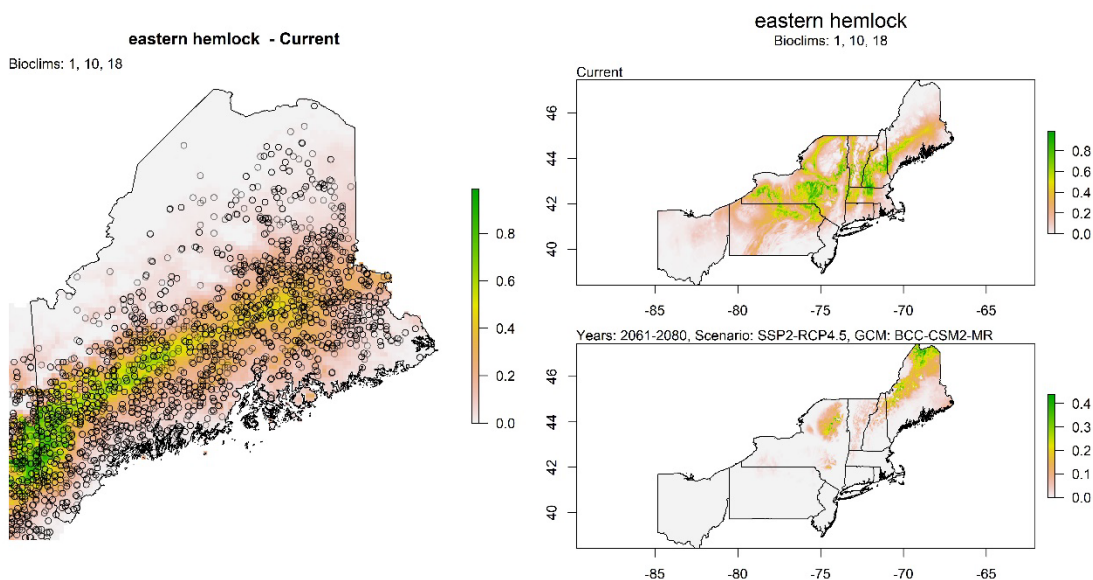


ISSUES: DATA SET

Experimenting with Species Distribution Models to Predict Tree Distributions

Noah D. Charney, Kara SantaLucia, Monica Miles, Ian Adelman, and DeCorey Bolton, Jr
Department of Wildlife, Fisheries, and Conservation Biology, University of Maine, Orono,
ME 04469

Corresponding author: Noah D. Charney (noah.charney@maine.edu)



Output of script for Eastern Hemlock, focused on current distribution in Maine, predicted habitat across the northeast in current and future time period.

THE ECOLOGICAL QUESTION:

How are tree distributions predicted to respond to climate change and how do the choice of species distribution model parameters influence these predictions?

FOUR DIMENSIONAL ECOLOGY EDUCATION (4DEE) FRAMEWORK

- **Core Ecological Concepts:**
 - Communities
 - Landscapes
 - Biosphere
- **Ecology Practices:**
 - Quantitative reasoning and computational thinking

TIEE

Teaching Issues and Experiments in Ecology - Volume 20, December 2024

- Working collaboratively
- **Human-Environment Interactions:**
 - Human impacts on the environment from local to global scales
- **Cross-cutting Themes:**
 - Space & Time

WHAT STUDENTS DO:

Use R to project future tree distributions across the Northeastern US using Forest Inventory and Analysis (FIA) and Worldclim data.

STUDENT-ACTIVE APPROACHES:

Each student or student group can choose different parameters to experiment with and different species, each exploring their own questions, then students share with the class to compare their findings.

STUDENT ASSESSMENTS:

Student presentations in a slideshow, including figures output by R that students annotate and explain.

CLASS TIME:

Approximately one 75-minute class period.

COURSE CONTEXT:

Mid-level (300s) course aimed for students with prior coursework in ecology. No prior R experience is necessary; however, it is necessary to have R and R Studio installed prior to class on whatever computers will be used for the lab. It is recommended that a standard environment such as a computer lab be used.

