

DISSERTATION

ASSESSING PLANT DIVERSITY TO ENABLE
CONTINENTAL-SCALE MONITORING AND FORECASTING

Submitted by

David T. Barnett

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2017

Doctoral Committee:

Advisor: Thomas Stohlgren

Paul Evangelista
Patrick Martin
Jeffery Morisette

Copyright by David T. Barnett 2017

All Rights Reserved

ABSTRACT

ASSESSING PLANT DIVERSITY TO ENABLE CONTINENTAL-SCALE MONITORING AND FORECASTING

The Earth System is dynamic. It influences and is influenced by physical, chemical, and geological processes, but it may be the least understood of these systems. The biosphere interacts with the physical Earth System on diurnal and seasonal scales, and over decades and centuries. The living system interacts with itself and other systems at a variety of scales. At large, continental scales, exchange between biotic elements and the atmosphere and surface water control climate, hydrology, and productivity. At small scales plants interact with each other and exchange energy and matter with the atmosphere and soil. Understanding the Earth System requires comparable methods and analysis across scales and over decades. This is particularly true given that the Earth System is increasingly facing changes in climate and disturbances, the redistribution of species, and land-use change.

The National Ecological Observatory Network (NEON) is a platform designed to enable an understanding of the causes and consequences of change on ecology. By simultaneously measuring the drivers of change and ecological responses – organisms, atmosphere, and soil – it will enable the ecological community to better understand the nature of interactions and support forecasts of future states. This work describes questions, analysis, and testing for the development of the plant diversity observations to be made by NEON.

Models and forecasts require information from each of the sites that comprise NEON. The study design that directs spatial distribution of plots for sampling diversity relies on a random design that is stratified by land cover with replication intended to detect differences in trends between sites over thirty years. A classic power analysis that relied on prototype data and satellite imagery to parameterize

temporal and spatial variability indicated that a sample size of 30 plots per site would sufficiently differentiate trends across sites. Results from multiple sites collecting data according to the design demonstrated that patterns of spatial variation were higher than expected and that a larger sample size would be required to satisfy the specified test.

Plant diversity data collected according to the design also must be comparable within and across sites. Variations in level of effort challenge the statistical comparison of plant species richness data. Comparing richness where the coverage - as defined by slope of the species accumulation curve - provides a statistically rigorous and biologically meaningful point of comparison. To sample such that species accumulation curves terminated at a slope of seven, plots were allocated proportional to the square-root of the strata area within each site. When comparing plant species richness data collected according to the proposed allocation from six it was found that only 30% of the within-site species accumulation curves terminated at a slope of seven, and only 33% of the species accumulation curves at the scale of the site terminated at a slope of seven.

Ensuring the creation of a design that generates data capable of describing extant status and future states will require iteration and continued evaluation. A method for ensuring plots are located such that change will be detected was evaluated by generating species distribution models of two invasive plant species, *Pennisetum clandestinum* and *Holcus lanatus* as predicted by topography and extant and future climate data. The models suggested that suitable habitat for *Pennisetum clandestinum* may decrease in extent while suitable habitat for *Holcus lanatus* may expand at the site over time. To adequately document and improve understanding of the causes and consequence of habitat expansion, additional sampling plots could be placed in areas vulnerable to by *Holcus lanatus* in the future. Similarly, any resources available for the control of plant species invasion may be better expended on *Holcus lanatus*. This is one example of the many uses of NEON data to assist land managers.

ACKNOWLEDGEMENTS

I would like to thank all of my family who supported me and made significant life changes to accommodate these efforts. I'm grateful for the many, many friends who offer support and humanity. I would like to thank Tom Stohlgren for being like a second father and a mentor. I am extremely grateful to the other members of my committee, Paul Evangelista (and also for being a true friend for years), Patrick Martin, and Jeff Morisette for their support and guidance. Guidance and insight was also provided by many at NEON and those who have supported NEON, including Dave Schimel, Paul A. Duffy, Rachel Krauss, Elena Azuaje, Kathi Irvine, Frank Davis, Alan Gelfand, Andrea Thorpe, David Gudex-Cross, Michael Patterson, Jalynda McKay, Joel McCorkel, Courtney L. Meier, Peter Adler, Jim Clark, Bob Peet, Brian Enquist, Debra Peters, James Grace, Mark Vellend, Susan Harrison, and Ben Chemel. Finally I'm grateful to the many members of the ecological community who have provided insight and support to the NEON project over many years.

I also would like to thank teachers: David Mitchell, Ann Pratt, Lucile Clemm, Chris Balch, Jerry Wooding, Jim Serach, Arthur Karp, Jim Enderson, Tas Kelso, Jim Ebersole, Tom Wolf, Peter Blasenheim, Dan Binkley, John Weins, Tom Hobbs, Dave Schimel, Paul Duffy, and Jim Clark.

TABLE OF CONTENTS

ABSTRACT..... ii

ACKNOWLEDGEMENTS..... iv

TOOLS AND INSIGHTS FOR UNDERSTANDING LONG-TERM PATTERNS OF PLANT DIVERSITY 1

 Introduction 1

 Quantifying Patterns to Understand Change..... 3

 A sampling system to detect trends 4

 Comparability to facilitate understanding 5

 Integrating data and models to sample change 6

 Conclusion..... 7

REFERENCES..... 9

The terrestrial organism and biogeochemistry spatial sampling design for..... 13

the National Ecological Observatory Network..... 13

 Introduction 13

 Design Criteria..... 15

 Sampling Design for the Terrestrial Observation System 19

 Sampling frame 21

 Randomization 22

 Stratification..... 27

 Minimum sample size 32

 Sample allocation..... 41

 Data analysis with variance estimators 43

 Testing the study design: plant diversity 45

 Methods..... 45

 Results..... 46

 Discussion..... 47

 Iterating and optimizing the study design 48

 Conclusion..... 49

REFERENCES..... 51

Strategies for comparing plant diversity in a national network of sites..... 60

 Introduction 60

 Methods..... 65

 Study sites 65

Field collection of data.....	66
Models and Analyses	69
Results.....	71
Tests of species accumulation curve slope	71
Comparisons of species richness	75
Discussion.....	78
Testing assumptions of the design.....	79
Comparisons across sites and strata.....	83
Applications.....	83
Future directions.....	84
References	86
Tools Planning for climate change when designing invasive plant species studies ³	95
Introduction	95
Local questions and baseline data relevant to invasive plant species management	96
An iterative framework for evaluating spatial and temporal hypotheses.....	98
Hypothesis 1: Species distributions and potential habitat suitability change predictably in space .	100
Hypothesis 2: Species distributions and potential habitat suitability change predictably in time...	102
Hypothesis 3: Spatial and temporal trends in invasion are best measured with a sampling design that captures biotic and abiotic gradients	105
Caveats.....	106
Improvements.....	108
References	110
Conclusions about the NEON sample design for Plant Diversity	119
Introduction	119
The terrestrial organism and biogeochemistry spatial sampling design for the National Ecological Observatory Network	119
Strategies for comparing plant diversity in a national network of sites.....	121
Planning for climate change when designing invasive plant species studies.....	122
Notes and recommendations for the next 30 years of NEON monitoring	123
A tension between designs	123
Additional considerations	124
References	127
Appendix 1. Sample Size calculations	129
Appendix 2: Sample Mean and Sample Variance with Design-based Estimators	134

TOOLS AND INSIGHTS FOR UNDERSTANDING LONG-TERM PATTERNS OF PLANT DIVERSITY

Introduction

The challenges of a changing world require investigations across space-time scales to trace the origination of causes and the consequences for ecological systems (Schimel and Keller 2015). That understanding requires the implicit study of cross-scale interactions (Peters et al. 2008), a focus on mechanistic studies that cross a variety of extents (Heffernan et al. 2014), synthesizing existing data, linking spatially distributed observation and sensor-based networks and experiments, and creating and maintaining “big data” programs that span continents (Soranno and Schimel 2014). The National Ecological Observatory Network (NEON), designed to facilitate a community-driven understanding of the causes and consequences of ecological change, is one such big data platform (Keller et al. 2008). This body of work describes the science-based approach – the theoretical development and testing – applied to design the plant diversity component of the NEON program.

Coordinated, long-term observations of plant diversity across the continental United States will provide insight to links between pattern and process at multiple spatial and temporal scales, and facilitate forecasting of patterns and ecosystem function into the future (Keller et al. 2008). Quantifying patterns with methods comparable at multiple spatial scales, across regions and continents, and through time allows an assessment of how plant diversity responds to a diversity of conditions and drivers of change. By targeting a diversity of environmental conditions, observations are likely to capture a greater diversity of species and species-environment assemblages. The iterative integration of these data with models will allow validation of space-time predictions and efficiently direct subsequent observations to areas vulnerable to change (Stohlgren 2007). This process is capable of informing an evolving understanding of how species-environment relationships respond to drivers of change and impact the functioning and dynamics of ecosystems (Schimel et al. 2011).

Observations of plant diversity have played a central role in the development of the theory and practice of ecology. Charles Darwin documented the distribution of plant species assemblages in his backyard (Magurran and McGill 2010) prior to defining theories that described species interactions and species-environment relationships during his subsequent global exploration (Darwin 1859). The science has evolved, furthering the understanding of the processes that govern the interactions of species and species-environment relationships (Vellend 2010). Investigation of these processes and resulting patterns drive contemporary ecology. Understanding species distribution, fecundity, and persistence dominate population ecology (Clark et al. 2004). Community ecology focuses on the interactions of two or more species and the resulting impact on species composition in time and space. Other approaches to studying plant diversity focus on the importance of regional species pools, and the relationship between environmental factors and the distribution, occurrence, and abundance of species (Stohlgren 2007).

Plant species comprise much of the structure of ecosystems and are an important strata for processes such as the cycling of water, carbon, nitrogen, and phosphorous (Hooper and Vitousek 1998, Diaz et al. 2003). Common and unique species (e.g., nitrogen fixers, invasive species) dominate ecosystem function. The traits – phenotypic characteristics that influence species performance and/or ecosystem function (Grime 1973, Weiher 1999) - associated with these species, such as leaf nitrogen content and canopy height, contribute to the functioning of ecosystems by controlling photosynthesis, respiration and other processes. The contribution of subdominant species in a system was thought to be minimal until field-based experiments and observations recognized that systems simultaneously carry out multiple functions (Hooper et al. 2005, Cardinale et al. 2011). Evidence of the importance of species richness to functional diversity and ecosystem multifunctionality has increased with coordinated investigation across continental scales (Maestre et al. 2012).

Plant diversity is sensitive to change; changes in climate (Ibanez et al. 2006, Magurran and Dornelas 2010), species invasion (Vitousek 1996, Mack et al. 2000) land use change, and disturbance (Dornelas 2010). Paleo records demonstrate the influence of shifting climate on species distributions (Wagner and Lyons 2010). Since natural selection is influenced by natural and anthropogenic-induced climate change, species not suited to emerging conditions will be forced to adapt or track change through a combination of dispersal and adaptation to novel conditions and interactions (Clark et al. 2012). Even without directional changes in climate, plant species composition will change as species migrate and adapt, alter resource availability, interact with other species (e.g., herbivores, soil biota), and respond to disturbance (Stohlgren 2007). Land use may drive the most pronounced changes. Disturbance to the structure of soil and species, changing disturbance regimes, and inputs to systems have direct and indirect impacts on plant diversity (Pickett and White 1985, Pickett et al. 1989). Collectively, many factors influence the direction and magnitude of changes in plant diversity including changes in genetic diversity, species composition and abundance, and distribution and interactions of other species in a complex environment.

Quantifying Patterns to Understand Change

What design considerations assure observations of plant diversity will describe long-term trends? How can plant diversity sampling adequately describe local landscapes while enabling comparison across sites? How can monitoring data be efficiently describe change and guide management? These questions will be explored through a series of related papers that: (1) defines a sampling strategy that directs the collection of data capable of detecting space-time trends and is suitable for integrating resulting observations of plant diversity and other taxonomic groups and soil with drivers of change, (2) describes and tests a framework that accounts for site-scale differences in plant species richness and observation effort with species-accumulation curves to make observations comparable across sites, and (3) presents tools for leveraging observations of plant species-environment

relationships iteratively position monitoring to better detect change and contribute long-term management strategies.

A sampling system to detect trends

A continental observatory for monitoring plant diversity and a variety of other terrestrial organisms and soil requires a system for directing the distribution and intensity of sampling within sites such that resulting data is capable of detecting and comparing trends across space and through time. The paper “The terrestrial organism and biogeochemistry spatial sampling design for the National Ecological Observatory Network” for a special edition of *Ecosphere* describes the sample design that is the foundation of the data collection effort. The primary goal was a design that supported the NEON mission (Schimel et al. 2011); it was, and will remain, necessary to ensure that decades of funding will result in insightful information. The design must enable the detection and comparison of trends and the integration of plant diversity data a variety of other data streams: soil, organisms, climate, atmosphere, and remote sensing imagery.

Sample design typically requires a specific question and analyses. This requirement presented a challenge as high-level NEON questions are broad and the ecological community will work with NEON data to answer numerous questions with a diversity of analytical approaches. Several design factors were incorporated to address these unknowns. Samples were distributed randomly within sites to both ensure unbiased characterizations, but to also provide data suitable to a variety of analyses. Samples were stratified to increase efficiency and to focus observations landscape characteristics characterized by other NEON data collection platforms such as the tower-based sensors that collect describe many of the factors likely to drive, and be influenced by, changes in plant diversity (e.g., temperature, precipitation, net ecosystem exchange). This guiding principle of observing plant diversity with the same design and methods at sites subjected to divergent trends in these forcing factors resulted in a question capable of parameterizing a power analysis for sample sizes. Is there a difference in temporal trends in

plant diversity between two sites? A model appropriate to this question provided a framework for the analysis that was parameterized with existing data. Early NEON data collected according to the design provide the chance to assess capacity of the design to detect and differentiate trends and point towards opportunities for design iteration and optimization.

Comparability to facilitate understanding

A comprehensive and general understanding of how plant species diversity is changing in response to a variety of forcing factors requires comparable observations. A paper titled “Strategies for comparing plant diversity in a national network of sites” for *Ecological Applications*, develops a framework for describing how prescribed sample sizes might be distributed within sites or optimized after initial collection by comparing plant diversity across large spatial scales. Drawing comparisons between status and trends in plant diversity, coupled with ancillary data capable of describing the drivers of these changes, may facilitate the comprehensive understanding of large-scale trends and the factors that govern patterns at local and continental scales. However, comparability is challenged by disparities in sampling effort, the abundance of species at local scales, and the density of individuals observed.

The information returned from observations of plant diversity might become comparable by standardizing effort with respect to plant diversity. Species accumulation curves describe the rate at which new species are added to a sampling effort (Gotelli and Colwell 2001). In the context of plot-based sampling, each plot captures a list of species. The sample-based species accumulation curve describes the rate at which unique species are added to the total pool of observed species with successive plot sampling (Barnett and Stohlgren 2003, Gotelli and Colwell 2011). Curves start steep, when few plots are included in the random sample, the probability that successive plots add new species to the total number recorded is high. The slope of the curve typically becomes less steep as continued plot sampling captures fewer new species. The inflection point of the species accumulation

curve – that point where many of the species captured in each plot were previously documented – may provide an attainable target for standardization. The number of species (y-axis) and the number of plots (x-axis) required to reach this point is descriptive of the landscape and should provide a diversity-based standard for comparison that can be achieved across sites. Plot-based plant diversity data collected from six NEON sites tested and evaluated comparability and differences across the continental U.S.

Integrating data and models to sample change

Random sample designs that guide plot-based plant diversity sampling efforts have the potential to miss ecological dynamics that are essential to the NEON goal of understanding changing ecological pattern and process. Rare plant species are unlikely to be detected and new invasive species can be missed when random sampling locations miss areas vulnerable to invasion such as disturbed areas or riparian corridors (Barnett et al. 2007). A paper titled “Planning for climate change when designing invasive plant species studies” for *Bioscience* will examine how integrating initial data collections with climate data to generate forecasts of potential change may become an essential tool for iterating sample designs. Invasive species – as both drivers and result of change - are central to the NEON mission and can constitute a significant component of plant diversity. Estimates of species distributions in space describe areas vulnerable to invasion and the natural resources that might be threatened. Independent variables relevant to forecasting models such as measures of landscape, land use, and climate, can provide important insight into the drivers of invasion. As these explanatory factors change through time, landscape patterns of invasion will also change. Incorporating estimates of future climatic and land use condition allows models to describe future patterns of potential plant species invasions. These estimates of future condition facilitate proactive sampling strategy such that plots that measure incidence of invasion and enable an understanding of impacts to native flora can be placed in parts of a site vulnerable to invasion based on model results.

Plot-based and species mapping data that informed NEON methods were collected at the Hakalau Forest National Wildlife Refuge in Hawaii that is immediately adjacent to a site initially targeted for inclusion in the NEON collection of sites. The resulting data focused on modeling distributions of two invasive plant species and the creation of an approach that could be incorporated into the design iteration component the NEON study design; a tool to augment the design to ensure change is detected. In addition to offering insight into changing species distributions and potential new plot locations, these data and models should also have direct implications for management strategies. Those species most likely to undergo rapid expansion can be aggressively controlled and efficient monitoring systems can be developed that evaluate control, and iteratively improve models.

Conclusion

This collection of papers will describe insights and techniques for understanding plant diversity at local to continental scales. The goal is to provide a platform capable of quantifying patterns and change. The design described herein directs the collection of the data from sites that will contribute to NEON's continental scaling objectives and inform the space-time models to forecast these changes (Figure 1.1). The data from these approaches are capable of integration with other streams of information to inform the causes and consequence of change; they are tools needed for responding to and managing change. The approaches and results will not represent perfection, and they will undoubtedly evolve over time. There are caveats associated with the modeling techniques and assumptions of the investigation - the number needed reach the asymptote of the species-accumulation curve for example. Each approach is designed to be iterative, almost like a hypothesis statement needing to be refined with the targeted and ongoing collection of more information. By improving techniques and adaptively sampling to capture the change, the goal is to improve the ability science and management to efficiently quantify and understand dynamic patterns of plant diversity and the ramifications for ecosystems.

