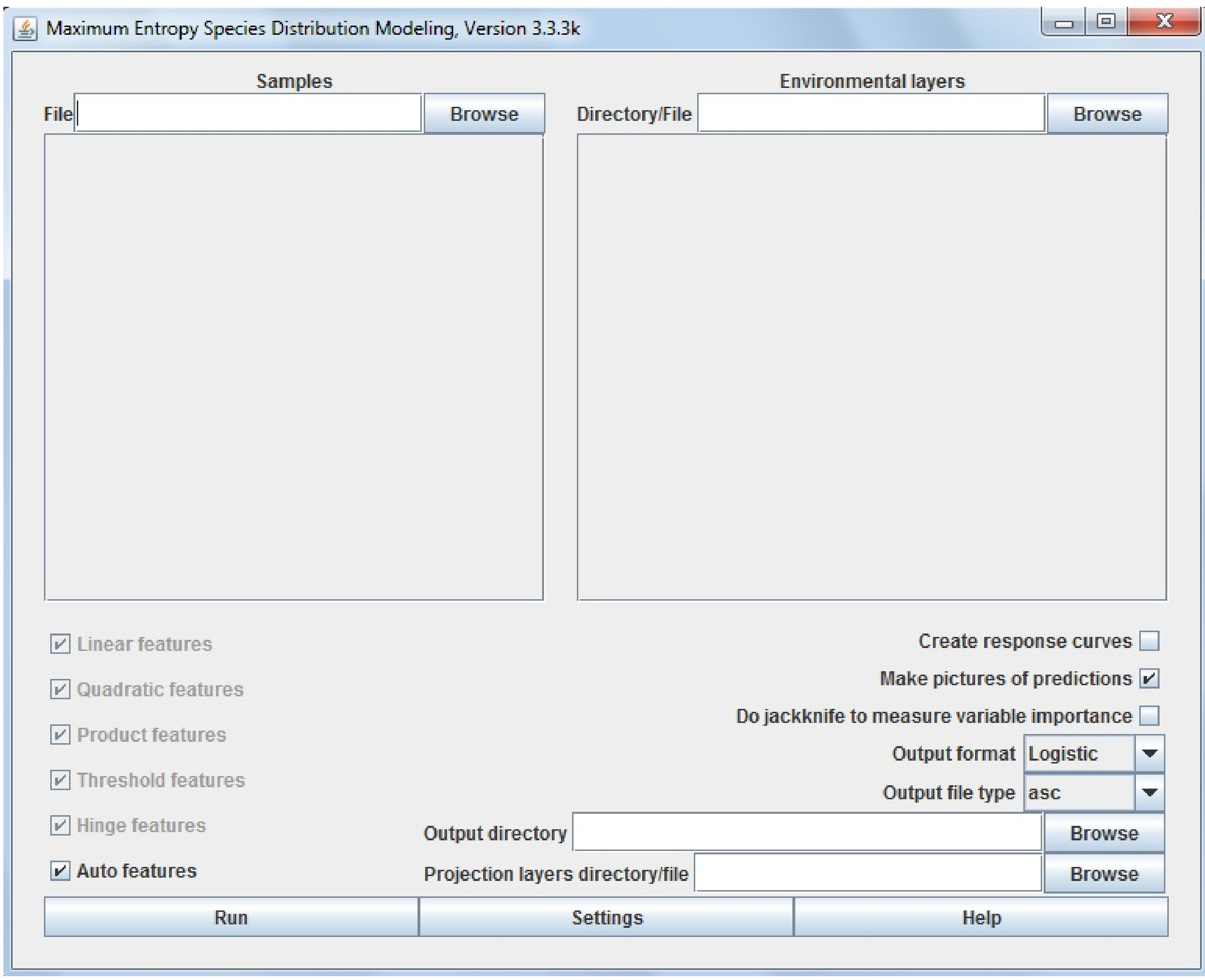
Modeling Software introduction, 30 minutes

After the questions about modeling have been answered, students will use Maxent software to model the current and future niches of *Ficus aurea*. Students will then answer questions about how the habitat ranges change through the next fifty years, how the amount of data directly impacts the accuracy of the prediction, and what environmental data were used to make the current and predictive models.

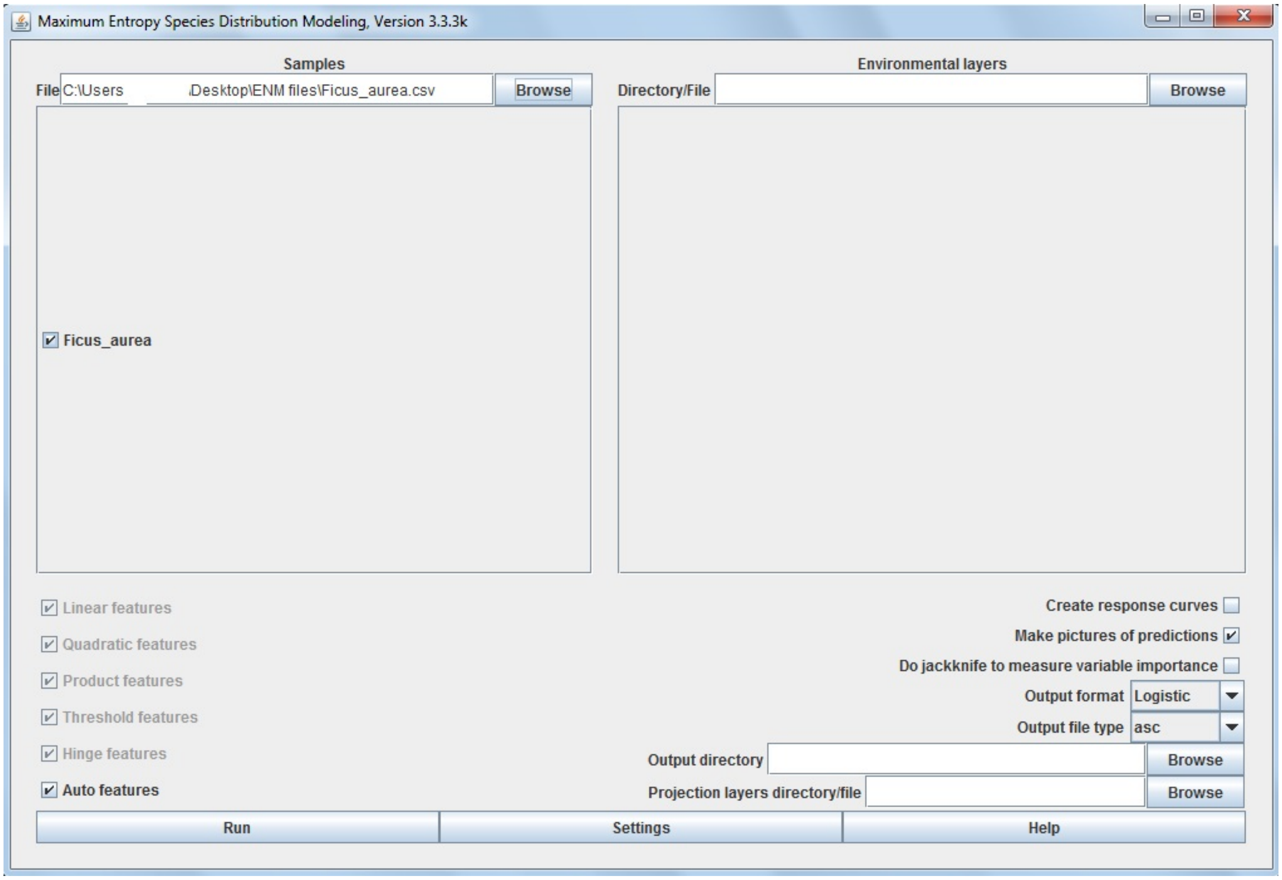
Procedure:

1. Make sure to have the file entitled "ENM Files" on each computer (if it is "zipped", make sure to "unzip" the file into a folder). Open the "ENM Files" folder to reveal the files inside.

2. Open the file called "maxent". This should open a program that looks like:

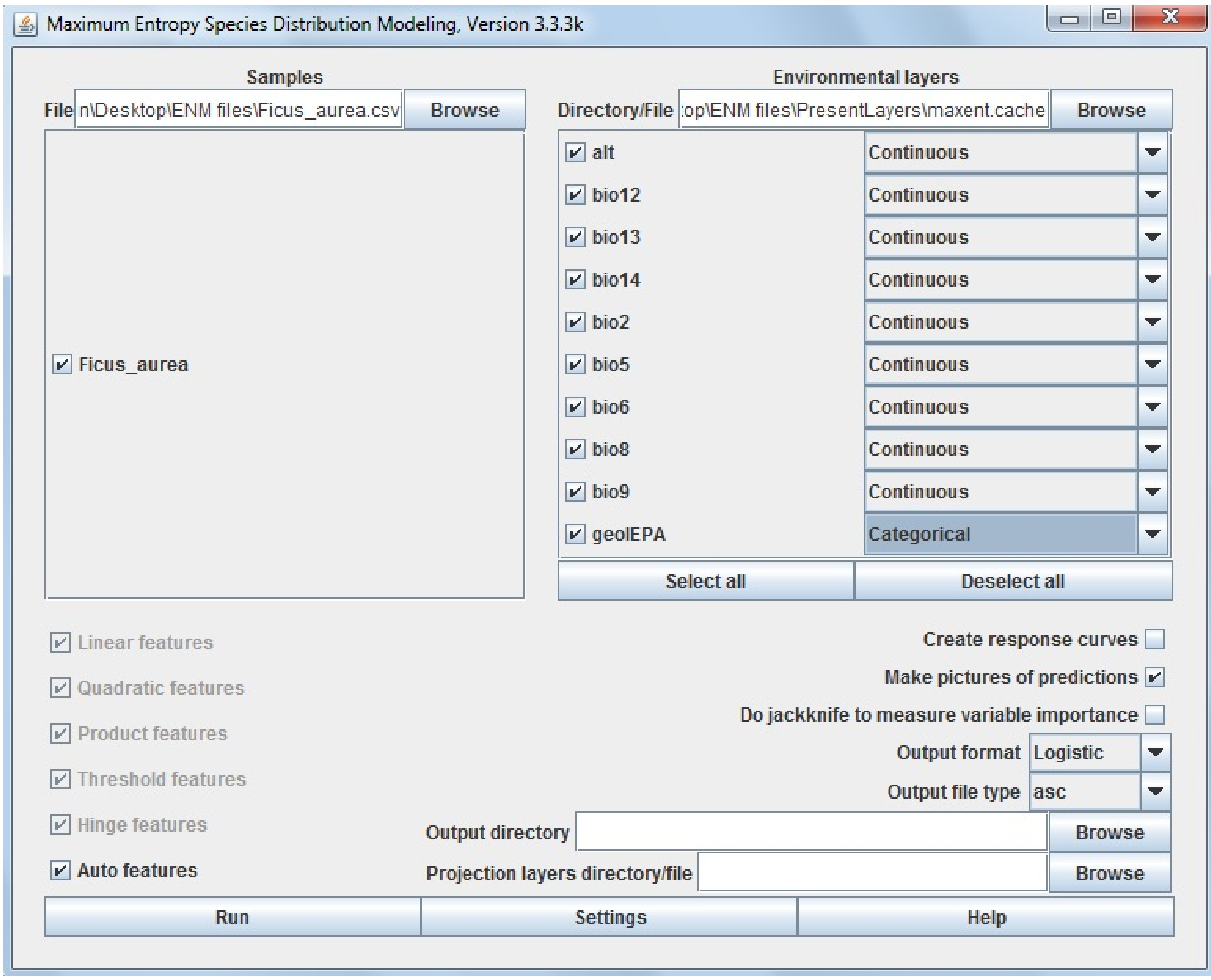


3. Students will need to select a file for both the “Samples” and “Environmental layers”. On the “Samples” side of the program, next to “File”, click “Browse”. This will open a new dialog box in order to pick the species of organism that will be modeled. Navigate to where the “ENM Files” folder is found and open this folder. Students will choose the file entitled “Ficus_aurea.csv” and then click “Open”. The program should now look like:

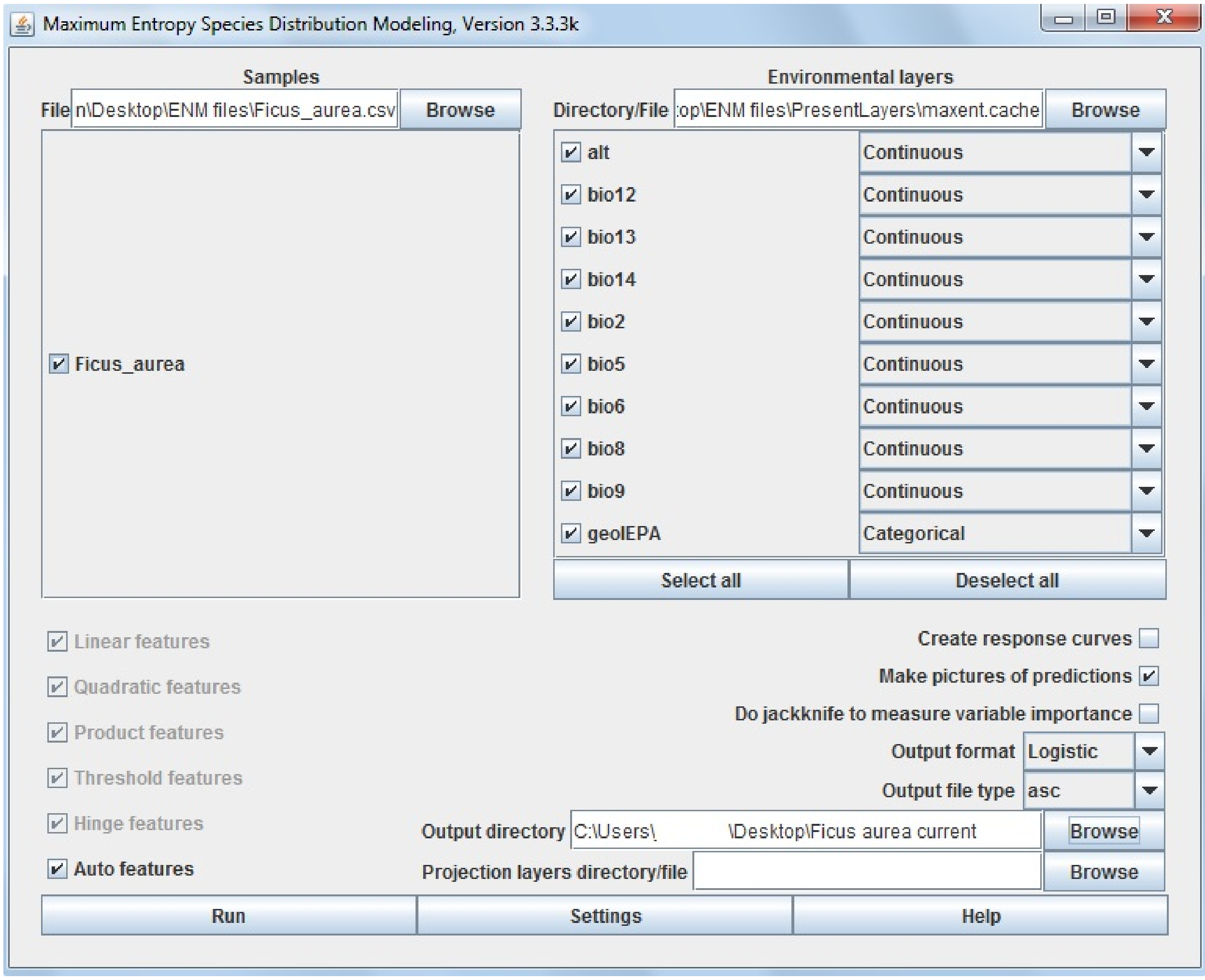


4. Next, the students will need to select the "Environmental layers" to create the model. On the "Environmental layers" side of the program, next to "File", click "Browse". This will open a new dialog box in order to pick the layers that will be used. Navigate to where the "ENM Files" folder is found and open this folder. Students will choose the file entitled "PresentLayers" and then click "Open". Click the file entitled "maxent.cache" and then click "open".

