TIEE

Teaching Issues and Experiments in Ecology - Volume 9, November 2013

EXPERIMENTS

Performing a population viability analysis from data students collect on a local plant

Noah Charney[1] and Sydne Record[2]

[1]School of Natural Science, Hampshire College, 893 West Street, Amherst, MA 01002; noah@alumni.amherst.edu

[2]Harvard University, Harvard Forest, 324 North Main Street, Petersham, MA 01266; srecord@fas.harvard.edu

ABSTRACT
During two lab periods, students collect demographic data on perennial plants and then use these data in a matrix model to perform population viability analyses. During the first lab, students tag and record data on individual plants in the field. During the second lab, students compile these data to build transition matrices and then use R to run simulations. In the first year that instructors run this experiment, simulated field data of subsequent years is used to complete the classroom exercises. In subsequent years, students use real data from previous years. No prior experience with R is necessary.

KEYWORD DESCRIPTORS

- **Ecological Topic Keywords:** Conservation biology, demography, extinction, life history, plant ecology, population ecology
- **Science Methodological Skills Keywords:** Collecting data, data analysis, field observation skills, field work, graphing data, quantitative data analysis, stochasticity
- **Pedagogical Methods Keywords:** Informal group work, participation, probe, cooperative group learning, questioning

CLASS TIME

30-45 minutes in lecture to explain matrix models, 3 hours in the field to collect data, 2-3 hours in lab to analyze data.

OUTSIDE OF CLASS TIME

30 minutes to enter data into a spreadsheet.

STUDENT PRODUCTS

Data sheets and lab worksheet. Instructors may expand the assignment to include writing a full research paper.

SETTING

Instructor must identify an appropriate perennial species, and sampling should occur where it is abundant enough and when reproductive stages can be determined. Ideally there would be at least 100 plants spread out over an area where different groups of students in the class will be able to sample portions of the population. Analysis takes place in a computer lab with R open-source software.

COURSE CONTEXT

We have used this lab in 5 different upper-level classes with small lab sections (6-15 students per section): Rare Species Conservation, Conservation Biology, Plant Biology, and Terrestrial Ecology.

INSTITUTION

We have used this lab at three small liberal arts undergraduate four-year colleges. Two of these colleges were private, and one was public.

TRANSFERABILITY

This exercise could be performed by majors or non-majors in intro or upper division college courses or in pre-college environments, although the skills taught might be most relevant to college majors. This exercise is easily transferable across geographies and different plant species. Students with physical disabilities may encounter challenges in the field work portion of the lab depending on access to the chosen field site.

ACKNOWLEDGEMENTS

This exercise would not be possible without support from the open source community of contributors to R statistical software. In particular the code for this exercise depends on the ‘popbio’ package written by C.J. Stubben and B.G. Milligan and the ‘gplots’ package written by Gregory R. Warnes, Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz, and

Bill Venables. We thank Dr. Christopher Beck and two anonymous reviewers for assistance with the manuscript. We also thank Dr. Charles Ross for helping to troubleshoot problems when he used this lab in his class.

SYNOPSIS OF THE EXPERIMENT

Principal Ecological Question Addressed

How can demographic matrix models be used to inform management decisions for plant species?

What Happens

First, the students are introduced to demographic matrix models and the particular plant study system. This introduction could be in the form of a 20-45 minute lecture, or students could complete pre-lab readings and questions on their own followed by presentations at the start of lab. Second, students collect one time point of data in the field. Note that if it is not logistically possible to collect data in the field, then the provided data may be used. However, students will likely gain the most out of this activity if the field portion of the exercise is included. Third, students spend 2-3 hours in lab using computers to model the field data while working through a lab handout.

During the first year that instructors run this experiment, they provide simulated data of what students would get next year in the field (see comment on first year simulations below). In subsequent years, students use real data from previous years. Thus, the classroom exercises for the first year are essentially identical to subsequent years.

Experiment Objectives

At the end of this lab, students should:

1. Understand demographic models and life-history strategies
2. Gain familiarity with R.
3. Understand stochasticity and the challenges in making predictions

Equipment/ Logistics Required

- Instructor must identify appropriate species and field locations
- Aluminum plant tags (e.g SKU#79260 at www.forestry-suppliers.com)
- Stakes for marking plot corners. Rebar with safety caps preferred for long term.
- Optional: metal detector for finding plants in subsequent years
- Meter tapes, field flagging

- Clipboards, pens/pencils
- Small rulers
- Field forms for data entry (modified from forms we provide)
- Computers
- R Statistical Software (available for free at www.r-project.org).

Summary of What is Due

Students will be assessed on the field data collection portion of the exercise by entering their data into a .csv file that is sent to the instructor prior to the computing portion of the exercise. Students will also be assessed on their completion of the exercise handout that they work on during the computing portion of the exercise. Instructors may also ask students to complete a full research paper.