SOURCE:

- Forest Inventory and Analysis (FIA) data (<https://research.fs.usda.gov/programs/fia>)
- WorldClim Data (<https://www.worldclim.org/>)

ACKNOWLEDGEMENTS:

NDC originally developed the code and datasets for teaching Advanced Conservation Biology in the Department of Wildlife, Fisheries, and Conservation Biology at the University of Maine in 2021. This work was supported in part by McIntire-Stennis funding, project award no. ME0-42312, from the U.S. Department of Agriculture's National Institute of Food and Agriculture.

OVERVIEW OF THE ECOLOGICAL BACKGROUND

Species Distribution Models (SDM) are a common but fraught tool used by ecologists to understand how species respond to climate (Franklin 1995, Guisan & Zimmermann 2000, Charney et al. 2021). Here, students will use an SDM approach with the Bioclim algorithm in R to forecast how the future distributions of tree species may change as the climate changes (Booth et al. 2013, Hijmans et al. 2013). In so doing, students become familiar with the basic components of an SDM, and are asked to explore some of the choices that go into these models, including:

(1) **Shared Socioeconomic Pathway (SSP)**. Here, students may consider 4 different possible political/economic future scenarios put forward by the IPCC (<https://climatescenarios.org/primer/mitigation/>).

(2) **General Circulation Model (GCM)**. Different research groups around the globe use different biophysical models of the earth and its climate. Here, students may choose between two of the dozens of models out there to see (https://www.worldclim.org/data/cmip6/cmip6_clim2.5m.html).

(3) **Bioclimatic Variables (BioClims)**. These serve as the predictors in the SDM model, with 19 different bioclim variables to choose from, all derived from monthly temperature and precipitation data. The goal in modeling is to choose the variables that are biologically meaningful for the species you are studying, while picking as few as possible so as not to overfit the data. Students are asked to think about what makes sense given their species (<https://www.worldclim.org/data/bioclim.html>).

(4) **Tree Species**. Students can choose from any species in the Forest Inventory and Analysis (FIA) dataset, and are encouraged to consult available resources about their species biology and outside climate risk assessments (e.g. <https://forestadaptation.org/learn/resource-finder/climate-change-projections-individual-tree-species-new-england-and-northern>).

In making these predictions, students should be encouraged to not simply take their models for granted. In fact, there is a broad consensus among modelers that SDMs should not be used for the purpose of making such predictions, even though they commonly are used this way by practitioners (Charney et al. 2021). One of the easiest shortcomings for students to grasp here is that these models only capture the climate envelope in which species could survive – it does not incorporate dispersal limitations, forest succession, or other biotic interactions.

LEARNING OBJECTIVES

1. Understand the Concept of Species Distribution Modeling (SDM):
 - Explain the principles of SDM and its significance in conservation biology and climate change ecology.

- Understand and articulate the choices that go into an SDM, including:
 - i. General Circulation Models
 - ii. Future Scenarios in IPCC framework of SSPs
 - iii. Choice of model input variables, in this case the 19 bioclim variables
 - Interpret output color heat maps produced by SDMs.
 - Become aware of the limitations of this modeling approach.
2. Understand tree species distribution ecology.
 3. Familiarity with Publicly Available data Sets
 - Interpret maps of climate data sourced from the CMIP6 project and understand its relevance in modeling future climate scenarios.
 - Identify and describe the 19 bioclimatic variables used as predictors in SDM analyses.
 - Explore the structure and content of Forest Inventory and Analysis (FIA) data.
 4. Implement Species Distribution Modeling in R:
 - Understand basic operation of R, including objects, scripts, and console.
 - Navigate RStudio to access and run scripts for species distribution modeling using the provided dataset.
 5. Apply Critical Thinking and Problem-Solving Skills:
 - Critically evaluate the strengths and limitations of the SDM approach, considering factors such as model assumptions, data quality, and uncertainties.
 - Formulate hypotheses and research questions related to the effects of climate change on tree species distributions, and design appropriate SDM experiments to address them.
 6. Communicate Findings Effectively:
 - Present findings from SDM analyses in clear and coherent written and oral formats, incorporating data visualization techniques to enhance understanding.
 - Engage in discussions with peers to communicate insights, exchange ideas, and solicit feedback on SDM methodologies and interpretations.

DATA SETS

Core data was originally derived from Forest Inventory and Analysis (FIA) program (<https://research.fs.usda.gov/programs/fia>) and WorldClim Data (<https://www.worldclim.org/>) and packaged into two zipped files for download:

- [SDM.zip](#) unzips into a directory called “SDM,” which contains:
 - *WorldClim*: subdirectory containing future bioclim data used for the model, cropped to the northeast U.S.