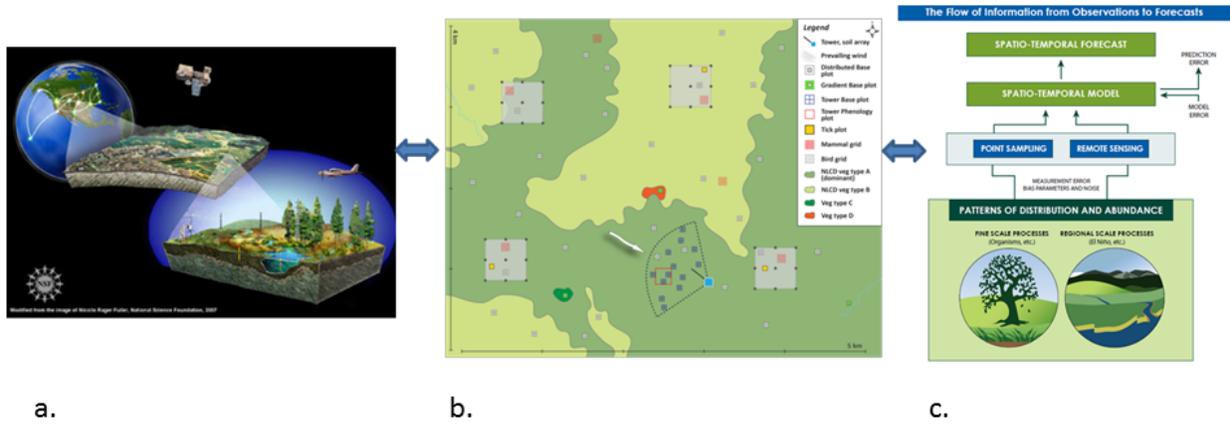


Figure 1.1. The foundation of the NEON design and the effort come from data collected at individual sites. These plant diversity data are collected according to a statistically robust sample design that produces data comparable within and across sites, and is easily adjusted and optimized (a). These data will support the NEON effort to scale patterns and understanding to regions and the continent (b), and provide the point-sampling data needed for space-time models that will forecast future ecological states (c).

REFERENCES

- Barnett, D. T. and T. J. Stohlgren. 2003. A nested-intensity design for surveying plant diversity. *Biodiversity and Conservation* 12:255-278.
- Barnett, D. T., T. J. Stohlgren, C. S. Jarnevich, G. W. Chong, J. A. Ericson, T. R. Davern, and S. E. Simonson. 2007. The art and science of weed mapping. *Environmental Monitoring and Assessment* 132:235-525.
- Cardinale, B. J., K. L. Matulich, D. U. Hooper, J. E. Byrnes, E. Duffy, L. Gamfeldt, P. Balvanera, M. I. O'Connor, and A. Gonzalez. 2011. The functional role of producer diversity in ecosystems. *American Journal of Botany* 98:572-592.
- Clark, J. S., D. M. Bell, M. Kwit, A. Stine, B. Vierra, and K. Zhu. 2012. Individual-scale inference to anticipate climate-change vulnerability of biodiversity. *Philosophical Transactions of the Royal Society B-Biological Sciences* 367:236-246.
- Clark, J. S., S. LaDeau, and I. Ibanez. 2004. Fecundity of trees and the colonization-competition hypothesis. *Ecological Monographs* 74:415-442.
- Darwin, C. 1859. *The Origin of Species*. Reprinted by Penguin Books, London.
- Diaz, S., A. J. Symstad, F. S. Chapin, D. A. Wardle, and L. F. Huenneke. 2003. Functional diversity revealed by removal experiments. *Trends in Ecology & Evolution* 18:140-146.
- Dornelas, M. 2010. Disturbance and change in biodiversity. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365:3719-3727.
- Gotelli, N. J. and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4:379-391.

- Gotelli, Nicholas J.; Colwell, R.K., 2011. Estimating species richness. In B. J. Magurran, Anne E.; McGill, ed. *Biological diversity frontiers in measurement and assessment*. Oxford University Press, pp. 39–54.
- Grime, J. P. 1973. Control of species density in herbaceous vegetation. *Journal of Environmental Management* 1:151-167.
- Heffernan, J. B., P. A. Soranno, M. J. Angilletta Jr., L. B. Buckley, D. S. Gruner, T. H. Keitt, J. R. Kellner, J. S. Kominoski, A. V. Rocha, J. Xiao, T. K. Harms, S. J. Goring, L. E. Koenig, W. H. McDowell, H. Powell, A. D. Richardson, C. A. Stow, R. Vargas, and K. C. Weathers. 2014. *Marosystems ecology: understanding ecological patterns and processes at continental scales*. *Frontiers in Ecology and Environment* 12(1):5-14.
- Hooper, D. U., F. S. Chapin, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, S. Naeem, B. Schmid, H. Setälä, A. J. Symstad, J. Vandermeer, and D. A. Wardle. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* 75:3-35.
- Hooper, D. U. and P. M. Vitousek. 1998. Effects of plant composition and diversity on nutrient cycling. *Ecological Monographs* 68:121-149.
- Ibanez, I., J. S. Clark, M. C. Dietze, K. Feeley, M. Hersh, S. LaDeau, A. McBride, N. E. Welch, and M. S. Wolosin. 2006. Predicting biodiversity change: Outside the climate envelope, beyond the species-area curve. *Ecology* 87:1896-1906.
- Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and Environment*. DOI: 10.1890/1540-9295(2008)6[282:ACSFTN]2.0.CO;2
- Mack, R. N., D. Simberloff, M. Lonsdale, H. Evans, M. Clout, and F. Bazzaz. 2000. Biotic invasions: causes, epidemiology, global consequences and control. *Issues in Ecology* 5:1-20.

- Maestre, F. T., J. L. Quero, N. J. Gotelli, A. Escudero, V. Ochoa, M. Delgado-Baquerizo, M. Garcia-Gomez, M. A. Bowker, S. Soliveres, C. Escolar, P. Garcia-Palacios, M. Berdugo, E. Valencia, B. Gozalo, A. Gallardo, L. Aguilera, T. Arredondo, J. Blones, B. Boeken, D. Bran, A. A. Conceicao, O. Cabrera, M. Chaieb, M. Derak, D. J. Eldridge, C. I. Espinosa, A. Florentino, J. Gaitan, M. G. Gatica, W. Ghiloufi, S. Gomez-Gonzalez, J. R. Gutierrez, R. M. Hernandez, X. W. Huang, E. Huber-Sannwald, M. Jankju, M. Miriti, J. Monerris, R. L. Mau, E. Morici, K. Naseri, A. Ospina, V. Polo, A. Prina, E. Pucheta, D. A. Ramirez-Collantes, R. Romao, M. Tighe, C. Torres-Diaz, J. Val, J. P. Veiga, D. L. Wang, and E. Zaady. 2012. Plant Species Richness and Ecosystem Multifunctionality in Global Drylands. *Science* 335:214-218.
- Magurran, A. E. and M. Dornelas. 2010. Biological diversity in a changing world. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365:3593-3597.
- Magurran, A. E. and B. McGill. 2010. Challenges and opportunities in the measurement and assessment of biological diversity. *in* A. E. Magurran and B. McGill, editors. *Biological diversity: Frontiers in measurement and assessment*. Oxford University Press, New York.
- Peters, D. P. C., P. M. Groffman, K. J. Nadelhoffer, N. B. Grimm, S. L. Collins, W. K. Michener, and M. A. Huston. 2008. Living in an increasingly connected world: a framework for continental-scale environmental science. *Frontiers in Ecology and Environment* 6(5):229-237.
- Pickett, S. T. A., J. Kolasa, J. J. Armesto, and S. L. Collins. 1989. The ecological concept of disturbance and its expression at various hierarchical levels. *Oikos* 54:129-136.
- Pickett, S. T. A. and P. S. White, editors. 1985. *The Ecology of Natural Disturbance and Patch Dynamics*. Academic Press, New York, NY.
- Schimel, D. S. and M. Keller. 2015. Big questions, big science: meeting the challenges of global ecology. *Oecologia* 177:925-934.

- Soranno, P. A. and D. S. Schimel. 2014. Macrosystems ecology: big data, big ecology. *Frontiers in Ecology and the Environment* 12(1):1-3.
- Stohlgren, T. 2007. *Measuring plant diversity, lessons from the field*. Oxford University Press, New York.
- Vellend, M. 2010. Conceptual synthesis in community ecology. *Quarterly Review of Biology* 85:183-206.
- Vitousek, P. M., C. M. D'Antonio, L.L. Loope, and R. Westbrooks. 1996. Biological invasions as global environmental change. *American Scientist* 84:468-478.
- Wagner, P. J. and S. K. Lyons. 2010. Estimating extinction with the fossil record. *in* A. E. Magurran and B. McGill, editors. *Biological diversity: Frontiers in measurement and assessment*. Oxford University Press, New York.
- Weiher, E. 1999. The combined effects of scale and productivity on species richness. *Journal of Ecology* 87:1005-1011.

THE TERRESTRIAL ORGANISM AND BIOGEOCHEMISTRY SPATIAL SAMPLING DESIGN FOR
THE NATIONAL ECOLOGICAL OBSERVATORY NETWORK¹

Introduction

The National Ecological Observatory Network (NEON) is designed to improve understanding and forecasting of ecological change at continental scales over decades (Schimel et al. 2011). Insight into ecological cause and effect will result from integrating systematic observations of the drivers of change and ecological response at as many as 47 terrestrial sites throughout the continental United States and Alaska, Hawaii, and Puerto Rico for thirty years (Vitousek 1997, Keller et al. 2008, Luo et al. 2011). Sites encompass wildlands and cross a variety of gradients (e.g., land-use, species invasion, nitrogen deposition) to address regional and continental-scale ecological questions. Within sites, measurements of atmosphere, soil, water, select organisms and disease, and airborne observations yield freely available data to enable a new paradigm in ecological science with insights for education and direction for policy.

Automated sensors and observations will describe the ecological status and future trends NEON is designed to detect with a suite of measurements that span spatial and temporal scales. Fixed-wing aircraft census vegetation at landscape scales ($\sim 400\text{km}^2$) with high-resolution remote sensing at annual time steps and tower-based sensors capture temporally continuous fluxes over smaller spatial extents ($\sim 0.5\text{km}^2$). However, neither a census nor temporally continuous measurements are appropriate for understanding patterns of terrestrial biogeochemistry and organisms at the scale of a NEON site ($\sim 5\text{-}60\text{km}^2$). A complete census of organisms and biogeochemistry is biologically and financially impractical –

¹Additional authors: Paul A. Duffy, David S. Schimel, Rachel E. Krauss, Elena I. Azuaje, Kathi M. Irvine, Frank W. Davis, Alan E. Gelfand, Andrea S. Thorpe, David Gudex-Cross, Michael Patterson, Jalynda M. McKay, Joel T. McCorkel, and Courtney L. Meier

microbes are ubiquitous and birds are mobile. Likewise, measurement of these ecological responses at sensor-like temporal frequencies is impossible, and even frequent observations at local scales would likely provide redundant information or, due to financial constraints, be limited in spatial extent. Hence, terrestrial organisms and soil will be collected in the field by crews trained in standardized protocols measured at discrete temporal and spatial units by people making field-based observations (Kao et al. 2012, Thorpe et al. 2015).

The diversity of biogeochemistry and organismal measurements that will be made by the NEON Terrestrial Observation System (Thorpe et al. 2016) presents a formidable challenge to the coordinated collection of data for the Observatory. Measurements include biodiversity, phenology, biomass, stoichiometry, prevalence of disease, and genomics of soil and organisms with a range of life histories and phylogenetic traits (Keller et al. 2008, Schimel et al. 2011, Kao et al. 2012, Thorpe et al. 2015). Components of each will be targeted for observation with a sample design that directs the spatial location at which populations and states of interest shall be sampled (Thompson 2012). The design must collect data that capture spatial variability, facilitate the integration of observations, enable analysis with a diversity of analytical approaches, and contribute to ecological insight at large spatio-temporal scales. The strategy is described herein: guided by NEON principles and requirements, the Terrestrial Observation System sampling design provides a data collection framework that is statistically rigorous, operationally efficient, flexible, and readily facilitates integration with other data to advance the understanding of the drivers of and responses to ecological change. It should be noted that while this document provides the rationale and details of the NEON sample design for terrestrial organisms and soil, the description, justification and study design specifics for the taxonomic groups and soil sampled are described elsewhere (Barnett et al. in prep, Hinckley et al. 2015, Hoekman et al. in prep(a), Hoekman et al. in prep(b), Meier et al. in prep, Springer et al. 2015, Thibault et al. in prep, Thorpe et al. 2016).

Design Criteria

NEON will enable understanding and forecasting of the impacts of climate change, land-use change and invasive species on continental-scale ecology by providing infrastructure and consistent methodologies to support research and education (Keller et al. 2008). The traceable links between this high-level NEON mission statement and the data the Observatory produces provide a framework for the NEON design. The scope of the NEON mission is generally defined by the Grand Challenges in environmental science identified by the National Research Council (2001). High-level requirements synthesize the mission, Grand Challenges, and theoretical basis for measurements into formalized statements that describe the fundamental aspects and guiding architecture of the NEON strategy ((Schimel et al. 2011); Table 2.1). The sample design for organisms and soil is part of this requirements-driven hierarchical structure; high-level requirements “upstream” requirements and “downstream” data products provide context and constraints under which sample design specific requirements and details were developed.

The sample design for observations at local, site-specific scales must deliver data that optimally informs continental-scale ecology. Adopting the requirements framework allows traceability to elements of the continental sampling strategy and the high-level requirements that constrain the spatial observation at discrete landscapes across the continent (Table 2.1). A set of lower-level requirements specific to the sample design captures these objectives and provides a direct link to the high-level NEON requirements (Table 2.1).

Table 2.1. Connections between NEON high-level requirements and the requirements that guide the local, site-specific sample design for the terrestrial organism and soil observations.

NEON mission and high-level requirements from the NEON Science Strategy	Guiding principles and requirements of the Terrestrial Sampling Design
<ul style="list-style-type: none"> · NEON shall address ecological processes at the continental scale and the integration of local behavior to the continent, and shall observe transport processes that couple ecosystems across continental scales (i.e. continental-scale ecology). · NEON will allow extrapolation from the observatory’s local 	

<p>sites to the nation. NEON will integrate continental-scale data with site-based observations to facilitate extrapolation from the local measurements to the national observatory.</p> <ul style="list-style-type: none"> · NEON’s spatial observing design will systematically sample national variability in ecological characteristics, using an a priori division of the nation to allow extrapolation from limited intensive sampling of core wildland sites back to the continental scale. 	<ul style="list-style-type: none"> ✓ Direct the collection of the raw material for continental ecology
<ul style="list-style-type: none"> · NEON’s goal is to improve understanding and forecasting of ecological change at continental scales. · NEON shall detect and quantify ecological responses to and interactions between climate, land use, and biological invasion, which play out over decades. · NEON observing strategies will be designed to support new and ongoing ecological forecasting programs, including requirements for state and parameter data, and a timely and regular data delivery schedule. 	<ul style="list-style-type: none"> ✓ Efficiently capture landscape-scale pattern and trend
<ul style="list-style-type: none"> · NEON shall observe the causes and consequences of environmental change in order to establish the link between ecological cause and effect. · NEON’s measurement strategy will include coordinated and co-located measurements of drivers of environmental change and biological responses. 	<ul style="list-style-type: none"> ✓ Provide infrastructure that co-locates terrestrial measurements and links observations to other NEON data streams
<ul style="list-style-type: none"> · NEON shall provide infrastructure to scientific and education communities, by supplying long-term, continental-scale information for research and education, and by supplying resources so that additional sensors, measurements, experiments, and learning opportunities can be deployed by the community. · The NEON infrastructure shall support experiments that accelerate changes toward anticipated future conditions. · NEON will enable experiments that accelerate drivers of ecological change toward anticipated future physical, chemical, biological, or other conditions to enable parameterization and testing of ecological forecast models, and to deepen understanding of ecological change. 	<ul style="list-style-type: none"> ✓ Facilitate spatial integration of NEON data with community-driven investigation
<ul style="list-style-type: none"> · The NEON data system will be open to enable free and open exchange of scientific information. Data products will be 	

designed to maximize the usability of the data. The NEON sites will be designed to be as amenable to new measurements and experiments as possible in order to effectively provide NEON infrastructure to scientists, educators, and citizens.	✓ Anticipate the need for design flexibility
· NEON infrastructure and observing system signal-to-noise characteristics will be designed to observe decadal-scale changes against a background of seasonal-to-interannual variability over a 30-year lifetime.	✓ Optimize the design through iterative observation and evaluation of spatial and temporal variability

A more detailed explanation of the requirements associated with the terrestrial sample design provides further guidance for the design:

- **Direct the collection of the raw material for continental ecology.** Site-specific observations provide the foundation of the continental observatory (Urquhart et al. 1998). The deployment of an unbiased and consistent sample design will provide comparable ecological response metrics across sites and domains (Olsen et al. 1999, Lindenmayer and Likens 2010). Efforts to scale patterns to larger areas will be aided, for example, by optimizing of the links to NEON remote sensing observations, adequately characterizing landscape features that dominate at regional scales, and by sampling with methods comparable to other network, agency, and other science and monitoring efforts.
- **Efficiently capture landscape-scale pattern and trend.** Organisms and soil should be measured with intensity sufficient to detect the presence of a trend over the life of the Observatory (Legg and Nagy 2006, Lindenmayer and Likens 2009). The design must contribute to accurate, precise, and unbiased descriptions of local landscapes. Sample number and location will be directed by the sample design (Urquhart et al. 1998, Thompson 2012) while trend detection will depend on a diversity of community-derived analytical approaches applied to the data. Given the variety of approaches likely to be employed and the diversity of questions to be addressed with NEON data products, the sample design framework must be applicable to classical, contemporary, and future statistical approaches

- that characterize patterns in space and through time (Cressie et al. 2009, Cressie and Wickle 2011).
- ***Provide infrastructure that co-locates terrestrial measurements and links observations to other NEON data streams.*** The terrestrial measurements must be co-located to provide a more complete picture of processes associated with targeted observations and trends across the groups to be sampled (Fancy et al. 2009). Point-based observations must also be readily integrated with the spatially continuous NEON remote sensing platform and temporally continuous sensor measurements (Sacks et al. 2007, Sun et al. 2010). The evaluation of correlative relationships through the iterative combination of models and data (Luo et al. 2011) will provide insight into mechanistic links between the cause and response of ecological change. These relationships can then be further explored and tested with rigorous experiments by the ecological community (Keller et al. 2008, Lindenmayer and Likens 2010).
 - ***Facilitate spatial integration of NEON data with community-driven investigation.*** The terrestrial sampling design must provide a framework that encourages the scientific community to conduct experiments and other observations that integrate with NEON data to synergistically and efficiently deepen understanding of ecological processes (Lindenmayer and Likens 2010).
 - ***Anticipate the need for design flexibility.*** The sample design must accommodate changes as NEON responds to unexpected and/or emerging patterns and contribute to questions contemporary ecology has not yet considered (Overton and Stehman 1996).
 - ***Optimize the design through iterative observation and evaluation of spatial and temporal trends and variability.*** The number and spatial-temporal distribution of samples reflects assumptions about variability of response, landscape characteristics, and budget

constraints. Early data will serve to evaluate these assumptions and provide guidance for the reallocation of sampling to better address NEON questions (Hooten et al. 2009, Lindenmayer and Likens 2009). Additionally, the unprecedented characterization of NEON sites by the airborne observation platform will allow the identification of gradients, disturbance, and/or other landscape features that might be measured to better understand spatial-temporal patterns over the life of the Observatory.