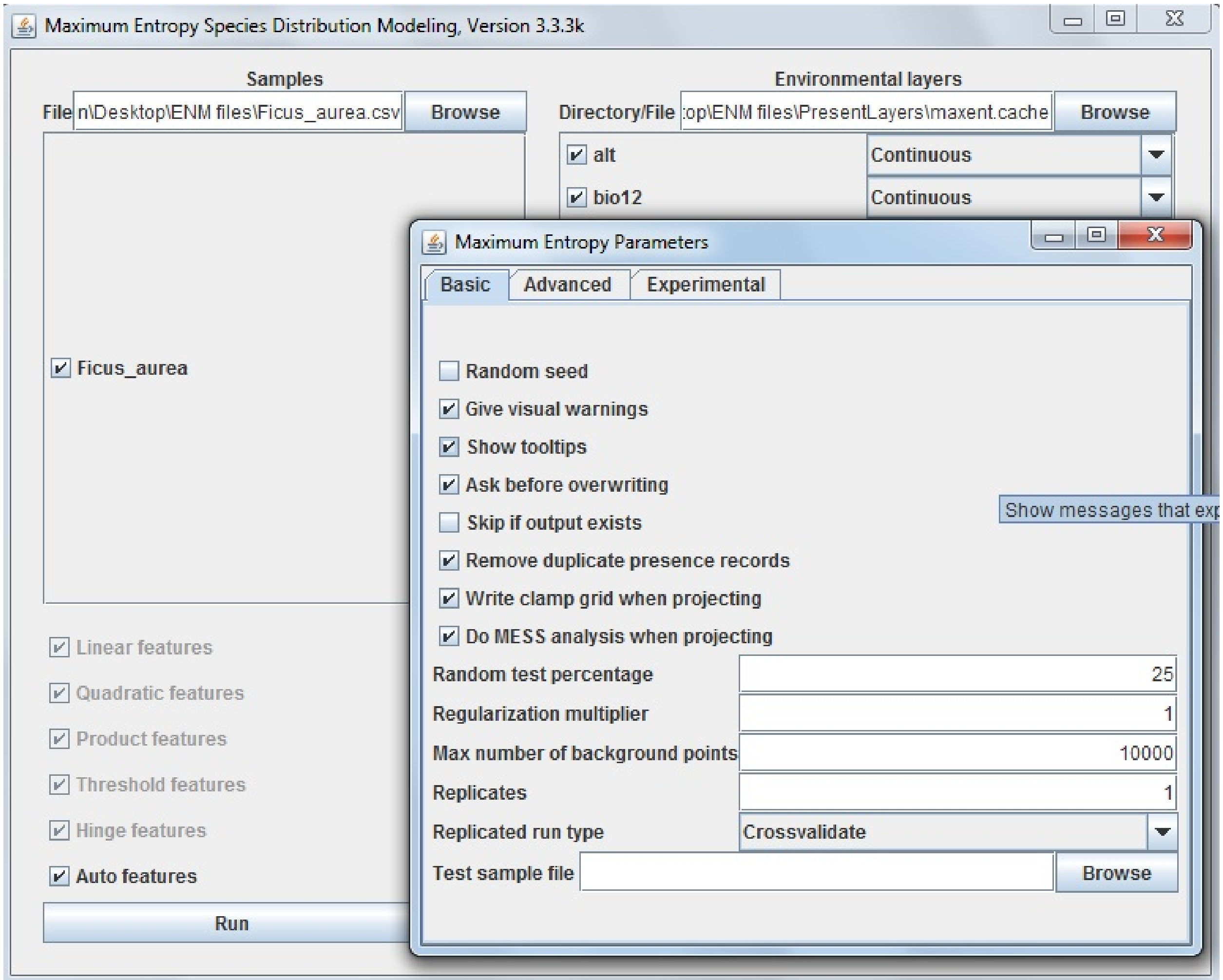
5. Once the layers have been selected, change the layer called "geolEPA" to "categorical" in the drop-down menu next to its name. The program should now look like:



6. Now that both the species and layers have been selected, a location to output the data to must be created. This can be done by first creating a new folder on the desktop that could be called “Ficus aurea current”. Once this folder has been created, click on the “Brouse” button located next to “Output directory” and navigate to the location of the folder that you just created, and then click “Open”. The program should now look like:

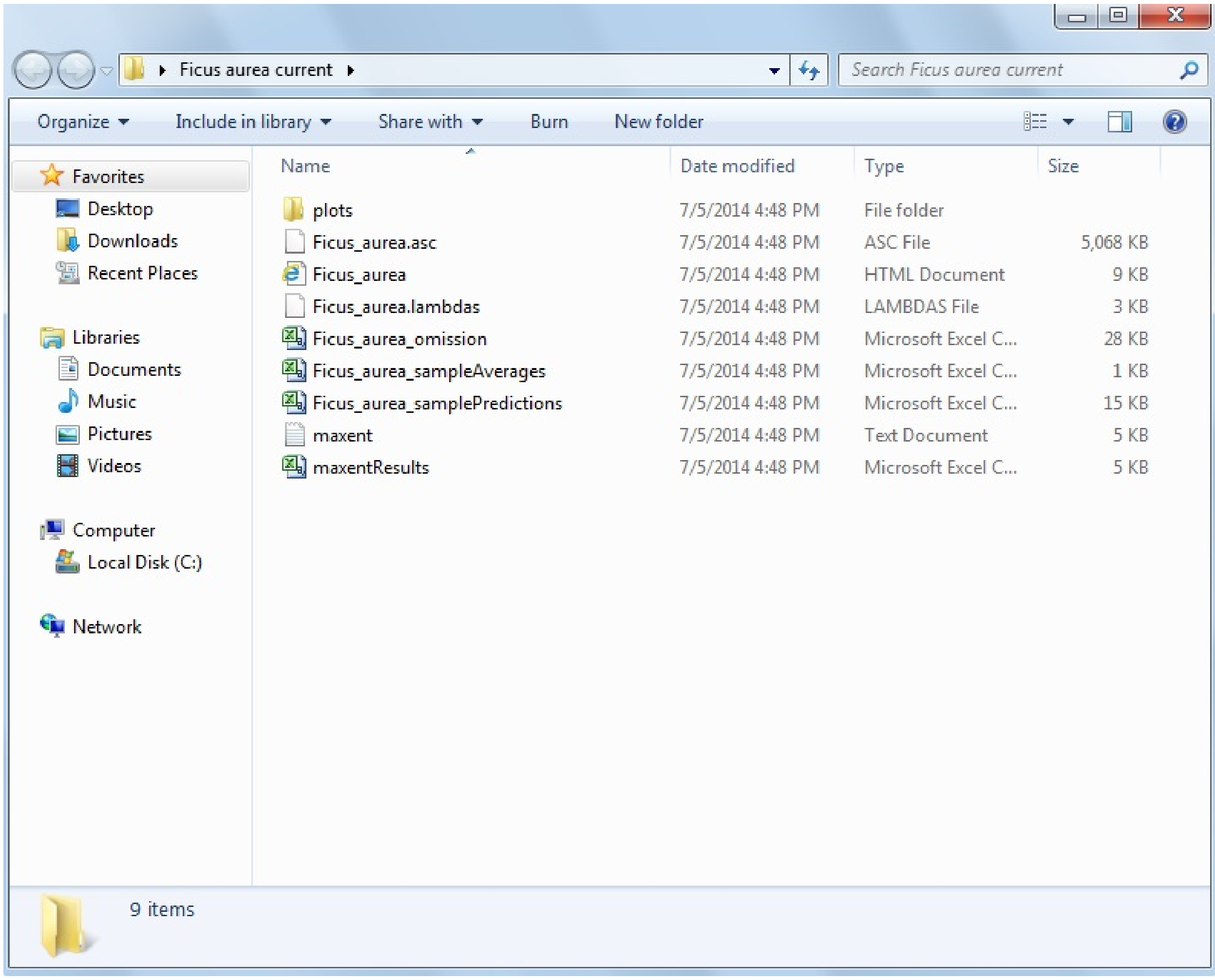


7. One last step is required. Click on the “Setting” button and change the “Random Test Percentage” to “25”. This will use 25% of the raw data to test the software’s current hypothesis and 75% of the data to create the hypothesis. The program should now look like:



8. Exit out of the “Settings” window by clicking the red “X” in the top right corner. The program is now ready to create a model. Click “Run”. If a dialog box pops up stating that some environmental data is missing, then click “Suppress similar visual warnings”.

9. The program should then process the data and put the newly created models in the folder created earlier entitled "Ficus aurea current". Navigate to the folder and open it. The newly created files should look like:

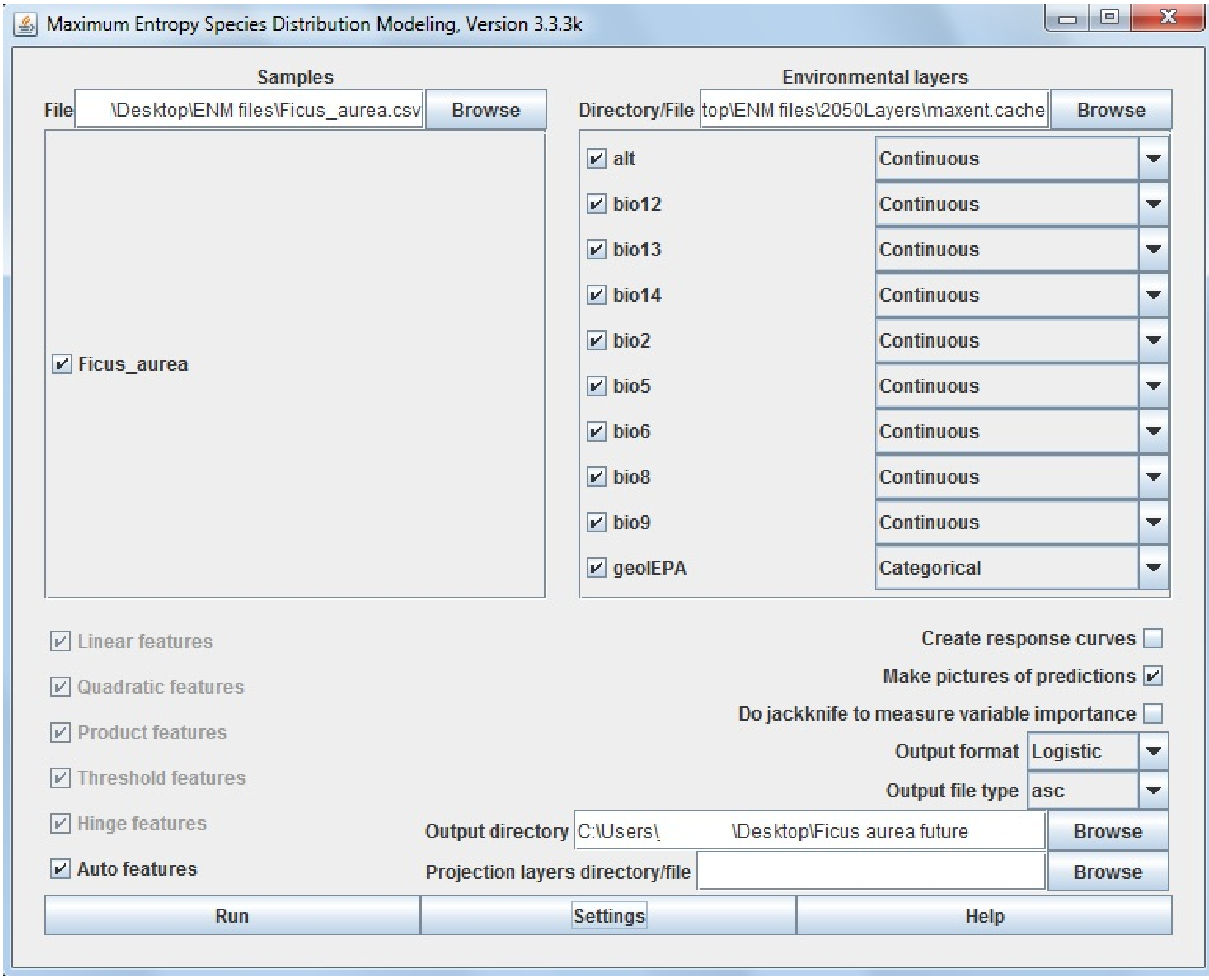


10. Click on the file entitled "Ficus_aurea" (the only one that is an HTML file). A browser window should open with all of the modeled data in one place.

11. Students will then need to go back to "Step 4" and change the environmental layers in order to model the future niches. On the "Environmental layers" side of the program, next to "File", click "Browse". This will open a new dialog box in order to pick the layers that will be used. Navigate to where the "ENM Files" folder is found and open this folder. Students will choose the file entitled "2050Layers" and then click "Open". Click the file entitled "maxent.cache" and then click "open".

12. Once the layers have been selected, change the layer called "geolEPA" to "categorical" in the drop-down menu next to its name.

13. Now that both the species and layers have been selected, a new location to output the data to must be created. This can be done by first creating a new folder on the desktop that could be called "Ficus aurea future". Once this folder has been created, click on the "Brouse" button located next to "Output directory" and navigate to the location of the folder that you just created, and then click "Open". The program should now look like:



14. The program is now ready to create a model. Click "Run". If a dialog box pops up stating that some environmental data is missing, then click "Suppress similar visual warnings".

15. The program should then process the data and put the newly created models in the folder created earlier entitled "Ficus aurea future". Navigate to the folder and open it.

16. Click on the file entitled "Ficus_aurea" (the only one that is an HTML file). A browser window should open with all of the modeled data in one place.

17. Students will then switch the "Sample" file to "Ficus aurea truncated" and will need to create a new output folder, such as "Ficus aurea truncated". The same "Environmental layers" can be used.

18. The program is now ready to create a model. Click "Run". If a dialog box pops up stating that some environmental data is missing, then click "Suppress similar visual warnings".

19. Students will then walk through and answer questions regarding the generated data. This will provide students a very basic understanding of ENM.

Key concepts found in these data:

Map comparison:
It takes very close inspection of each of these maps to identify predicted distribution differences. Ideally, each map needs to be next to the other and enlarged to the maximum size allowed by the computer monitor. Subtle differences can be seen between each map showing how *Ficus aurea* is predicted to become more frequent in the state of Florida- especially South Florida.

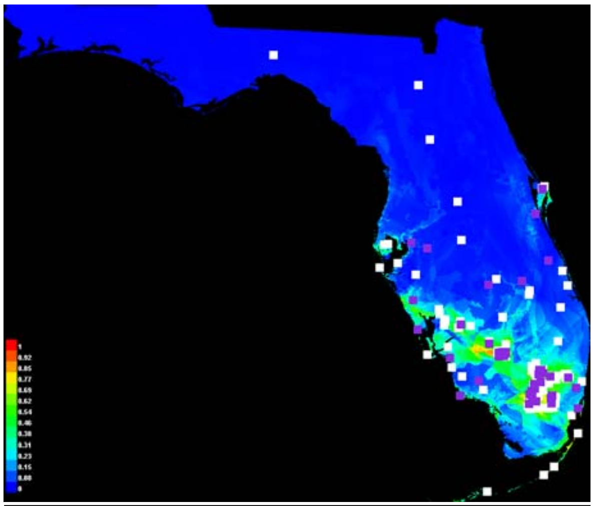


Figure 1: Predicted current distribution of *Ficus aurea.*

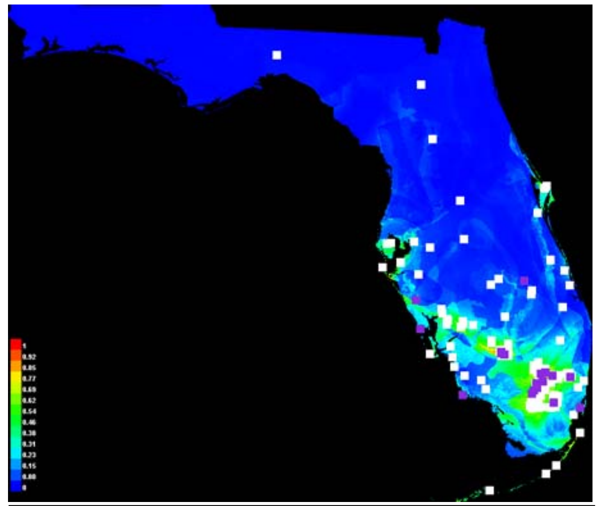


Figure 2: Predicted future (2050) distribution of *Ficus aurea.*

Which data were most used to make these predictions:
Based on the current distribution of *Ficus aurea*, the software determines which environmental layers are most important in predicting the future distribution. See "Figure 3". The environmental layers include the following (from http://worldclim.org/bioclim).

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
Alt= Altitude
GeolEPA= Geological data

Variable	Percent contribution	Permutation importance
bio6	37.8	46
bio13	29.4	29.5
geolEPA	15.9	5
bio2	5.4	1.2
bio12	4.7	8
bio14	2.4	5.2
bio9	1.6	0.6
alt	1.1	0.6
bio8	1.1	2.4
bio5	0.6	1.5

Figure 3: Environmental layers contribution in order to make the prediction of the distribution of *Ficus aurea* in 2050. It is in order of contribution. IE: The environmental layer "bio6" (Min Temperature of Coldest Month) was most useful in making this prediction.

Importance of data quantity

The amount of data used to make the predictions is very important. Graphs generated in the data can tell us if enough data were used to make useful predictions (see figures 4 and 5) and the generated maps clearly show major differences in distributions (see figures 6 and 7). Species data should have more than 30 data points in order to make useful predictions, but the more the better.

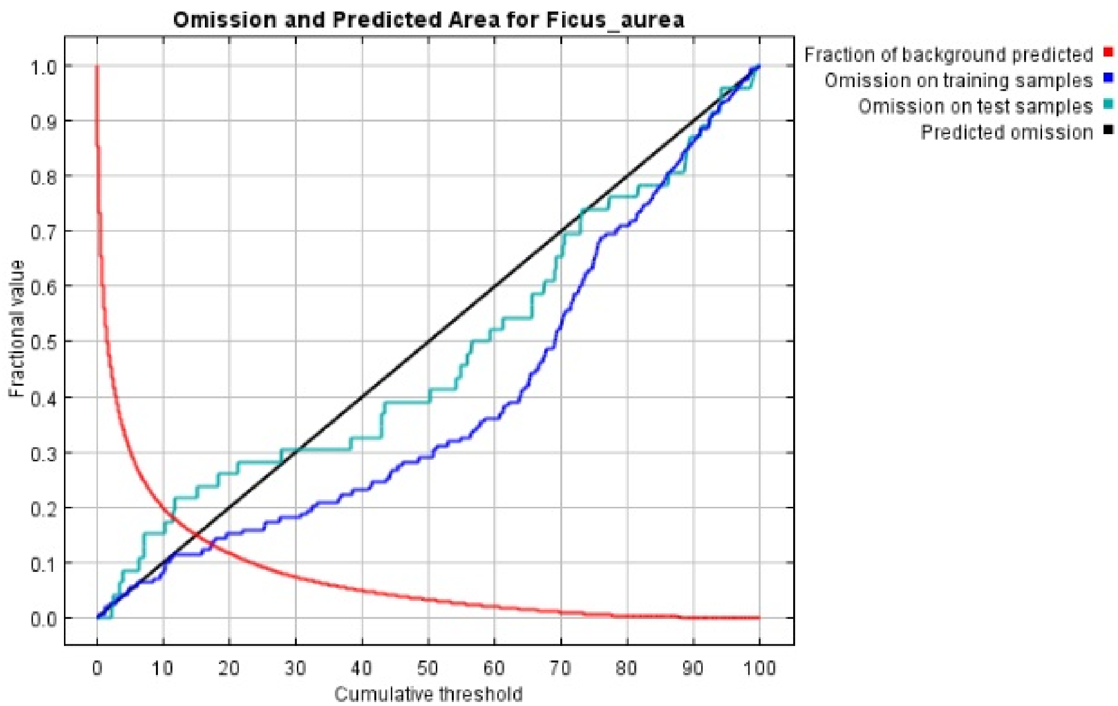


Figure 6: Omission and Predicted Area for *Ficus aurea* using 226 data points. Notice how the blue (omissions on training samples) and green (omissions on test samples) lines are close together and close to the black line (predicted omissions).