- *Figures*: this subdirectory will appear once the model is run and will contain output figures from the script.
- *bioclim_cur_northeast.grd* and *bioclim_cur_northeast.gri*: raster stack of the current bioclim layers derived from WorldClim data.
- *FIA_northeast.csv*: a spreadsheet with FIA data on tree occurrences at plots across the northeast U.S.
- *NortheastStates.dbf*, *NortheastStates.shp*, *NortheastStates.shx*: a shapefile of the political boundaries of northeast states
- *state_codes*: an object for loading into R that contains the two-letter state codes for all U.S. states and provinces.
- *lower48*: an object for loading into R that contains the two-letter state codes for the lower 48 U.S. states.
- *SDM_source.R*: the R file containing functions and environmental variables used in this lab
- **SDM.R: the R Script file that students need to run the lab. This is the only file students need to look at or open**, everything else in this list can be safely ignored, except for inspecting the outputs in the *Figures* folder.
- (Optional): [global.zip](#) is a zipped folder for classes that wish to look at a different region other than the northeast. If desired, this folder should be unzipped and place the resulting “global” directory within the SDM folder above. Further instructions for using this global data set are within the main SDM.R script. Note that FIA data is only for the United States, so “global” might sound misleading, it really just means all available here. Inside the global directory, you will find:
 - *WorldClim*: subdirectory containing global future bioclim data used for the model.
 - *bioclim_cur_global.grd* and *bioclim_cur_global.gri*: raster stack of the current bioclim layers derived from WorldClim data.
 - *FIA_global.csv*: a spreadsheet with FIA data on tree occurrences at plots across the entire U.S.
 - *US_Canada.dbf*, *US_Canada.shp*, *US_Canada.shx*: a shapefile of the political boundaries of the U.S. and Canada.

STUDENT INSTRUCTIONS

Species Distribution Models (SDM) are a common but fraught tool used by ecologists to understand how a species responds to climate. Although an easy way to understand and map species’ relationships to current climate over large areas, there is a consensus that the typical approach for using standard SDMs to extrapolate into the future violates model assumptions and is likely to lead to erroneous predictions (Charney et al. 2021). Nonetheless, it is a commonly used technique and there is active research aimed at improving these models. Here,

TIEE

Teaching Issues and Experiments in Ecology - Volume 20, December 2024

you will choose one or more tree species to build an SDM and use this to predict future ranges of that species.

SDMs are a type of **Climate Envelope Model**, meaning they aim to capture the range of possible climate conditions in which a species could survive. Like any model, an SDM includes both **predictor** and **response** variables. Here, the response variable is whether or not a species was found at a sampling location. The predictor variable is the climate at that location. We can break today's modeling into three parts:

- (1) First, we fit the SDM using data on observed trees growing at field plots and the historical climate data at those observation locations.
- (2) Second, we use this fitted SDM and the region-wide historical climate data to map the suitability of the climate for that species at all locations across the entire region, including the spots in between our field plots.
- (3) Last, we replace the historical climate data with future climate data and re-run the SDM to make a prediction map of where the species is likely to survive in the future.

To complete this 3-step model, we must rely on several data sources and make several choices. Your task today is to investigate how these choices influence your model output. You will be working through a pre-written R Script with a numbering system that corresponds to the instructions on this sheet. This script will guide you through:

- (A) Introduction to R (lines 13 - 44)
- (B) Setting up the workspace for today's lab (lines 47- 68)
- (C) Load the data we will be using in our model, including:
 - (a) FIA Data (lines 71 - 89)
 - (b) Climate Data (lines 91 - 107)
- (D) Parameterizing your model by choosing:
 - (a) Tree species (line 117)
 - (b) Emissions Scenario (line 127)
 - (c) General Circulation Model (line 134)
 - (d) Future Year (line 137)
 - (e) Predictor Variables (line 140)
- (E) Run the SDMs (lines 144 - 168)

A. Introduction to R (lines 13 - 44).

R Statistical Software is an open-source programming language used by ecologists and other scientists for a variety of tasks. All of the code needed for this lab has been written for you in a **script** file, which you should open in R Studio. If you do not already have R Studio on your computer, you must first download and install it. Then, make sure you have downloaded and unzipped the compressed folder for today's lab.