The high-level NEON requirements capture the essence of the NEON mission and Grand Challenge, creating direction and context for actionable design of Observatory components. The sample design requirements outlined above stem from high-level design elements and provide further direction and constraints in the face of specific design needs: how observations should be distributed in space at the scale of NEON sites.

Sampling Design for the Terrestrial Observation System

Two principles guide the site-scale terrestrial organismal sampling design: randomization and robustness. Randomizing sample locations is possible in – and facilitates comparability of data across – a diversity of biomes (Carpenter 2008), guards against the collection of data that are not representative of the populations of interest (Thompson 2012), and yields data suitable to a diversity of analytical approaches (Cressie et al. 2009). The design must be robust in the sense that it is capable of performing under a diversity of conditions, and accommodating a variety of data types and questions (Olsen et al. 1999).

Terrestrial observations range from microbes to long-lived trees. NEON science questions will be addressed with hundreds of data products. The ecological community will ask untold additional questions and tease answers from data with a range of analytical techniques. And, these techniques will evolve over decades (Cressie and Wickle 2011). Intended to detect patterns across a diversity of spatial conditions (Carpenter 2008) and elucidate temporal trends by meeting the demands of contemporary

and future ecological paradigms (Cressie et al. 2009) in support of a long-term observatory, the sample design for terrestrial organisms and biogeochemistry includes the following elements:

- **The sample frame** is the area from which observations are made (Reynolds 2012).
- **Random sampling** allows an unbiased description of the landscape (Thompson 2012), facilitates integration with other data, supports design-based inference (Sarndal 1978), and provides data that can be assimilated into numerous model-based approaches to inference and understanding.
- **Stratification** increases efficiency (Cochran 1977) and provides a framework for describing the variability of landscape characteristics targeted by the NEON design.
- **Sample size determination** ensures that NEON will contribute to ecology over the life of the Observatory by providing sufficient data to support key questions (Thompson 2012, page 30).
- **Sample allocation** allows a distribution of sampling effort appropriate to particular observations and NEON questions.
- **Data analysis with variance estimators** provides a solution for analysis of data with design-based inference (Stehman 2000).
- **Iteration** allows optimization of the sample design (Di Zio et al. 2004).

Furthering the understanding of ecological change requires an emphasis on integration and collocation of observations with a design not optimized for any particular taxonomic group. The spatial and temporal resolution and extent at which the design resolves ecological patterns will vary across responses and is ultimately constrained by scientific feasibility within an envelope of logistics and funding. Hence, the proposed design represents a multitude of compromises from competing priorities and a primary focus on implementing continental-scale ecology at local scales.

Sampling frame

The sampling frame defines the area from which observations are made to characterize variables of interest (Reynolds 2012). At the scale of NEON sites, the sampling frame depends on the type of plot (see Thorpe et al. 2015) and taxonomic group of interest. In the case of many of the vegetation and soil observations (Thorpe et al. 2015), the frame typically corresponds to an associated management or ownership boundary (Figure 2.1). This typically includes the location of the tower-based sensor measurements and the aquatic measurements at some sites (Thorpe et al. 2015). Design constraints limit the spatial extent of some observations. Mosquito sampling occurs within 45 m of roads, and small mammal sampling occurs within 300 m of roads due to the frequency of visit and equipment required for sampling.

The size of the sampling frames is variable, from small landscapes (e.g., an agricultural site in Sterling, Colorado < 5 km²) to larger wildland sites (e.g., part of Oak Ridge National Lab 67 km²). At several sites, the area available for sampling is too large to be sampled given budget and travel constraints or some sections of the site are not available for sampling (e.g., Oak Ridge National Lab). In these cases, a subset of the areas is targeted for sampling based on discussions with site hosts, local scientists, and logistical constraints. These truncated sites generally result in a 15 – 80 km² sampling frame.

NEON's tower-based sensors measure physical and chemical properties of atmosphere-related processes such as solar radiation, ozone, and net ecosystem exchange. Tower Plots (Thorpe et al. 2015) sample that part of the landscape reflected in the sensor data to allow calibration and comparison of temporal trends. That sample space – the airsheds and in some cases the landscape in-between – constitutes the sample frame for those observations (Figure 2.1).



Figure 2.1. NEON's Domain1 is located in the south-east United States. The site at the Ordway-Swisher Biological Station in central Florida is managed as a research station by the University of Florida and includes a diversity of pine on sandy soils, broadleaf forests on wetter soils, and wet marshes. The site boundary encompasses a 34 km² area. The NEON tower (in white) supports sensors that measure fluxes from primary and secondary airsheds (in yellow). Airsheds, or in some cases, the complete 360-degree area defined by the primary airshed radius, define the sample frame for vegetation and soil designed to help inform flux observations.

Randomization

The unbiased sample associated with randomization (Cochran 1977, Thompson 2012) is the foundation of the NEON sample design. Randomly sampling from the frame eliminates potential bias associated with subjective sampling and affords the assumption that the statistical bias, the difference between the sample mean and true mean, is zero (Cochran 1977, Gitzen and Millspaugh 2012).

This unbiased sampling of target response variables is essential to a probabilistic sample design. Probability sampling mandates that each randomly selected sample location have a known, non-zero chance of being selected for observation (Thompson 2012). The principles of randomization allow the design-based inference of population parameters from points to the unsampled landscape by integrating data and inclusion probabilities – the chance of each sample location being selected for observation - with design-based estimators (Sarndal 1978, Stehman 2000). Appropriate estimators can be determined by structure of the data and particular sample design (Stevens and Olsen 2004).

Contemporary ecology relies on a variety of alternative sampling approaches. For example, systematic sampling locates observations according to a uniform grid (Cochran 1977, Thompson 2012).

By forcing sampling effort across the landscape, systematic sampling minimizes spatial autocorrelation and can capture landscape heterogeneity (Fortin et al. 1989, Theobald et al. 2007). However, the uniform distribution of sampling limits the opportunity to capture spatial patterns that might exist in the data (Fortin et al. 1989). Systematic designs that incorporate an element of randomization (e.g. spatially balanced designs) vary the spatial distance between sample locations, allowing the design to better describe the impact of spatial patterns associated with underlying processes. Other designs include stratified (Cochran 1977, Overton and Stehman 1996), spatially balanced sampling (Stevens and Olsen 2004), cluster sampling (Cochran 1977, Stehman 2009), variable density designs (Stevens 1997), and two-stage designs (McDonald 2012). Not all designs support design-based inference. Sampling areas thought to be representative of a site – subjective sampling - assumes a near-perfect *a priori* understanding of the landscape (Stoddard et al. 1998, McDonald 2012) and does not allow for the detection of unexpected patterns across a landscape (Lindenmayer et al. 2010). The lack of fundamental randomization results in a sample that is not unbiased and is incompatible with design-based inference to the unsampled population (McDonald 2012).

Model-based sample designs (Albert et al. 2010, Smith et al. 2012) are becoming increasingly popular for specific research and monitoring questions, but they are not sufficiently general with respect to the design requirements for the variety of organisms, soil, and questions NEON hopes to address. Relying on models, instead of design-based inference for the description of unsampled landscapes and populations, frees the sample design from constraints of randomization imposed by a probability-based design (Sarndal 1978). Statistically-rigorous modeling techniques allow for the distillation of patterns from a sample. Basic approaches explain variability in the response variable with traditional frequentist statistical models, typically linear statistical analyses with corresponding necessary and sufficient conditions. More complex techniques focus on the spatial structure of data, rely on machine-learning algorithms to understand non-linear relationships between multiple variables (Elith et al. 2010), allow

parameters to be defined as probabilities (Wikle and Royle 1999, Fuentes et al. 2007), or describe patterns from data measured through time and across space (Cressie and Wickle 2011). These model-based approaches to inference can be optimized by specific sampling efforts. Data can be collected according to a stratified, non-random design that targets the spatial structure of a population (ver Hoff 2002), captures the complete dynamic range of particular variables (Di Zio et al. 2004), or focuses on particular gradients and patterns (Chao and Thompson 2001). However, a sample design optimized for a specific question or parameter fails the test of generality required to sample many organisms and address a diversity of ecological questions (Bradford et al. 2010).

By relying on randomization, the NEON sample design will produce data suitable to a variety of analytical techniques, from design-based inference to model-based approaches (Cressie et al. 2009). This process of teasing patterns and understanding from data is crucial to the success of NEON. Facilitating the integration of disparate data and identifying the mechanisms that underlie observed patterns (Levin 1992) is key to understanding the causes and consequences of change over the life of the Observatory.

Randomization at NEON sites

The design requirements collectively provide a strong case for explicit emphasis on the characterization of spatial patterns. Despite the benefits provided by the randomization of a simple random sample, these benefits can be tempered by a lack of spatial coverage. The NEON design satisfies these constraints by sampling with a spatially-balanced sampling framework that also provides randomization. Spatially-balanced sampling results in a probability-based study design, with low to moderate variance, and is both simple and flexible (Stevens and Olsen 2004). The Reversed Random Quadrat-Recursive Raster (RRQRR; Theobald et al. 2007) approach is similar to the Generalized Random Tessellation Stratified (GRTS) algorithm implemented by several existing long-term ecological monitoring efforts (Larsen et al. 2008, Fancy et al. 2009). The principle difference is that RRQRR achieves spatial

balance in a Geographic Information System (GIS) environment and produces a complete sample instead of a defined sample size. Implementation in GIS facilitates the incorporation of site boundaries, identifies barriers to sampling (e.g., roads, lakes), allows visualization of the study design, and provides design flexibility and redundancy to assign alternative locations should a plot be unsuitable for sampling (Theobald et al. 2007).

The complete sample associated with the RRQRR algorithm allows design flexibility that is critical to logistical efficiency and scientific success. Every sample unit (a 30 x 30 m pixel in the case of the NEON design) receives a potential plot location that is numbered in a spatially-balanced framework, addressed – assigned a named location, randomized, and ordered such that sampling according to a one-dimensional list provides a random, spatially-balanced design allocation across the site (Theobald et al. 2007). Should a particular plot be unsuitable for sampling, the next unassigned, sequential plot on the list can be included in the sample. Other reasons to add plot locations may arise. Results from initial sampling will provide data to direct iterative observations that might require different sample size and distribution. Additionally, independent Principal Investigator-driven science may more efficiently address questions beyond the scope of the NEON design by leveraging the NEON data stream and utilizing sample locations specified by this design approach. The availability of sampling locations from the NEON terrestrial study design will facilitate this integration.

Generation of the spatially-balanced design is accomplished with the RRQRR function that maps 2-dimensional space into 1-dimensional space. RRQRR employs Morton ordering (Theobald et al. 2007), a hierarchical quadrant-recursive ordering. Morton ordering creates a recursive, space-filling address by generating “N” shaped patterns of 2x2 quads that are composed of lower-left, upper-left, lower-right, and upper-right cells numbered and nested at hierarchical scales (Figure 2.2). The pattern maximizes 2-dimensional proximal relationships when converting to 1-dimensional space such that 1-dimensional ordered addresses are close together in 2-dimensional space (Theobald et al. 2007).

The NEON sampling design as a random sample

The spatially balanced, random sampling locations generated by the RRQRR algorithm provide the sample design with flexibility. While the NEON design does optimize sampling according to a stratified-random design (see below) by selecting a subset of available points from particular strata, a subset of the data can be treated as a random sample. The initial generation of sample locations in the random, spatially balanced and ordered list conforms to assumptions (Theobald et al. 2007) that allow a subset of the sample locations and resulting data to function as a random sample. This number of sample locations and the fraction of the total sampling effort that can be considered random depends on site size, heterogeneity, and in the evenness of selected strata. All of the sample locations can be considered random at homogenous sites, while those sites represented by a variety of strata result in a relatively smaller sample size available to any analysis and assumptions dependent on a random sample (Table 2.2). A list of plots that can be used in the context of a random design by site will be available through the NEON data portal. These alternatives to sampling make the data more broadly available to a variety of NEON data consumers, ecological questions, and statistical applications. Tradeoffs and preferences abound in the ecological community.

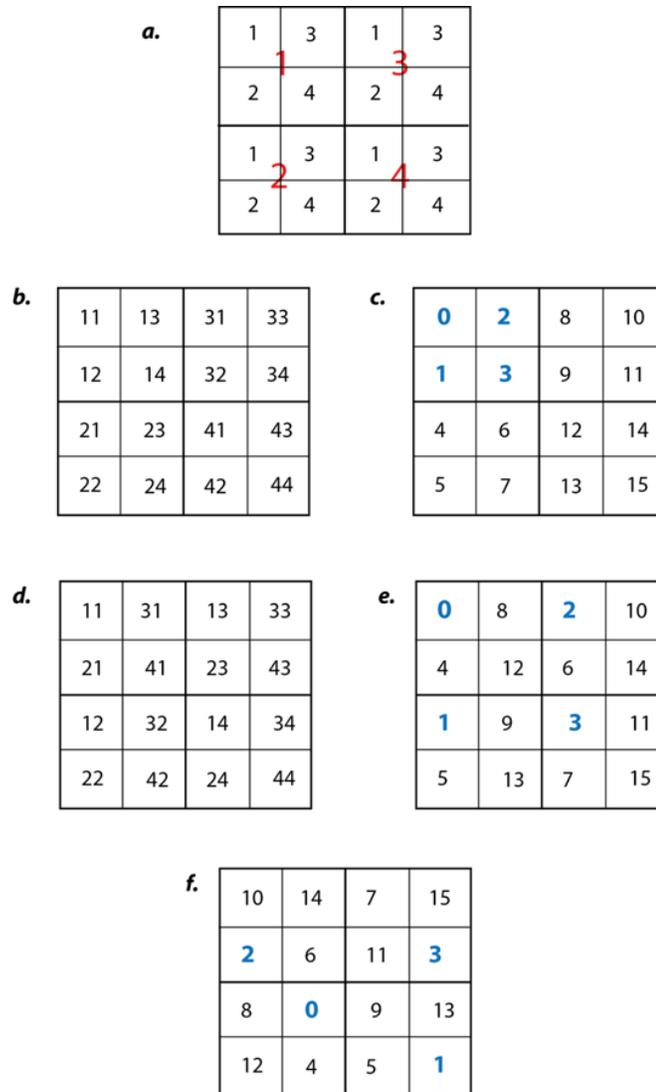


Figure 2.2. The spatially balanced RRQRR design for locating sample plots across NEON sites. RRQRR assigned Morton addresses to a very large number of cells in a raster. The steps to create a spatially balanced list based on the RRQRR design include (a) the recursive order formation of the Morton Address on a two dimensional frame of coordinates into quadrant levels, the numbers in red represent one quadrant level and numbers in black represent another quadrant level; (b), the Morton addresses representing the recursive order; (c) an assigned sequential Morton Order; (d) the Morton Address is reversed to create a uniform systematic pattern; (e) a new systematic Morton Order pattern is created; (f) and randomization is generated at each quadrant level. After Theobald et al. (2007).

Stratification

Stratification divides the landscape of interest into non-overlapping subareas from which sample locations are identified (Cochran 1977, Johnson 2012). The approach provides value when the ecological

measurements of interest are more similar within a stratum than among strata (Johnson 2012). Specifically, from the perspective of design-based inference, stratification aims to reduce the variance (Nusser et al. 1998, Scott 1998) of parameter estimates under the condition that the average variation of an estimator within a stratum is less than the average variation among strata (Michaelsen et al. 1994). The increase in precision typically results in greater efficiency; fewer observations describe the within-stratum variability of parameter estimates and patterns of interest across the entire sampling frame (Cochran 1977).

The NEON terrestrial sample design stratifies by land cover type in a manner consistent with the guiding principles of the domain delineation, to facilitate comparison within and across NEON sites, and to ensure the design captures a variety of environmental gradients at each site. Stratification according to the National Land Cover Database (Fry et al. 2011) provides a continuous land cover classification across the United States including Puerto Rico, Alaska, and Hawaii, allowing consistent and comparable stratification across the diversity of NEON sampling frames. This stratification satisfies multiple design requirements and objectives.

First, stratification is an integral part of the NEON design at multiple scales, and when applied to the terrestrial sample design, stratification provides consistency and ensures observations describe local landscape characteristics essential to the continental-scale observatory. NEON domains – essentially a stratification of the continent – were derived from eco-climatic factors (Hargrove and Hoffman 2004) that contribute to large-scale patterns of vegetation (Figure 2.3). Within each domain, NEON sites are selected to represent the dominant vegetation type (Schimel et al. 2011). At each NEON site, tower-based sensors were positioned to measure these dominant vegetation types. The sensors measure ecosystem properties that drive ecological response (Chapin et al. 2012, Clark et al. 2012, Sala et al. 2012). Observing terrestrial biogeochemistry and organisms in this dominant vegetation type at each

Table 2.2. A subset of Distributed Plots can be used as a random sample. Three example sites, Konza Prairie Biological Station (KONZ), Talladega National Forest (TALL), and the Jornada (JORN) suggest that a greater number of samples function as part of a random sample at sites with fewer strata. Greater within-site heterogeneity with respect to number and relative size of strata results in a smaller number of plots that can be considered part of a random sample.