DETAILED DESCRIPTION OF THE EXPERIMENT

Introduction

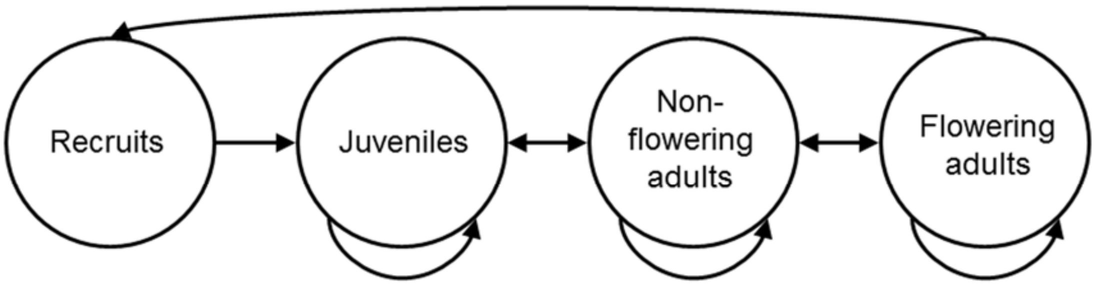
There are many instances in which an ecologist might be interested in the dynamics of a population. For example, if a non-native species arrives in an area, ecologists may be interested in whether or not populations of that organism are increasing rapidly. This knowledge of the population dynamics of the non-native organisms would help ecologists to understand whether or not the organism may be invasive in its new home and would help to inform the management of the organism.

Conservation biologists preserving populations of rare and endangered species often want to know if vulnerable populations are at risk of plummeting to extinction. In order to quantify population dynamics, ecologists have developed a number of statistical modeling tools called demographic models, also known as population viability analyses (PVA) when applied to rare and endangered species. In a PVA framework, researchers can simulate the impact of varying management techniques such as harvest quotas, head-starting programs for young, or habitat restoration. Thus, demographic modeling is an important tool for choosing appropriate conservation strategies.

Demographic models vary in complexity. Simple, unstructured models treat all individuals in a population equally, whereas more complex, structured models allow the fates of individuals of different classes to be different (Gotelli 2001). Examples of classes that can be used in a model include size or age. For many organisms, structured demographic models are more biologically realistic. For instance, a salamander may have different survival rates in its egg, juvenile, and adult forms.

This activity focuses on a widely used type of structured population model, the demographic matrix model (Leslie 1945, Caswell 2001). In a demographic matrix model, each individual in a population is categorized into a particular class. In order to understand the population dynamics of this population over time, you would need to collect data on the fates of individuals of each of these classes over several time steps (e.g., years, seasons). The length of the time steps over which you would collect data would depend on what time interval is most relevant to the life history of your study organism. For instance, annual plants that live for only one year are often censused several times within the growing season of a single year, whereas perennial plants that live for multiple years are often censused 1-2 times per year (Morris and Doak 2002).

After you have collected data for a minimum of two time points, you are ready to begin modeling the fates of individuals within the population over time. Here, let’s consider a population of Northern Pitcher Plants (*Sarracenia purpurea* L.) where individuals are categorized into four stage classes (recruits, juveniles, non-flowering adults and flowering adults) (Fig. 3, Gotelli and Ellison 2002, 2006). It is helpful when thinking about the population dynamics of a species to draw out a life cycle diagram such as Figure 3. Drawing a life cycle diagram is a first step in visualizing how you will build your model. This diagram indicates how an individual transitions between classes throughout its life and whether an individual may return to a class that it has already passed through. From this diagram, you will begin to build the transition matrix, **A**, of the demographic matrix model. Each entry of the transition matrix indicates the probability that an individual in one class at time t will become an individual of the same class or another class at time t+1. A zero value in the transition matrix indicates that it is not possible for an individual in one class at time t to transition to another class at time t+1. For instance, in Figure 3 individuals of the recruit stage class are not capable of growing to the flowering adult stage class over one time step. Allowed (non-zero) transitions within **A** are calculated by dividing the total number of individuals in a class at time t by the number of individuals that transition to a given class by time t+1. There is also often a row in transition matrices that indicates fecundity, which is the number of offspring that an individual contributes to the next generation.



Stage at year t

Stage at year t+1	Recruit	Juvenile	Non-flowering adult	Flowering adult
Recruit	0.0000	0.0000	0.0000	4.0000
Juvenile	0.1000	0.9540	0.0900	0.0000
Non-flowering adult	0.0000	0.0360	0.7010	0.8375
Flowering adult	0.0000	0.0000	0.1802	0.1610

Figure 3. (Top) Life cycle diagram for Northern Pitcher Plant (*Sarracenia purpurea*) (Gotelli and Ellison 2002, 2006). Circles indicate stage classes and arrows indicate allowed transitions between classes. (Bottom) Corresponding transition matrix where the top row indicates fecundities.