You will then open the script file, “SDM.R.” *Note: it is critical that you run the script from the unzipped folder, otherwise you will get errors.*

When you open the script file, it will appear in a window within R Studio. In that script window, you will see notes (preceded by a #) and code. When you put your cursor on a line with code and hit “ctrl” (or “command” on a Mac) and “enter” at the same time, the code will run in the **console** window. Alternatively, you can click the “Run” button in the menu. Part A will introduce you to how R works as an object-based language.

B. Setting up the workspace for today’s lab (lines 47 - 68)

The power of R is that researchers from around the world contribute bits of code in **packages** that anyone else can download and use. This lab relies on two packages, “dismo” and “sf” (lines 56 - 57). Although the script you are running has all the code you need to see, it relies on a series of functions that are in the **source** file (line 67) and were built specifically for this classroom activity – you do not need to look at this, but you can if you want to.

Once you set the working directory (lines 59 - 62), all you need to do to run this lab through once is to hit “ctrl” and “enter” over and over again until you get to the bottom. However, you should read the notes as you go, and watch for errors in the console window. If you get any errors, usually it is because something failed to run earlier in the script, such as the working directory wasn’t set (lines 59 - 62), or the source file wasn’t loaded (line 67).

C. (C) Load the data we will be using in our model:

- a. **FIA Data** (lines 71 - 89) Today’s lab is built on field observations of trees at plots in the US Forest Service’s Forest Inventory and Analysis (FIA) program, which provides data on the distribution and condition of America’s forests in every state. The program, started about 80 years ago, monitors 355,000 permanent field inventory plots across the country. Individual plots are resampled every 5-10 years for tree species, tree size, overall tree condition, understory vegetation, soil information and more. This nationwide dataset can be used to create maps of current and past forest cover and species distribution or, as in this activity, as a starting condition for a species distribution model. The default for this lab is to look only at data from 29,663 plots in the northeastern US. For a nice intro, watch the short YouTube Video “[Forest Inventory and Analysis And You](#)” produced by the forest service, and/or visit their website: www.fs.usda.gov/research/programs/fia.

- b. **Climate Data** (lines 91 - 107). Historical and Future climate data has already been downloaded from [Worldclim](#) and is included in your folders. These data include monthly precipitation and temperature estimates averaged over 20-year time periods for each future scenario and general circulation model. From these data, biologists often derive more ecologically relevant climate metrics. Today, we will be choosing from the commonly used set of 19 bioclimatic (**bioclim**) variables, which capture annual averages, seasonality, and extreme factors. In this section of code, you can look at maps of each of these variables, and you will ultimately be tasked with choosing one or more of these for your SDM:

BIO1: Annual Mean Temperature

BIO2: Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3: Isothermality (BIO2/BIO7) ($\times 100$)

BIO4: Temperature Seasonality (standard deviation $\times 100$)

BIO5: Max Temperature of Warmest Month

BIO6: Min Temperature of Coldest Month

BIO7: Temperature Annual Range (BIO5-BIO6)

BIO8: Mean Temperature of Wettest Quarter

BIO9: Mean Temperature of Driest Quarter

BIO10: Mean Temperature of Warmest Quarter

BIO11: Mean Temperature of Coldest Quarter

BIO12: Annual Precipitation

BIO13: Precipitation of Wettest Month

BIO14: Precipitation of Driest Month

BIO15: Precipitation Seasonality (Coefficient of Variation)

BIO16: Precipitation of Wettest Quarter

BIO17: Precipitation of Driest Quarter

BIO18: Precipitation of Warmest Quarter

BIO19: Precipitation of Coldest Quarter

D. Parameterizing your model by choosing:

- a. **Tree species** (Line 117). Using their life history characteristics, researchers have rated tree species' adaptability, abundance, and habitat change capability. Species differ in their moisture requirements, cold hardiness, seed dispersal, and more, resulting in different distribution patterns. You are encouraged to research the characteristics and habitat needs of the species you choose to think about how it might respond to climate change (see [USDA's Silvics of North America](#)).

