Project description

Introduction

Unifying relative and absolute scales. A central problem in ecology is the effect of spatial and temporal scales on the study of pattern-generating mechanisms (ecological processes)[1]. However, much of ecological theory (metapopulation theory, neutral theory, etc.) has been developed using relative measures of scale (e.g. local vs. regional, demographic parameters). While theory is developed in relative scale for greater generality across taxa, conservation planning and research studies need to operate on absolute scales (square kilometers, years), creating a need for a synthesis between relative and absolute scales in ecology.

This project aims to address three major gaps in our ecological understanding of spatial and temporal scale. I will **1)**explore the relationship between absolute and relative scale [2] across taxa. Identifying the relationship between relative and absolute scale will allow me to **2)** derive a model for unifying absolute and relative scales across taxa and **3)** identify the scales over which space-for-time substitutions can be made. This project will also demonstrate the utility and viability of fully remote, reproducible scientific collaboration using computational and ecoinformatics approaches.

Research objectives, methods, and significance

Objective 1

Explore the relationship between absolute and relative scale across taxa.

Objective 2 Derive a model for unifying absolute and relative scales across taxa.

Objective 3

Identify the scales over which space-for-time substitutions can be made.

Methods Data-driven approaches to ecological questions (ecoinformatics) require the use of computationally intensive methods, but have been an increasingly valuable tool for biodiversity/ecological research [3]. Ecoinformatics/macroecological approaches use large quantities of ecological data, often aggregated from the literature or from large survey efforts to address ecological questions using a comparative statistical approach [4, 5], allowing ecologists to address questions at broader spatial or temporal scales and for a greater range of taxa than possible with a traditional field experimental approach.

Data

While there are challenges of using synthetic and publicly available datasets [6], using a diverse array of publicly available datasets allows ecologists to leverage existing data to answer questions using a comparative approach.

Abundance dataset I will use spatially and temporally explicit community abundance data covering a wide array of taxa. There is currently a diverse array of data suitable for testing, from large, publicly available datasets including birds [7], plants[8], plankton [**???**], and microbial [9] dataset, in addition to other smaller publicly available datasets [10] and datasets that can be obtained through memoranda of understanding. Community abundance data will allow me to explore two intraspecific macroecological patterns

Species-specific data I will also use datasets that provide information on species parameters at different scales, such as dispersal distance, home range sizes, maximum longevity, generation times, and geographic range size. While there are pre-existing datasets for some of these parameters, the largest home-range dataset is >600 vertebrate species, a relative small subset of vertebrate diversity. Additional data will be collected from the literature to add to the currently available data.

Analysis While a comparative approach can be a powerful approach for addressing broad ecological questions, the data needed to address these types of questions does not lend itself well to traditional hypothetico-deductive methods [11].

The most complete species-specific data is for terrestrial vertebrates. I will begin my initial analyses with the terrestrial vertebrate groups and will generate a general model of the relationship between relative and absolute scale across taxa using species-specific data. I will then try to extend the model using additional data from other taxa, and see how far it goes before the model becomes overly complex or non-predictive.

Objective 1: Explore the relationship between absolute and relative scale across taxa. Unification of spatial and temporal scaling relationships across taxa relative to an organism and the relative position in the ecological hierarchy (organism, population, community, etc.) will provide a framework for developing more accurate predictive models of ecological pattern and processes that are spatially/temporally explicit.

Objective 2: Derive a model for unifying absolute and relative scales across taxa.
Deriving a relationship between relative and absolute scale across taxa will provide guidelines for assessing data-deficient species and provide additional information for conservation planning.

Objective 3: Identify the scales over which space-for-time substitutions can be made.
Although space-for-time substitutions have been used in ecology with varying degrees of efficacy[12], there

has been little work done on when space-for-time substitutions can be reasonably used, and over what scales, if any, the relationship between space and time is reliable.