Once you have constructed the transition matrix, it is possible to project the population dynamics forward in time by multiplying a vector, N_t, containing the numbers of individuals in each class at time *t* by the transition matrix, **A**, to get the population size at time *t*+1:

$$N_{t+1} = \mathbf{A}N_t \ .$$

Using simple multiplication, you would always get the same answer if you began with the same initial population and transition matrix. This is a deterministic model. This model can be translated into a stochastic model if you interpret the entries in the transition matrix as probabilities. If you draw from probability distributions when projecting the model forward, you will get different answers each time, even though you begin with the same initial conditions. Using this method, the simulations can incorporate demographic stochasticity. Over the years, as more data are collected, it is possible to build multiple transition matrices for each year and then randomly select from those matrices as the population dynamics are projected forward in time to incorporate environmental stochasticity into the estimates of population size over time. In a population viability analysis, stochastic models are used to see what percentage of the simulated population projections fall below a threshold population size in the future. Managers can then compare the relative probability of population persistence under different scenarios. However, you should recognize that projections of the model are biased by the time period over which observations were collected. For instance, if you happen to collect data during years with extreme weather (e.g. during a 50-year flood), the model may not be representative of longer term population dynamics.

In this exercise, you will have first-hand experience in collecting data to parameterize a demographic model (a.k.a. population viability analysis) using R open source statistical software (R Development Core Team 2013) (Figure 4). Note that in this lab, you are only collecting one time point of data. Your instructor will provide the other years' worth of data. The simulations used in this lab incorporate demographic stochasticity, but not environmental stochasticity. The simulations in this lab also include a rudimentary carrying capacity. While there are more sophisticated ways to model a carrying capacity, the code for this model simply kills off any plants that exceed the carrying capacity. You could effectively remove the carrying capacity from the model by setting it to an arbitrarily high number. However, when there are more plants in the population, the simulation takes longer to run. If your population is growing rapidly and your carrying capacity is very large, the simulation could take hours to run.

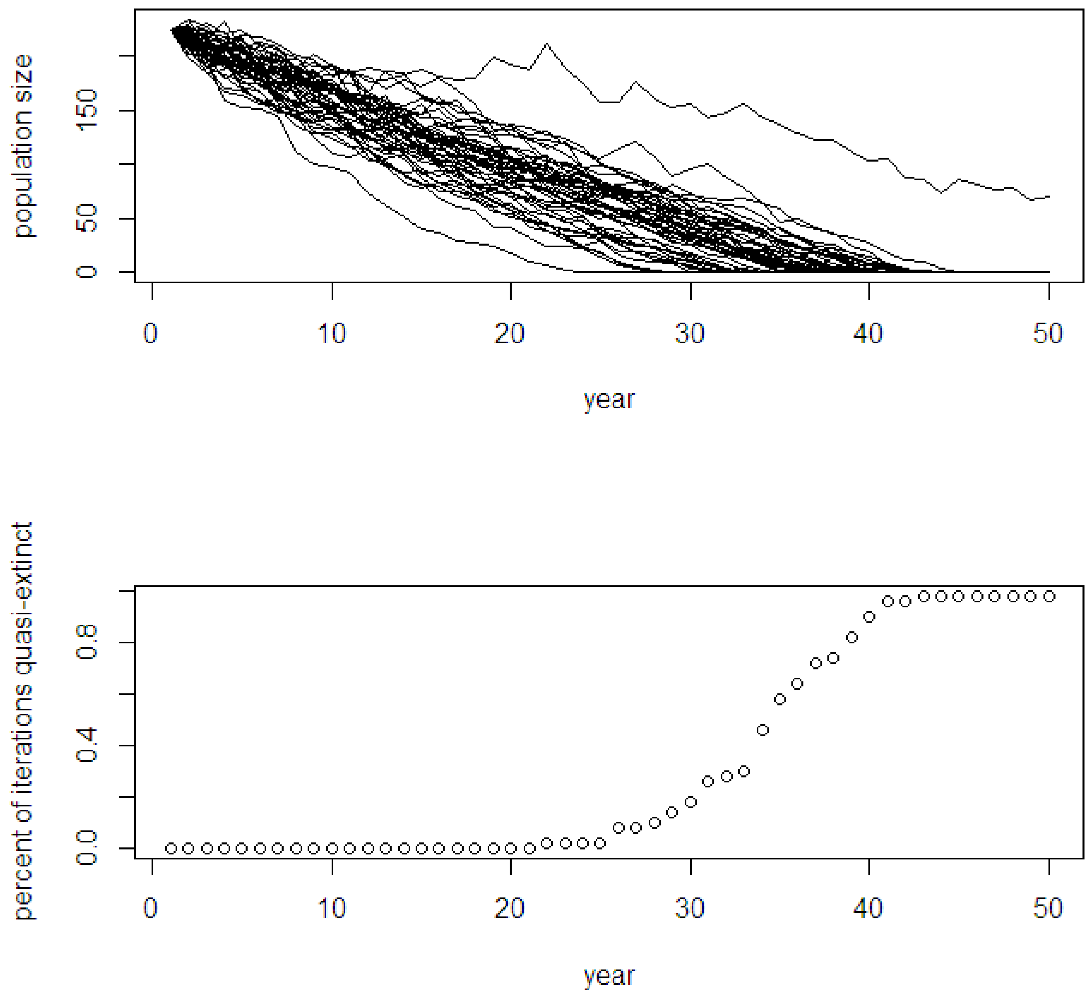


Figure 4. Screenshot of simulation output from this lab. The top graph shows simulated population trajectories, and the bottom graph shows the percentage of populations falling below the quasi-extinction threshold versus time.