- b. **Emissions Scenario** (Line 127) This activity provides the choice between four different future emissions scenarios, selected from [scenarios developed by the IPCC](#) to capture the different possible future political, economic, and cultural choices humans make. Here, we consider the scenarios below, where your choice of 1, 2, 3, or 4 in R corresponds to:
1. **SSP 1 with RCP 2.6: "Taking the Green Road"**- A continuous global effort to increase sustainability, both in terms of consumption and development, to keep us on track with the ideal environmental boundaries proposed by the IPCC.
 2. **SSP 2 with RCP 4.5- "Middle of the Road"**- A future where emissions patterns are very similar to the past.
 3. **SSP 3 with RCP 7.0- "A Rocky Road"**- A rise in nationalism and international tensions, with many countries opting to focus only on solving their own issues rather than approaching it from a global perspective.
 4. **SSP 5 with RCP 8.5- "Taking the Highway"**- A future where the growth of capital and technological progress is taken to be the solution to sustainable development. With this rapid and strong push for economic growth, fossil fuel resources become heavily exploited.

For more, see this [Primer To Climate Scenarios](#).

- c. **General Circulation Model** (Line 134) General Circulation Models (GCM) combine mathematical representations of the atmosphere and ocean to depict the climate within a three-dimensional grid over the surface of the earth (Baede et al. 2001). Dozens of different GCMs have been developed independently by different research teams to create increasingly complex models with different choices of how and what physical and chemical processes and feedback mechanisms to include, such as water vapor, clouds, land surface, biogeochemistry, water flow, soil processes, plant growth and more (IPCC 2024). GCMs can be used to simulate the response of the climate at specific geographical points to increases in greenhouse gasses, but each model has its strengths and weaknesses based on the input parameters, spatial resolution, and assumptions.

In this activity, there are two different GCMs to choose from that incorporate your choices for bioclim variables and future emissions scenarios as additional parameters for predictions of future climate conditions in the study area. The choices include MIROC6 (choice #1), developed by a community of collaborative modelers in Japan

and IPSL-CM6A-LR (choice #2) developed by the Institut Pierre-Simon Laplace Climate Modelling Centre in France (Tatebe et al. 2019, Boucher et al. 2020). Both GCM options incorporate complex components that model the atmosphere, the land surface, and the ocean, but inputs, outputs, and assumptions differ across the models.

- d. Future Year (line 137). As time goes on, we anticipate increasingly dramatic shifts in the climate and species. Here, note that our data is broken down into 20-year increments (2021-2060, 2061-2080, 2081-2100) so do not expect to see differences between 2065 and 2066, or even 2065 and 2075, for instance.
- e. Predictor Variables (line 140). This is where you choose which bioclimatic variables to include in your model. Generally, it is important to include as few variables as necessary, so as not to **overfit** your data. Here, you should think a bit about your tree species and what bioclim variables might be most appropriate.

In general, all trees need light, water, and soil to grow and reproduce, but individual species have evolved over time with specific adaptations for certain conditions. Species differ in their average lifespan, shade tolerance, moisture requirements, cold hardiness, rate of growth, age of maturity, seed dispersal and germination, and more. These different characteristics result in different distributions for each species. For example, quaking aspen (*Populus tremuloides*) is widely distributed and abundant due to its high tolerance of a wide range of temperatures, precipitation, and soil types. In comparison, red spruce (*Picea rubens*) has a narrower range limited to the northeast. It grows best in a cool and moist climate and well-drained acidic soils. Seed germination and establishment is particularly sensitive to moisture. As the climate changes, these same life history characteristics will influence a species' ability to adapt and cope with changes or effectively migrate to more suitable locations. A standard reference to consult for tree characteristics in the US is the [USDA's Silvics of North America](#).

- E. Run the SDMs. Here, you run the SDM and inspect the output maps. The code calls the Bioclim algorithm, a classic SDM that looks at the climates across all the locations where a species was reported as occurring and uses this to predict occurrences of the species outside of the observation

TIEE

Teaching Issues and Experiments in Ecology - Volume 20, December 2024

points. As you run the script, first, R will create a map focused on the state of your choosing (line 161) where each black circle is an FIA plot where the species was actually observed and the colors represent the modeled prediction of how suitable the climate is for that species based on the historical data. By comparing the colors and points in this plot, you should be able to assess how well your SDM fits the observed data in the current time period. Next (line 162), R will generate maps of the climate suitability for your species across the whole region for the current and future time periods. The maps will show up in other windows and will be saved to the hard drive. After you have run through the code once without changing anything, now it is up to you to go back and tinker with the parameters in order to answer the questions below.