Training objectives

Training in ecological software development and spatial and temporal analytical techniques will complement and increase my current skill set and areas of expertise. My doctoral research focused on using maximum likelihood methods with Python programming to test competing models of species abundance distributions. I have limited experience with spatial and temporal analysis or programming in R, the primary statistical programming language used by ecologists. This project would give me strong training in temporal and spatial analysis, experience in developing software in a programming language (R), and experience with a wide variety of data that will make it easier for me to work as a research scientist through developing my skills in programming, statistics, and ecoinformatics. I will receive training in these skills through textbooks, computational resources, and mentoring.

Future directions/career development

Despite challenges of chronic illness, I am able to remain an active researcher through computational tools that allow me to perform research and collaborate with other scientists remotely. Besides contributing to the scientific process through research, I have been working and will continue to work to increase accessibility and awareness for chronically ill/disabled ecologists and ecologists from other under-represented groups.

Career development

The combination of computational, ecoinformatics, and quantitative skills is a rare skill set among ecologists, and one that is particularly well suited to working remotely. Developing these skills further while engaging in remote collaborations with scientists at other institutions will provide a clear demonstration that telecommuting and remote collaboration should be a more accepted method for employment in the sciences, much as it is in the computer/tech fields. This paradigm shift supports not only scientists with disabilities/chronic illness, but also scientists who are unable to attend events in person for a variety of reasons. In addition, developing these skills also provides a way for me to continue doing science without formal institutional support if necessary.

Justification of sponsor(s) and host institution(s)

Timeline	Scientific Objectives	Broadening Participation	Communication of Results
Year 1	Phase 1: Design/compile home range database from literature, prepare and process additional datasets. Write analysis coUe/begin preliminary analysis for research objectives.	Begin development and testing of remote workshop access with Software / Data Carpentry. Develop/blog about implementation of tools/technology to facilitate remote collaboration. Serve as AccessSTEM DO-IT mentor.	Present results at scientific meetings. Submit initial publications. Code, data, presentations, preprints will be fully open (GitHub, figshare, arXiv/bioRxiv).
Year 2	Phase 2: Finalize analyses for all objectives, disseminate results through publication and presentation of results.	Remotely teach Software/Data Carpentry workshops to remote participants. Continue blogging about ways to work remotely. Continue serving as AccessSTEM DO-IT mentor.	Submit remaining publications. Archive publication versions of code via Zenodo. Communicate results at meetings.

Timothée Poisot (University of Montreal)-computational ecology

Poisot's research interests and skills make him an ideas sponsor for both the intellectual development and implementation of this project as well as an excellent mentor for developing and refining the computational skills that are an essential component of my training objectives. **Poisot** is an Assistant Professor in the Department of Biological Sciences and is a member of the Quebec Centre for Biodiversity Science. He is a strong advocate for open and reproducible science and his lab is heavily involved in creating well designed software for ecological analyses. In additional to Poisot's computational expertise that will facilitate remote mentorship, the University of Montreal has extensive remotely accessible computational resources and offers webinar training for using the high-performance computing resources.

Due to my physical limitations, I will not be able to be physically present at the host institutions; however, I will be able to interact with colleagues on-site through a remote presence via a Kubi teleconferencing robot (https://www.revolverobotics.com/), which will allow me to have control over my movement on site, leading to more natural interactions. Additionally, all the sponsoring scientists are well-versed in the computational tools and workflow that make it possible for me to effectively collaborate with others remotely.

Goals and benchmarks

Timeline

Broader impacts

Overview

Working to improve opportunities and address structural inequalities to broaden participation for under-represented groups is, unfortunately part of being a member of an under-represented group or groups. Having to self-advocate to improve structural inequalities and to obtain access as a member of an under-represented group robs researchers of valuable and limited time and energy that they could spend doing science, and many scientists belong to more than one category of under-represented group. One category of under-representation that is frequently overlooked is scientists who are facing short or long-term disability due to health conditions, either mental or physical.

A lack of familiarity with disability/chronic illness means that accessibility accommodations tend to be post-hoc, and cause the vast majority of current efforts to provide accessibility to chronically ill/disabled scientists to be inadequate or non-existent. I have been active in bringing awareness to problems associated with being a scientist with a disability/chronic illness and providing solutions to include scientists with disabilities/chronic illness through blog posts, social media, and reaching out to organizations in my field. I propose to use currently existing tools and technology to improve access for many categories of under-represented group through creating a framework for remote access tools for collaboration as well as remote options for learning to use those tools.