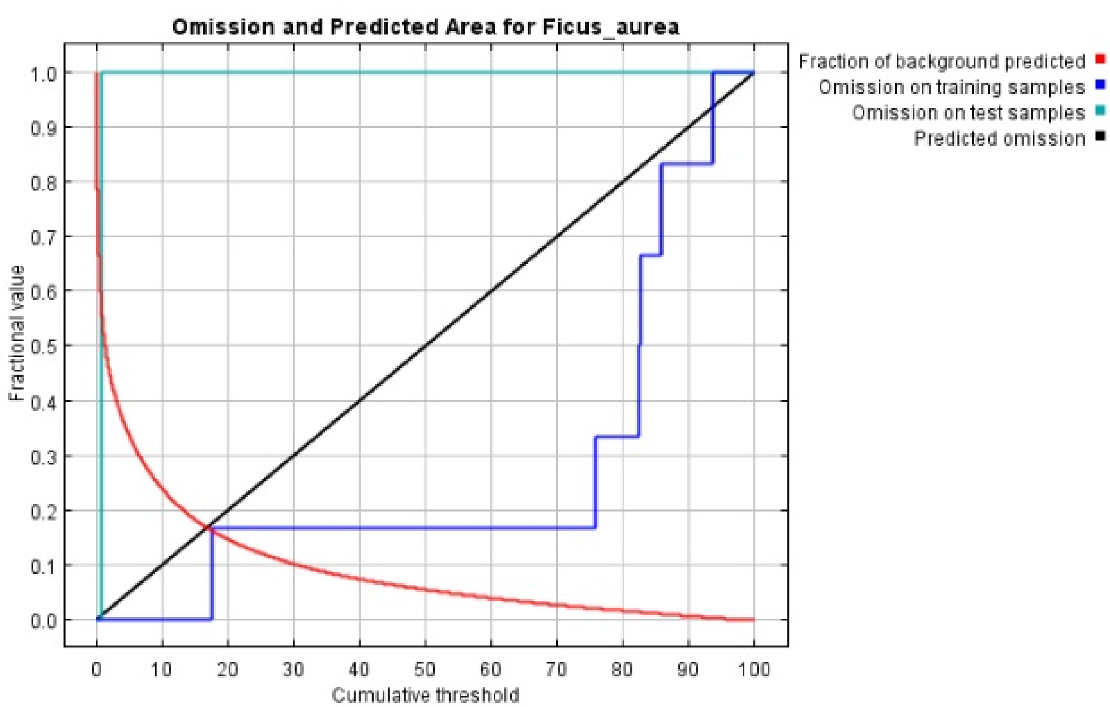


Figure 7: Omission and Predicted Area for *Ficus aurea* using 10 data points. Notice how the blue (omissions on training samples) and green (omissions on test samples) lines are not close together and not close to the black line (predicted omissions). This indicates not enough data were used to generate the prediction.

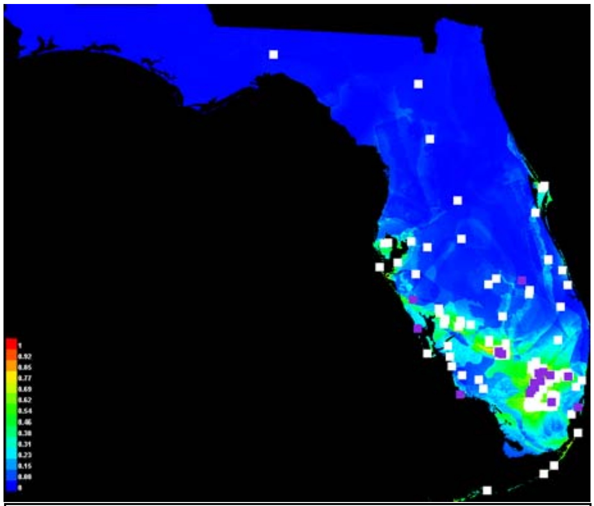


Figure 6: Predicted future (2050) distribution of *Ficus aurea* using 226 data points.

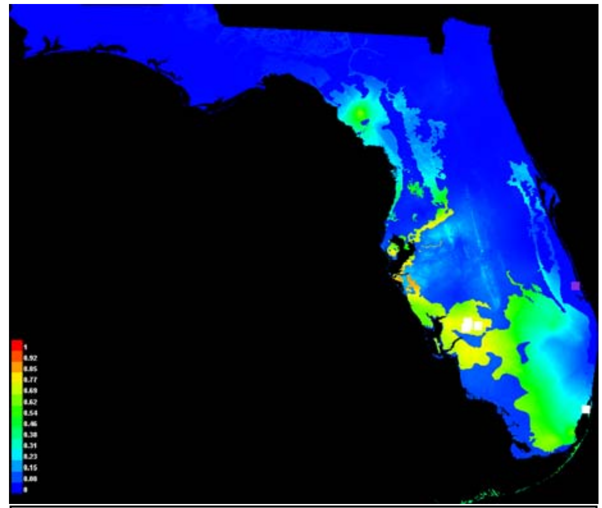


Figure 7: Predicted future (2050) distribution of *Ficus aurea* using 10 data points.

ASSESSMENT SUGGESTIONS:
By completing all of the questions included with “Lesson 1”, students will be able to:
- Use modeling software to predict the current population distribution of *Ficus aurea* in the state of Florida
- Use modeling software to predict the future population distributions of *Ficus aurea* in the state of Florida and how this species will evolve over the next 35 years
- Analyze data sets to determine if enough data is present to make accurate predictions with the ENM software
- Graphically depict how a niche can change over time by comparing current and future predicted distributions of *Ficus aurea*
- Determine the most important factor in the evolution of *Ficus aurea* population’s distribution is “bio6” (Min Temperature of Coldest Month)

EXTENSIONS:
Other plant species can be found in the “ENM” file folder. Students can explore the projected distribution of these different species.

RESOURCES/REFERENCES:

1. Grinnell, J. (1917). "The niche-relationships of the California Thrasher". *The Auk.,* **34**: 427–433.
2. Shugart, H. (1998). **Terrestrial Ecosystems in Changing Environments**. Cambridge: Cambridge University Press.
3. Elton, C. (1927). **Animal Ecology**. Sidgwick and Jackson, London.
4. Hutchinson, G. (1957). Concluding remarks. Cold Spring Harbor Symp. **22**:415–427.
5. Holt, R. (2009). Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. Proceedings of the National Academy of Sciences of the United States of America **106**:19659-19665.

STUDENT PAGES: Background Information

Use the website http://scienceasaverb.wordpress.com/2010/10/29/introduction-to-ecological-niche-modeling-environmental-niche-modeling-species-distribution-modeling-part-two-what-is-niche-modeling/ to answer the following questions in order to gain a fundamental understanding of what niche modeling is and how it is used.

1. What type of locality data (exact location where the organism is found) is preferred in this day and age?

2. Why are these types of data preferred?

3. Look at the figure of the state of California. What does each green dot represent?

4. It is impossible to know ALL of the environmental factors present at each of these locations, but we do have maps showing some environmental trends across large areas. If we put both the location data of a found species and some of the recorded environmental data together, what can be done?

5. What type of environmental data can be used to make these niche models?

6. "Very simple heuristics" is listed as a modeling method along with "highly complex" algorithms that feature "machine learning". What is a key difference between these two modeling methods?

7. What do the different colors in the state of California represent?

8. In reference to the amount of data points, how can more accurate models be created?

9. What two things drastically affect our ability to estimate a niche?

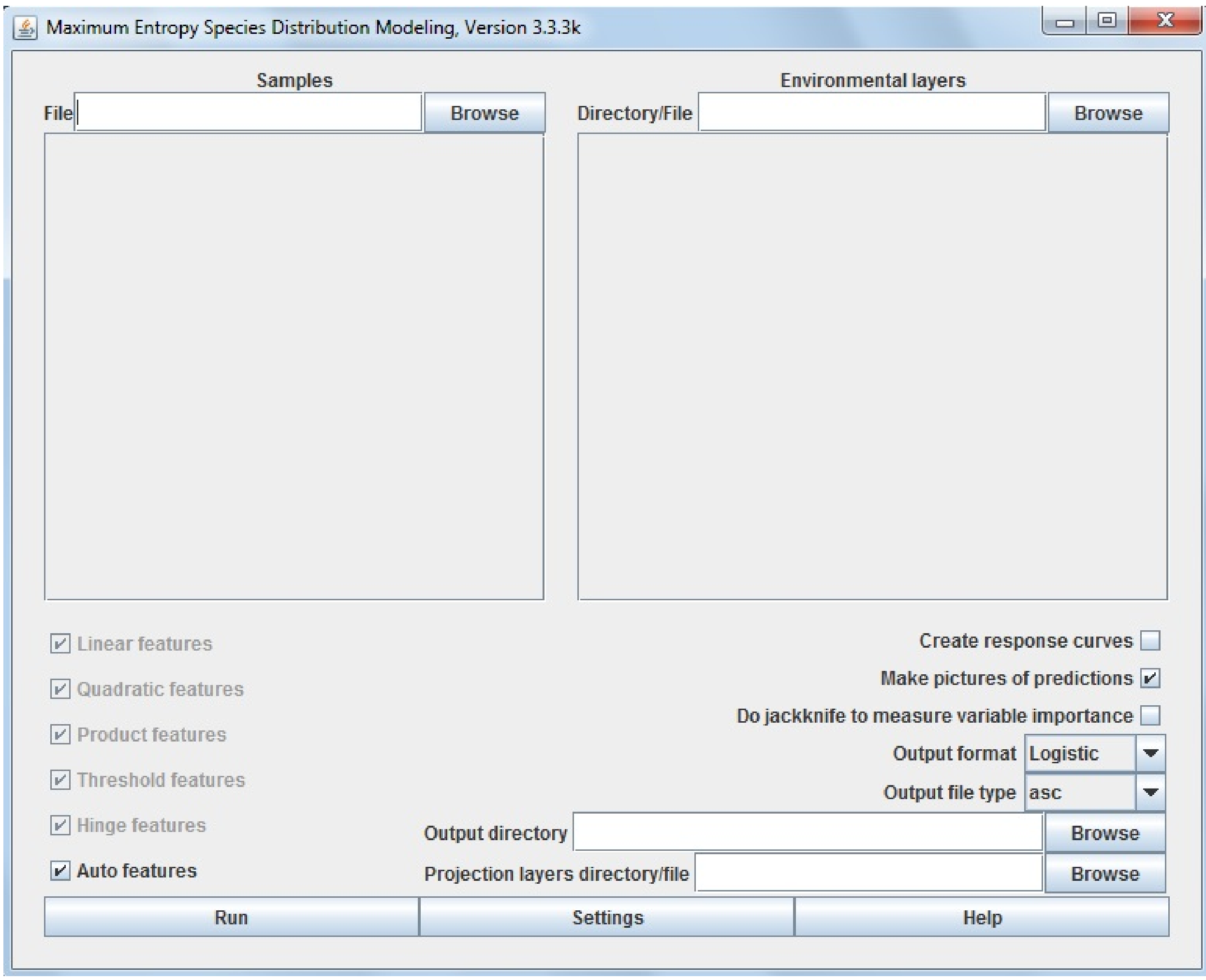
STUDENT HANDOUT: Laboratory Methodologies and Questions

Use Maxent software to model the current and future niches of *Ficus aurea* then answer questions about how the habitat ranges change through the next fifty years, how the amount of data directly impacts the accuracy of the prediction, and what environmental data were used to make the current and predictive models.

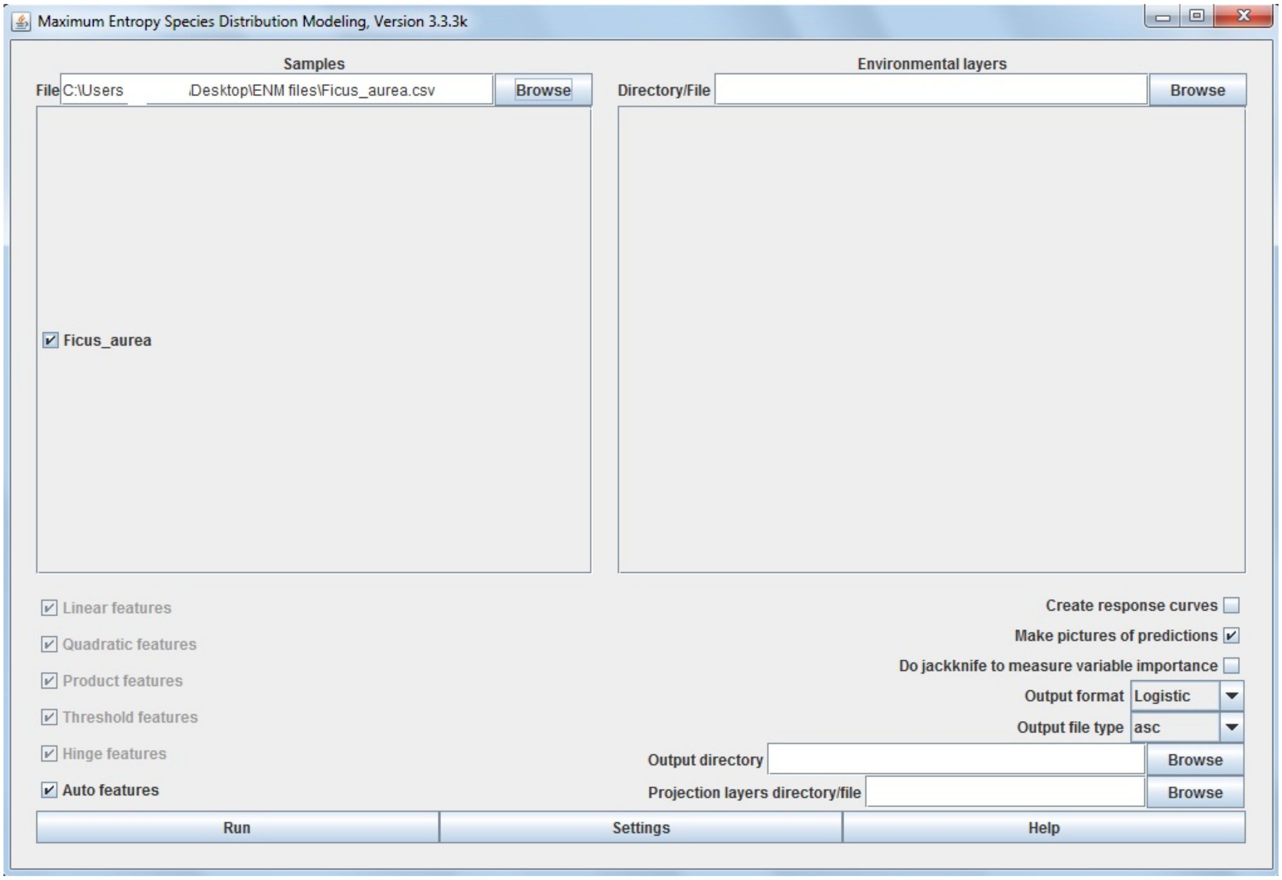
Procedure:

1. Make sure to have the file entitled "ENM Files" on your computer (if it is "zipped", make sure to "unzip" the file into a folder). Open the "ENM Files" folder to reveal the files inside.

2. Open the file called "maxent". This should open a program that looks like:



3. You will need to select a file for both the “Samples” and “Environmental layers”. On the “Samples” side of the program, next to “File”, click “Browse”. This will open a new dialog box in order to pick the species of organism that will be modeled. Navigate to where the “ENM Files” folder is found and open this folder. Students will choose the file entitled “Ficus_aurea.csv” and then click “Open”. The program should now look like:

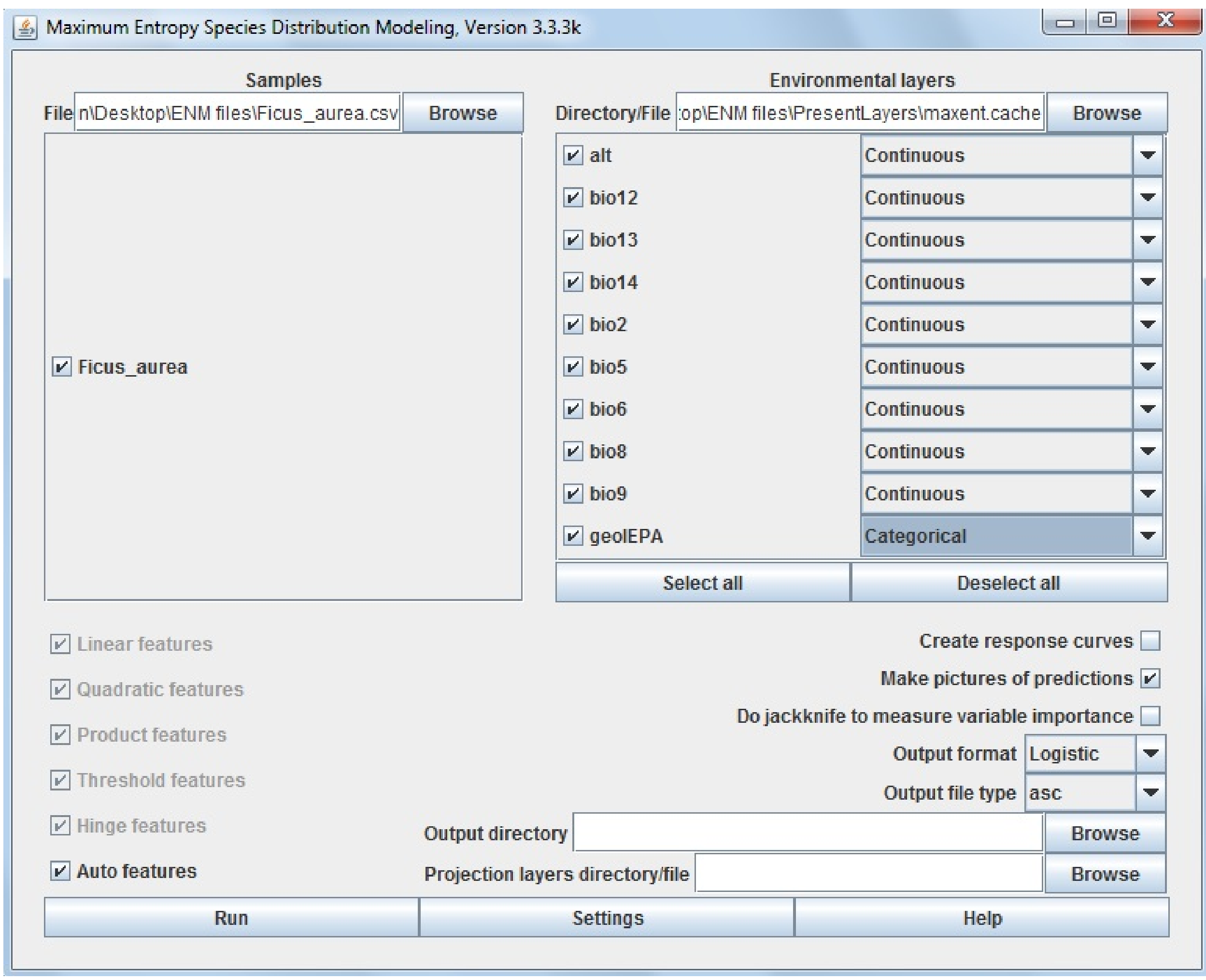


4. Next, you will need to select the "Environmental layers" to create the model. On the "Environmental layers" side of the program, next to "File", click "Browse". This will open a new dialog box in order to pick the layers that will be used. Navigate to where the "ENM Files" folder is found and open this folder. Choose the file entitled "PresentLayers" and then click "Open". Click the file entitled "maxent.cache" and then click "open". The list (from http://worldclim.org/bioclim) of usable layers is found below, but you will only be using several of these in addition to geologic layers.