Materials and Methods

Study Site(s)

The study site will depend on the species identified by the instructor. An ideal site should contain 50 to 300 individual plants spread over sufficient space for students to work, but not too far spread out such that it is difficult to locate plants. We have conducted this study using striped pipsissewa (*Chimaphila maculata*) in three 15-m by 15-m plots in a forest and using greater celendine (*Chelidonium majus*) spread along the edge of a paved bicycle path.

Overview of Data Collection and Analysis Methods

Students should answer these questions before the start of lab:

1. If we are going to carry out a PVA for our plant species, what data do we need to collect for our plant population?
2. What information do we need about each individual plant in the population?
3. How might we systematically and efficiently collect the data to make sure we measure all plants, don't measure a plant twice, can keep track of which data goes with which plant, avoid any mix-ups, and so students in other years can easily follow our system?
4. What else do we need to do if we want to continue to track this population in future years?
5. How do we go from the data that we collect to a transition matrix?

Questions for Further Thought and Discussion:

1. How do you interpret the jagged lines plotted on the graphs?
2. What is "quasi-extinction," and why do we use that instead of regular extinction?
3. Which transition probability is the population trajectory most sensitive to?
4. If you add a few plants, or remove a few plants each year (by fixing the supplemental matrix with positive or negative numbers), at which stage does adding/removing plants have the biggest impact on the populations?
5. If you were to harvest five flowering plants each year, how long would the population persist?
6. What if you were to sprinkle seed capsules into the forest, how many would you need to put out there each year to offset the harvest? Note, you'll need the germination rate – unless otherwise instructed, let's assume one plant establishes for each seed capsule put out there.
7. Having worked through matrix multiplication either on the handout or with your instructor and plotting several years of growth or decline in a population, what is the difference between deterministic and stochastic plots of population growth? Why? How does the computer create stochastic models, and how could we execute this same process without a computer ?

8. What assumptions went into this model? Are there any simplifications we made that might not match reality well? How is reality different from the model, and how would it affect the predictions?
9. Do you trust this model? How would you improve it?
10. What applications of Population Viability Analyses such as this are appropriate, and what applications are inappropriate?
11. Have there been other population viability analyses done on this plant? Describe the life history strategy of the closest plant relative to our species for which there is a published demographic model. How does the study design from the article you found compare with our study design? Can we use their data to help inform our model?

References and Links:

Caswell, H. 2001. *Matrix Population Models: Construction, Analysis and Interpretation*, 2nd Ed. Sinauer Associates, Sunderland, MA.

Conard, J. M. 2011. Demographic vital rates and population growth: an introduction to projection matrices and elasticity analysis. Teaching Issues and Experiments in Ecology, 7: Practice #3. (http://tiee.esa.org/vol/v7/issues/data_sets/conard/abstract.html)

Crouse, D. T., L. B. Crowder, and H. Caswell. 1987. A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation. Ecology 68(5): 1412-1423.

Gotelli, N.J. 2001. *A Primer of Ecology*, 3rd Ed. Sinauer Associates, Sunderland, MA.

Gotelli, N.J. and A.M. Ellison. 2002. Nitrogen deposition and extinction risk in the northern pitcher plant, *Sarracenia purpurea*. Ecology 83: 2758-2765.

Gotelli, N.J., and A.M. Ellison. 2006. Forecasting extinction risk with nonstationary matrix models. Ecological Applications 16: 51-61.

Groom, M. J., G. K. Meffe, C. R. Carroll. 2006. Principles of Conservation Biology. Sunderland, Massachusetts, USA: Sinauer Associates.

Leslie, P.H. 1945. On the use of matrices in certain population mathematics. Biometrika 33:1489-1502.

Morris, W. F., & Doak, D. F. 2002. *Quantitative conservation biology*. Sunderland, Massachusetts, USA: Sinauer Associates.

Lanza, J. 2012. Demography from physical cemeteries, "virtual cemeteries," and census data. *Teaching Issues and Experiments in Ecology* 8.

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Tools for Assessment of Student Learning Outcomes:

Rubric for student evaluation:
Pts Activity

Pre-lab
10 Answer pre-lab questions, 2 pts per question.

Week 1
20 Participation (Focus and Cooperation) in gathering field data during lab
- Presence (+5 pts)
- Focus (+5 pts)
- Working with partner and group (+5 pts)
- Answering 1-on-1 questions (+5 pts)

20 Entering data into spreadsheet and submitting via email by the assigned date
- Data entered and emailed to instructor on time (+10 pts)
- Data entered correctly - formatting, all cells filled in (+5 pts)
- Original data sheets turned in, written completely and legibly (+5 pts)

Week 2
20 Participation in working through problems, and during discussion
- Presence (+4pts)
- Focus (+4pts)
- Working with partner (+4pts)
- Answering 1-on-1 questions (+4pts)
- Talking during discussion (+4pts)

4 Worksheet (see attached) Question A.

Clear interpretation of plot (4/4 pts)
Unclear, or incorrect interpretation (2/4 pts)
No attempt (0/4 pts)