Site	Subtype	Stratified-random plots			Number of random plots
		NLCD cover type	Area (km ²)	Number of plots	
KONZ	Base plot	Grassland/herbaceous	29.8	23	19
		Deciduous forest	3.3	<u>7</u> Total: 30	
KONZ	Mosquito point	Grassland/herbaceous	4.9	9	10
		Deciduous forest	0.3	<u>1</u> Total: 10	
KONZ	Mammal grid	Grassland/herbaceous	28.2	6	5
		Deciduous forest	3.1	<u>2</u> Total: 8	
KONZ	Tick plot	Grassland/herbaceous	29.8	4	3
		Deciduous forest	3.3	<u>2</u> Total: 6	
KONZ	Bird grid	Grassland/herbaceous	29.8	9	7
		Deciduous forest	3.3	<u>3</u> Total: 12	
TALL	Base plot	Deciduous forest	16.6	10	10
		Evergreen forest	18.2	11	
		Mixed forest	13.8	<u>9</u>	
				Total: 30	
TALL	Mosquito point	Deciduous forest	1.8	3	1
		Evergreen forest	3.1	4	
		Mixed forest	1.6	<u>3</u>	
				Total: 10	
TALL	Mammal grid	Deciduous forest	15.4	3	3
		Evergreen forest	15.9	3	
		Mixed forest	12.4	<u>2</u>	
				Total: 8	
TALL	Tick plot	Deciduous forest	16.6	2	5
		Evergreen forest	18.2	2	
		Mixed forest	13.8	<u>2</u>	
				Total: 6	
TALL	Bird grid	Deciduous forest	16.6	5	4
		Evergreen forest	18.2	5	
		Mixed forest	13.8	<u>5</u>	
				Total: 6	
JORN	Base plot	Shrub/scrub	45.7	30	30
JORN	Mosquito point	Shrub/scrub		10	10
JORN	Mammal grid	Shrub/scrub		6	6
JORN	Tick plot	Shrub/scrub	45.7	6	6
JORN	Bird grid	Shrub/scrub	45.7	10	7

NEON site will quantify the relationship between state factors – variables that control characteristics of soil and ecosystems (Chapin et al. 2012) – and ecological response. Through time these observations will provide insight into the causes and consequences of change at NEON sites which, due to the scalable design, will further understanding at larger spatial scales.

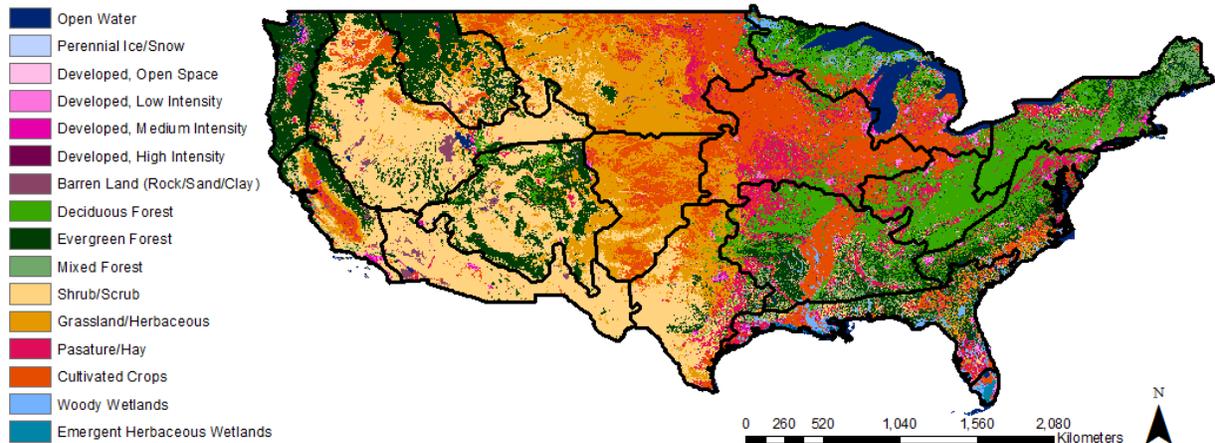


Figure 2.3. NEON domains layered on top of land cover types as described by the National Land Cover Database.

Second, stratification by land cover allows differential allocation of resources and sampling effort across cover types. In addition to facilitating a focus on the dominant vegetation type as described above, stratification provides a means to facilitate comparison. Sampling with an initial allocation that makes assumptions about patterns of the variability associated with an ecological response across the landscape allows for a distribution of observations that will stabilize variance of estimators among strata. Approximately equal patterns of variability facilitates comparison of ecological response across vegetation types within a site and, crucial to the success of a the continental Observatory, comparison among NEON sites as well.

Caveats associated with stratification by cover type merit recognition, and alternative schemes exist. Vegetation will change over time (Scott 1998). NEON hopes to capture this change, but the choice of a dynamic strata will complicate design-based inference (Fancy et al. 2009). As such, NEON will

develop adjustments to design-based estimators and the inclusion probability of each sampling stratum (Stevens and Olsen 2004). Other long-term monitoring units either do not stratify, or select immutable strata (Reynolds 2012). Elevation might be suitable at sites where vegetation changes reflect significant topography and relief (Li et al. 2009); however much of the biological variability across the continent responds to other factors. Soil type is less likely to change in a meaningful way over the life of the observatory and continental-scale maps exist across the continent. However, many soil maps were created according to inconsistent standards at the county level, are not highly accurate, and interpolation between dispersed sampling reflects vegetation captured by aerial photography. These and other unchanging strata might be appropriate for a local study or to optimize for a particular question or taxonomic group (Fancy and Bennetts 2012). Stratification by vegetation represents a compromise that emphasizes a consistent approach to continental-scale ecology that can be implemented in a consistent way across all domains.

Stratification at NEON sites

The land cover vegetation strata were described by the National Land Cover Database (Fry et al. 2011). The NLCD is created through a partnership that includes the US Geological Survey, the Environmental Protection Agency and other federal partners. The categories are general and describe high-level and coarse descriptions of landcover (Figure 2.4). In the context of the RRQRR sample design, stratification is achieved by iteratively intersecting points from the sample list with each land cover type by assigning an inclusion probability of one to areas associated with the target vegetation type and zero for non-target types. In other words, the one-dimensional list developed by the RRQRR remains unchanged; selecting points within a particular land cover type filters that list. The result is a random, spatially-balanced sample design that is stratified by vegetation (Figure 2.4).

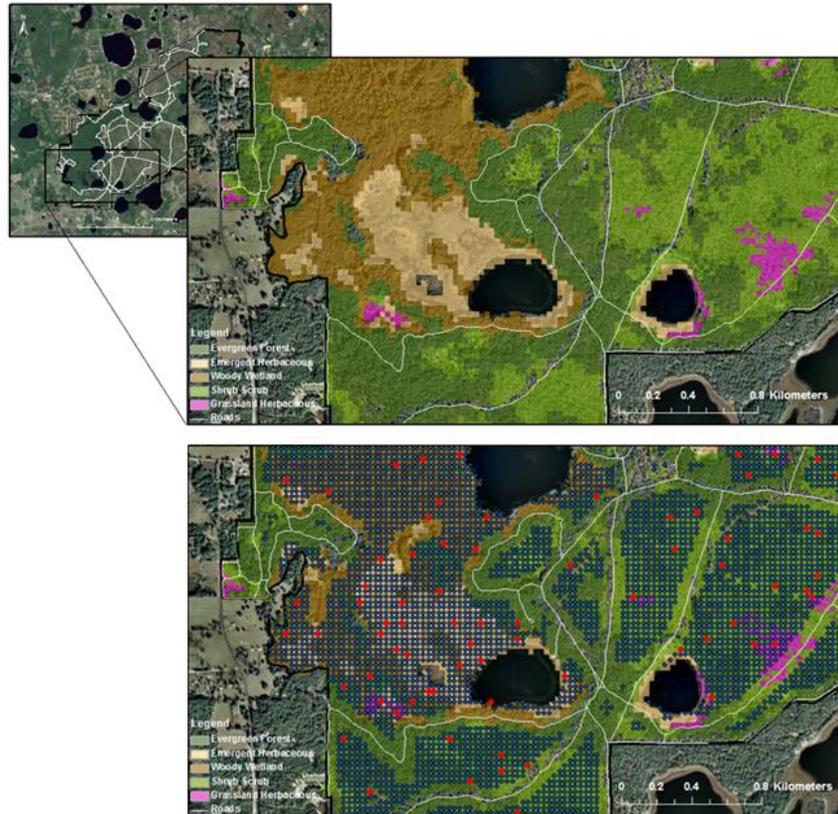


Figure 2.4. Stratification by the National Land Cover Database at the Ordway-Swisher Biological Station (a). Blue dots represent potential sampling locations from the spatially balanced and randomized sample, and red points indicate hypothetical sample locations selected from the complete sample (b).

Minimum sample size

An overarching requirement of the design is that minimally sufficient data be collected within each stratum where samples are allocated. This ensures that the NEON effort will provide tangible contributions to conceptual models of the interactions between species and environmental drivers over the life of the observatory. Simply put, if data will be collected in a given vegetation class, it is necessary to ensure that after thirty years, these data are sufficient to understand local patterns and, ultimately, inform the NEON Grand Challenges (Legg and Nagy 2006). Much like the need for a generalized sample design that is robust to observations of biogeochemistry and multiple biological groups, the sample sizes

must be sufficient to answer an array of questions (Gitzen and Millspaugh 2012) across a number of disparate ecological response variables.

Quantitative sample size calculations are most often performed against the backdrop of a classical hypothesis test and corresponding power analysis. These analyses are constrained by a number of factors including: a question of interest, a corresponding hypothesis test regarding a parameter of interest in a statistical model, assumptions regarding the error tolerances (i.e., power) and estimates of parameter values for the population of interest (Hoenig and Heisey 2001). In order to characterize minimally sufficient sample sizes for the design, several key questions that are derived from the design requirements are considered.

As an initial case, a question representative of the large-scale, long-term science NEON will enable was considered to provide context for the analysis of sample size: is there a difference in temporal trends of a given response of interest between two populations of interest? Examples of specific questions enabled by NEON data might include:

- Are trends in tree canopy height in the deciduous forest cover type different between a wildland site and a site managed for timber harvest in Domain 5?
- How do trends in invasive plant species richness differ between a wildland site and a site managed for cattle grazing in Domain 12?
- How do temporal patterns of plant diversity vary three across sites distributed across an elevation gradient in Domain 17?

The described sample size analysis considered a test of the difference in the magnitude of trends between any two NEON sites. One way to account for the diverse range of ecological response that will be sampled is to characterize the range of variability (across these disparate populations of responses) in parameters that need to be specified in order to constrain the sample size. This approach does not provide a unique solution; rather it provides a range of minimal sample sizes that correspond to the

range of parameter values that are considered. In this way, the differences in minimal sample size as a function of the populations considered can be accounted for when utilizing this information to constrain the overall sample design. The result of this design constraint provides a guideline for sample size rather than a definitive threshold. The analysis incorporated the capability to assess the impact of varying parameters that must be specified *a priori*. Once several years of data are collected, the design can be reassessed, and iteratively optimized with alternative methods using data from the initial sampling results.

A classical power analysis (Hoenig and Heisey 2001, Thompson 2012) guided the estimation of sample size. A linear mixed effects model with repeated measures was used to represent differences in trends between two sites. These analyses can be applied to any test of a difference between the slopes, which respectively quantify change through time at each site where repeat measurements are taken on the same sampling units within each group. In general, the sampling units correspond to the spatial extent across which the response of interest is measured. In this context, the sampling units are the pixels (i.e., units) within the RRQRR grid at each site. Values for the parameters in the statistical model that have relevance to these sample size calculations – within site spatial variability of the response variable, temporal variability of the response variable, and temporal correlation structures of the response variable - must be informed by evidence from previous studies or prototype data. The model accommodates both compound symmetric and first order autoregressive temporal correlation structures for the repeated measures component of the variance calculations. In practice, the values associated with the parameters will vary across each of the response variables and across sites.

Initial sample size calculation

In addition to the sample variance, the magnitude of the correlation associated with the repeated measures, and the temporal correlation structure, sample size calculations that utilize a power constraint require specification of acceptable error tolerances for each of the two types of decision

error, minimum detectable difference associated with the type II error, and estimates of relevant parameters for (co)variance (Thompson 2012). This specific application also requires the number of repeat measurements – initially assumed one annually - within the course of the study. The notation presented here generally follows Searle (Searle 1971) and utilizes the approach of Yi and Panzarella (Yi and Panzarella 2002) to specify the relationship between the specified significant difference in slopes through time (i.e., the location in the alternative parameter space where the power of the test is constrained), as well as the treatment of the variance associated with the slopes depicting changes in trends through time at sites to be compared. Hence, consider the following repeated measures model with mixed effects:

$$Y_i = \mu_0 + \mu_{0i} + \alpha_1 * time + \beta_{1i} * time + \alpha_2 * site + \beta_{int} * (site * time) + \varepsilon_i \quad [1]$$

where the following symbolic definitions hold:

- Y_i is a vector representing observations through time t (i.e. the number of repeat measurements) at the i^{th} sampling location,
- with respect to measurement i , μ_{0i} is a random intercept, β_{1i} is a random slope of time for the i^{th} sampling location,
- α_1 is the mean trend for Y_i ,
- α_2 is the difference between the overall means from the groups of observations taken from the two different sites or sampling frames,
- β_{int} is the difference in trends through time between the groups of observations taken from two different sites or sampling frames. It is a hypothesis test regarding this parameter that constrains the sample size calculations presented here.
- ε_i is a vector representing errors through time t (i.e. the number of repeat measurements) at the i^{th} sampling location.

The parameters (Equation 1) can be grouped according to their consideration as representing either random or fixed effects. The random effect parameters were denoted as, $\lambda_i = (\mu_{0i}, \beta_{1i})$ and the fixed effect parameters were denoted $\tau = (\mu_0, \alpha_1, \alpha_2, \beta_{int})$. Using this grouping of the parameters, the equation 1 can be re-written as,

$$Y_i = X_i\tau + M_i\lambda_i + \varepsilon_i \quad [2]$$

Where, X_i is a design matrix with t rows and p columns, and M_i is a matrix with t rows and q columns. Here $q \leq p$ and the columns of M_i are also columns of X_i .

This formulation (Equation 2) is convenient for the expression of the sampling distribution of the parameter of interest, β_{int} . Using both the Wald test and an appeal to the asymptotic normality of β_{int} allows for the following approximation of the test statistic of interest (Yi and Panzarella, 2002).

$$\frac{\hat{\beta}_{int}}{\sqrt{var(\hat{\beta}_{int})}} \sim N(0, 1) \quad [3]$$

Under the assumption that the sample sizes between populations are equal, we can use equation 3 to arrive at the following formula for sample sizes,

$$n = \frac{\left(Z_{(1-\frac{\alpha}{2})} + Z_{\beta} \right) * (X_1^T V^{-1} X_1 + X_2^T V^{-1} X_2)^{-1}}{\beta_{int}^2} \quad [4]$$

where,

- Z represents the quantile from the standard normal distribution corresponding to the desired error rate for the type I and type II errors,
- X_1 is the design matrix corresponding to samples of one population of interest,
- X_2 is the design matrix corresponding to the samples of the other population of interest,
- V is the covariance matrix for the observed data Y.

Initial minimum sample size at NEON sites

Ranges for the relevant parameter values in the sample size calculations were considered since the nature of the exact response across sites and variables of interest is unknown. Population variance was estimated across the groups of organisms to be sampled by NEON from a review of literature (Knapp and Smith 2001, Eisen et al. 2008, Cardenas and Buddle 2009) that included LTER publications and data archives (Cedar Creek, Hubbard Brook, Jornada, Sevilleta, USGS NAWQA Program) and from initial data collection at NEON sites. Ultimately, four levels of population variance were assessed (Table 2.3).

In the absence of time series data, temporal parameters were estimated with ten years of MODIS-derived Normalized Difference Vegetation Index (NDVI) that was assumed to be an adequate high-level descriptor of ecosystem variability. These data provide nine observations for the lag-1 interannual correlation of this signal, which integrates across space (i.e. the core site footprint) and time as constrained to NDVI peak greenness (Figure 2.5). Correlations of these NDVI data informed the range of temporal correlations initially specified in the sample size calculations (Figure 2.5, Table 2.3). The form of the temporal correlation structure was also characterized with these NDVI data. The analyses across the twenty core sites suggested that a compound symmetric correlation structure was appropriate for the 20 sites tested, but sample calculations are included for the first order autoregressive process as it is likely some of the other 17 sites will actually display trends more closely aligned with an autoregressive framework.

Type I error tolerance was assessed for levels of 0.05 and 0.10. In order to impose a constraint on the power curve for this test, it is necessary to specify the significant difference between slopes at which the power is set to 0.80. For these analyses, a significant difference was determined to exist if the slopes were great than 20% different from one another.

In the case of the compound-symmetric temporal specification there was a monotonic, yet non-linear relationship between the number of samples, the temporal correlation, the population variance, and collection of data through time (Figure 2.6). The impact of changing the type I error rate from 0.1 to 0.05 was less than the range of values corresponding to changes in correlation and population variance.