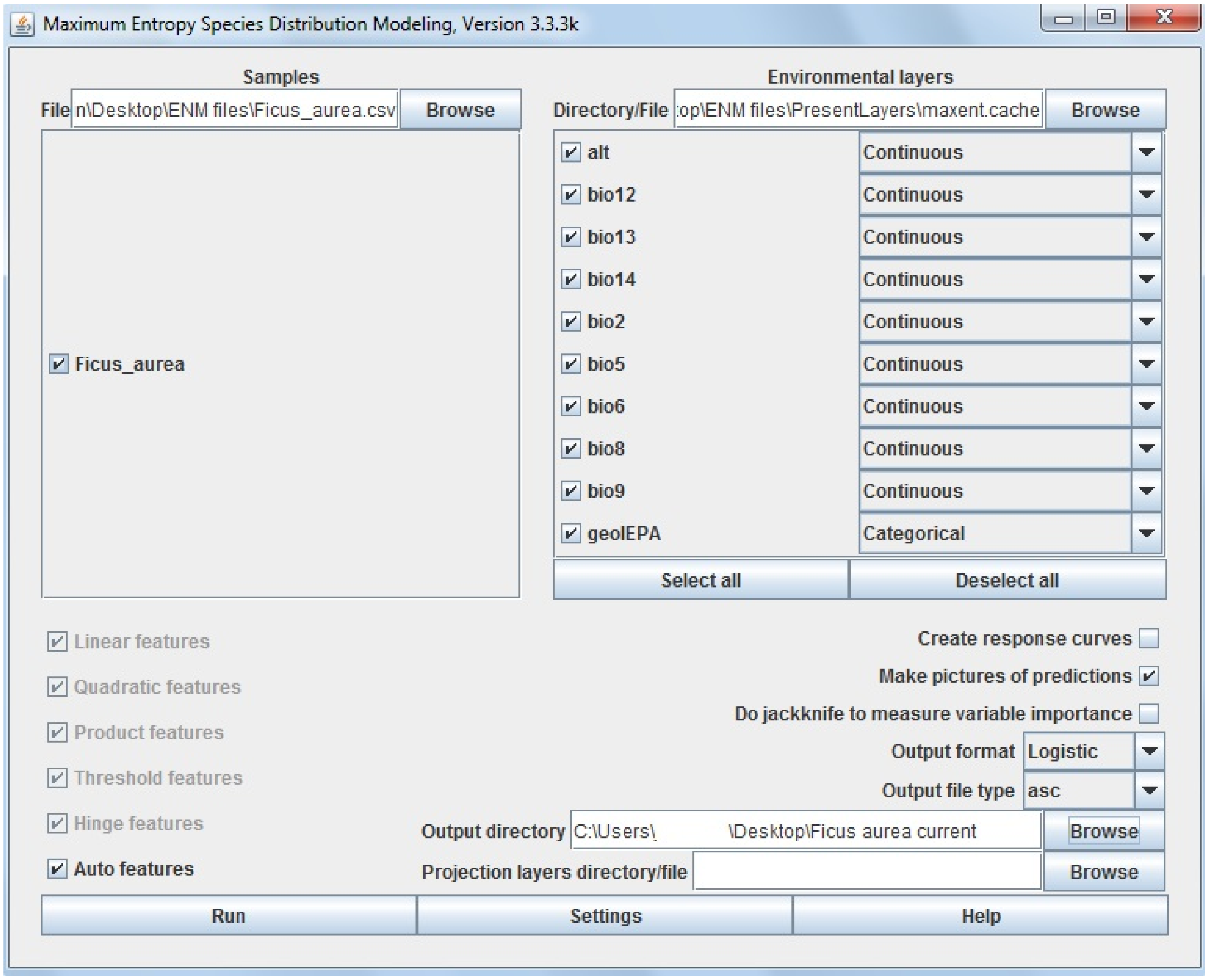
Possible Environmental Layers:

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
Alt= Altitude
GeolEPA= Geological data

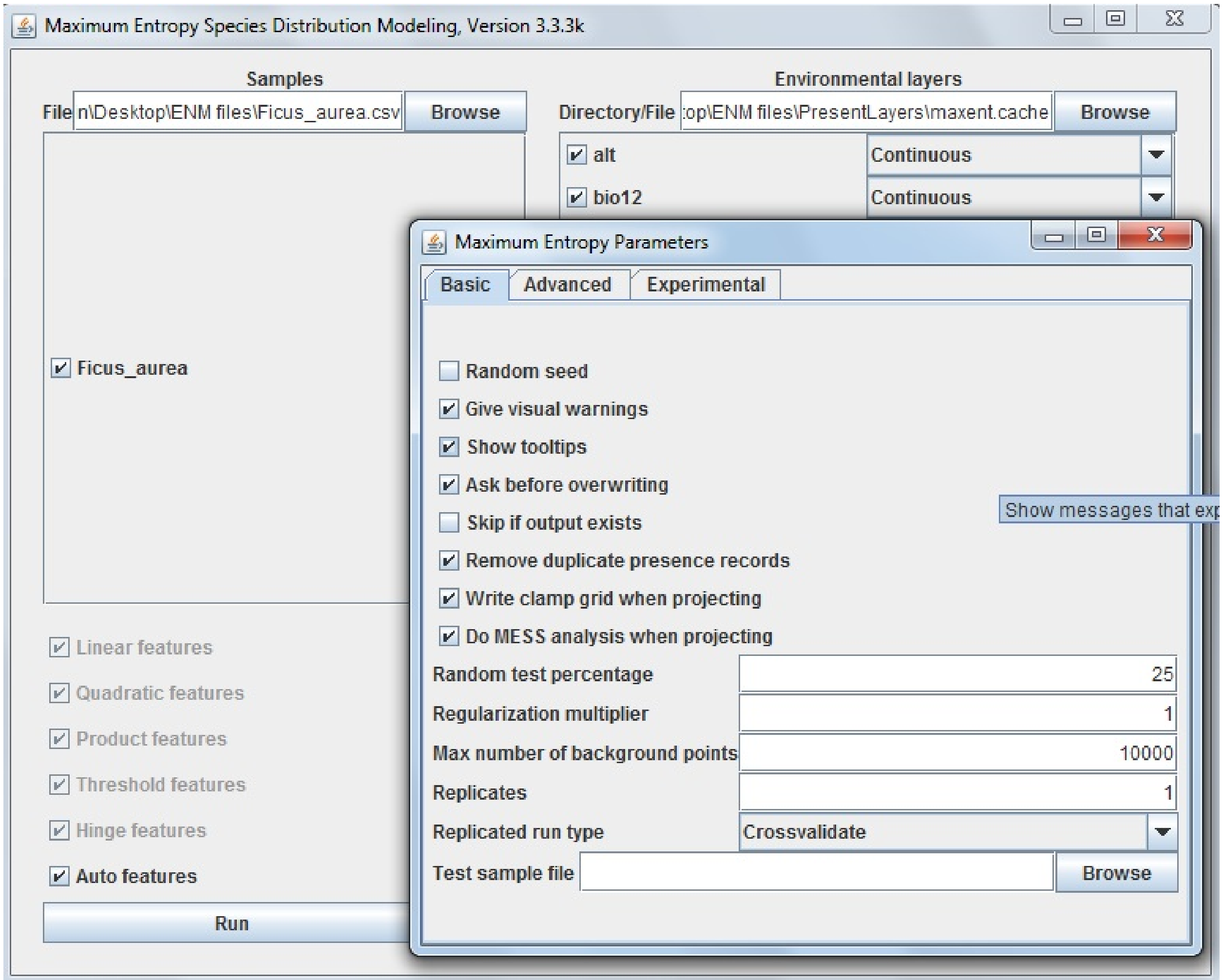
5. Once the layers have been selected, change the layer called "geolEPA" to "categorical" in the drop-down menu next to its name. The program should now look like:



6. Now that both the species and layers have been selected, a location to output the data to must be created. This can be done by first creating a new folder on the desktop that could be called “Ficus aurea current”. Once this folder has been created, click on the “Browse” button located next to “Output directory” and navigate to the location of the folder that you just created, and then click “Open”. The program should now look like:

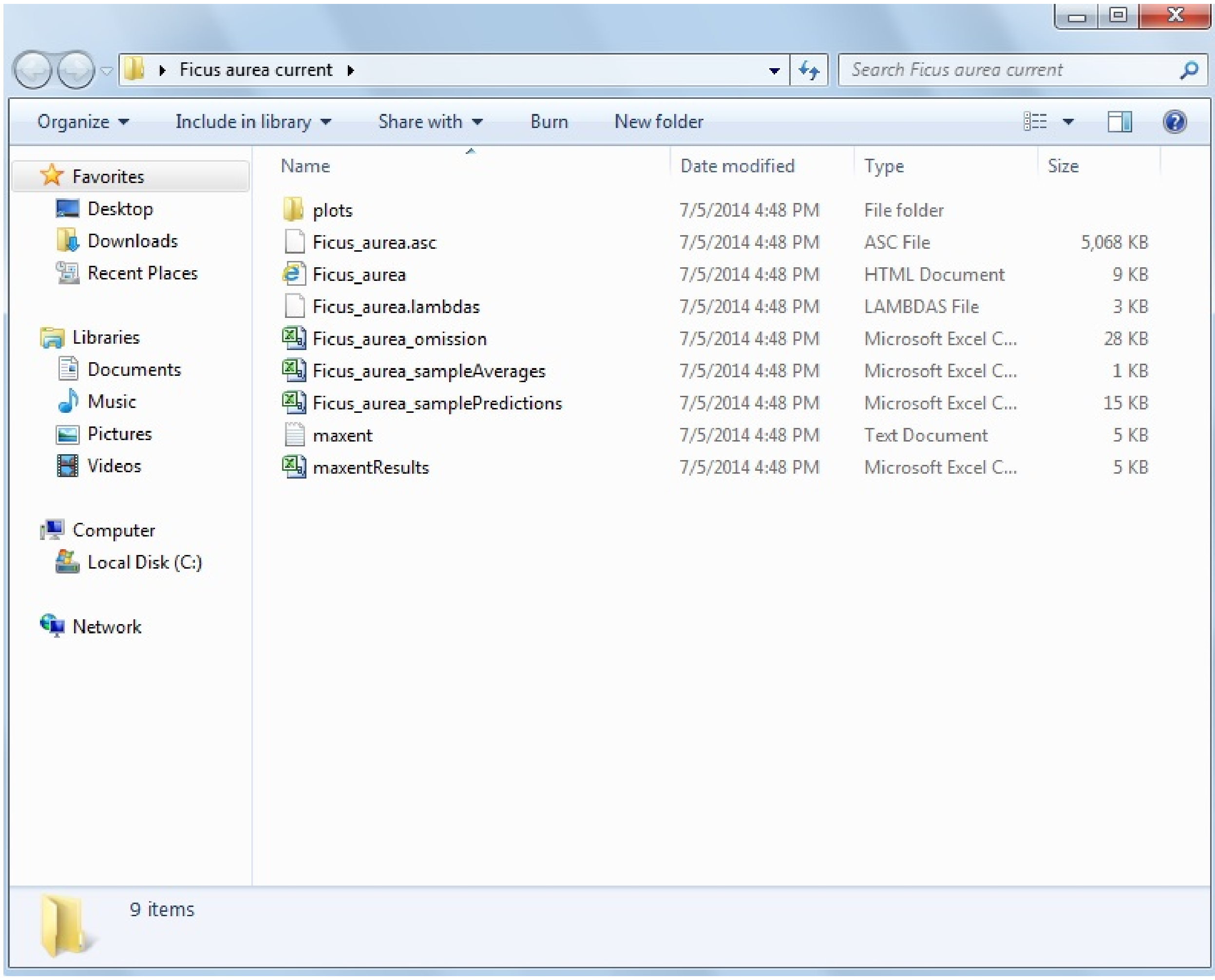


7. One last step is required. Click on the “Setting” button and change the “Random Test Percentage” to “25”. This will use 25% of the raw data to test the software’s current hypothesis and 75% of the data to create the hypothesis. The program should now look like:



8. Exit out of the “Settings” window by clicking the red “X” in the top right corner. The program is now ready to create a model. Click “Run”. If a dialog box pops up stating that some environmental data is missing, then click “Suppress similar visual warnings”.

9. The program should then process the data and put the newly created models in the folder created earlier entitled "Ficus aurea current". Navigate to the folder and open it. The newly created files should look like:

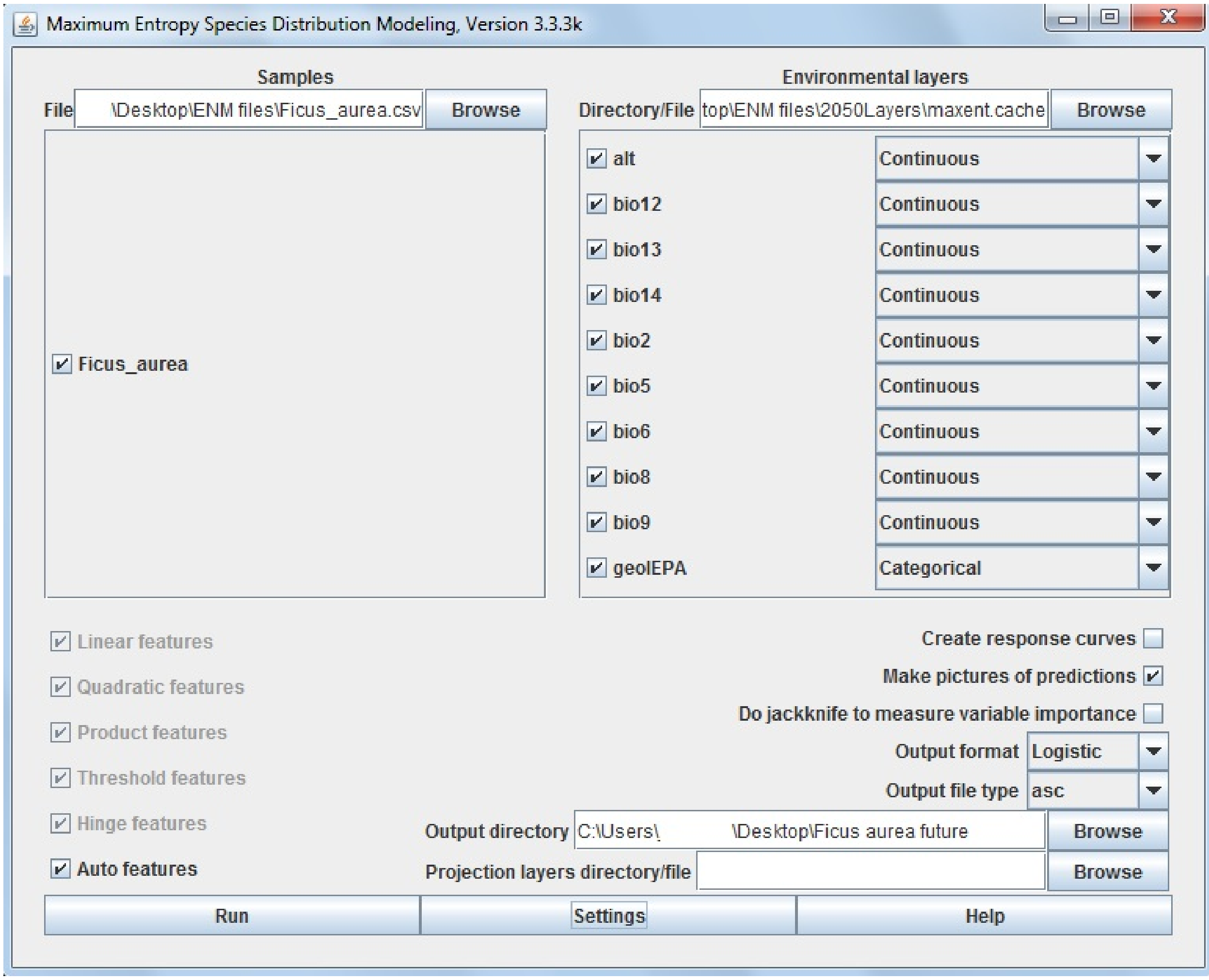


10. Click on the file entitled "Ficus_aurea" (the only one that is an HTML file). A browser window should open with all of the modeled data in one place.

11. You will then need to go back to "Step 4" and change the environmental layers in order to model the future niches. On the "Environmental layers" side of the program, next to "File", click "Browse". This will open a new dialog box in order to pick the layers that will be used. Navigate to where the "ENM Files" folder is found and open this folder. Choose the file entitled "2050Layers" and then click "Open". Click the file entitled "maxent.cache" and then click "open".