Deliverables

Make a slide on the shared Google Slides document that captures your findings in this lab. In predicting future distributions of tree species habitat, explore one or more of the following questions for a species (or set of species) of your choosing:

- How do mid-century predictions compare to end-of-century predictions?
- Does the choice of which GCM you use change the predictions?
- How do the 4 different SSPs differ in their predictions?
- Does the choice of which Bioclimes you include change the predictions, and what do you think are the right ones to include for your species?
- Do all (or many) tree species respond the same to your choice of parameters above (GCM, SSP, Bioclim, Year) or do different species respond differently to each parameterization?

For a given tree species or set of species, does the future predicted range match with expectations given what you know about biology and changing climate? Think about this both for the local areas in the state you chose and regionally across the northeast.

Discussion Question

- What do these maps actually mean?
- What are the limitations to this modeling approach?
- Given that this is just a model, what mechanisms in the real world limit the reality of these forecasts? Think in particular about the difference between fundamental and realized niche.
- What is the best use of these models for conservation?
- Depending on your species and region chosen, your inputs might not encompass the full range of a species' current distribution. For instance, if you are modeling northern red oak using only the northeast states, your model didn't see that northern red oak also currently thrives across much of the southeast. What would this mean for the model?

- Here, we primarily explored uncertainty in predictions based on how we parameterize our models, but other sources of uncertainty exist, such as observation uncertainty in our input data. What specific sources of uncertainty can you describe and how might they impact our models?

References

- Baede, A.P.M., E. Ahlonsou, Y. Ding, and D. Schimel. 2001. The Climate System: an Overview. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: New York.
- Booth, T.H., H.A. Nix, J.R. Busby, and M.F. Hutchinson, 2014. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity and Distributions* 20:1-9.
- Boucher, O., J. Servonnat, A.L. Albright, O. Aumont, Y. Balkanski, V. Bastrikov, S. Bekki, R. Bonnet, S. Bony, L. Bopp, and P. Braconnot. 2020. Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems* 12:e2019MS002010.
- Charney, N.D., S. Record, B. Gerstner, C. Merow, P. Zarnetske, and B.J. Enquist. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. *Frontiers in Ecology and Evolution* 9:689295.
- Franklin, J. 2010. Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions* 16:321-330.
- Guisan, A and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147-186.
- Hijmans, R.J., S. Phillips, J. Leathwick, and J. Elith. 2023. `_dismo: Species Distribution Modeling_`. R package version 1.3-14, <<https://CRAN.R-project.org/package=dismo>>.
- IPCC, 2023: Summary for Policymakers. In: *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 1-34, doi: 10.59327/IPCC/AR6-9789291691647.001
- PCC-TGICA, 2007: General Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment. Intergovernmental Panel on Climate Change, Task Group on Data and Scenario Support for Impact and Climate Assessment. 2:66.

- Record, S. and N.D. Charney. 2016. Modeling species ranges. *Chance* 29:31-37.
- Riahi, K. D.P. van Vuuren, E. Kriegler, J. Edmonds, B.C. O'Neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, O. Fricko, W. Lutz, A. Popp, J. Crespo Cuaresma, Samir KC, M. Leimbach, L. Jiang, T. Kram, S. Rao, J. Emmerling, K. Ebi, T. Hasegawa, P. Havlik, F. Humpenöder, L. Aleluia Da Silva, S. Smith, E. Stehfest, V. Bosetti, J. Eom, D. Gernaat, T. Masui, J. Rogelj, J. Strefler, L. Drouet, V. Krey, G. Luderer, M. Harmsen, K. Takahashi, L. Baumstark, J.C. Doelman, M. Kainuma, Z. Klimont, G. Marangoni, H. Lotze-Campen, M. Obersteiner, A. Tabeau, and M. Tavoni. 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change* 42:153–168.
- Tatebe, H., T. Ogura, T. Nitta, Y. Komuro, K. Ogochi, T. Takemura, K. Sudo, M. Sekiguchi, M. Abe, F. Saito, and M. Chikira. 2019. Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development* 12:2727-2765.