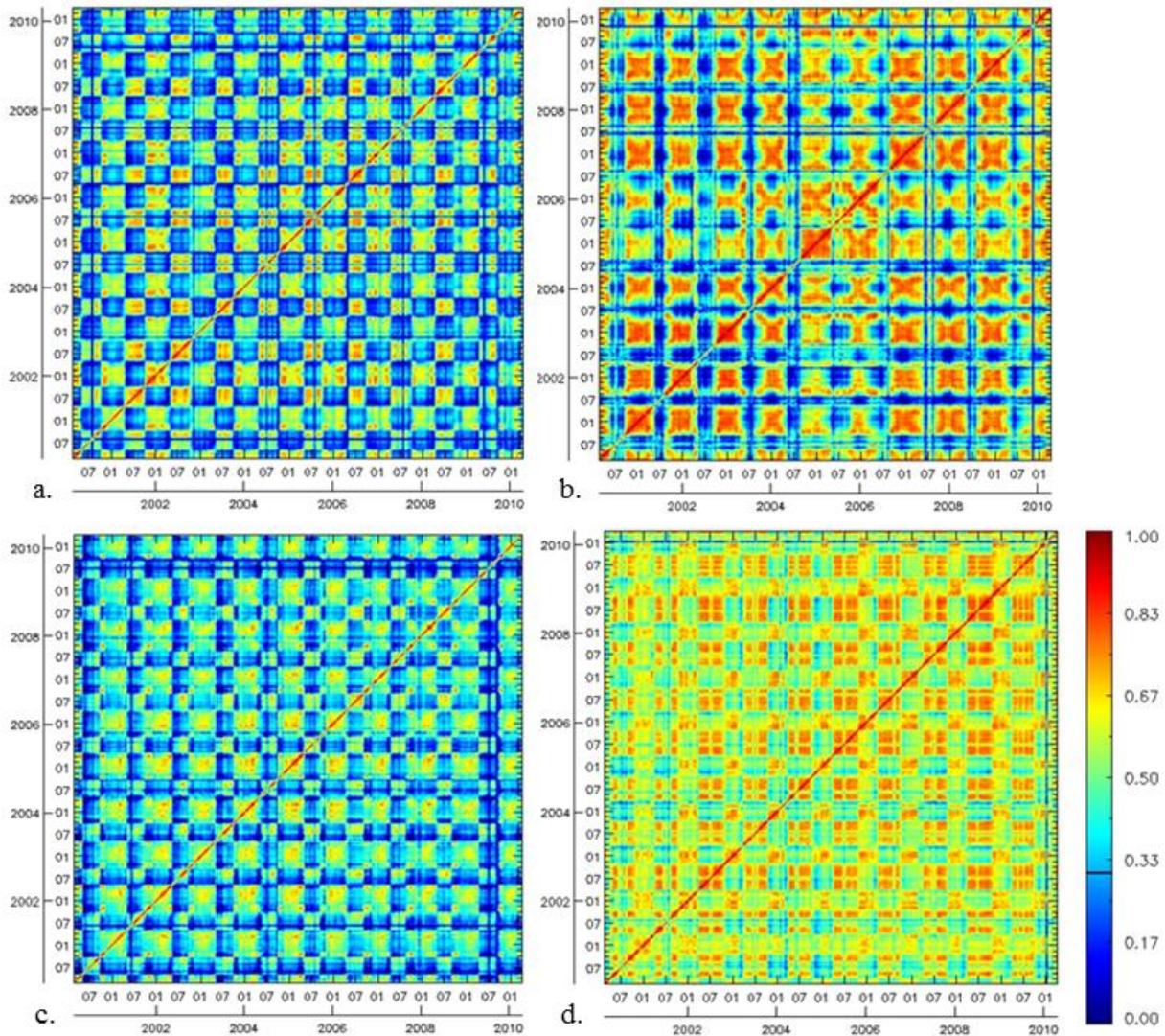


Figure 2.5. Annual temporal correlations from 2000 - 2010 of normalized difference vegetation index (NDVI) at Harvard Forest (a), Ordway-Swisher Biological Station (b), University of Notre Dame Environmental Research Center (c), and Oak Ridge National Lab (d). The lack of a consistent decay in temporal correlation at these sites through time over any consecutive number of years suggests that a compound symmetric form is an appropriate correlation structure of the sample size results.

After thirty years, the minimum number of samples needed across the range of values considered in both the compound symmetric and auto-regressive case was 10-189, with the lower number corresponding to the high correlation, low variability case, and the larger number of samples needed for the low correlation, high variability case (Table 2.3). The magnitude of the correlation associated with the autoregressive process demonstrated a lack of monotonicity between the number of samples and both the number of years data are collected (Figure 2.6).

An important assumption that was made but not assessed quantitatively in the context of the sensitivity of the results was that of equal sample allocation between sites. The calculations presented here are likely to be robust with respect to minor deviations from this assumption of equal allocation. For this work, the assumption that the sample sizes are equal between sites was made for the sake of simplicity. This interpretation could be relaxed to accommodate different sample sizes should that be necessary given the variability in size and heterogeneity across all NEON sites.

Another assumption was the specification of the significant difference at which the power constraint is imposed. The parameter in the statistical model that was used to build the test for the sample size calculations considered the slope of the interaction between site and time. In order to impose a constraint on the power curve for this test, it was necessary to specify the significant difference between slopes at which the power is set to 0.80. For these analyses, a significant difference was determined to exist if the slopes were greater than 20% different from one another.

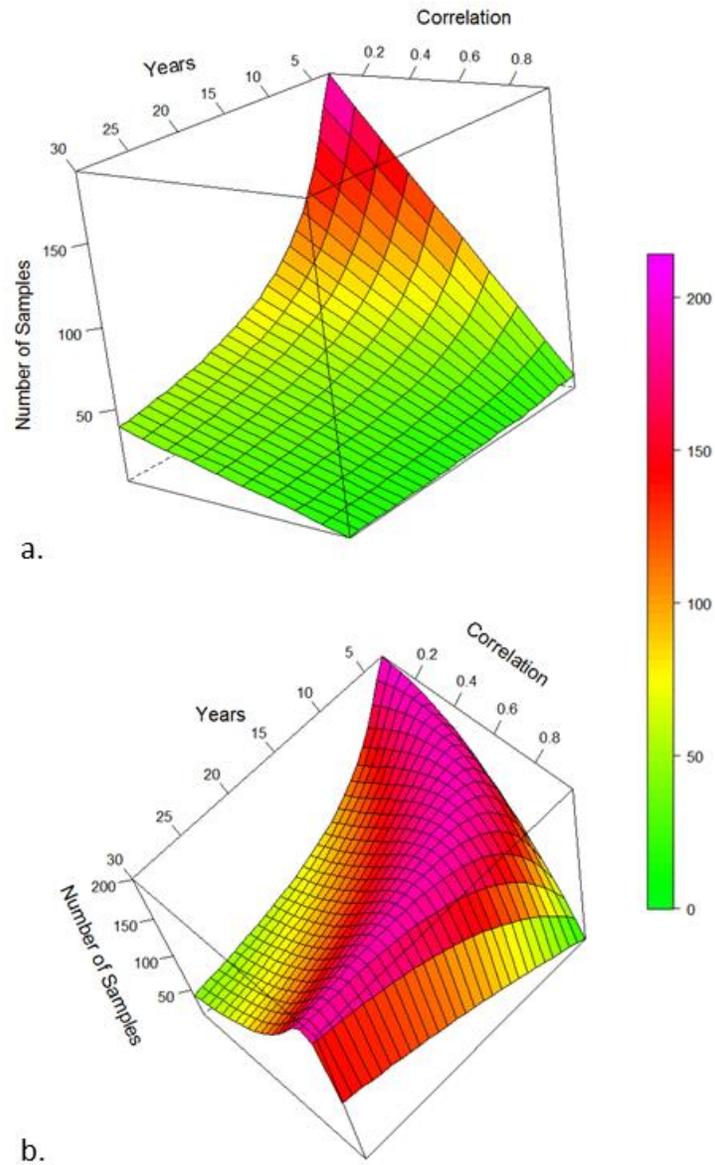


Figure 2.6. Minimum sample size as a function of years and temporal correlation for the compound symmetric correlation structure (a) and the autoregressive structure (b) with the type I error set at 0.1.

Table 2.3. Minimum sample sizes associated with the compound symmetric form of the repeated measures, mixed model for a range of correlation (ρ), population variance (σ^2), and years.

Type I error is fixed at 0.10				Type I error is fixed at 0.05			
$\sigma^2 = 1.00$				$\sigma^2 = 1.00$			
Year	$\rho=0.25$	$\rho=0.50$	$\rho=0.75$	Year	$\rho=0.25$	$\rho=0.50$	$\rho=0.75$
10	40	28	16	10	51	35	20
20	24	17	10	20	30	21	13
30	17	13	8	30	22	16	10
$\sigma^2 = 2.00$				$\sigma^2 = 2.00$			
10	76	52	28	10	97	66	35
20	44	30	17	20	56	39	21
30	30	22	13	30	40	28	16
$\sigma^2 = 3.00$				$\sigma^2 = 3.00$			
10	113	76	40	10	143	97	51
20	64	44	24	20	81	56	30
30	45	31	17	30	57	40	22
$\sigma^2 = 4.00$				$\sigma^2 = 4.00$			
10	149	101	52	10	189	128	66
20	84	57	30	20	107	73	39
30	59	41	22	30	75	51	28

Sample allocation

The distribution of sampling effort - the sample allocation - must balance logistical constraints and science goals. Constraining the sample to dominant landscape characteristics reduces cost and focuses sampling on continental ecology. An allocation that standardizes effort across landscape variability will facilitate comparison within and across sites throughout the observatory (Olsen et al. 1999).

Initial sampling will largely be limited to dominant cover types (greater than 5% coverage of the sampling frame) within each site boundary. This extends the guiding principle that if an ecological response is to be measured, the data must be meaningful in the context of NEON objectives. NEON sites, and the tower-based sensors, were selected to represent dominant vegetation types across the

NEON domains. Plant diversity and other co-located terrestrial measurements will focus on quantifying variability of these types in an effort to better understand relationships between pattern and process at local scales, as well as to contribute to the description of biological patterns at larger scales (Urquhart et al. 1998). The design examined the implications of constraining sampling to cover types greater than both five and ten percent of aerial coverage. Given a fixed sampling effort, there is a trade-off in selecting the level for inclusion of vegetation classes between five and ten percent; sampling vegetation types less than ten percent (but greater than five percent) pulls samples away from the more representative vegetation classes.

Excluding rare vegetation is not without tradeoffs. Disproportionate numbers of species may be endemic to rare vegetation types (Stohlgren et al. 1998), and rare vegetation types might be differentially susceptible to environmental change (Stohlgren 2007, Suding et al. 2008). These rare types, riparian corridors or ecotones for example, may be targeted in iterative sampling efforts or by efforts organized by members of the ecological community.

Landscapes are patchy, and land cover provides one metric to describe that site-scale variability. Increasing the sample size in strata with greater variability standardizes the sampling effort and facilitates comparison. It also increases total sample size, which is costly. Science goals must be balanced by the expense of field-based observations. In the case of plant biodiversity, a design optimized for comparability across cover types and sites would standardize according to measured variance, or better yet, would standardize effort with respect to diversity as indicated by the inflection of the species accumulation curve. In the absence of data from many sites, it was assumed that area can serve as a proxy for variability. Placing a larger number of plots in cover types with a larger area, but relatively more in smaller cover types across the landscape was a means to that end. Samples of other taxa and biogeochemistry were distributed proportionally to the NLCD land cover.

Data analysis with variance estimators

A goal is to collect data according to a design robust to a variety of design estimation and modeling techniques (Sarndal 1978, Cressie et al. 2009). As discussed, design-based inference requires data collected according to a probabilistic design (Reynolds 2012). Various modeling approaches might benefit from the collection of data according to specific stratification, but most can also ingest data based on principles of randomization.

Under the assumption of a stratified-random design, the appropriate design-based estimator (Stevens and Olsen 2004, Thompson 2012) was identified to ensure rigor of the sample design (Lindenmayer and Likens 2009). A spatially-balanced design stratified by vegetation type is equivalent to a stratified-random sample (i.e., within each strata each sample of a given size has an equal probability of selection). Estimators have been developed for the computation of the stratified sample mean and variance when data are collected according to a stratified random sample design (Thompson 2012). The estimator of the sample mean is given by,

$$\bar{y}_{strat} = \frac{1}{N} \sum_{i=1}^S N_i \bar{y}_i \quad [5]$$

where,

- \bar{y}_i = is the sample mean from the i^{th} stratum
- N_i = the number of units in the i^{th} stratum
- N = the number of units across all strata
- S = the number of strata

An unbiased estimator of the variance for this estimator is given by,

$$\widehat{Var}(\bar{y}_{strat}) = \sum_{i=1}^S \left(\frac{N_i}{N}\right)^2 \left(\frac{N_i - n_i}{N_i}\right) \frac{s_i^2}{n_i} \quad [6]$$

where,

- s_i^2 = the sample variance from the i^{th} stratum
- n_i = the number of units in the sample from the i^{th} stratum

The area is computed using the 30m² spatial resolution that corresponds to the NLCD delineation within the footprint of the site. These pixels are considered the sampling units in these calculations. In situations where the sample sizes within strata are sufficiently large (allowing for more comfortable assumption of normality via the central limit theorem), approximate confidence intervals can be formed using the following

$$\bar{y}_{strat} \pm Z_{(\alpha/2)} * \left(\widehat{Var}(\bar{y}_{strat})\right)^{1/2} \quad [7]$$

where,

- $Z_{(\alpha/2)}$ = is the value from normal distribution corresponding to a 100(1- α)% confidence interval.

Few of the sites in the initial implementation will have strata with sufficiently large samples that allow this approximation (Equation 7). For strata with sample sizes smaller than 30, Thompson (1992) suggests using a t-distribution with degrees of freedom approximated using Satterthwaite's method

$$d = \frac{(\sum_{i=1}^S a_i s_i^2)^2}{\left[\sum_{i=1}^S (a_i s_i^2)^2 / (n_i - 1) \right]} \quad [8]$$

and,

$$a_i = N_i(N_i - n_i)/n_i \quad [9]$$

Testing the study design: plant diversity

The described study design has been deployed across the NEON network, directing initial sampling efforts of multiple taxonomic groups and soil at more than thirty sites. The data from the design provided the opportunity to evaluate temporal assumptions and sample size required to enable comparison of trends in plant diversity at sites across the United States. For the purposes of evaluating the design, plant diversity data – available from the NEON portal - from were evaluated from four NEON sites: Harvard Forest, Ordway-Swisher Biological Station, the University of Notre Dame Environmental Research Center, and Oak Ridge National Lab. Two aggregations of data can be evaluated with NEON's plot-based approach to sampling plant diversity: the observed species richness across plots within a site which must be compared with species accumulation curves where sampling effort or coverage are equal (Barnett et al. in prep, Gotelli and Colwell 2001, Chao et. al 2014), or the mean number of species per plot (Stohlgren 2007). To incorporate a test of the design-based estimators, which are not appropriate for lists of species accumulated across plots, this evaluation assessed the ability of the design to detect cross-site differences in trends in the mean plant species richness per plot.

Methods

As described, determining sample size requires specification of parameters that influence the capacity to differentiate trends: spatial and temporal variability. Because the Observatory is young and data were available from a single time step only, the critical component of the design that could be evaluated against the expectation was the observed sample variance (s^2) as an estimate of the population variance (σ^2). Preliminary plot sampling efforts in 2011 at three NEON sites – Harvard Forest ($s^2= 2.02$), Ordway-Swisher Biological Station ($s^2= 4.09$), and Yellowstone National Park ($s^2= 3.12$) - were relied on for initial estimates (Table 2.3). Temporal variation was estimated from the mean correlation of the NDVI values during peak greenness at each site (Figure 2.5, Table 2.4). With these parameters, and assuming a Type I error fixed at 0.10, it was assumed that a sample size of 30 would sufficiently

contribute to test in question after 20 years of the 30 year Observatory; 22 plots would be needed if the requirement was minimally sufficient data after 30 years. With the same Type I error, 57 plots would be required at after 20 years at Ordway-Swisher Biological Station while 41 plots would be sufficient after 30 years of sampling. Given funding constraints and with the preliminary statistical constraint of equal sample size across sites 30 plots were targeted for sampling. Sample locations were defined by implementation of the described design at each site, and the sample was allocated proportional to square-root of the area (Barnett et al. in prep) of all NLCD classes that occupied more than 5% of the site.

Data were collected in the summer of 2014 from four sites: Harvard Forest in Massachusetts, the University of Notre Dame Environmental Research Center in Michigan, and Oak Ridge National Lab in Tennessee. At each sample location, plant species richness and abundance were observed in multi-scale, 400-m² square plots. The incidence and cover of species were recorded in eight 1m² subplots, and the species were documented in eight 10m² and four 100m² subplots (Peet et al. 1998, Barnett et al in prep.). The data considered reflects the total number of species recorded across each 400-m² plot.

Results

At the scale of the site – data from all plots and strata – there was no significant difference in the species richness per plot (Table 2.4). The sample variance (s^2) ranged from 4.10 at Harvard Forest to 8.42 at Ordway-Swisher Biological Station (Table 2.4). When matrixed with estimates of temporal correlation (ρ) within and across years (Figures 2.5 and 2.6, Tables 2.3 and 2.4), it becomes clear that the patterns of spatial variation were larger than anticipated based on preliminary data. For example, at Harvard Forest the sample variance ($s^2 = 4.10$, Table 2.4) coincided with the largest projected population variance (Table 2.3) and the NDVI-estimated temporal correlation ($\rho = 0.48$, Table 2.4, Figure 2.5) indicate that a larger sample size – an estimate of 41-51 plots depending on the Type 1 error – would be required to contribute to a robust differentiation in trends of plant species richness per plot even after

30 years. Similarly, temporal correlation as described by NDVI was expected to be higher (less variability through time) at Oak Ridge National Lab ($\rho= 0.71$) but the spatial variation – larger than was estimated in sample size calculations based on preliminary data – suggested that more plots would be needed at the site to contribute data capable of differentiating differences in trend between sites over 30 years.

Discussion

Ensuring the sample design is capable of differentiating trends in plant diversity will require further testing and design modification. It is possible that continued collection of plant diversity data will indicate that NDVI values, typically considered an indicator of vegetation greenness (Carlson and Ripley 1997, Kerr and Ostrovsky 2003) and not necessarily reflective of plant species richness, demonstrate greater variation in variability than temporal patterns of plant species richness. This is likely to be particularly true in systems dominated by large-stature, long-living plant species where individuals and species are less likely to respond to intra- and inter-annual variations in factors that control vegetation greenness over the nine to ten years these data were collected (Scanlon et al. 2005, Liu et al. 2011). However, if this is not the case, and it is determined that the described test is essential to meeting NEON requirements and high-level science objectives, more plots must be added either by securing additional funds or reallocating resources from other components of the design as part of the design iteration.

Table 2.4. Results from the plot-based plant diversity data collected at four NEON sites that provide critical information about the capacity of the NEON sampling design to compare trends at different sites through time including mean species per plot, the sample variance, and the mean NDVI correlation between two time steps during peak greenness.