12. Once the layers have been selected, change the layer called "geolEPA" to "categorical" in the drop-down menu next to its name.

13. Now that both the species and layers have been selected, a new location to output the data to must be created. This can be done by first creating a new folder on the desktop that could be called "Ficus aurea future". Once this folder has been created, click on the "Browse" button located next to "Output directory" and navigate to the location of the folder that you just created, and then click "Open". The program should now look like:



14. The program is now ready to create a model. Click "Run". If a dialog box pops up stating that some environmental data is missing, then click "Suppress similar visual warnings".

15. The program should then process the data and put the newly created models in the folder created earlier entitled "Ficus aurea future". Navigate to the folder and open it.

16. Click on the file entitled "Ficus_aurea" (the only one that is an HTML file). A browser window should open with all of the modeled data in one place.

17. Switch the "Sample" file to "Ficus aurea truncated" and create a new output folder, such as "Ficus aurea truncated". The same "Environmental layers" can be used.

18. The program is now ready to create a model. Click "Run". If a dialog box pops up stating that some environmental data is missing, then click "Suppress similar visual warnings".

19. The program should then process the data and put the newly created models in the folder created earlier entitled "Ficus aurea truncated". Navigate to the folder and open it.

20. Click on the file entitled "Ficus_aurea" (the only one that is an HTML file). A browser window should open with all of the modeled data in one place.

You should have generated three different niche models (*Ficus aurea* current, *Ficus aurea* future, and *Ficus aurea* truncated). Answer the following questions about the three different models.

1. Compare the predicted distribution maps for *Ficus aurea* current and *Ficus aurea* future. It will take close side-by-side inspection of these maps to notice distribution differences. Blue and "cold colors" represent areas that are predicted to have a low frequency of *Ficus aurea.* Red and "warm colors" represent areas that are predicted to have a high frequency of *Ficus aurea.* The dots are the sample locations (the different colored dots represent the locations used to make the hypothesis versus the locations used to test the hypothesis). How will the distribution of *Ficus aurea* change over the next thirty-five years?

2. Was the change of distribution more or less drastic than you expected?

3. Hypothesize why there is predicted to be this change in the distribution of *Ficus aurea* over the next fifty years.

4. Find the area of the analysis entitled "Analysis of variable contributions" for both *Ficus aurea* current and *Ficus aurea* future. This section tells us what environmental layers were most useful in making the distribution prediction. What were the two most important layers for *Ficus aurea* current and *Ficus aurea* future (see "step 4" from above for the list of layers)?

5. The "*Ficus aurea* future" file used 226 data points to make its prediction while the "*Ficus aurea* truncated" file used only 10 data points. Both used layers to predict the distribution of the species in 2050. Which prediction do you think will be more accurate and why?

6. Compare the predicted distribution maps of both “*Ficus aurea* future” and “*Ficus aurea* truncated”. How are they different?

7. The generated data can also tell us if enough data points are being used to create useful predictions. Find the graph at the top of the analysis labeled “Omission and Predicted Area for Ficus aurea”. Ideally, the blue, green, and black line will be very close together on the graph if enough data is used. How do the graphs vary between “*Ficus aurea* future” and “*Ficus aurea* truncated” and what does it indicate?

It is recommended to have > 30 data points in order to create a useful prediction.

8. How is the theory of evolution supported by the future models of distribution?

9. What possible mechanisms are causing the species to be more abundant in the new areas predicted by the future distribution? Explain each possible mechanism.

TEACHER PAGES: *KEY* FOR "BACKGROUND INFORMATION"

1. What type of locality data (exact location where the organism is found) is preferred in this day and age? GPS data
2. Why are these types of data preferred? More accurate and can be used to estimate environmental tolerances and preferences.
3. Look at the figure of the state of California. What does each green dot represent? The specific location of a found organism.
4. It is impossible to know ALL of the environmental factors present at each of these locations, but we do have maps showing some environmental trends across large areas. If we put both the location data of a found species and some of the recorded environmental data together, what can be done? Mathematical data can be generated to make predictions of distribution.
5. What type of environmental data can be used to make these niche models? Isothermality, mean temperature, annual precipitation, etc…
6. "Very simple heuristics" is listed as a modeling method along with "highly complex" algorithms that feature "machine learning". What is a key difference between these two modeling methods? The algorithms take into account where the organism is NOT found as well as where it is found to predict distributions.
7. What do the different colors in the state of California represent? Estimated suitability of a habitat for a species.
8. In reference to the amount of data points, how can more accurate models be created? More data points the more accurate the prediction.
9. What two things drastically affect our ability to estimate a niche? 1. Number of occurrence points. 2. "Distribution of environmental variables in the study region."

TEACHER PAGES: *KEY* FOR "Laboratory Methodologies and Questions"

1. Compare the predicted distribution maps for *Ficus aurea* current and *Ficus aurea* future. It will take close side-by-side inspection of these maps to notice distribution differences. Blue and "cold colors" represent areas that are predicted to have a low frequency of *Ficus aurea.* Red and "warm colors" represent areas that are predicted to have a high frequency of *Ficus aurea.* The dots are the sample locations (the different colored dots represent the locations used to make the hypothesis versus the locations used to test the hypothesis). How will the distribution of *Ficus aurea* change over the next fifty years? This species will be more common in Florida, especially South Florida. Additionally, it will be less common in parts of North Florida.

2. Was the change of distribution more or less drastic than you expected? This answer will be based on student opinion.

3. Hypothesize why there is predicted to be this change in the distribution of *Ficus aurea* over the next fifty years. Although students might come up with varying answers, most will probably include climate change (possibly warming). This specific distribution model predicts that the minimum temperature of the coldest month is the most important factor in the change in distribution (climate change) but students will get to that specific answer shortly.

4. Find the area of the analysis entitled "Analysis of variable contributions" for both *Ficus aurea* current and *Ficus aurea* future. This section tells us what environmental layers were most useful in making the distribution prediction. What were the two most important layers for *Ficus aurea* current and *Ficus aurea* future (see "step 4" from above for the list of layers)?

> *Ficus aurea* current-1. Bio6, lowest temperate of the coldest month. 2. Precipitation of wettest month.
>
> *Ficus aurea* future- 1. Bio6, lowest temperate of the coldest month. 2. Precipitation of wettest month.

5. The “*Ficus aurea* future” file used 226 data points to make its prediction while the “*Ficus aurea* truncated” file used only 10 data points. Both used layers to predict the distribution of the species in 2050. Which prediction do you think will be more accurate and why? “*Ficus aurea* future” because (as noted above in the background) more data points will give us a more accurate prediction.

6. Compare the predicted distribution maps of both “*Ficus aurea* future” and “*Ficus aurea* truncated”. How are they different? Besides the obvious fewer dots on the map, the gradients are more extreme (almost “all or nothing” in the context of the species being found in a specific area).

7. The generated data can also tell us if enough data points are being used to create useful predictions. Find the graph at the top of the analysis labeled “Omission and Predicted Area for Ficus aurea”. Ideally, the blue, green, and black line will be very close together on the graph if enough data is used. How do the graphs vary between “*Ficus aurea* future” and “*Ficus aurea* truncated” and what does it indicate?
For “*Ficus aurea* future”, the lines are all relatively close together indicating enough data was used. For “*Ficus aurea* truncated”, the lines are all very far apart indicating that not enough data was used.

8. How is the theory of evolution supported by the future models of distribution? The range of distributions is changing

9. What possible mechanisms are causing the species to be more abundant in the new areas predicted by the future distribution? Explain each possible mechanism.

Natural selection: Those who can survive, survive and pass on their traits

Genetic drift: The drift in allele frequency that is often caused by chance

Gene flow: The physical movement of genes from one location to another

LESSON TWO: "HOW TO GET NICHE, FAST!"

KEY QUESTION(S): How do we obtain and prepare data that can be used to make distribution predictions? How can we use software to predict the current and future distributions (evolution) of a specific Florida species?

KEY SCIENCE CONCEPTS: Biological niche, evolution, databases, data collection, Microsoft Excel®, scientific predictions, environmental change, distribution modeling.

OVERALL TIME ESTIMATE: One day (55 minutes).

LEARNING STYLES: Visual, logical.

VOCABULARY:
ENM software- Ecological Niche Modeling software. Software that uses environmental information and georeferenced species data to model the possible distribution of an organism.
Niches- Any location in the environment in which the organisms can survive. See "Lesson 1" for more information about "niche" definition.
Macroevolution- Change to the frequency of an allele within a population over time

LESSON SUMMARY: Students will choose a specific Florida species in order to investigate distribution changes (evolution) through the next thirty-five years. Data collection through the use of online databases will be used to compile coordinate information that can then be used in the ENM software to make these predictions.

STUDENT LEARNING OBJECTIVES:
Students will be able to:
- Use a database to collect data
- Use Microsoft Excel® to manipulate data into a usable form
- Use modeling software to predict current population distributions of a chosen species
- Use modeling software to predict future population distributions and the evolution of that species
- Graphically depict how a niche can change over time
- Determine the most important factor(s) in the evolution of that population's distribution
- Predict what could happen to their chosen species if any apparent trends were to continue
- Decide if these types of data could be important for determining policies in the future

STANDARDS
<u>Next Generation Sunshine State Standards</u>

SC.912.L.15.1	**SC.912.L.15.3**
SC.912.L.17.13	**SC.912.L.17.15**
SC.912.L.17.4	**SC.912.L.17.8**
SC.912.N.1.1	**SC.912.N.1.3**
SC.912.N.1.4	**SC.912.N.3.5**
SC.912.N.4.1	

<u>Advanced Placement (AP) Biology Learning Outcomes(LO)/Science Practices (SO)</u>*:*

LO 1.1	**LO 1.2**	**LO 1.3**
LO 1.4	**LO 1.5**	**LO 1.9**
LO 1.12	**LO 1.13**	**LO 1.22**
LO 1.25	**LO 1.26**	

Science Practices 1,4,5,6

Advanced International Certificate of Education (AICE) Biology Learning Outcomes:

K. a **P.d** **P. h**

MATERIALS:

Computers with Internet access, required .zip file (see "Lesson 1"), and Microsoft Excel®.

BACKGROUND INFORMATION:

There is little doubt in the scientific community that the climate is changing (see "Ecological and evolutionary responses to recent climate change", Parmesan, 2006 for a thorough review of this topic). It is a logical conclusion that if the climate is changing, then the distribution of organisms within these changing climates will also shift.[1,2] Because of the advancement in both computer hardware and software, scientists have moved to computer simulations to predict **niches**.[3,4] It is now possible to use environmental data to model current and future distributions of organisms.[5-11] Using computer simulations of how distributions could change (evolve) in the future will most likely become more prevalent as the digitization of data is become more widespread. For further information in how the use of computer simulations is permeating ecological niche modeling, see "Species Distribution Modeling and the Challenge of Predicting Future Distributions" (Graham et al., 2011).

Climate data is freely available to the public, but finding it in a format that is easy to incorporate into curricula can be a challenge. In this module, predefined climate data is provided in "layers". The software will compile and overlay the different layers in order to make its predictions. These specific data were provided by the Soltis Lab at the University of Florida and will only work for specimens found within the state. The layers that are specifically included within this module are (from http://worldclim.org/):

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
Alt= Altitude
GeolEP= Geology

Other geographic regions and environmental layers are available from http://worldclim.org/ but their use is beyond the scope of this module.

Although students have already walked through the use of the **ENM software** (if "Lesson 1" was completed), there was nothing new or groundbreaking actually discovered by the student. In this lesson, students will be modeling the niche of an organism of their *own* choosing. It is possible that the species that they will choose has never been modeled before, either for current conditions or future conditions. Because of this, the data used to complete the model must be compiled by the students. Students will use the website http://www.gbif.org/ in order to do this. This database houses the largest collection of biodiversity information freely available on the Internet. Although this database contains over 400 million records (as of July, 2014), many species have not been added either because of lack of species information or deficiency of digitized records (see "Lesson 3" of this module for how students can get involved in the digitization process). For this module, students will need to choose a species that has been documented in abundance (>30 georeferenced data points) in the state of Florida. Once a species has been chosen (and enough records are confirmed) the georeferenced data can be downloaded (see "Procedure and Discussion").

Microsoft Excel® will then be used to slightly modify the data and put them in a usable format for the ENM software. The use of software that features spreadsheet manipulation is a vital skill that students should have when entering their desired fields of study. This module does not linger on its usage in the classroom, but instead allows students to be exposed to how a spreadsheet is set up and how to create a new spreadsheet. Once the data have been put into a usable format, students will use these data in the ENM software to predict current and future distributions of their chosen organism within the state of Florida.

There is a very good chance that the predicted distribution will change over the next 35 years. This change is **macroevolution**. It can be hard to visualize (and teach) how specific populations evolve through time, but students using this software are doing just that.

ADVANCE PREPARATION:
Make sure that all students have access to the Internet, required .zip file, and Microsoft Excel®.

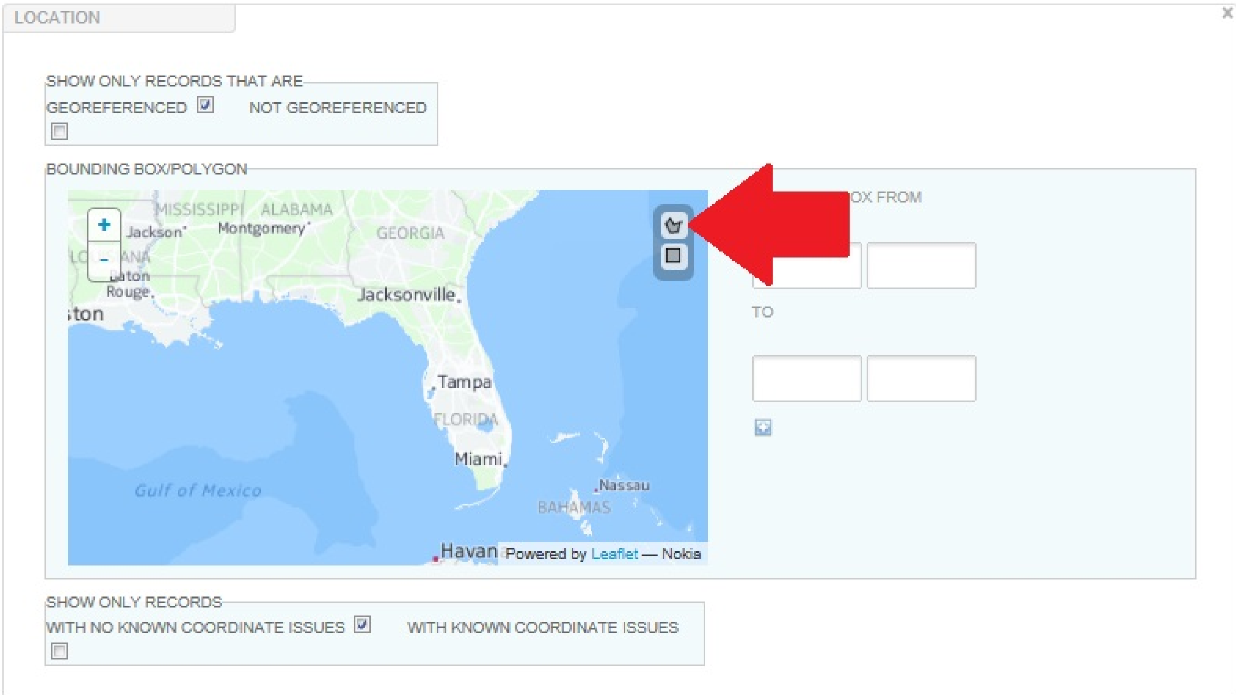
Required Laboratory Equipment:
Computers with Internet access, required .zip file, and Microsoft Excel®.

PROCEDURE AND DISCUSSION QUESTIONS WITH TIME ESTIMATES:

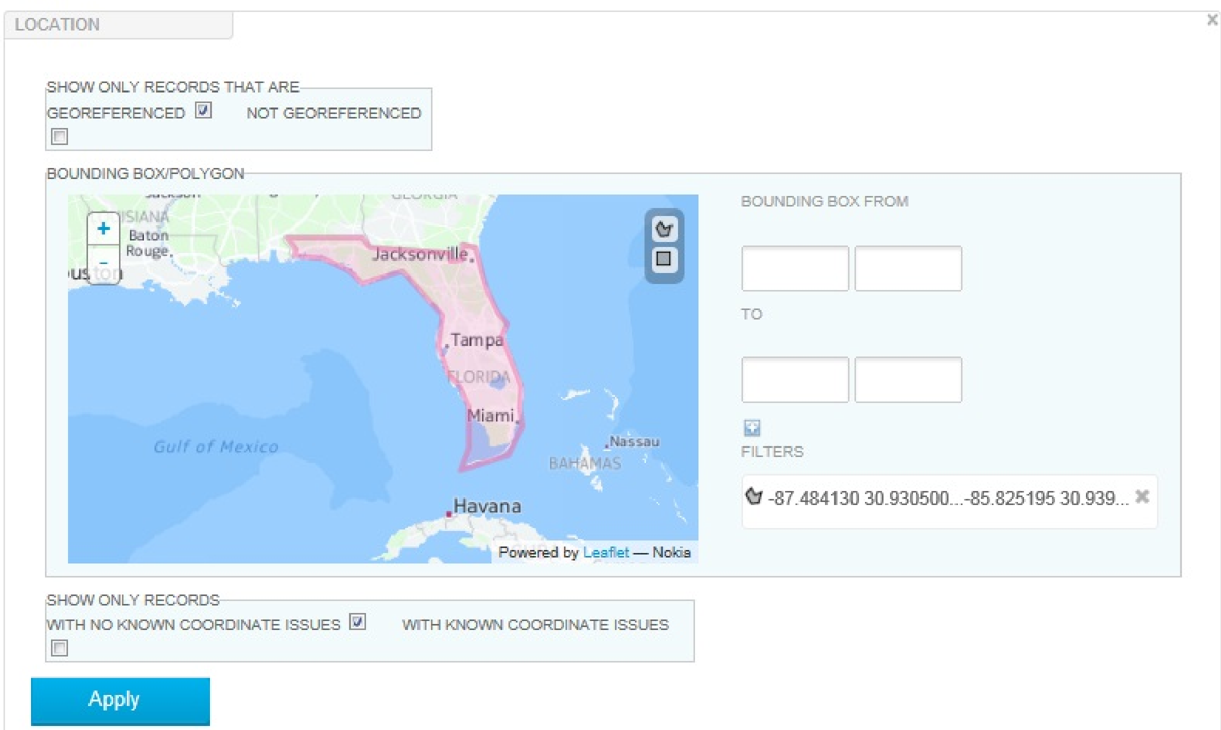
Students choose Florida organism, 5 minutes
Although any native Florida organism can be chosen for this project, students need to make sure that enough data is available to model the distributions of it (will be checked in the next step "Data collection via GBIF"). For ease, a website that has a modest list of common Florida species can be found at http://www.swfwmd.state.fl.us/recreation/species/ and will be recommended for students to choose their organism. It is up to the discretion of the instructor if a student wants to use a different organism not in these field guides. Students will then answer a small set of questions about their chosen organism.