4 Question B.
Correct definition of quasi-extinction and justification for use (4/4 pts)
Correct definition, no justification or justification without definition (2/4 pts)
No attempt (0/4 pts)

8 Question C.
Correct transition probability, clear and complete explanation of process (8/8 pts)
Incorrect transition, clear and complete explanation of process (6/8 pts)
Correct transition probability, incomplete explanation of process (4/8 pts)
No attempt (0/8 pts)

4 Question D.
Correct stage identified (4/4 pts)
Incorrect stage (2/4 pts)
No attempt (0/4 pts)

4 Question E.
Correct estimate (4/4 pts)
Incorrect estimate (2/4 pts)
No attempt (0/4 pts)

4 Question F.
Correct number of seed capsules (4/4 pts)
Incorrect number (2/4 pts)
No attempt (0/4 pts)

12 Question G.
≥ 3 assumptions identified & explained, impact on model analyzed (12/12 pts)
2 assumptions identified & explained, impact on model analyzed (10/12 pts)
1 assumptions identified & explained, impact on model analyzed (8/12 pts)
3 assumptions identified, but not explained (8/12 pts)
2 assumptions identified, but not explained (6/12 pts)
1 assumption identified, but not explained (4/12 pts)
No attempt (0/12 pts)

(110) Total

NOTES TO FACULTY

Comments on Challenges to Anticipate and Solve:

Challenge #1. Making sure the models run smoothly may be the instructor's biggest challenge. From a teaching perspective, we prefer that, when projected forward, the average of all simulations is a constant population size. In matrix multiplication terms, this occurs when λ is one. We also prefer that some transition probabilities have a dramatically different impact on the population growth rate than other probabilities. Our philosophy stems from the consensus that PVA's are meant to be used more for understanding the relative contribution of different classes and management strategies, rather than for making predictions. Thus, a model that predicts rapid decline or rapid growth is less interesting than a model balanced at a steady population size. Instructors have considerable control over how the model behaves when defining stage classes and when deciding how to calculate recruitment. Instructors may choose to run labs differently, and it is the time in between the two lab periods when instructors can see how the models will behave. You will need to set aside a substantial amount of time after the students turn in their field data and before the second week's meeting – particularly during the first year (see comment on first year simulations below). You might consider leaving an extra week of space in between the two lab sessions. During this time, you will compile their data, decide on appropriate stage classes, run the models, simulate next year's data, and troubleshoot any problems that arise in an iterative process. Working this through with mock data beforehand when selecting your study species will help. If you decide on stage classes before data collection begins, it may make your task easier, but if it turns out that the stage classes you selected don't provide good teaching models, it will be harder to correct. During this time, we highly recommend testing the scripts on the actual computers that the students will be using in lab, in case there are any compatibility issues. One stumbling block may be encountered on the first line of the code by users of Macs with newer versions of R, where students have to enter the path directory exactly. For this reason, the lab will be easier on PCs.

Challenge #2. Being overwhelmed by R. Both the instructor and the students are likely to feel confused and possibly upset at first sight of R and the script. This is normal. Don't worry, it's much easier than it looks. The key is to just go methodically one line at a time from top to bottom - don't try to skip ahead. On the lines with a "#," just read what it says. On the lines without a "#," just hit "Ctrl-R," (on a PC) and the code will automatically run. Students don't actually have to do much at all.

Challenge #3. Physical challenges of field work. Many students have difficulty locating plants, keeping track of them, handling the small tags and staying physically comfortable, especially if weather conditions aren't ideal. Allow plenty of time for the field work, bring extra warm clothes if the weather seems bad, encourage teamwork and help collect the data yourself. Subdividing the study area into more small areas than there are groups allows the faster groups to do multiple sections while the slower groups concentrate on completing their single section.

One issue that arises occasionally is that some students may be absent from the field day, possibly because they have a physical limitation. For example, one of our students sprained her leg the week of the field work and was unable to hike to the field site. That year, instead of each student entering their own data into spreadsheets for homework, we assigned the task of data entry to the student with the sprained leg, which allowed her to participate and gave her familiarity with the data collection scheme.

Challenge #4. Tallying the counts (step 2.1 above in "Overview of Data Collection and Analysis Methods") can take much longer than you might expect and could result in incorrect numbers - although it can also go quickly and smoothly depending on the students. Encourage an efficient division of labor. If this task seems too tedious, you are short on time, or you have students who take longer for such tasks, consider skipping this step and providing pre-tallied data. Step E of the instructor R script will give you these data.

Comments on Introducing the Experiment to Your Students:

Ideally, you are aiming to collect long-term publishable data with this lab. It is easy to find plants with no, or few, long-term demographic studies, and so let the students know that they are conducting real science that will be published. As a class, they have a job to complete together. The science depends on the job being done well. You may also extend the lab by adding more time up front in which students think about what data to collect and design the field protocols.

Give the students a sense for the conservation importance of studying the particular species. For striped pipsissewa, the species is endangered in parts of its range. By studying the plant in the core of its range, we can contribute to conservation elsewhere. Greater celendine is considered invasive in many places, so with this plant it is a matter of understanding why plants are successful invaders and how to control them, so it is useful to study it in the core.