Site	Mean species per plot	Sample variance (s^2)	NDVI peak greenness	Mean and standard deviation NDVI correlation (ρ)
Harvard Forest	37	4.10	Jun 13 – Oct 19	0.48 (0.12)
Ordway-Swisher Biological Station	36	8.42	Jan 7 – Dec 9	0.52 (0.04)
University of Notre Dame Environmental Research Center	34	7.93	Jun 13 – Oct 13	0.53 (0.06)
Oak Ridge National Lab	44	6.03	May 15 – Oct 19	0.71 (0.04)

Iterating and optimizing the study design

The first several years of NEON will provide data to inform the design. Those data will test design assumptions, evaluate the ability of the design to detect spatial and temporal trends within and across NEON sites, and direct adjustments to the design (Wikle and Royle 1999).

Prior to optimization, the distribution and number of plots associated with each NEON site may require adjustment as a result of logistic constraints, alterations or advancements of scientific methods and information, and an improved understanding of site-specific population variability. Some of the proposed plot locations may be unavailable for NEON sampling for reasons such as:

- The host institution or landowner may reject the a proposed plot due to ecological concerns (presence of endangered species or other long-term research) or other logistical reasons (road construction).
- Plots may intersect buildings, roads, or other developments or natural features such as rock formations that are not suitable for NEON sampling.
- The location may be inaccessible due to steep slopes or other natural features that pose danger to field technicians.
- The time to travel to remote locations may make the observation too costly. NEON is committed to a design that can allow inference to the target study area, but a design with travel time that exceeds allocated funding may require alterations that reduce the number of locations or alters the sampling frame.
- NLCD calssification error will result in plot locations that do not land in the target vegetation type.

Linking continuous surfaces with ground-based point measurements will provide new ways to measure ecological pattern and trend (Ollinger et al. 2008). Where remote-sensing proxies for ground measurements are robust, or there is a 1:1 comparison between a ground measurement and a

remotely-sensed measurement, the airborne data approximates a complete census of variables of interest at a given point in time (Asner et al. 2008). This information changes the notion of, and in some instances the need for, a ground-based sampling approach. In the case of the many variables that cannot be directly measured with a remote approach (e.g., disease, microbial functional groups, insects, small mammals), the airborne imagery will provide information (e.g., the structure of small mammal habitat) that might direct a reallocation of sampling effort.

NEON is designed to provide data sufficient to understand relationships between forcing drivers of change and ecological response at multiple scales (Schimel et al. 2011). For many processes, NEON will not be able to determine if the study design and associated observations are able to detect the nature of the functional relationships between drivers and ecological response until more is known about trends, temporal variability, and uncertainty associated with measurements (Chao and Thompson 2001, Fuentes et al. 2007). Data collected over the first several years of observations will define the measurement accuracy and precision, and sampling intensity and frequency needed to detect trends (Di Zio et al. 2004). The site-specific study design will likely require alterations to sufficiently inform local-scale allocation.

Conclusion

As a continental-scale observatory, NEON will provide comprehensive data that will allow scientists to address the impacts of change on ecological patterns and processes. Detecting change, or ecological trends, at regional and continental scales requires specific long-term observation at local scales. The sample design provides a scientifically rigorous framework that directs the spatial location of local observations. It is an integral component of the larger NEON strategy which is guided by the assimilation of science questions, guiding principles and requirements, multiple observing platforms with specific protocols, products, analyses, and mechanisms for sharing the results. This sample design is a fundamental component of the ecological observatory.

Specification of a sample design suitable to a long-term, continental-scale ecological observatory faces several general challenges which must subsequently be translated into specific design constraints. The design must be appropriate for sampling multiple taxonomic groups and processes, and also be capable of sampling such that cohesive integration of drivers and response can be achieved. Resulting data will be public and confronted by ecological community with very different methods for addressing untold ecological questions. The sample design must accommodate these different analytical paradigms. Finally, the design must provide sufficient information for the detection and quantification of continental-scale trends in ecological responses. These conditions collectively constrained the development of the site-scale sample design. The design is randomized and stratified by vegetation. Guidelines for minimum sample size, analysis of data, and optimization are considered. These design efforts will provide an unbiased data product that can be assimilated into design and model-based approaches to inference for the efficient detection of trends that are scalable within the context of the NEON design.

REFERENCES

- Albert, C. H., N. G. Yoccoz, T. C. Edwards, C. H. Graham, N. E. Zimmermann, and W. Thuiller. 2010. Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography* 33:1028-1037.
- Asner, G. P., R. F. Hughes, P. M. Vitousek, D. E. Knapp, T. Kennedy-Bowdoin, J. Boardman, R. E. Martin, M. Eastwood, and R. O. Green. 2008. Invasive plants transform the three-dimensional structure of rain forests. *Proceedings of the National Academy of Sciences of the United States of America* 105:4519-4523.
- Bradford, J. B., P. Weishampel, M. L. Smith, R. Kolka, R. A. Birdsey, S. V. Ollinger, and M. G. Ryan. 2010. Carbon pools and fluxes in small temperate forest landscapes: Variability and implications for sampling design. *Forest Ecology and Management* 259:1245-1254.
- Barnett, D., P. Adler, D.C. Peters, B. Enquist, J. Grace, S. Harrison, R. Peet, T. Stohlgren, and M. Vellend. *In review*. The plant diversity sampling design for the National Ecology Observatory Network. *Ecosphere*.
- Cardenas, A. M. and C. Buddle. 2009. Introduced and native ground beetle assemblages (Coleoptera: Carabidae) along a successional gradient in an urban landscape. *Journal of Insect Conservation* 13:151-163.
- Carpenter, S. 2008. Emergence of ecological networks. *Frontiers in Ecology and the Environment* 6:228.
- Chao, C. T. and S. K. Thompson. 2001. Optimal adaptive selection of sampling sites. *Environmetrics* 12:517-538.
- Chapin, F. S., P. A. Matson, and P. M. Vitousek. 2012. *Principles of terrestrial ecosystem ecology*. Springer, New York.

- Clark, J. S., D. M. Bell, M. Kwit, A. Stine, B. Vierra, and K. Zhu. 2012. Individual-scale inference to anticipate climate-change vulnerability of biodiversity. *Philosophical Transactions of the Royal Society B-Biological Sciences* 367:236-246.
- Cochran, W. G. 1977. *Sampling Techniques*. Third edition. John Wiley and Sons, New York.
- Cressie, N., C. A. Calder, J. S. Clark, J. M. V. Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* 19:553-570.
- Cressie, N. and C. K. Wikle. 2011. *Statistics for Spatio-Temporal Data*. Wiley, New York.
- Di Zio, S., L. Fontanella, and L. Ippoliti. 2004. Optimal spatial sampling schemes for environmental surveys. *Environmental and Ecological Statistics* 11:397-414.
- Eisen, L., B. G. Bolling, C. D. Blair, B. J. Beaty, and C. G. Moore. 2008. Mosquito species richness, composition, and abundance along habitat-climate-elevation gradients in the northern Colorado Front Range. *Journal of Medical Entomology* 45:800-811.
- Elmendorf, S.C., K.D. Jones, B.I. Cook, J.M. Diez, C.A.F. Enquist, M.O. Jones, R.A. Hufft, S.J. Mazer, A.J. Miller-Rushing, D. Moore, M. D. Schwartz, and J.F. Weltzin. *in press*. The Plant Phenology sampling design for the National Ecological Observatory Network. *Ecosphere*.
- Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1:330-342.
- Fancy, S. G. and R. E. Bennetts. 2012. Design and analysis of long-term ecological monitoring studies. *in* R. A. Gitzen, M. J. J., C. B. Cooper, and D. S. Licht, editors. YBP Print DDA. Cambridge University Press, Cambridge, UK.
- Fancy, S. G., J. E. Gross, and S. L. Carter. 2009. Monitoring the condition of natural resources in US national parks. *Environmental Monitoring and Assessment* 151:161-174.

- Fortin, M. J., P. Drapeau, and P. Legendre. 1989. Spatial Auto-Correlation and Sampling Design in Plant Ecology. *Vegetatio* 83:209-222.
- Fry, J. A., G. Xian, S. M. Jin, J. A. Dewitz, C. G. Homer, L. M. Yang, C. A. Barnes, N. D. Herold, and J. D. Wickham. 2011. National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering and Remote Sensing* 77:859-864.
- Fuentes, M., A. Chaudhuri, and D. M. Holland. 2007. Bayesian entropy for spatial sampling design of environmental data. *Environmental and Ecological Statistics* 14:323-340.
- Gitzen, R. A. and J. J. Millsbaugh. 2012. Ecological monitoring: the heart of the matter. *in* R. A. Gitzen, M. J. J., A. B. Cooper, and D. S. Licht, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge.
- Hargrove, W. W. and F. M. Hoffman. 1999. Using multivariate clustering to characterize ecoregion borders. *Computing in Science & Engineering* 1:18-25.
- Hargrove, W. W. and F. M. Hoffman. 2004. Potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management* 34:S39-S60.
- Hinckley, E. S., G. B. Bonan, G. J. Bowen, B. P. Colman, P. A. Duffy, C. L. Goodale, B. Z. Houlton, E. Marin-Spiotta, K. Ogle, S. V. Ollinger, E. A. Paul, P. M. Vitousek, K. C. Weathers, D. G. Williams. *in press*. *Ecosphere*.
- Hoekman, D, G. Ball, R. A. Browne, R. Davidson, T. Erwin, J. LaBonte, K. Ober, B. Knisley, J. Lundgren, D. Maddison, W. Moore, J. Niemelä, D. Pearson, J. Spence, K. Will, and T. Work. *in prep*. Design for ground beetle abundance and diversity sampling within the National Ecological Observatory Network. *Ecosphere*.
- Hoekman, D., Y. P. Springer, C. M. Barker, R. Barrera, M. S. Blackmore, W. E. Bradshaw, D. H. Foley, H. S. Ginsberg, M. H. Hayden, C. M. Holzapfel, S. A. Juliano, L. D. Kramer, S. L. LaDeau, T. P. Livdahl, C.

- G. Moore, R. S. Nasi, W. K. Reisen, and H. M. Savage. *in prep*. Design for mosquito abundance, diversity, and phenology sampling within NEON. *Ecosphere*.
- Hoenig, J. M. and D. M. Heisey. 2001. The abuse of power: The pervasive fallacy of power calculations for data analysis. *American Statistician* 55:19-24.
- Hooten, M. B., C. K. Wike, S. L. Sheriff, and J. W. Rushin. 2009. Optimal spatio-temporal hybrid sampling designs for ecological monitoring. *Journal of Vegetation Science* 20:639-649.
- Johnson, D. 2012. Monitoring that matters. *in* R. A. Gitzen, M. J. J., A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies. Cambridge University Press, Cambridge.
- Kao, R. H., C. M. Gibson, R. E. Gallery, C. L. Meier, D. T. Barnett, K. M. Docherty, K. K. Blevins, P. D. Travers, E. Azuaje, Y. P. Springer, K. M. Thibault, V. J. McKenzie, M. Keller, L. F. Alves, E. L. S. Hinckley, J. Parnell, and D. Schimel. 2012. NEON terrestrial field observations: Designing continental-scale, standardized sampling. *Ecosphere* 3(12):115.
- Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and the Environment* 6:282 - 284.
- Knapp, A. K. and M. D. Smith. 2001. Variation among biomes in temporal dynamics of aboveground primary production. *Science* 291:481-484.
- Larsen, D. P., A. R. Olsen, and D. L. Stevens. 2008. Using a master sample to integrate stream monitoring programs. *Journal of Agricultural Biological and Environmental Statistics* 13:243-254.
- Legg, C. J. and L. Nagy. 2006. Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management* 78:194-199.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* 73:1943-1967.

- Li, J., D. W. Hilbert, T. Parker, and S. Williams. 2009. How do species respond to climate change along an elevation gradient? A case study of the grey-headed robin (*Heteromyias albispecularis*). *Global Change Biology* 15:255-267.
- Lindenmayer, D. B. and G. E. Likens. 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. *TRENDS in Ecology & Evolution* 24:482-486.
- Lindenmayer, D. B. and G. E. Likens. 2010. The science and application of ecological monitoring. *Biological Conservation* 143:1317-1328.
- Lindenmayer, D. B., G. E. Likens, C. J. Krebs, and R. J. Hobbs. 2010. Improved probability of detection of ecological "surprises". *Proceedings of the National Academy of Sciences of the United States of America* 107:21957-21962.
- Luo, Y. Q., K. Ogle, C. Tucker, S. F. Fei, C. Gao, S. LaDeau, J. S. Clark, and D. S. Schimel. 2011. Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications* 21:1429-1442.
- McDonald, T. 2012. Spatial sampling designs. *in* R. A. Gitzen, M. J. J., A. B. Cooper, and D. S. Licht, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge, UK ;.
- Meier C.M., H. Muller-Landau, J. Lutz, M. Friedl, M. Mack, R. Birdsey, S. Schnitzer, and T. Fahey. *in prep.* The Plant Biomass, Productivity, and Leaf Area Index Sampling Design for the National Ecological Observatory Network. *Ecosphere*.
- Michaelsen, J., D. S. Schimel, M. A. Friedl, F. W. Davis, and R. C. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science* 5:673-686.
- Nusser, S. M., F. J. Breidt, and W. A. Fuller. 1998. Design and estimation for investigating the dynamics of natural resources. *Ecological Applications* 8:234-245.

- Ollinger, S. V., A. D. Richardson, M. E. Martin, D. Y. Hollinger, S. E. Frolking, P. B. Reich, L. C. Plourde, G. G. Katul, J. W. Munger, R. Oren, M. L. Smithb, K. T. P. U, P. V. Bolstad, B. D. Cook, M. C. Day, T. A. Martin, R. K. Monson, and H. P. Schmid. 2008. Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Sciences of the United States of America* 105:19336-19341.
- Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. Statistical issues for monitoring ecological and natural resources in the United States. *Environmental Monitoring and Assessment* 54:1-45.
- Overton, W. S. and S. V. Stehman. 1996. Desirable design characteristics for long-term monitoring of ecological variables. *Environmental and Ecological Statistics* 3:349-361.
- Reynolds, J. H. 2012. An overview of statistical considerations in long-term monitoring. *in* R. A. Gitzen, M. J. J., A. B. Cooper, and D. S. Licht, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge.
- Sacks, W. J., D. S. Schimel, and R. K. Monson. 2007. Coupling between carbon cycling and climate in a high-elevation, subalpine forest: a model-data fusion analysis. *Oecologia* 151:54-68.
- Sala, O. E., L. A. Gherardi, L. Reichmann, E. Jobbagy, and D. Peters. 2012. Legacies of precipitation fluctuations on primary production: theory and data synthesis. *Philosophical Transactions of the Royal Society B-Biological Sciences* 367:3135-3144.
- Sarndal, C. E. 1978. Design-based and model-based inference in survey sampling. *Scandinavian Journal of Statistics* 5:27-52.
- Schimel, D., M. Keller, S. Berukoff, R. Kao, H. Loescher, H. Powell, T. Kampe, D. Moore, W. Gram, D. Barnett, R. Gallery, C. Gibson, K. Goodman, C. Meier, S. Parker, L. Pitelka, Y. Springer, K. Thibault, and R. Utz. 2011. 2011 science strategy: Enabling continental-scale ecological forecasting.

- Scott, C. T. 1998. Sampling methods for estimating change in forest resources. *Ecological Applications* 8:228-233.
- Searle, S. R. 1971. Topics in variance component estimation. *Biometrics* 27:1-7.
- Smith, D. R., Y. Lei, H. A. Walter, and J. A. Young. 2012. Incorporating predicted species distribution in adaptive and conventional sampling designs. *in* R. A. Gitzen, M. J. J., A. B. Cooper, and D. S. Licht, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge.
- Springer, Y.P., D. Hoekman, P.T.J. Johnson, P. A. Duffy, R. A. Hufft, D.T. Barnett, B.F. Allan, B.R. Amman, C.M. Barker, R. Barrera, C.B. Beard, L. Beati, M. Begon, M.S. Blackmore, W.E. Bradshaw, D. Brisson, C.H. Calisher, J.E. Childs, M.A. Diuk-Wasser, R.J. Douglass, R.J. Eisen, D.H. Foley, J.E. Foley, H.D. Gaff, S.L. Gardner, H.S. Ginsberg, G.E. Glass, S.A. Hamer, M.H. Hayden, B. Hjelle, C.M. Holzapfel, S.A. Juliano, L.D. Kramer, A.J. Kuenzi, S.L. LaDeau, T.P. Livdahl, J.N. Mills, C.G. Moore, S. Morand, R.S. Nasci, N.H. Ogden, R.S. Ostfeld, R.R. Parmenter, J. Piesman, W.K. Reisen, H.M. Savage, D.E. Sonenshine, A. Swei, M.J. Yabsley. *In press*. Continental-scale surveillance of infectious agents: Tick-, mosquito-, and rodent-borne parasite sampling designs for the National Ecological Observatory Network (NEON).
- Stanish, L.F., J.J. Parnell, R. Gallery, K. Docherty, G. King, M. Allen, C. Blackwood, L. Zeglin, L. Kinkel, N. Fierer, D. Nemergut, J. Tiedje, and J. Gilbert. *in prep*. The science design behind soil microbial sampling for the National Ecological Observatory Network. *Ecosphere*.
- Stehman, S. V. 2000. Practical implications of design-based sampling inference for thematic map accuracy assessment. *Remote Sensing of Environment* 72:35-45.
- Stehman, S. V. 2009. Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing* 30:5243-5272.