Data collection via GBIF, 20 minutes
Once the organism has been chosen, students will use the Global Biodiversity Information Facility (GBIF) to determine if their chosen species has enough data points available to accurately model its distribution. Remember, more than 30 data points should be available to make an accurate distribution prediction. Greater amounts of data points will give more accurate predictions. For the sake of this procedural walkthrough, this author will choose the Florida organism *Hyla squirella*, or the Squirrel Treefrog. The steps for data collection are found below:

1. Go to http://www.gbif.org/
2. Create a new account. Make sure students use an email address that is accessible at the school location (if the school does not allow any type of email access, then these steps of choosing an organism and data collection via GBIF will have to be completed at home).
3. Find the dropdown menu entitled "Data" and click on "Explore species".
4. Enter the scientific name of the organism into the search box; in this case, *Hyla squirella.* Four results are shown, with the most probable match being listed first. Click on the top link to be taken to that organism's information page.
5. This new webpage has an overview of the organism, taxonomy, georeferenced data, and description. We are most concerned with the georeferenced data. This map shows the locations of the georeferenced data and it is possible to zoom in move around the map.
6. In the georeferenced data area, find where it says "VIEW RECORDS" and click on the blue text that states how many total records were found. In this case, 2,803 total records were found for *Hyla squirella.*
7. This new webpage now shows all of the georeferenced data and needs to be filtered to show just results found in Florida. Click the blue text that says "Add a filter". Once the drop-down menu appears, select "Location".
8. A map will show up allowing you to select your location. Zoom into Florida and select the tool that is shown below next to the red arrow:

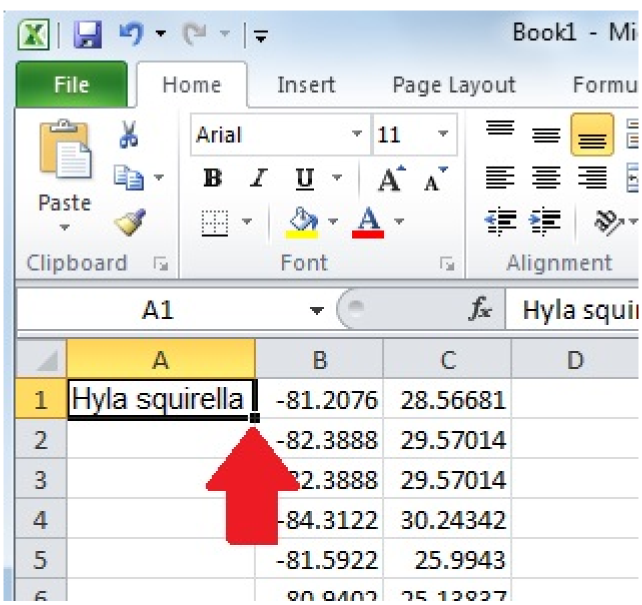


9. This tool allows you to encompass the correct area. Because Florida is oddly shaped, this tool is necessary, but tricky to use. It will most likely take a couple tries to encompass the area that you desire, but try to keep the area INSIDE of Florida. The map should then look something like this:

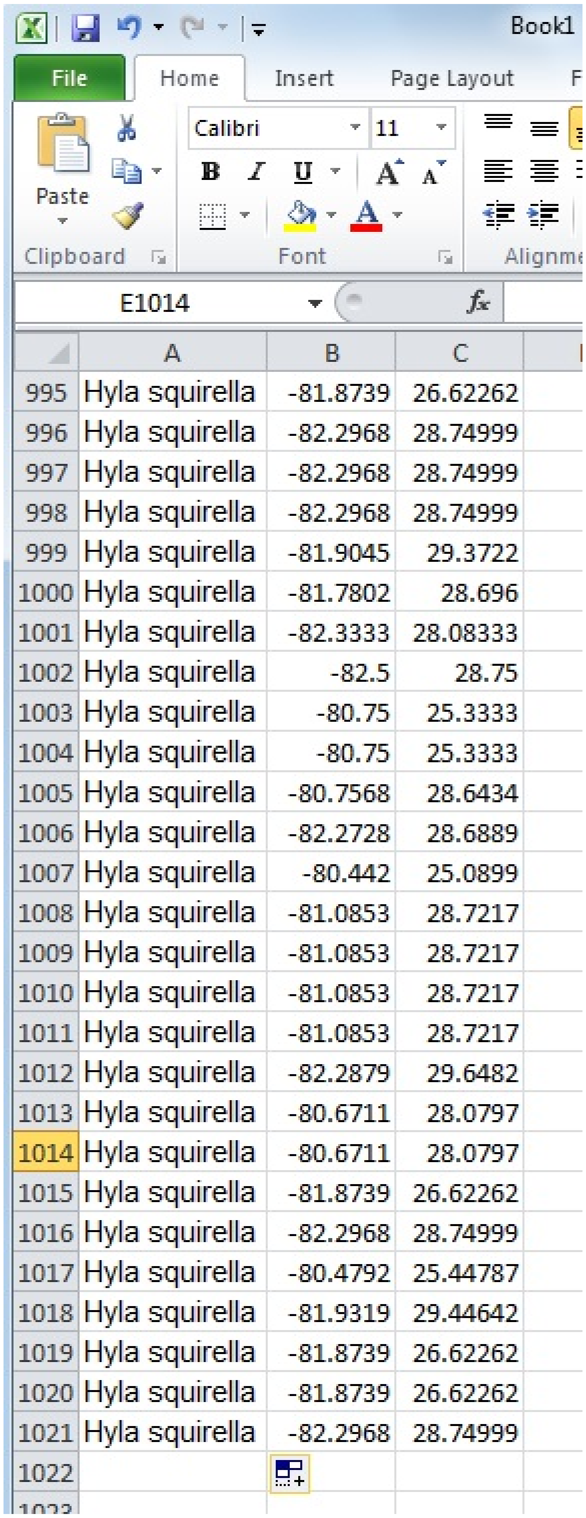


10. Click "Apply". The data will now be filtered to show only the georeferenced data inside of this area. Ensure that more than 30 georeferenced data points are still present after the results are filtered.
11. The results now need to be downloaded. It will prompt the user to login, if not already logged in. Proceed with the download. The link to the needed data will be emailed to the student's email used to set up the account. The amount of time that this takes will vary, but it is usually in a couple minutes (but will depend on how many data points were present).
12. Download and save the file using the link that was emailed to you.
13. Open the downloaded file and find the file entitled "occurrence". Copy and paste this single file to your desktop.

Data preparation with Microsoft Excel®, 10 minutes
1. Open Microsoft Excel®
2. In Microsoft Excel®, go to "File" and then "Open". Navigate to where you placed the copy and pasted file entitled "occurrence" (the user might have to make sure that the dialog window is showing ALL files, not just Excel files).
3. Depending on the version of Microsoft Excel® being used, you will most likely have to click "Okay" several times to load the data. There should be a large amount data that is shown in many different columns and row. The only needed data are found in individual columns and are entitled "decimalLatitude" and "decimalLongitude". Once these data have been found, a new spreadsheet needs to be created. This new spreadsheet will allow the data to be put in a usable format for the ENM software.
4. Open a new spreadsheet in Microsoft Excel® (it will be easier to open a new file than to add tabs to the current spreadsheet).
5. Three columns will be used in this new spreadsheet. WITHOUT titling column "B", copy and paste all of the longitudes into column "B".
6. WITHOUT titling column "C", copy and paste all of the latitudes into column "C". There should be an equal number of longitudes and latitudes on the spreadsheet.
7. WITHOUT titling column "A", write in the scientific name of the organism into row "1". Once it is in row "1", click on the cell. This should put a box around cell "A1". On the bottom of the box, there should be a small area that when you mouse over it a "+" shows up, as seen here:



8. Click (and hold) while dragging to the bottom of the data. This should place the scientific name into all cells in column “A”, as shown here:



	A	B	C
995	Hyla squirella	-81.8739	26.62262
996	Hyla squirella	-82.2968	28.74999
997	Hyla squirella	-82.2968	28.74999
998	Hyla squirella	-82.2968	28.74999
999	Hyla squirella	-81.9045	29.3722
1000	Hyla squirella	-81.7802	28.696
1001	Hyla squirella	-82.3333	28.08333
1002	Hyla squirella	-82.5	28.75
1003	Hyla squirella	-80.75	25.3333
1004	Hyla squirella	-80.75	25.3333
1005	Hyla squirella	-80.7568	28.6434
1006	Hyla squirella	-82.2728	28.6889
1007	Hyla squirella	-80.442	25.0899
1008	Hyla squirella	-81.0853	28.7217
1009	Hyla squirella	-81.0853	28.7217
1010	Hyla squirella	-81.0853	28.7217
1011	Hyla squirella	-81.0853	28.7217
1012	Hyla squirella	-82.2879	29.6482
1013	Hyla squirella	-80.6711	28.0797
1014	Hyla squirella	-80.6711	28.0797
1015	Hyla squirella	-81.8739	26.62262
1016	Hyla squirella	-82.2968	28.74999
1017	Hyla squirella	-80.4792	25.44787
1018	Hyla squirella	-81.9319	29.44642
1019	Hyla squirella	-81.8739	26.62262
1020	Hyla squirella	-81.8739	26.62262
1021	Hyla squirella	-82.2968	28.74999
1022			

9. Make sure the data is formatted as above and then save the file. Go to “file”, “save”, give the file a name of your choice, and then save as **“CSV (Comma delimited)”.** Click “save”. A window might pop up asking if this is how you want to save the files. Click “OK” and/or “Yes”.

Predictions of current and future distributions with ENM software, 10 minutes
Now that the data are in the proper format, the ENM software will be used in exactly the same way as it was used in “Lesson 1”, except that a different species data will be used. Students might need to refer back to “Lesson 1”, but eventually the predictions should be created as completed in the earlier lesson.

Analysis of created data, 10 minutes
Students will analyze the created distribution predictions in a very similar manner to “Lesson 1”.

ASSESSMENT SUGGESTIONS:
By completing all of the questions included with “Lesson 2”, students will be able to:
- Use GBIF to collect data on a specific species
- Use Microsoft Excel® to manipulate raw data into a usable form for the ENM software
- Use modeling software to predict current population distributions of a chosen species
- Use modeling software to predict future population distributions and the evolution of that species
- Graphically depict how a niche can change over time
- Determine the most important factor(s) in the evolution of that population’s distribution
- Predict what could happen to their chosen species if any apparent trends were to continue
- Decide that these types of data could be important for determining policies in the future

EXTENSIONS:
1. Students can research multiple different species and see if certain distribution trends show up.
2. Students can save the maps, print them, and create posters for their chosen organisms which show how the distribution of it will change over the next ~35 years. This extension can then be incorporated into public campaigns or used on Earth Day to raise awareness around the school or community.

RESOURCES/REFERENCES:

1. La Sorte, F., & F. Thompson (2007). Poleward shifts in winter ranges of North American birds. *Ecology* **88**:1803-1812.
2. Moritz, C., J. Patton, C. Conroy, J. Parra, G. White, and S. Beissinger (2008). Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science* **322**:261-264.
3. Peterson, T. (2001). Predicting species' geographic distributions based on ecological niche modeling. *Condor* **103**, 599–605.
4. Peterson, T., M. Ortega-Huerta, J. Bartley, V. Sanchez-Cordero, J. Soberon, R.Buddemeier, and D. Stockwell (2002). Future projections for Mexican faunas under global climate change scenarios. *Nature* **416**:626-629.
5. Anciaes, M., & A. Peterson (2006). Climate change effects on neotropical manakin diversity based on ecological niche modeling. *Condor* **108**:778-791.
6. Soberón, J. & A. Peterson (2005). Interpretation of models of fundamental ecological niches and species’ distributional areas. *Biodiversity Informatics*, **2**: 1-10.
7. Pounds, J., M. Fogden, and J. Campbell (1999). Biological response to climate change on a tropical mountain. *Nature* **398**:611–615.
8. Peterson, A., V. Sánchez-Cordero, J. Soberón , J. Bartley , R. Buddemeier, A. Navarro-Sigüenzae (2001). Effects of global climate change on geographic distributions of Mexican Cracidae. *Ecological Modelling* **144**:21-30.
9. Peterson, A., E. Martínez-Meyer, C. González-Salazar, P. Hall (2004). Modeled climate change effects on distributions of Canadian butterfly species. *Can. J. Zool*. **82**:851–858.
10. Root, T., J. Price, K. Hall, S. Schneider, C. Rosenzweig, J. Pounds (2003). Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57-60.

11. Roura-Pascual, N., A. Suarez, C. Gómez, P. Pons, Y. Touyama, A. Wild, and A. Peterson (2005). Geographic potential of Argentine ants (*Linepithema humile Mayr*) in the face of global climate change.Proceedings of the Royal Society of London. B **271**:2527-2535.

Graham , C., B. Loiselle , J. Velasquez , and F. Cuesta (2011). Species distribution modeling and the challenge of predicting future distributions. In S. K. Herzog , R. Martínez , P. M. Jørgensen , and H. Tiessen [EDS.], Climate Change and Biodiversity in the Tropical Andes. IAI-SCOPE, São José dos Campos, Brazil, pp 295–310.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics* **37**:637-669.

STUDENT PAGES: Background Information

There is little doubt in the scientific community that the climate is changing (see "Ecological and evolutionary responses to recent climate change", Parmesan, 2006 for a thorough review of this topic). It is a logical conclusion that if the climate is changing, then the distribution of organisms within these changing climates will also shift.[1,2] Because of the advancement in both computer hardware and software, scientists have moved to computer simulations to predict niches.[3,4] It is now possible to use environmental data to model current and future distributions of organisms.[5-11] Using computer simulations of how distributions could change(evolve) in the future will most likely become more prevalent as the digitization of data is become more widespread. For further information in how the use of computer simulations is permeating ecological niche modeling, see "Species Distribution Modeling and the Challenge of Predicting Future Distributions" (Graham et al., 2011).

Climate data is freely available to the public, but finding it in a format that is easy to use can be a challenge. In this module, predefined climate data is provided in "layers". The software will compile and overlay the different layers in order to make its predictions. These specific data were provided by the Soltis Lab at the University of Florida and will only work for specimens found within the state. The layers that are specifically included within this module are (from http://worldclim.org/):

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
Alt= Altitude
GeolEP= Geology

In this lesson, you will be modeling the niche of an organism of your own choosing. It is possible that the species that you will choose has never been modeled before, either for current conditions or future conditions. Because of this, the data used to complete the model must be compiled. You will use the website http://www.gbif.org/ in order to do this. This database houses the largest collection of biodiversity information freely available on the Internet. Although this database contains over 400 million records (as of July, 2014), many species have not been added either because of lack of species information or deficiency of digitization. For this module, you will need to choose a species that has been documented in abundance **(>30 georeferenced data points**) in the state of Florida. Once a species has been chosen (and enough records are confirmed) the georeferenced data will be downloaded.

You will then use Microsoft Excel® to slightly modify the data and put them in a usable format for the ENM software. Once the data have been put into a usable format, you will use these data in the ENM software to predict current and future distributions of your chosen organism within the state of Florida.

There is a very good chance that the predicted distribution will change over the next 35 years. This change is macroevolution. It can be hard to visualize how specific populations evolve through time, but by using this software, you are doing just that.

References

1. La Sorte, F., & F. Thompson (2007). Poleward shifts in winter ranges of North American birds. *Ecology* **88**:1803-1812.
2. Moritz, C., J. Patton, C. Conroy, J. Parra, G. White, and S. Beissinger (2008). Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science* **322**:261-264.
3. Peterson, T. (2001). Predicting species' geographic distributions based on ecological niche modeling. *Condor* **103**, 599–605.
4. Peterson, T., M. Ortega-Huerta, J. Bartley, V. Sanchez-Cordero, J. Soberon, R.Buddemeier, and D. Stockwell (2002). Future projections for Mexican faunas under global climate change scenarios. *Nature* **416**:626-629.
5. Anciaes, M., & A. Peterson (2006). Climate change effects on neotropical manakin diversity based on ecological niche modeling. *Condor* **108**:778-791.
6. Soberón, J. & A. Peterson (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**: 1-10.
7. Pounds, J., M. Fogden, and J. Campbell (1999). Biological response to climate change on a tropical mountain. *Nature* **398**:611–615.
8. Peterson, A., V. Sánchez-Cordero, J. Soberón , J. Bartley , R. Buddemeier, A. Navarro-Sigüenzae (2001). Effects of global climate change on geographic distributions of Mexican Cracidae. *Ecological Modelling* **144**:21-30.
9. Peterson, A., E. Martínez-Meyer, C. González-Salazar, P. Hall (2004). Modeled climate change effects on distributions of Canadian butterfly species. *Can. J. Zool*. **82**:851–858.
10. Root, T., J. Price, K. Hall, S. Schneider, C. Rosenzweig, J. Pounds (2003). Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57-60.
11. Roura-Pascual, N., A. Suarez, C. Gómez, P. Pons, Y. Touyama, A. Wild, and A. Peterson (2005). Geographic potential of Argentine ants (*Linepithema humile Mayr*) in the face of global climate change.Proceedings of the Royal Society of London. B **271**:2527-2535.