Make sure they have some rudimentary grasp of how stage-based matrix models work before collecting the field data. We like to do a simple matrix multiplication example on the board with the class just before going out. This way, they can think about how to categorize the plants they are handling in the field so that they fit into the stage classes in the matrix on the board. You may also have them work through the TIEE exercise provided by Conard (2011). Another option would be to apply this lab, or at least the deterministic portion of it in another software setting, such as the Populus software or Microsoft Excel.

Before the computer lab, we like to give students an appreciation for the place of R in ecology. It is free, open-source software with packages contributed by ecologists all over the globe. This particular simulation relies on packages created by several different groups of people. They could write and publish their own library for free for everyone else to use. That is why it is so powerful and adaptable. You can run simulations, analyze data, make figures, use it as a GIS platform, and all sorts of other things.

This lab complements another TIEE experiment well. Lanza (2012) provides a study of demographic data in cemeteries. The cemetery lab does not use matrix models, but instead focuses on survivorship curves. These two experiments together would make a nice unit on population growth.

Below is our suggested outline of the data collection and analysis methods process

1. First Week: Collect Field Data.
    - 1.1. Establish the bounds of the sampling area. Place rebar or flagging in the ground at the four corners of the plots.
    - 1.2. Divide the plots into subsections for student pairs to work on. Placing a grid of string or measuring tapes helps.
    - 1.3 Working in pairs, locate all individual plants within assigned areas and flag their locations.
    - 1.3. Working in pairs, mark all plants with numbered aluminum tags. Each pair should be pre-assigned a set of numbers to avoid overlap. For example, in the first year of the study, student pair one would number their plants with ID numbers between 1000 - 1099, pair two 1100 - 1199, etc... In the second year of the study, any new plants found would be numbered between 2000 and 2999.
    - 1.4. Working in pairs, fill out field sheets, recording the ID number and size of each plant. Your instructor will tell you how to measure size.

   - 1.5. Remove all temporary flagging and meter tapes, but make sure to leave permanent markers for next year.
   - 1.6. As homework, each student pair must enter their data into a template excel spreadsheet provided by the instructor, and email their data to the instructor. If there are pre-defined stage classes, students may classify plants according to stage classes before entering data.

2. Second Week: Construct Matrix Models in Computer Lab.
   - 2.1. Tally all of the transitions between stage classes provided by instructor. Enter these counts into the observed transition table on the provided worksheet. The remainder of the tasks will be done in pairs.
   - 2.2. Convert the observed counts into transition probabilities on the worksheet.
   - 2.3. Estimate recruitment rates.
   - 2.4. Follow worksheet instructions for opening R, entering data, and running models.
   - 2.5. Answer worksheet questions using both R and through discussions.

Comments on the Data Collection and Analysis Methods:

Up Front Investment. To adapt this lab, instructors must be willing to make a serious one-time investment of thought up front into choosing the right study species and becoming familiar with the R scripts we have written. In the first year that you run the experiment, you will use the script 'PVA_instructor_script-singleyear.R' to simulate a second year's set of data based on a stable-stage distribution. In subsequent years, you will use the script 'PVA_instructor_script-multiyear.R' to analyze the data. Once you get it going, you will be collecting more and more data each year, which should ultimately be publishable demographic data.

Choosing the Species. The ideal species would be a local perennial plant that lives less than 10 years, does not reproduce clonally, has a fruiting or flowering stage during the semester, has seeds for which germination/establishment can be easily quantified, has easily measured traits to distinguish stage classes, and has some conservation interest in some part of its range. However, the plants we used did not fit this description in several ways, and we used this as a teaching opportunity to explore model assumptions and limitations with the students. Instructors may consider partnering with local conservation organizations, or selecting species to monitor in areas undergoing habitat restoration to extend the service portion of this project.

Using the R Scripts. The instructions for using the R scripts are contained within the scripts. To prepare your computer you should first download and install R (www.R-project.org). Then create a folder on your computer, and place in it all the files we have provided. Open R, then open the script 'PVA_instructor_script-singleyear.R.' This will have all the instructions in it. Probably the hardest part of this script is in step C, where we have automatically translated measured plant data into stage classes. In our data, we examined histograms to develop and code the rules. You will need to inspect your own data, and code your own rules. Alternatively, you could complete this step by hand or have the students do it, although you will lose flexibility in tweaking your stages to get a model that works well for your purposes.

Making Student Handouts. You will need two handouts: data entry sheets and lab worksheet. We have provided the versions we use as Microsoft Office Documents: "pva_field_sheet_spring2013.xlsx" and "PVA_student_handout.docx." Modify them to fit your species. The provided handout is from the **third** year of our experiment. For the first and second years, you will only have one transition matrix. We have made some gross approximations when it comes to recruitment and seedling establishment in our species. Think through how to estimate recruitment for your species, with literature searches, direct measurements, or the process we used.