NOTES TO FACULTY

This activity is primarily intended for use in a 300-level course. Students do not need a background in R to complete this activity, but some basic knowledge may be helpful. There is no code writing involved, other than customizing the variables of the existing code; rather, students simply run through each line of the program sequentially. The scripts have been set up with the goal of minimizing the amount of actual code students need to look at, with many of the parameters set as global variables used in the source file.

We recommend that faculty read through the program and become familiar with it prior to using it in a class. This activity is best suited for a standardized computer lab with multiple desktop computers. Depending on the number of available computers and the number of students in the class, students may split up into groups of 2 or 3 and work together. It is possible to run this lab on student laptops, as we did once in a zoom-based course in 2021. However, we found that our population of students had such variable computers (hard drive space, processor speed, platform) that we preferred the more controlled computer-lab environment.

The setup of this lab is meant to have students explore different questions, and come to different insights, and we recommend the format of having students prepare slides on a shared Google Slides document for everyone to share, and possibly for class presentations or part of a discussion in a subsequent lecture class.

If you would like to do a region other than the northeast, we have provided the code and instructions to do this at the end of the script. It requires downloading the larger global zipped directory (15 GB) and unzipping into your SDM directory. You may want to create the region before class by modifying the defaults in the source script to set this region as your variable and adjust the main script accordingly. Then, you can re-zip the whole SDM package without the global directory before distributing to your students.

Time is often limited within the lab period for students to get far into learning about tree species ecology, so a good homework assignment beforehand may be to ask students to research the ecology of a few selected species, or instructors might find other ways to integrate this knowledge into the course plan.

The assignment under deliverables is left intentionally open-ended, and instructors may wish to be more specific. Students might choose a single species and answer a single question, or they might choose a set of species and answer all of the questions. We find that in our classes with a very diverse range of preparedness, some students struggle to get one question answered while others can fly through all of them. Thus, we prefer flexible targets and

TIEE

Teaching Issues and Experiments in Ecology - Volume 20, December 2024

individualized prompting of students, but depending on the student population, a more standard goalpost can be set.

For the deliverable, instructors may wish to ask students to create two slides – the first on background about the natural history of the species, variables selected, etc..., and the second on the results, or possibly one slide for each question answered.

In the discussion, we encourage consideration of the core limits of SDMs as correlative models. Even though a climate variable correlates with occurrence of a tree, the actual cause may be something totally separate – for instance, bedrock geology – which is why it is generally not a good idea to forecast with SDMs (Record and Charney 2016, Charney et al. 2021). There is a recent push to move to more mechanistic models by incorporating species interactions, dispersal, population dynamics, etc (Franklin 2010).

Students often are intimidated by R, and do not understand that to run the code the first time through all they have to do is hit ctrl-enter over and over again and it will self-execute without any need for them to think, even though they should be reading the notes and processing as they go. Students sometimes feel compelled to re-type the code into the console or do some other unnecessary steps that slow them down and confuse them. We recommend showing R to students casually on the overhead in lecture multiple times prior to asking them to use it in a lab setting.

Students often encounter errors because either:

- a. They have not run all of the code in order, such as not setting the working directory, not loading the source file, or not loading the libraries, or
- b. They are running the R Script from the zipped (compressed) folder, not from the uncompressed folder.

Students often struggle with interpreting the colors and points on the maps. Plan to spend some time walking through this part with them.

If instructors find it confusing that option #4 is SSP5, the source code could be modified so that option #5 is SSP5 and option #4 throws an error, or so that entering either 4 or 5 will both use SSP5. See lines 40-43 in the source code, these are written out there.

If you wish to include more GCMs in the model, you can download into the “Global” folder the 16 “bc” files from Worldclim associated with that GCM for each SSP and each 20-year period, using the 2.5 minute spatial resolution (https://www.worldclim.org/data/cmip6/cmip6_clim2.5m.html). Then, run the `create.region()` function again. Finally, modify line 37 of the source file (where the object `all.models` is defined) to add in your new GCM as an option.

References

- Charney, N.D., S. Record, B. Gerstner, C. Merow, P. Zarnetske, and B.J. Enquist. 2021. A test of species distribution model transferability across environmental and geographic space for 108 western North American tree species. *Frontiers in Ecology and Evolution* 9:689295.
- Franklin, J. 2010. Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions* 16:321-330.
- Record, S. and N.D. Charney. 2016. Modeling species ranges. *Chance* 29:31-37.