Graham , C., B. Loiselle , J. Velasquez , and F. Cuesta (2011). Species distribution modeling and the challenge of predicting future distributions. In S. K. Herzog , R. Martínez , P. M. Jørgensen , and H. Tiessen [EDS.], Climate Change and Biodiversity in the Tropical Andes. IAI-SCOPE, São José dos Campos, Brazil, pp 295–310.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics* **37**:637-669.

STUDENT HANDOUT: Laboratory Methodologies and Questions

Choose an organism!

Although any native Florida organism can be chosen for this project, you need to make sure that enough data is available to model the distributions of it (will be checked in the next step "Data collection via GBIF"). For ease, a website that has a modest list of common Florida species can be found at http://www.swfwmd.state.fl.us/recreation/species/ .

Scientific name of chosen organism: ____

Common name of chosen organism: ____

Data collection via GBIF

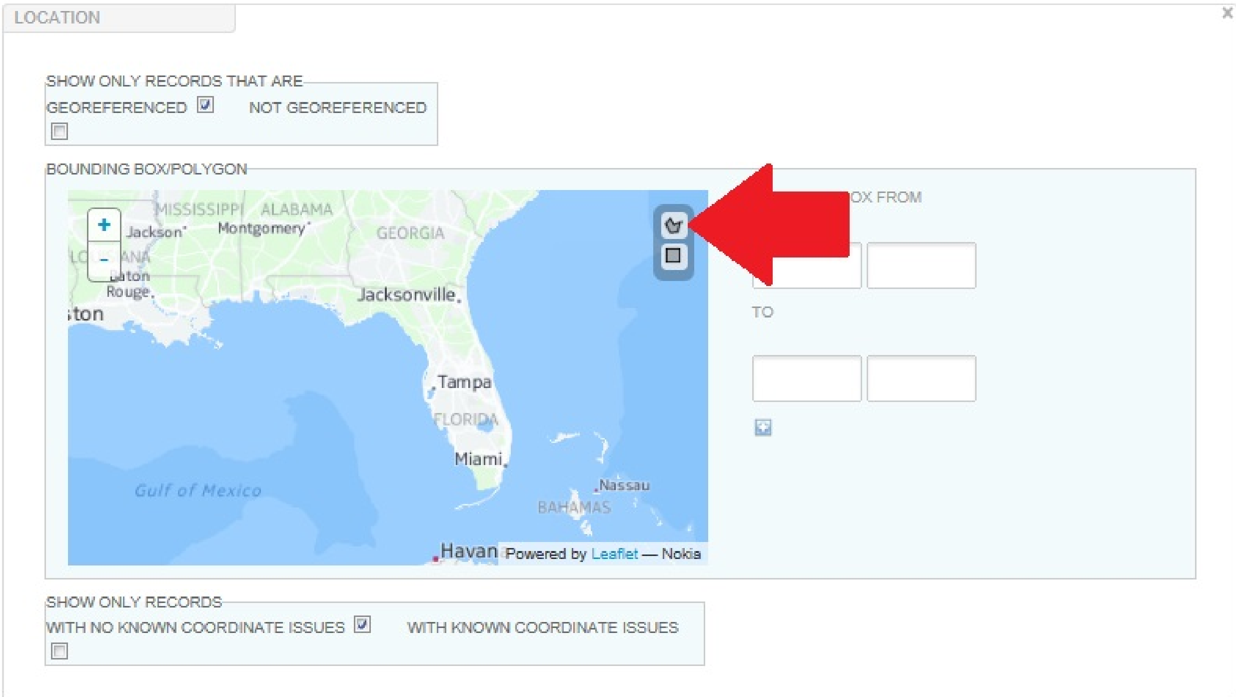
Once the organism has been chosen, you will use the Global Biodiversity Information Facility (GBIF) to determine if your chosen species has enough data points available to accurately model its distribution. Remember, more than 30 data points should be available to make an accurate distribution prediction. Greater amounts of data points will give more accurate predictions. The steps for data collection are found below:

1. Go to http://www.gbif.org/
2. Create a new account. Make sure you use an email address that is accessible at the school location (if the school does not allow any type of email access, then these steps of choosing an organism and data collection via GBIF will have to be completed at home).
3. Find the dropdown menu entitled "Data" and click on "Explore species".
4. Enter the scientific name of your organism into the search box. Results are shown, with the most probable match being listed first. Click on the top link to be taken to that organism's information page.
5. This new webpage has an overview of the organism, taxonomy, georeferenced data, and description. We are most concerned with the georeferenced data. This map shows the locations of the georeferenced data and it is possible to zoom in move around the map.
6. In the georeferenced data area, find where it says "VIEW RECORDS" and click on the blue text that states how many total records were found.

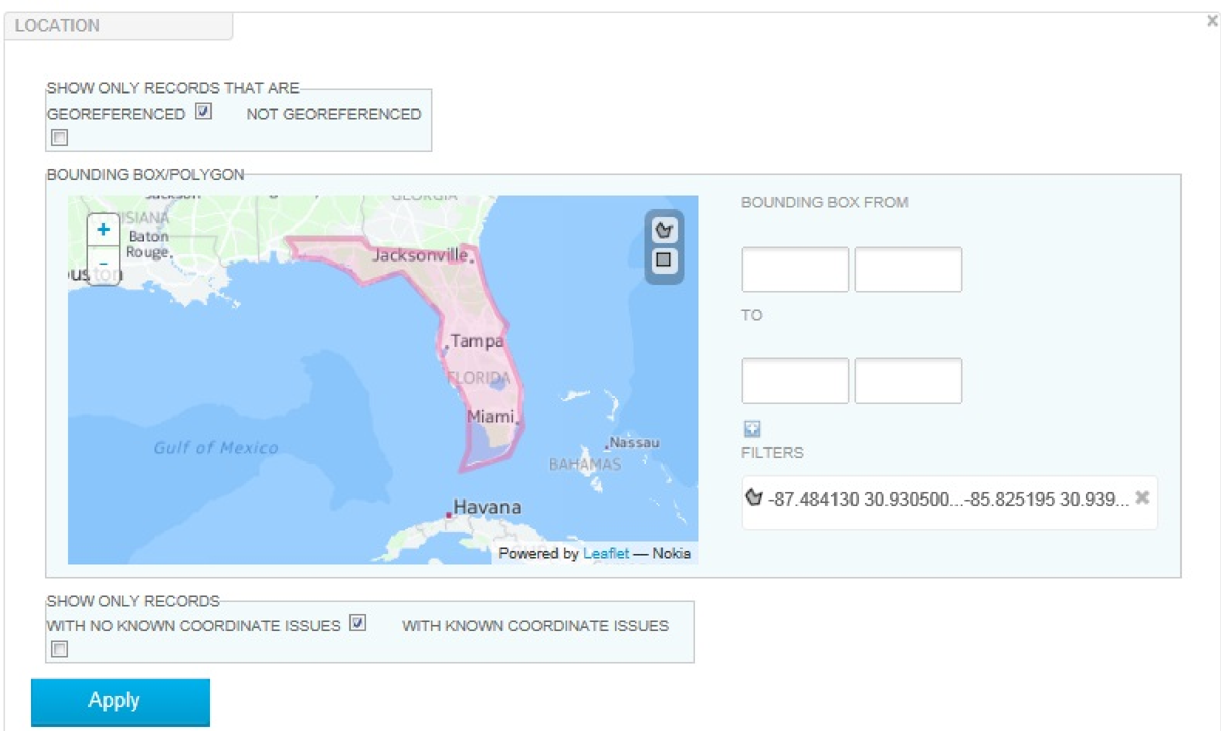
How many total records are found? ____________

If it is less than 30, then choose a different organism.

7. This new webpage now shows all of the georeferenced data and needs to be filtered to show just results found in Florida. Click the blue text that says "Add a filter". Once the drop-down menu appears, select "Location".
8. A map will show up allowing you to select your location. Zoom into Florida and select the tool that is shown below next to the red arrow:



9. This tool allows you to encompass the desired area (we will only be using environmental data for the state of Florida. Because Florida is oddly shaped, this tool is necessary, but tricky to. It will most likely take a couple tries to encompass the area that you desire, but try to keep the area INSIDE of Florida. The map should then look something like this (you will have to zoom in further to adequately select an area):



10. Click “Apply”. The data will now be filtered to show only the georeferenced data inside of this area. **Ensure that more than 30 georeferenced data points are still present after the results are filtered.**

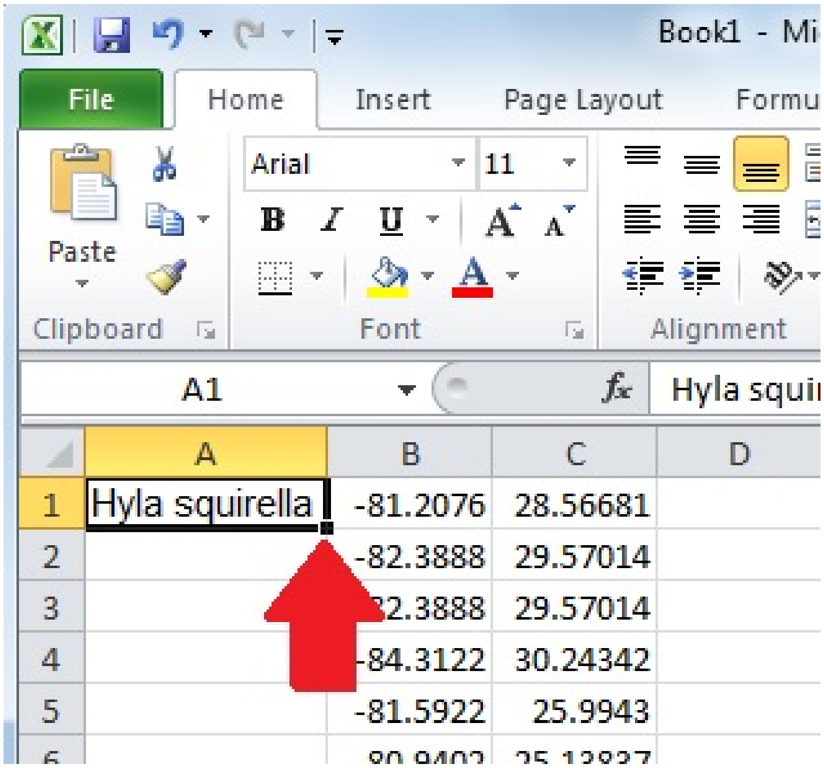
How many georeferenced records are found? ____________

If it is less than 30, then choose a different organism.

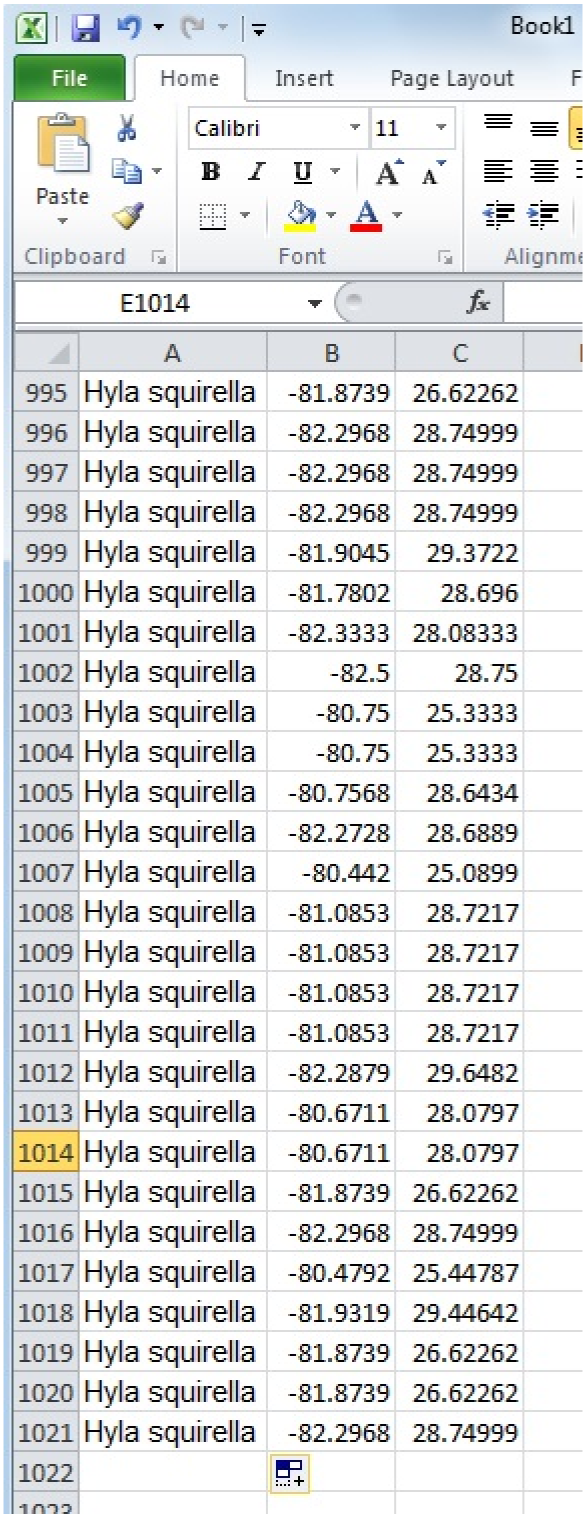
11. The results now need to be downloaded. It will prompt you to login, if not already logged in. Proceed with the download. The link to the needed data will be emailed to your email used to set up the account. The amount of time that this takes will vary, but it is usually in a couple minutes (but will depend on how many data points were present).
12. Download and save the file using the link that was emailed to you.
13. Open the downloaded file and find the file entitled “occurrence”. Copy and paste this singular file to your desktop.

Data preparation with Microsoft Excel®
1. Open Microsoft Excel®
2. In Microsoft Excel®, go to “File” and then “Open”. Navigate to where you placed the copy and pasted file entitled “occurrence” (you might have to make sure that the dialog window is showing ALL files, not just Excel files).
3. Depending on the version of Microsoft Excel® being used, you will most likely have to click “Okay” several times to load the data. There should be a large amount data that are shown in many different columns and rows. The only needed data are found in individual columns and are entitled "decimalLatitude” and “decimalLongitude”. Once these data have been found, a new spreadsheet needs to be created. This new spreadsheet will allow the data to be put in a usable format for the ENM software.
4. Open a new spreadsheet in Microsoft Excel® (it will be easier to open a new file than to add tabs to the current spreadsheet).
5. Three columns will be used in this new spreadsheet. WITHOUT titling column “B”, copy and paste all of the longitudes into column “B”.
6. WITHOUT titling column “C”, copy and paste all of the latitudes into column “C”. There should be an equal number of longitudes and latitudes on the spreadsheet.

7. WITHOUT titling column "A", write in the scientific name of the organism into row "1". Once it is in row "1", click on the cell. This should put a box around cell "A1". On the bottom of the box, there should be a small area that when you mouse over it a "+" shows up, as seen here:



8. Click (and hold) while dragging to the bottom of the data. This should place the scientific name into all cells in column “A”, as shown here:



	A	B	C
995	Hyla squirella	-81.8739	26.62262
996	Hyla squirella	-82.2968	28.74999
997	Hyla squirella	-82.2968	28.74999
998	Hyla squirella	-82.2968	28.74999
999	Hyla squirella	-81.9045	29.3722
1000	Hyla squirella	-81.7802	28.696
1001	Hyla squirella	-82.3333	28.08333
1002	Hyla squirella	-82.5	28.75
1003	Hyla squirella	-80.75	25.3333
1004	Hyla squirella	-80.75	25.3333
1005	Hyla squirella	-80.7568	28.6434
1006	Hyla squirella	-82.2728	28.6889
1007	Hyla squirella	-80.442	25.0899
1008	Hyla squirella	-81.0853	28.7217
1009	Hyla squirella	-81.0853	28.7217
1010	Hyla squirella	-81.0853	28.7217
1011	Hyla squirella	-81.0853	28.7217
1012	Hyla squirella	-82.2879	29.6482
1013	Hyla squirella	-80.6711	28.0797
1014	Hyla squirella	-80.6711	28.0797
1015	Hyla squirella	-81.8739	26.62262
1016	Hyla squirella	-82.2968	28.74999
1017	Hyla squirella	-80.4792	25.44787
1018	Hyla squirella	-81.9319	29.44642
1019	Hyla squirella	-81.8739	26.62262
1020	Hyla squirella	-81.8739	26.62262
1021	Hyla squirella	-82.2968	28.74999
1022			
1023			

9. Make sure the data is formatted as above and then save the file. Go to “file”, “save”, give the file a name of your choice, and then save as **“CSV (Comma delimited)”**. After clicking “save”, a window might pop up asking if this is how you want to save the files. Click “OK” and/or “Yes”.

Predict current and future distributions of your species with ENM software
Now that the data are in the proper format, the ENM software will be used in the exactly same way as it was used in “Lesson 1”, except that a different species data will be used. You might need to refer back to “Lesson 1”, but eventually the predictions should be created as completed in the earlier lesson.

Analysis of prediction data
Begin by opening the browser file for your newly processed distribution prediction for CURRENT conditions.

1. Are there any distribution trends that can be seen on the map?

2. According to the software, what was the most important layer used to create this current prediction?

3. If there are specific locations where your organism were found more commonly, then propose a hypothesis why this is (you might have to do a small amount of research about your organism to determine what type(s) of habitats it prefers to live in):

Open the browser file for your newly processed distribution prediction for FUTURE (2050) conditions.

4. Compare the maps (current vs. future) and determine how the distribution of your organism will change over the next ~35 years. Keep in mind, the changes will often be subtle, so make the maps as large as possible and look closely!

5. According to the software, what was the most important layer used to create this future prediction?

6. If the distribution of your chosen species is increasing, then what could this lead to in relation to species diversity and why (if the trend continues)?

7. If the distribution of your chosen species is decreasing, then what could this lead to in relation to species diversity and why (if the trend continues)?