First Year Simulations. In the first year, you will not yet have collected two time-steps, so we have developed a work-around to allow the lab to run in the same manner. In the "PVA_instructor_script-singleyear.R" script, there is code to simulate next year's data based upon this year's. In running this lab, we like to give the students the data and have them pretend it is real data from next year. We have not in the past told students many details about the process for generating that data, because it was not relevant to our teaching goals. The process for generating next year's data works by assuming that the plants are in a stable stage distribution, and the population is at equilibrium. The code relies on an optimization function to fit a transition matrix to the data given these assumptions. As noted in the R code, this optimization process is sensitive to the range of values over which you allow the functions to search for optimal transition probabilities. If you choose to narrowly constrain the search limits of certain transition probabilities, you can control the resultant stable matrix. However, improper limits may also cause the optimization procedure to fail to converge.

Determining the Classes. Determining the number of classes for the transition matrix is one of the first things you need to do when analyzing the demographic data you collect. First you want to consider the natural history of the life cycle of you particular study species. If you want to quantify the life cycle using a continuous measure (e.g., size or age), then you will need to determine the size or age limits for each class. This can be a tricky problem. There needs to be a large enough number of classes within the life cycle that individuals within the same class are similar demographically. However, there also need to be a sufficient number of individuals within each class, so that you have reasonable confidence in the estimates of the transition probabilities. Ultimately, you have to take these two points into consideration and strike a balance between the two. To classify the stages, you may examine histograms of the plant sizes, and attempt to get a good distribution of plants in each class. We have provided the code we used to do this in the instructor script, although you may need to modify it some to fit your particular species. Alternatively, this can be done by hand or with more familiar software, such as Excel.

Comments on Questions for Further Thought:

[Note: Questions 1 – 6 and 8 also appear on our sample worksheet. They may be presented in whichever format works best for the class. Instructor may also consider having the students generate their own list of questions about the lab first as a basis for discussion, prior to running them through these questions.]

1. *How do you interpret the jagged lines plotted on the graphs?*
   There is a line in the code in which students can enter the number of iterations of the model. Point them to that line and have them change that. They should realize that as the number of iterations changes, so does the number of lines. Each line is the trajectory of one simulated population going into the future as it stochastically grows and shrinks due to demographic processes.

2. *What is "quasi-extinction," and why do we use that instead of regular extinction?*
   Students should be interpreting the quasi-extinction graph output in R. We consider quasi-extinction in conservation biology because we are concerned not only with total extinction (0 individuals), but population sizes that are lower than a certain threshold number (like 10 individuals). Connect to other parts of your course where you may have talked about demographic and genetic risks associated with small population sizes.

3. *Which transition probability is the population trajectory most sensitive to?*
   Have students play with each transition probability and see what happens. Connect this to the classic paper on sea turtle demography (Crouse et al. 1987).

4. *If you add a few plants, or remove a few plants each year (by fixing the supplemental matrix with positive or negative numbers), at which stage does adding/removing plants have the biggest impact on the populations?*
   Have students play with the "supplemental matrices" in the R code and see what happens. Connect this to management techniques like harvest quotas and headstarting.

5. *If you were to harvest five flowering plants each year, how long would the population persist?*
   Connect to wildlife management concepts of maximum sustainable yield, etc.

6. *What if you were to sprinkle seed capsules into the forest, how many would you need to put out there each year to offset the harvest? Note, you'll need the germination rate – unless otherwise instructed, let's assume one plant establishes for each seed capsule put out there.*
   Have students think through germination rates, seedling establishment, and how you would estimate these values.

7. *Having worked through matrix multiplication either on the handout or with your instructor and plotting several years of growth or decline in a population, what is the difference between deterministic and stochastic plots of population growth? Why? How does the computer create stochastic models, and how could we execute this same process without a computer?*
   This is your chance to get them to really understand the difference between deterministic and stochastic models, and to understand demographic vs environmental stochasticity. The answer here is that the computer includes demographic stochasticity (not environmental), and the physical analogy is to perform a series of weighted coin-flips. Or you could use a 20-sided die: if a plant's transition probabilities are 0.25, 0.25, and 0.5, then numbers 1-5 are transition 1, 6-10 are transition 2, and 11-20 are transition 3.

8. *What assumptions went into this model? Are there any simplifications we made that might not match reality well? How is reality different from the model, and how would it affect the predictions?*
   Think about any ways that your species deviates from the ideal species for this model. Does it reproduce clonally? Is the habitat stable? What if a tree

falls above? What about forest succession? How would climate change impact things?

9. *Do you trust this model? How would you improve it?*
Isn't this the best we can do? Think about how costly it would be to really do this right - and what if this was a critically endangered species that we didn't have time to study for decades?

10. *What applications of Population Viability Analyses such as this are appropriate, and what applications are inappropriate?*
PVAs should not be projected far into the future or used to assess the absolute size or trajectory of populations. Rather, it is preferable to use them to understand the relative importance of different life history stages. See chapter 1 of Morris and Doak (2002), which is reprinted on pages 433-435 of Groom et al (2006).

11. *Have there been other population viability analyses done on this plant? Describe the life history strategy of the closest plant relative to our species for which there is a published demographic model. How does the study design from the article you found compare with our study design? Can we use their data to help inform our model?*
Consider introducing students to Bayesian statistical approaches – e.g. informative priors.