Remote scientific collaborations have been a way to facilitate participation by under-represented groups for a long time (for example, women corresponding with male scientists through letters, only using their first initial). Having a better framework to guide remote collaboration and participation can increase participation for scientists with disability/chronic illness, scientists from primarily undergraduate institutions, and scientists who lack the means to travel. Remote options are also available for mentoring chronically ill/disabled students who are interested in science. I will also become involved with the AccessSTEM DO-IT (Disabilities, Opportunities, Internetworking, and Technology) program at the University of Washington as a mentor for disabled K-12 students interested in pursuing a career in a STEM field.

Broadening participation through tools and technology

Open science is also a powerful tool for increasing the accessibility of science for many different underrepresented groups, not only chronically ill/disabled researchers. For example, people at primarily undergraduate institutions or who are unaffiliated with an institution often lack access to the funding and resources that larger institutions have. An open approach to science allows access and sharing of educational tools and lessons, data, and papers that would be difficult or impossible to acquire otherwise. Remote work, combined with open science approaches helps to equalize the conditions for primarily undergraduate institutions and researchers doing computational work, allowing them to remain more competitive with researchers at larger institutions, as well as create opportunities for communicating science to under-represented groups at the K-12 levels through developing lessons and outreach tools.

Computational tools and techniques becoming increasingly more important in science, and having the opportunity to learn these skills has, so far, been primarily restricted to in-person workshops, classes, or events at larger institutions. At my urging, both the Data Carpentry and Software Carpentry organizations plan to develop and provide remote accessibility to workshops, as well as making certain that disability accessibility for workshops are an essential part of workshop planning and available accommodations are included in workshop announcements. Having remote accessibility means that scientists who are unable to travel due to financial, physical, or other limitations can still participate in learning the tools that they need to collaborate and remain competitive. Remote options also allow scientists who cannot travel to participate in sharing their skills, and train the next generation of scientists.

1. Levin, S. A. "**The Problem of Pattern and Scale in Ecology: the Robert H. MacArthur Award Lecture**" *Ecology* 73, no. 6 (1992): 1943–1967.

2. Meentemeyer, V. "**Geographical Perspectives of Space, Time, and Scale**" *Landscape ecology* 3, no. 3-4 (1989): 163–173.

3. Michener, W. K. and Jones, M. B. "**Ecoinformatics: Supporting Ecology as a Data-Intensive Science**" *Trends in ecology & evolution* 27, no. 2 (2012): 85–93.

4. Brown, J. H. "**Macroecology**" (1995):

5. Kelling, S., Hochachka, W. M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., and Hooker, G. "**Data-Intensive Science: a New Paradigm for Biodiversity Studies**" *BioScience* 59, no. 7 (2009): 613–620.

6. Reichman, O., Jones, M. B., and Schildhauer, M. P. "**Challenges and Opportunities of Open Data in Ecology**" *Science* 331, no. 6018 (2011):

7. Pardieck, K. L., Ziolkowski Jr, D. J., and Hudson, M.-A. "**North American Breeding Bird Survey Dataset 1966 - 2013, Version 2013.0**" (2014): Available at http://www.pwrc.usgs.gov/BBS/RawData/

8. USDA Forest Service. "**Forest Inventory and Analysis National Core Field Guide (Phase 2 and 3). Version 4.0.**" (2010):

9. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., and others. "**The Metagenomics RAST Server–a Public Resource for the Automatic Phylogenetic and Functional Analysis of Metagenomes**" *BMC bioinformatics* 9, no. 1 (2008): 386.

10. Baldridge, E. "**Community Abundance Data**" (2013): Available at http://dx.doi.org/10.6084/m9.figshare.769251

11. Blackburn, T. M. "**Method in Macroecology**" *Basic and Applied Ecology* 5, no. 5 (2004): 401–412.

12. Pickett, S. T. "**Space-for-Time Substitution as an Alternative to Long-Term Studies**" *Long-term studies in ecology* (1989): 110–135.