- Stevens, D. L. 1997. Variable density grid-based sampling designs for continuous spatial populations. *Environmetrics* 8:167-195.
- Stevens, D. L. and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association* 99:262-278.
- Stoddard, J. L., C. T. Driscoll, J. S. Kahl, and J. P. Kellogg. 1998. Can site-specific trends be extrapolated to a region? An acidification example for the northeast. *Ecological Applications* 8:288-299.
- Stohlgren, T. 2007. *Measuring plant diversity, lessons from the field*. Oxford University Press, New York.
- Stohlgren, T. J., M. Lee, K. A. Bull, Y. Otsuki, and C. A. Villa. 1998. Riparian zones as havens for exotic plant species in the central grasslands. *Plant Ecology* 138:113-125.
- Suding, K. N., S. Lavorel, F. S. Chapin, J. H. C. Cornelissen, S. Diaz, E. Garnier, D. Goldberg, D. U. Hooper, S. T. Jackson, and M. L. Navas. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Global Change Biology* 14:1125-1140.
- Sun, J. L., S. P. Oncley, S. P. Burns, B. B. Stephens, D. H. Lenschow, T. Campos, R. K. Monson, D. S. Schimel, W. J. Sacks, S. F. J. De Wekker, C. T. Lai, B. Lamb, D. Ojima, P. Z. Ellsworth, L. S. L. Sternberg, S. R. Zhong, C. Clements, D. J. P. Moore, D. E. Anderson, A. S. Watt, J. Hu, M. Tschudi, S. Aulenbach, E. Allwine, and T. Coons. 2010. A multiscale and multidisciplinary investigation of ecosystem-atmosphere CO₂ exchange over the Rocky Mountains of Colorado. *Bulletin of the American Meteorological Society* 91:209-230.
- Theobald, D. M., D. L. Stevens, D. White, N. S. Urquhart, A. R. Olsen, and J. B. Norman. 2007. Using GIS to generate spatially balanced random survey designs for natural resource applications. *Environmental Management* 40:134-146.

- Thibault, K. M., G. N. Cameron, R. A. McCleery, W. J. McShea, R. J. Rowe, and R. K. Swihart. *in prep.* The sampling design for small mammal abundance and diversity for the National Ecological Observatory Network. *Ecosphere*.
- Thibault, K. M., R.B. Chandler, T. Gardali, A.H. Hurlbert, K.L. Purcell, and S.K. Skagen. *in prep.* The Breeding Landbird Sampling Design for the National Ecological Observatory Network. *Ecosphere*.
- Thompson, S. K. 2012. *Sampling*. Wiley, New Jersey.
- Thorpe, A. S., D. T. Barnett, S. C. Elmendorf, E. S. Hinkley, D. Hoekman, K. D. Jones, K. E. LeVan, C. L. Meier, L. F. Stanish, and K. M. Thibault. 2016. Introduction to the sampling designs of the National Ecological Observatory Network Terrestrial Observation System. *Ecosphere* 7(12): article e01627. DOI: 10.1002/ecs2.1627.
- Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. *Ecological Applications* 8:246-257.
- Vitousek, P. M. 1997. Human domination of Earth's ecosystems. *Science* 278:21-21.
- Wikle, C. K. and J. A. Royle. 1999. Space-time dynamic design of environmental monitoring networks. *Journal of Agricultural Biological and Environmental Statistics* 4:489-507.
- Yi, Q. L. and T. Panzarella. 2002. Estimating sample size for tests on trends across repeated measurements with missing data based on the interaction term in a mixed model. *Controlled Clinical Trials* 23:481-496.

STRATEGIES FOR COMPARING PLANT DIVERSITY IN A NATIONAL NETWORK OF SITES²

Introduction

Plant species richness is a simple metric, applicable to theory (Darwin 1859, MacArthur and Wilson 1967, Hubbell 2001, Maestre et al. 2012) and practical for planning (Ahern 2013) and assessing causes and consequences of global change (Sala et al. 2000, Chapin III et al. 2000, Stohlgren et al. 2011). Species represent an intuitive unit of organization (Gotelli, Nicholas J.; Colwell 2011) defined by reproductive capacity resulting in distinct genetic organisms that interact with the environment in similar ways (Mauer and McGill 2011). While the collective co-occurrence of different species reflects the environmental constraints and process of a particular system (Case 1983, Kraft et al. 2015), environmental constraints on species are unique (Gleason 1939); species divergent historical responses to past climate variability and species-specific functional characteristics that mediate interactions (Diaz and Cabido 1997, Cardinale et al. 2006). As a result, changes in climate or other environmental parameters or processes such as disturbance regimes and land use have the capacity to differentially impact species and patterns of species richness.

As these drivers and constraints change, plant diversity changes (Vitousek et al. 1997, Hooper and Vitousek 1997, Chapin III et al. 2000). Documenting and understanding this change requires a consistent commitment to comparable observations (Stohlgren 2007, Mauer and McGill 2011). Through time observations provide insights to changing relative abundance, species composition, and plant species invasion (Magurran 2011). Measurement across space facilitates insights into species turnover, range shifts, and impacts of invasion (Stohlgren 2007). Integrating observations of plant diversity across space – at the scale of the United States - and time – over decades – with the range of directional forcing factors that persist at these large scales should facilitate the untangling of complex pattern-

² Additional authors: David S. Schimel, Thomas J. Stohlgren, and Paul A. Duffy

process relationships (Peters et al. 2014, Soranno and Schimel 2014, Collins 2016) and enable forecasting of future conditions (Keller et al. 2008).

However, plant diversity is an elusive (May 1988) and statistically complex quantity to measure (Gotelli and Colwell 2001). In most every system few species are common and many species are rare (Rozenweig 1995) and samples represent a downward biased estimate – not all species will be detected – of the true species richness at the scale of interest (Gotelli, Nicholas J.; Colwell 2011). As a result, comparing plant diversity across space and time presents challenges due to:

- Differences in underlying species richness (Chao and Jost 2012)
- Differences in species abundance and richness that may also reflect sampling effort or observation conditions (Gotelli and Colwell 2001, Stohlgren 2007, Mauer and McGill 2011)
- Differences in effective area sampled (Gotelli and Colwell 2001, Stohlgren 2007, Gotelli, Nicholas J.; Colwell 2011)
- The spatial arrangement of the samples (Stohlgren 2007, Gotelli, Nicholas J.; Colwell 2011)
- Differences in shape of the relative abundance distribution (McGill 2011)
- Differences in number of individuals collected or counted that reflect biologically meaningful patters of resource availability or growth conditions (Gotelli and Colwell 2001)

This paper explores methods that circumvent these challenges by incorporating extant understanding of plant species distributions and leveraging techniques for standardized sampling to facilitate robust comparisons of plant diversity data.

Raw estimates of plant species richness can only be compared when all of the species in a defined space have been detected. This level of completeness is rare when sampling plant species diversity in natural ecological systems (Gotelli, Nicholas J.; Colwell 2011). Alternatively, species accumulation curves provide a valid way to compare species richness from sampling two different assemblages (Sanders 1968, Hurlbert 1971, Gotelli and Colwell 2001, Chao and Jost 2012). Species

accumulation curves describe the increase in total number of species recorded (y-axis) as more plots (x-axis) are sampled. The probability of detecting new species with each successive sample depends on the number of species in the assemblage and their relative abundance. Systems more rich and with even abundance distributions result in a steeper curve compared to less rich assemblages characterized by uneven species abundance distributions (Gotelli, Nicholas J.; Colwell 2011). In all cases, the curves decrease monotonically with decreasing slope. The smoothed average represents the statistical expectation of the species accumulation curve, the variation between different orderings describes the variation in the number of species detected at any number of plots conditioned on the particular sample (Gotelli, Nicholas J.; Colwell 2011).

In addition to describing the composition and richness of a site or species assemblage, species accumulation curves allow statistically rigorous comparisons of species richness and diversity. To control for sampling effort comparisons are most frequently made where sample sizes are equal (Colwell and Coddington 1994, Gotelli and Colwell 2001, Chao et al. 2009, Chao and Jost 2012). In cases where a different number of samples were taken from each site or assemblage, the data can be rarefied – moving down the smoothed rarefaction curve – to the smaller sample size. Alternatively data from the site with the smaller sample size can be extrapolated to estimate species richness at a larger sample size (Colwell et al. 2012). However, comparison at a fixed sample size may not sufficiently characterize the richness of a particular assemblage (Chao and Jost 2012), biasing the degree of differences between the richness being considered. Because richness estimates based on fixed numbers compress the ratio of richness, comparison based on sample size are exposed to the risk of drawing conclusion more reflective of the sample size than the diversity characteristics of the community (Chao and Jost 2012, Colwell et al. 2012, Chao et al. 2014a, Hsieh et al. 2016a).

As an alternative, Chao et al. (Chao and Jost 2012, Chao et al. 2014b) demonstrated that species richness can also be compared when two or more sets of samples record the same proportion of the

total number of species represented in each assemblage (Chao and Jost 2012, Chao et al. 2015). This total number of species is unknown but approximated by the number of singletons and doubletons represented in the sample (Chao and Jost 2012, Chao et al. 2014b). 'Coverage' and corresponding plant species richness can be calculated for any sample size. The sampled species richness, expressed as a function of coverage, can be compared when coverage values across samples are equal, and, like comparisons standardized by effort, equivalent coverage and corresponding richness can be estimated by rarefaction and extrapolation as needed (Gotelli and Colwell 2001, Gotelli, Nicholas J.; Colwell 2011). The concept of coverage is not new. Ecologists frequently evaluate sample completeness by the terminal slope of traditional species accumulation curves (Colwell and Coddington 1994). The slope represents the expected increase in species when one plot is added to the sample; a steep slope suggests the presence of many undetected species, while a horizontal asymptote indicates that nearly all species in a constrained assemblage have been detected. Chao et al. (Chao and Jost 2012) prove that sampling such that the slopes of the species accumulation curves are equal is equivalent to a point where coverage is equal and supportive of a statistically robust standard for comparison of plant species richness.

The objective of this effort is to generate a framework that allows for the cost-effective comparison of plant diversity across unique sites and across habitat types within and across sites in the context of an ecological observatory. Sampling with sufficient intensity to reach the inflection point of the species accumulation curve as defined at a particular slope creates a diversity-based standardization that also protects against inefficient oversampling and prevents a small sample size incapable of differentiating assemblages. At small sample sizes, species richness tends to converge at small sample sizes and comparisons tend to reflect numbers of plots more than species richness (Colwell and Coddington 1994), and within assemblage variation can be large (Gotelli and Colwell 2001, Colwell et al. 2012). Conversely, sampling beyond the inflection point of the species accumulation curve results in

inefficient use of resources as few new species are detected with the addition of each new plot, and, particularly as the curve approaches an asymptote, further sampling does not increase the power to detect patterns and trends.

These goals and science considerations combined with logistical constraints for design and deployment of a top-down network of sites capable of delivering comparable plant species richness data. An inflection point only provides rigorous opportunities for comparison when defined by a particular target slope. A target slope of seven was selected as a point between the range of initial steep slopes (20 – 34 in many cases) and a slope of zero that would indicate all species in the assemblage were captured in the sample. At the scale of the site, samples numbers were allocated evenly. Environmental variation across sites may provide a wider diversity of habits to support a more diverse flora (Stein et al. 2014) that could benefit from more plots at some sites, but in the absence of data and to ease programmatic allocation of funds an initial equitable distribution of 30 plots were sampled in each site. Gotelli and Colwell (Gotelli, Nicholas J.; Colwell 2011) suggest a minimum of 20 plots, but recognize the numerous factors that make sites and sample size requirements unique. Within sites, plots were distributed according to a spatially balanced and stratified- random design (Stevens and Olsen 2004, Theobald et al. 2007). The National Land Cover Database (NLCD) served to stratify the sites to efficiently sample site-specific heterogeneity (Thompson 2012). To reach that inflection point at a slope seven on the species accumulation curve, more plots were placed in larger cover types to quantify the higher levels of species richness associated with the heterogeneity associated with more space (Arrhenius 1921, Preston 1962, Rozenweig 1995). Plots were allocated according to proportion of the square-root of the strata area, directing more plots to larger areas but creating a non-linear relationship to account for the disproportionate diversity that often persists in some rare cover types across landscapes (Stohlgren et al. 1997, 2003). Sampling would be best be informed by measured plant species richness

and variance, but in the absence of this data, the design provides a standard that targets the comparison and efficiency that can then be optimized towards a diversity-based sampling intensity.

This effort evaluates the following specific objectives and assumptions of the design:

1. To facilitate comparisons of plant species richness across sites, the 30 samples within each site will result in site-scale species accumulation curves that terminate at a slope of seven.
2. To sample such that cover types within a site are comparable, plots distributed according to the proportion of the square-root of the area of each strata by site result in species accumulation curves that terminate at a slope of seven.

Sampling plant diversity according to this plot allocation at sites across the continental United States as part of the National Ecological Observatory Network (Keller et al. 2008) provides the opportunity to test these assertions. The results will provide direction to further refine the design and improve the ability to compare plant species richness across sites and improve the understanding of the causes and consequences of changes in plant species richness.

Methods

Study sites

Six sites were selected for testing the capacity of the sample size and allocation to generate comparable plant species richness data (Table 3.1). The sites represent a diversity of biomes across the United States and varying degrees of within-site heterogeneity (Table 3.1), support common NLCD classes that can be compared across sites, and the data from early collection efforts is available from the National Ecological Observatory Network data portal (<http://data.neonscience.org/home>).

Table 3.1. Metrics of site characterization, heterogeneity, and sample size at six sites tested species accumulation curve-base comparisons of plant species richness.

Site	Site size (km ²)	Elevation (m)	NLCD types sampled	NLCD size (km ²)	Number of plots
Harvard Forest/Quabbin Reservation, MA	39.8	176 - 330	Deciduous forest	16.2	10
			Mixed forest	7.1	7
			Evergreen forest	11.1	9
			Woody wetlands	2.2	4
Oak Ridge National Lab, TN	67.1	258 - 362	Deciduous forest	43.0	15
			Evergreen forest	5.1	7
			Pasture/hay	2.5	5
Ordway-Swisher Biological Station, FL	36.9	32 - 35	Evergreen forest	11.9	11
			Woody wetlands	6.2	10
			Emergent herbaceous wetlands	3.4	7
Smithsonian Conservation Biology Institute, VA	7.4	238 - 454	Deciduous forest	7.4	18
			Pasture/hay	2.7	12
University of Notre Dame Environmental Research Center, WI	29.5	503 - 520	Deciduous forest	6.3	9
			Deciduous forest	6.3	10
			Mixed forest	2.5	6
			Woody wetlands	13.4	13
Woodworth/Chase Lake National Wildlife Refuge, ND	10.6	5720 - 589	Emergent herbaceous wetlands	1.5	10
			Grassland/herbaceous	6.6	20

Field collection of data

The spatial sampling design that directs the sampling within each site and the method for observing plant species richness at each of these locations are both critical to enabling comparisons of species richness with species accumulation curves (Stohlgren 2007). The emphasis on generating comparable data requires that the spatial sampling design satisfy assumptions of the analysis and the field method be appropriate for the diversity of sites across the network at which data will be collected.

Spatial sampling design. A primary requirement of comparing species accumulation curves is that data be collected randomly. An unbiased sample resulting from spatially balanced and random (Theobald et al. 2007) plot locations (Cochran 1977, Thompson 2012) is the foundation of the sample

design. The design eliminates the potential for bias and meets the assumptions of the species accumulation framework (Chao et al. 2015). The capacity of observations to describe plant species richness depends on the local patterns of heterogeneity and richness as well as logistical and financial constraints that govern sample sizes. Previous studies (Stohlgren 2007) and early input from the ecological community resulted in initial baseline funding for a sample size of 30 plots that will be distributed across each site.

To efficiently sample the landscape at each site and to be sure dominant cover types were adequately sampled by a network also focused on ecosystem ecology (Keller et al. 2008, Schimel et al. 2011), plots were stratified by landscape cover types within each site (Cochran 1977, Johnson 2012). The National Land Cover Database (Fry et al. 2011) provides cover classifications well suited to stratification for a national network, providing a continuous land cover classification across the United States including Puerto Rico, Alaska, and Hawaii. Sampling excluded the rarest NLCD cover types (< 5%) within each site to increase the likelihood of adequately characterizing plant diversity in each type sampled. However, by excluding rare cover types, species and trends associated with a component of native and non-native flora will go undetected (Stohlgren et al. 1999). A biodiversity-specific effort designed to record all cover types and all or nearly all of the species at each site and more rigorously track the species at continental scales would have required many more plots and sites.

Plot design

Plant species richness was recorded with a multi-scale plot design that borrows from techniques pioneered by Whittaker (Smida 1984), shares commonalities with modifications of his initial approach (Stohlgren et al. 1995, Stohlgren 2007), but more closely emulates techniques developed by the Carolina Vegetation Survey (Peet et al. 1998). Plant species will be documented with a 20 x 20m square plot comprised of four 10 x 10m subplots with nested subplots (Figure 3.1). Specifically, the following observations were made:

- The identity of each species according to naming conventions maintained by the US Department of Agriculture, Natural Resources Conservation Service PLANTS database (USDA, NRCS 2016) will be recorded in each subplot – ten 1-m², ten 10-m², four 100-m².
- Estimates of abundance are made with ocular estimates of cover within the 1-m² subplots.

The large, multi-scale plot design is modified from a method Whittaker (Smida 1984) developed with the origination of gradient analysis and ordination techniques (Whittaker 1960). These multiscale plots generate comparable data through time (Damschen et al. 2010) and across multiple US states (Stohlgren et al. 1998a, 1998b, 1999) to further general understanding of patterns and plant species diversity and species-environment relationships.

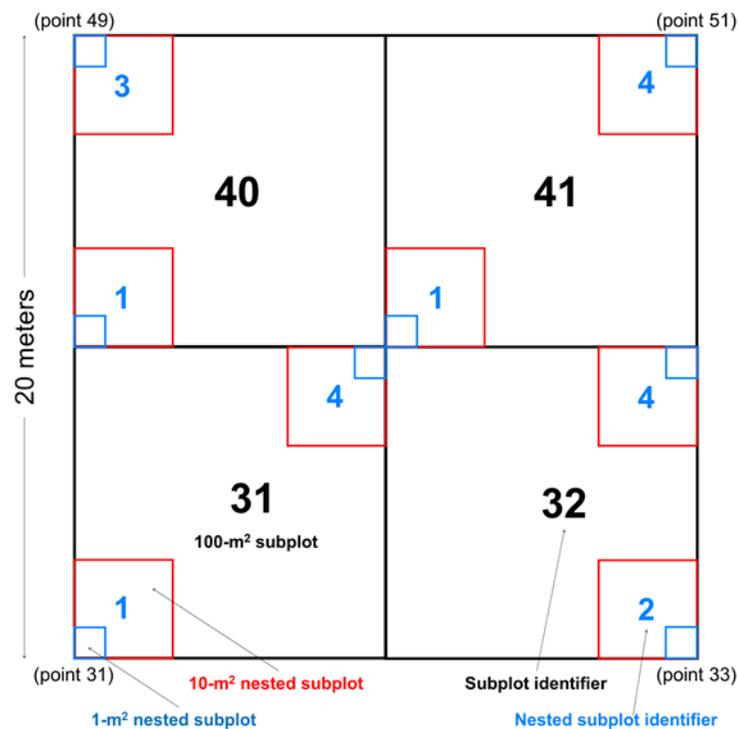


Figure 3.1. The multi-scale plot for recording observations of plant diversity at multiple sites.