8. How could the usage of the predicted distributions change environmental policy in the future?

9. What is the benefit of using computers to make predictions?

TEACHER PAGES: *KEY* FOR "Laboratory Methodologies and Questions"

Analysis of prediction data

1. Are there any distribution trends that can be seen on the map? This will vary depending on the organism chosen.

2. According to the software, what was the most important layer used to create this current prediction? This will vary depending on the organism chosen.

3. If there are specific locations where your organism were found more commonly, then propose a hypothesis why this is (you might have to do a small amount of research about your organism to determine what type(s) of habitats it prefers to live in): This will vary depending on the organism chosen, but most answers should relate back to question #2. The most important layer in predicting the distribution will partially dictate *why* the organism is predicted to be in certain areas.

4. Compare the maps (current vs. future) and determine how the distribution of your organism will change over the next ~35 years. Keep in mind, the changes will often be subtle, so make the maps as large as possible and look closely! This will vary depending on the organism chosen.

5. According to the software, what was the most important layer used to create this future prediction? This will vary depending on the organism chosen.

6. If the distribution of your chosen species is increasing, then what could this lead to in relation to species diversity and why (if the trend continues)? An increase in diversity because of increase in alleles present in the population. Possible speciation in the future.

7. If the distribution of your chosen species is decreasing, then what could this lead to in relation to species diversity and why (if the trend continues)? A decrease in diversity because of decrease in alleles present in the population. Possible extinction in the future.

8. How could the usage of the predicted distributions change environmental policy in the future? Policy could be created to protect species that will decline or policies to keep non-native species under control.

9. What is the benefit of using computers to make predictions? Can compile a huge amount of complex data and test complex hypotheses very quickly.

LESSON THREE: "SCIENTIFIC CITIZEN THROUGH CITIZEN SCIENCE"

KEY QUESTION(S): How can an average citizen contribute to the body of usable scientific information? Why is the contribution from all citizens so important?

KEY SCIENCE CONCEPTS: Citizen science, databases, data/information analysis.

OVERALL TIME ESTIMATE: One day (55 minutes)

LEARNING STYLES: Visual, logical, critical.

LESSON SUMMARY: Students will be introduced and participate in a "citizen science" program in order to see how their contributions can be used by researchers. Additionally, students will brainstorm possibly new citizen science programs.

STUDENT LEARNING OBJECTIVES:
Students will be able to:
- Understand what "citizen science" is and why it is important
- Actively participate in a citizen science project in order to contribute to the body of scientific knowledge
- Assess different types of databases in order to determine what the goal of specific citizen science project is
- Accurately analyze data of different types (the data type being analyzed will depend on the citizen science project they are most interested in)
- Brainstorm novel citizen science idea that can then be further perused by the individual

STANDARDS
<u>Next Generation Sunshine State Standards</u>

SC.912.L.17.15	**SC.912.N.1.1**
SC.912.N.1.3	**SC.912.N.1.4**
SC.912.N.4.1	

<u>Advanced Placement (AP) Biology Learning Outcomes(LO)/Science Practices (SO)</u>*:*
Science Practices 4,5,6

MATERIALS: Computers with Internet access.

BACKGROUND INFORMATION:
"Citizen Science" is a newly emerging term that can be defined as "partnerships between scientists and non-scientists in which authentic data are collected, shared, and analyzed".[1] Although the term might be new, the concept is not. The concept of average citizens collecting scientific data stretches back into the 1800s when lighthouse keepers were instructed to keep track of the birds that would hit the lighthouse.[2] There are numerous other projects that stretch back hundreds of years and rely on citizens to make observations.[2] In 1992, the National Science Foundation funded one of the first "official" citizen science projects that was created by the Cornell Lab of Ornithology and was entitled "Public Participation in Ornithology".[3] These projects required the public to collect and submit data in order to answer specific questions established by the Cornell research team.[3] The actual phrase "Citizen Science" (in context of using the public to collect data) was first coined by Rick Bonney in 1995 and is now a well-accepted term in the field of research.[4]

Originally, there was a great deal of skepticism about the accuracy of data collected and submitted by the public, but many studies have shown that basic data collection completed by the public is just as accurate as data generated by a technician.[2] In addition to this, the calculated cost savings can reach into the tens of thousands of dollars making once impossible projects now a reality.[2] The abundance of citizen science projects has grown over the last decade and is really only limited by the creativity of the authors and accessibility of the program.

The importance of these programs reaches beyond helping a researcher with their project. Multiple studies have shown that citizen science also benefits the participants in different ways including:[3,5-7]

- Increasing awareness and understanding of key scientific concepts related to the study
- Gaining knowledge in the process of science
- Increasing scientific inquiry skills
- Helping build community
- Increasing the knowledge of environmental regulation
- Increasing the engagement in science
- Helping develop interests
- Developing science skills such as identification of species, using measurement devices, collecting data, and following protocols
- Developing skills in reading, interpreting graphs, drawing conclusions from evidence, raising additional questions.

There is little question that we want our students to develop these skills as they enter our classroom and then continue them as contributing citizens when they exit our classroom. The citizen science projects presented in this module offer variation in difficulties, topics, and outcomes. Allow students to explore these different projects, contribute to the scientific body of knowledge, and maybe be inspired to start a project of their own.

ADVANCE PREPARATION:
None required.

Required Laboratory Equipment:
Computers with Internet access.

PROCEDURE AND DISCUSSION QUESTIONS WITH TIME ESTIMATES:

Introduction to "Citizen Science", 10 minutes
Briefly discuss what "Citizen Science" is after they read the "Student Background" information. Explain what the goal of each of the following programs/websites is (goals/backgrounds are excerpts taken from each respective website):

"Notes from Nature"- http://www.notesfromnature.org/
- "The Notes from Nature transcription project is a citizen science platform built to address this problem [lack of digitized data] by digitizing the world's biological collections one record at a time!".

"Foldit"- http://fold.it/portal/
- "Figuring out which of the many, many possible structures [of a protein] is the best one is regarded as one of the hardest problems in biology today and current methods take a lot of money and time, even for computers. Foldit attempts to predict the structure of a protein by taking advantage of humans' puzzle-solving intuitions and having people play competitively to fold the best proteins."

"Nanocrafter"- http://nanocrafter.org/
- "Nanocrafter is a scientific discovery game about synthetic biology. Use pieces of DNA to build everything from computer circuits to nanoscale machines, and help advance scientific research with your inventions!"

"Eyewire"- https://eyewire.org/signup
- "EyeWire is a game to map the brain from Seung Lab at MIT. Anyone can play and you need no scientific background. Over 130,000 people from 145 countries already do. Together we are mapping the 3D structure of neurons; advancing our quest to understand ourselves."

"Eterna"- http://eterna.cmu.edu/web/

- "By playing EteRNA, you will participate in creating the first large-scale library of synthetic RNA designs. Your efforts will help reveal new principles for designing RNA-based switches and nanomachines -- new systems for seeking and eventually controlling living cells and disease-causing viruses."

These are only several of the citizen science websites available that focus specifically on biology/ecology. If students know of (or are actively participating in) other citizen science websites, please have that student discuss what the focus of that project is. It is understood that not all students will be interested in biology. Because of this, part of this lesson includes visiting the website https://www.zooniverse.org/#all. This website is maintained by the Citizen Science Alliance and includes an unbelievable collection of citizen science projects in a variation of areas including space, climate, nature (Notes from Nature is a Zooniverse project), humanities, and biology. Students should be able to find at least one project in the "Zooniverse" that captures their imagination.

Students contribute to an area of their own choosing, 30 minutes

Students can choose a project from https://www.zooniverse.org/#all or from the list of aforementioned projects not associated with Zooniverse. Once they have chosen a project, have students begin their work within the project. Many of these projects will have introductory information after becoming a member, but this will vary depending on the project.

Students brainstorm new "Citizen Science" ideas, 10 minutes

Put students into groups of four after they have been individually exposed to different genera of citizen science. Have these groups brainstorm new ideas of "citizen science" outlets. Although some of these ideas might already exist, there might be some novel ideas that spark a student to take action and create something new and unique.

ASSESSMENT SUGGESTIONS:

By completing all of the questions included with "Lesson 3", students will be able to:

- Understand that "citizen science" is a partnership between scientists and non-scientists in which authentic data are collected, shared, and analyzed and that it is important for the researcher and citizen
- Actively participate in a citizen science project in order to contribute to the body of scientific knowledge
- Assess different types of databases in order to determine what the goal of specific citizen science project is
- Accurately analyze data of different types (the data type being analyzed will depend on the citizen science project they are most interested in)
- Brainstorm novel citizen science idea that can then be further perused by the individual

EXTENSIONS:

- The goal of this lesson is to have students take an initiative to engage in real science. Hopefully, by exposing students to these different projects, they will spend some of their own time outside of the classroom working within the project of their own choosing.
- Many other citizen science projects exist. Included in these projects are "apps" such as "Inaturalist" (https://www.inaturalist.org/) which allows students to upload pictures and data of different organisms. Students could create accounts and upload data found around the school, their home neighborhood, or any other places.

RESOURCES/REFERENCES:

1. Jordan, R, H. Ballard, T. Phillips (2012). Key issues and new approaches for evaluating citizen-science learning outcomes. *Frontiers in Ecology and the Environment* **10**(6): 307–309

2. Droege, S. (2007). Just because you paid them doesn't mean their data are better. In C. McEver, Bonney, R., Dickinson, J., Kelling, S., Rosenberg, K., and Shirk, J. (Eds.). Citizen science toolkit conference. Ithaca, NY: Cornell Laboratory of Ornithology. Ithaca, NY. [online] URL: http://www.birds.cornell.edu/citscitoolkit/conference/proceeding-pdfs/Droege%202007%20CS%20Conference.Pdf

3. Bonney, R., H. Ballard, R. Jordan, E. McCallie, T. Phillips, J. Shirk, C. Wilderman. (2009) Public Participation in Scientific Research: Defining the Field and Assessing its Potential for Informal Science Education. In A CAISE Inquiry Group Report, Center for Advancement of Informal Science Education (CAISE), Washington, DC (Technical Report).

4. Bonney, R. (1996). Citizen science: A Lab tradition. *Living Bird* **15**(4), 7–15.

5. Trumbull, D., R. Bonney, K. Bascom, A. Cabrel (2000). Thinking scientifically during participation in a citizen-science project. *Sci Educ*-Netherlands **84**: 265–75.
6. Jordan, R., F. Singer, J. Vaughan, A. Berkowitz (2009). What should every citizen know about ecology? *Front Ecol Environ* **7**: 495–500.
7. Jordan, R., S. Gray, D. Howe, W. Brooks, J. Ehrenfeld (2011). Knowledge gain and behavioral change in citizen-science programs. *Conserv Biol* **25**: 1148–54.

STUDENT PAGES: Background Information

"Citizen Science" is a newly emerging term that can be defined as a "partnerships between scientists and non-scientists in which authentic data are collected, shared, and analyzed".[1] Although the term might be new, the concept is not. The concept of average citizens collecting scientific data stretches back into the 1800s when lighthouse keepers were instructed to keep track of the birds that would hit the lighthouse.[2] There are numerous other projects that stretch back hundreds of years and rely on citizens to make observations.[2] In 1992, the National Science Foundation funded one of the first "official" citizen science projects that was created by the Cornell Lab of Ornithology and was entitled "Public Participation in Ornithology".[3] These projects required the public to collect and submit data in order to answer specific questions established by the Cornell research team.[3] The actual phrase "Citizen Science" (in context of using the public to collect data) was first coined by Rick Bonney in 1995 and is now a well-accepted term in the field of research.[4]

Originally, there was a great deal of skepticism about the accuracy of data collected and submitted by the public, but many studies have shown that basic data collection completed by the public is just as accurate as data generated by a technician.[2] In addition to this, the calculated cost savings can reach into the tens of thousands of dollars making once impossible projects now a reality.[2] The abundance of citizen science projects has grown over the last decade and is really only limited by the creativity of the authors and accessibility of the program.

Background Questions

1. What do you think the benefits of citizen science projects could be for the researcher?

2. What do you think the benefits of citizen science projects could be for the citizen?

3. Write a very simple description (one sentence) of what each of the following citizen science projects is attempting to accomplish:

"Notes from Nature"- http://www.notesfromnature.org/

"Foldit"- http://fold.it/portal/

"Nanocrafter"- http://nanocrafter.org/

"Eyewire"- https://eyewire.org/signup

"Eterna"- http://eterna.cmu.edu/web/

References

1. Jordan, R, H. Ballard, T. Phillips (2012). Key issues and new approaches for evaluating citizen-science learning outcomes. *Frontiers in Ecology and the Environment* **10**(6): 307–309
2. Droege, S. (2007). Just because you paid them doesn't mean their data are better. In C. McEver, Bonney, R., Dickinson, J., Kelling, S., Rosenberg, K., and Shirk, J. (Eds.). Citizen science toolkit conference. Ithaca, NY: Cornell Laboratory of Ornithology. Ithaca, NY. [online] URL: http://www.birds.cornell.edu/citscitoolkit/conference/proceeding-pdfs/Droege%202007%20CS%20Conference.Pdf
3. Bonney, R., H. Ballard, R. Jordan, E. McCallie, T. Phillips, J. Shirk, C. Wilderman. (2009) Public Participation in Scientific Research: Defining the Field and Assessing its Potential for Informal Science Education. In A CAISE Inquiry Group Report, Center for Advancement of Informal Science Education (CAISE), Washington, DC (Technical Report).
4. Bonney, R. (1996). Citizen science: A Lab tradition. *Living Bird* **15**(4), 7–15.

STUDENT HANDOUT: Laboratory Methodologies and Questions

1. In "Lesson 2" you might have originally chosen a species that had less than 30 georeferenced data points available. Additionally, in "Lesson 1" you saw that the more data points that exist, the more accurate the prediction will be. It is very likely that there are more than 30 of these select species living in Florida, so why do you think there such a shortage of georeferenced data available?

You most likely came up a valid hypothesis for why these data are not recorded, but there are collections of millions of specimens that can be found tucked away at natural history museums around the world that have georeferenced data included with them. Some of these collections are hundreds of years old and have hand-written data about the specimen. These samples present an overwhelming challenge for scientists- can this information be digitized and how should this be done? Someone could spend their lifetime digitizing samples and barely make a dent in the collections that can be found worldwide. Because of this, different institutions have outsourced this work to the public. Citizens such as yourself can help digitize data in order to add to the body of scientific knowledge available to researchers. This growing body of knowledge allows the scientific community to make more accurate predictions which can impact policy decisions and the preservation of endangered species.

Go to http://www.notesfromnature.org/. "Notes from Nature" is a project that utilizes the public to help with the digitization of preserved museum species. Set up an account and follow the very simple instructions to start digitizing data! Digitize three specimens and answer the following questions as you proceed:

2. What are the scientific names of the specimens that you digitized and where were they originally found?

a.

b.

c.

There are many citizen science projects available. Go to https://www.zooniverse.org/#all and explore some other projects that utilize the power of citizens to accomplish amazing things. Choose one of the Zooniverse projects or other projects found in the background information, go to their respective website and get working! You will be actively contributing to the scientific community!

3. What project(s) are you investigating any why?

4. What safeguards do you think these programs have in order to keep the user-submitted data as accurate as possible?

After investigating some of the citizen science programs available, get into a group of four individuals.

5. As a group, come up with a novel (new) citizen science program. It can be a modification of a project that already exists, but should not be a direct copy of any that you saw today. Keep in mind, it does not need to be biological, it can be cultural as well. Outline your group's program below:

Title:

Problem that the program is meant to solve:

How the problem could be solved by utilizing citizen science:

TEACHER PAGES: ***KEY*** FOR "BACKGROUND INFORMATION"

1. What do you think the benefits of citizen science projects could be for the researcher?

Data collection is accurate, cheap, massive amounts of data can be processed quickly.

2. What do you think the benefits of citizen science projects could be for the citizen?

- An increase in awareness and understanding of key scientific concepts related to the study
- A gain of knowledge in the process of science
- An increase in scientific inquiry skills
- Helps build community
- An increase in the knowledge of environmental regulation
- An increase in the engagement in science
- Helps develop interests
- Develops science skills such as identification of species, using measurement devices, collecting data, and following protocols
- Develops skills in reading, interpreting graphs, drawing conclusions from evidence, raising additional questions.

3. Write a very simple description (one sentence) of what each of the following citizen science projects is attempting to accomplish:

"Notes from Nature"- http://www.notesfromnature.org/
- Digitizing the world's biological collections

"Foldit"- http://fold.it/portal/
- Predict the structure of a protein

"Nanocrafter"- http://nanocrafter.org/
- Use pieces of DNA to build things

"Eyewire"- https://eyewire.org/signup
- Mapping the 3D structure of neurons

"Eterna"- http://eterna.cmu.edu/web/
- Create the first large-scale library of synthetic RNA designs

TEACHER PAGES: ***KEY*** FOR "Laboratory Methodologies and Questions"

1. In "Lesson 2" you might have originally chosen a species that had less than 30 georeferenced data points available. Additionally, in "Lesson 1" you saw that the more data points that exist, the more accurate the prediction will be. It is very likely that there are more than 30 of these select species living in Florida, so why do you think there such a shortage of georeferenced data available? Most students will probably state that the organism is not commonly seen or when it is seen it is not recorded. These are both plausible answers, but more complicated responses are possible.

2. What are the scientific names of the specimens that you digitized and where were they originally found?

a. Should be a genus and species name along with a location (either city, state, or GPS coordinate)

b. Should be a genus and species name along with a location (either city, state, or GPS coordinate)

c. Should be a genus and species name along with a location (either city, state, or GPS coordinate)

3. What project(s) are you investigating any why?

Allow students the freedom to choose a project that they are interested in.

4. What safeguards do you think these programs have in order to keep the user-submitted data as accurate as possible?

Most of the data are checked by multiple users in order to crosscheck the responses.

5. As a group, come up with a novel (new) citizen science program. It can be a modification of a project that already exists, but should not be a direct copy of any that you saw today. Keep in mind, it does not need to be biological, it can be cultural as well. Outline your group's program below:

Title: Unique to the group

Problem that the program is meant to solve: The problem should be relevant to science or society.

How the problem could be solved by utilizing citizen science: The method of using citizens to record data or make discoveries must include fairly simple steps that are doable by the average citizen.