Comments on the Assessment of Student Learning Outcomes:

We put a lot of weight into participation in this experiment, because it is about active learning. However, this can be difficult to assess. A particular challenge arises when the students are working in pairs at the computers, and one partner might be doing a lot more than the other. This could be because one student is much less familiar with computer programming, is feeling lost, or is unmotivated. A challenge to the way we have set up the R scripts is that students don't need to understand what's happening in the code in order to successfully run the models. Students that attempt to interpret the code will learn a lot more, but will take longer to complete the activity. Ask students questions about the code (see examples in formative evaluation comments below) to gauge their level of engagement.

We also put a lot of weight into the task of data entry. There is not a lot of learning involved in this step, however it is critical for you, the instructor, to get the data in a timely manner so that you can prepare for the second week of lab.

The high point value reflects the fact that the entire class depends on each student completing their portion of the task.

Comments on Formative Evaluation of this Experiment:

Pre-lab lecture: A quiz (either written, or orally as part of lecture) on students' ability to perform matrix multiplication and interpret the meanings of individual transition probabilities that you point out. Use this to determine whether you need to do more to explain matrix algebra and the concepts of matrix population models - perhaps give them a problem set to work through if they're not getting it. Also, use suggested pre-lab questions about designing the lab activities.

During Week 1 Field Work: Ask student groups individually questions about their numbering scheme and how they think the stages should be classified. Do they understand why they are putting tags on the plants?

Homework: Have the students entered the data correctly?

During Week 2 Computer Lab: Circulate through classroom and ask each pair of students questions about their progress. Have they ever used this type of software? Does it make sense what's happening from a software standpoint? How many iterations are they running? What transitions did they modify, and what did that mean? Point to a particular line on their screen and ask what it means. If many groups are missing these basic questions, consider drawing everyone's attention to the board and giving a mini lecture on the process.

In the discussion, and in the worksheet responses: Are there questions that students are consistently missing? Are students drawing connections to the other parts of your course?

Comments on Translating the Activity to Other Institutional Scales or Locations:

This experiment is best performed in groups of 20 or less. At institutions with large class sizes, this is most transferable to courses with smaller lab sections.

This study can be performed in any region where there are suitable plants accessible, see "choosing the species" in comments above. It is not as easily transferred to mobile organisms such as animals, as it does not include a mark-recapture component. The field collection must be performed during a season that the plant is observable, and is best performed during nice weather because measuring plants requires substantial amounts of time with little physical movement.

To understand this experiment, students must be able to perform matrix multiplication, and have some understanding of probability. Thus, high school students with such background should be able to complete and understand the material.

Students with physical disabilities may encounter challenges in the field work portion of the lab depending on access to the chosen field site. Instructors may use an inquiry-based approach by adding an extra week up front in which students design the field protocols. This would give disabled students the chance to participate in problem-solving with the class.

Students with other learning disabilities may have difficulty tallying the numbers or performing other pieces of the computer lab. Instructors may consider bridging some of the tedius portions of the computer lab by providing pre-tallied data, or giving students the completed transition matrices. The R scripts contain code for quickly performing these steps.

STUDENT COLLECTED DATA FROM THIS EXPERIMENT

The CSV files contain real data from students collected in three seasons on striped pipsissewa (*Chimaphila maculata*).

- multiyear_data_nostages.csv - a Comma-separated-values file containing the raw data from 3 surveys of our striped pipsissewa study as entered in by students.
- this_years_stages.csv - a Comma-separated-values file containing stages of each plant from the first survey of our striped pipsissewa study. Note that Excel might think that the file format is wrong because of the way we have coded the plant ID's in the first column. Ignore the warning in Excel.
- multiyear_data_shortform.csv - a Comma-separated-values file containing stages of each plant from three surveys of our striped pipsissewa study, organized into one column for each survey.
- multiyear_stages_longform.csv - a Comma-separated-values file containing stages of each plant from three surveys of our striped pipsissewa study, with all surveys in a single column.

The provided instructor scripts for R use these files as example data to run the simulations.

- PVA_instructor_script-singleyear.R - an R script for the instructor to set up the lab in the first year of a study.
- PVA_instructor_script-multiyear.R - an R script for the instructor to set up the lab in subsequent years.
- PVA_source.R - the R source file which the scripts above reference. This file contains all functions written specifically for this modeling exercise.
- PVA_student_script.R - the R script file which students are meant to open and use during the computer lab.

.

COPYRIGHT STATEMENT

Lastly, we request that you return your students' and your comments on this activity to the TIEE Managing Editor (tieesubmissions@esa.org) for posting at this site.

GENERIC DISCLAIMER

Adult supervision is recommended when performing this lab activity. We also recommend that common sense and proper safety precautions be followed by all participants. No responsibility is implied or taken by the contributing author, the editors of this Volume, nor anyone associated with maintaining the TIEE web site, nor by their academic employers, nor by the Ecological Society of America for anyone who sustains injuries as a result of using the materials or ideas, or performing the procedures put forth at the TIEE web site, or in any printed materials that derive therefrom.