Models and Analyses

Plant species richness within and across sites were compared with incidence-based – samples reflect the total number of species in each plot sampling plot or sampling unit - species accumulation curves (Gotelli and Colwell 2001). Species accumulation curves were generated in Program R (R Core Team, 2016) with the iNext package (Chao et al. 2015, Hsieh et al. 2016a). Several assumptions must be met to ensure statistical validity of tests associated with species accumulation curves:

- Samples must be randomly drawn from an assemblage (Colwell and Coddington 1994).
- The same method of observation must be implemented across assemblages (Chao and Jost 2012, Chao et al. 2014a).
- The community is closed (Gotelli and Colwell 2001).
- The total number of species and species abundance distributions must be stationary (Gotelli, Nicholas J.; Colwell 2011).
- Sampling must be with replacement or remain unchanged by removals (Gotelli, Nicholas J.; Colwell 2011).

The sampling design accounted for randomization and plot sampling method was deployed universally across sites. The dynamic nature of natural communities renders assumptions about closed communities and stationary species abundance distributions untenable. However immobile plant species were unlikely to change at time scales required for plot sampling, and these assumptions can be considered reasonable with the understanding that data and comparisons reflect patterns of plant species richness at the time of observation. This plant diversity observational sampling is not destructive and effectively satisfies the ‘with replacement’ assumption.

The order in which plots were observed is immaterial to the structure and patterns of diversity. The resulting species accumulation curves were generated by repeated random ordering that results in a smoothed (Gotelli, Nicholas J.; Colwell 2011, Colwell et al. 2012, Hsieh et al. 2016a) ‘rarefaction’ curves

diagnostic of extant patterns of plant diversity. These successive runs allow 95% confidence intervals associated with each curve capable of statistically rigorous assessments of slope and comparisons of species richness. Estimates of the unconditional variance – the true variation in species richness at any given sample size for the assemblage – is represented by the mean of numerous smoothed species accumulation curves and can be estimated from a single sample (Colwell et al. 2012, Chao et al. 2014a, Hsieh et al. 2016b). This allows for the estimation of variation through the extent of the species accumulation curve as opposed to variance conditioned on a particular sample that must, by definition, reduce to zero at the termination of the curve when all plots are included.

Evaluating the capacity to generate comparable data at a slope of seven was assessed in multiple ways:

- Site-specific species accumulation curves were evaluated against a decrease in the slope to seven.
- To evaluate the within-site NLCD allocation, the prescribed sample size was compared to a confidence interval created by the sample size where the slope of the upper and lower confidence curves were equal to seven.
- To further assess the assumption that larger cover classes would require a larger sample size to reach a slope of seven, the square root of the area was compared to the species richness where the slope of the line was equal to seven.

Cover classes within sites, common cover classes across sites, and sites were compared with species accumulation curves and associated confidence intervals. A multiple linear regression tested the hypothesis that larger, more heterogeneous sites at lower latitudes would be more diverse.

Results

Tests of species accumulation curve slope

The slope of most species accumulation curves decreased to a slope of seven or less. At the scale of sites, where the number of plots was intended to be equal (but was not quite equal, due to logistical challenges, Table 3.2), five of the six (83%) species accumulation curves described by all the plots decreased to a slope of seven or less. Extrapolation of one plot was required to reduce the slope of the curve at Harvard Forest to seven (Table 3.2). Within sites, where plots were allocated according to the proportion of the square root of the area of target NLCD cover classes, 82% of the NLCD species accumulation curves decreased to a slope of seven or less. The species accumulation curve describing the cover types that did not decrease to seven was extrapolated (two plots at emergent herbaceous wetland at Ordway-Swisher Biological Reserve, one plot at the pasture hay at Oak Ridge National Lab, and two plots at the woody wetlands at Harvard Forest) to facilitate comparison of species richness at this targeted slope (Table 3.2).

An assessment of species-accumulation curves terminating at a slope of seven – not over or under sampling – compared the number of sampled plots to the range of plots corresponding to the confidence interval curves were equal to seven. For example, at the woody wetlands cover type at Ordway-Swisher Biological Reserve, the ten plots sampled were significantly greater than the number of plots needed (4 ± 1) to reduce the slope of the species accumulation curve to seven (Figure 3.2). Two (12%) of the NLCD cover classes – the woody wetlands at Harvard Forest and the pasture/hay at Oak Ridge National Lab - were sampled with an insufficient number of plots, the sample size resulted in a slope within the confidence interval at five (30%) NLCD classes, and more plots than were required for a slope of seven were sampled at ten (59%) of the NLCD classes (Figure 3.3a). This oversampling was more pronounced at the site scale where four of six sites (67%) where the 28 to 30 plots sampled were significantly more than required to reach a slope of seven, and the number of plots fell within the range

associated with confidence-intervals that correspond to a slope of seven only at Harvard Forest and Ordway-Swisher Biological Station (Figure 3.3b). These results suggest that sampling intensity at most of the sites (83%) was greater than needed to decrease the slope to seven and the within-site allocation did not successfully result in a slope of seven at most (75%) sites.

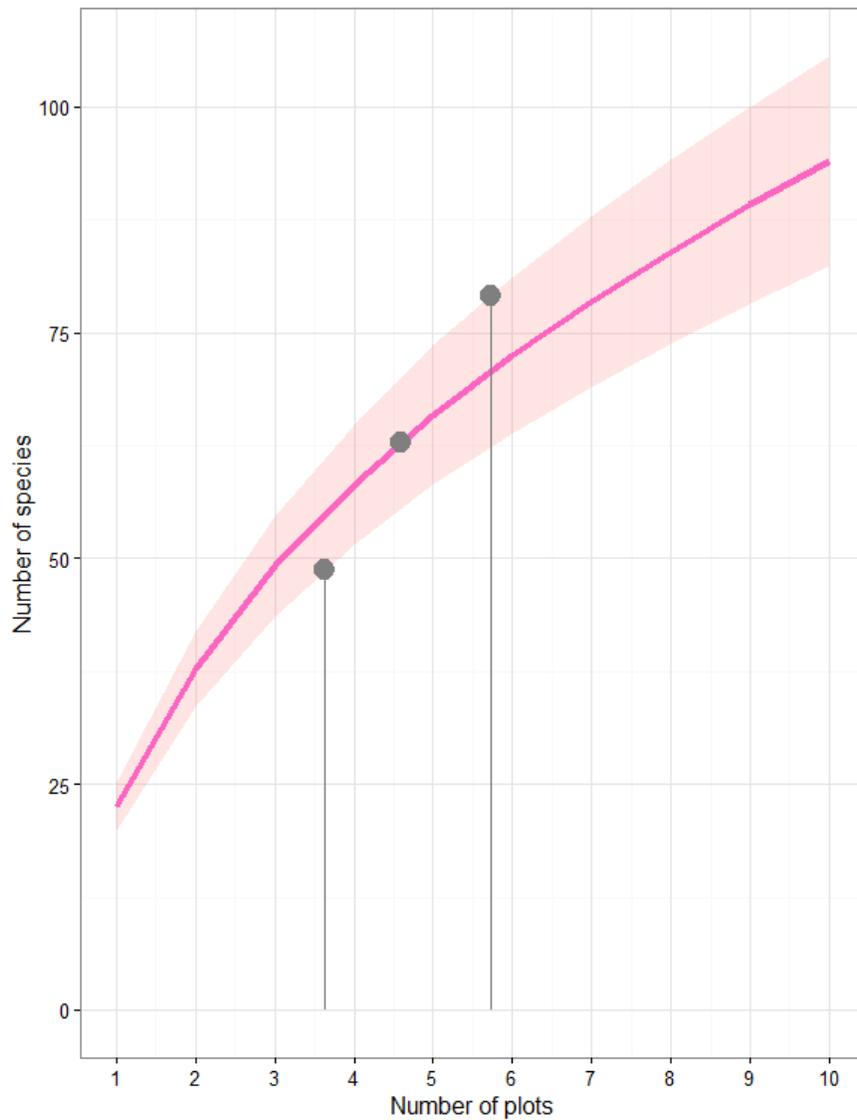
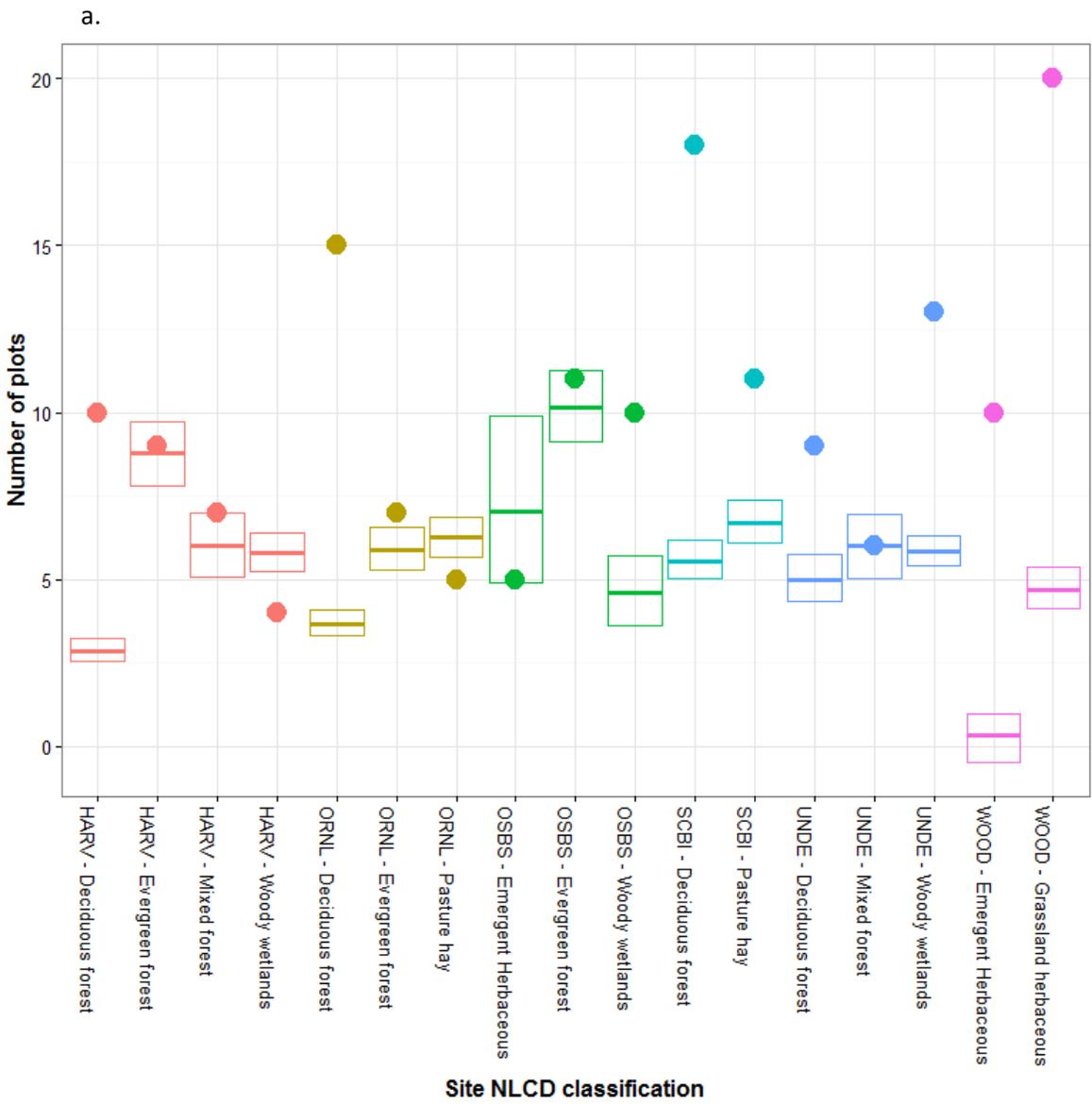


Figure 3.2. A confidence interval associated with the number of plots required for the slope of the species accumulation curve at the woody wetlands cover type at Ordway-Swisher Biological Reserve to decrease to a slope of seven. The confidence interval is described by the number of plots that correspond to a slope of seven for the lower (4 plots) and upper (6 plots) confidence interval curves. Sampling at the Ordway-Swisher Biological Reserve exceeded this confidence interval (ten plots at the terminus of the species accumulation curve).



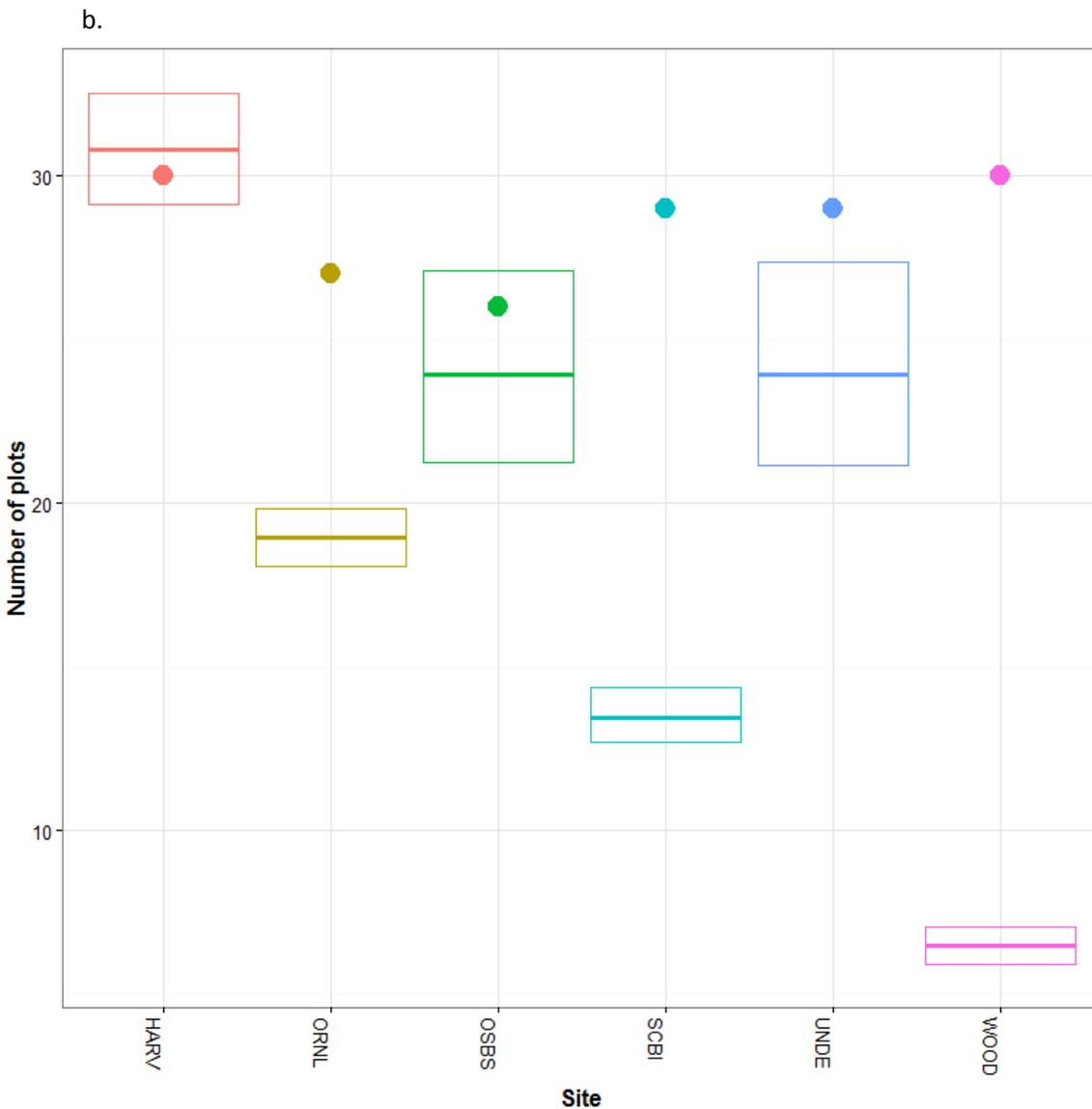


Figure 3.3. The number of plots and the confidence interval at which the slope of the curve was equal to seven at each NLCD class can be described with a boxplot. Overlaying the total number of plots sampled enables a visual portrayal of the statistically rigorous assessment of the prescribed sample size to sample to – but not beyond – a slope of seven at each NLCD class (a.) and within each of the six sites (b.).

The assumption was made that more plots would be required in larger NLCD classes to reach a slope of seven. To better understand why the results were not particularly robust to this allocation, the relationship between the area and the number of species detected when the slope of the species

accumulation curve was equal (for an equitable comparison of effort) to seven was evaluated. A line fitted to the data was neither positive nor significantly different from zero, suggesting that the larger NLCD sampled do not support greater species richness.

Comparisons of species richness

Comparisons of total species richness observed across sites, indicated that Harvard Forest, where extrapolation of one plot was required to reduce the species accumulation curve to a slope of seven, supported the greatest number of species. There was no significant difference between Oak Ridge National Laboratory, Ordway-Swisher Biological Station, and the University of Notre Dame Environmental Research Center, fewer species were observed at the Smithsonian Conservation Biological Institute, and still fewer species observed at Woodworth Station (Figure 3.4).

Comparisons of species richness within a single NLCD class across different sites, and across different NLCD classes within sites were also made with species accumulation curves. Cross-site comparisons of common - present at more than one site – NLCD classes, which required species accumulation extrapolation of one plot in the woody wetland at Harvard Forest, indicated that the woody wetland at Harvard Forest was significantly more species rich than the woody wetland NLCD class at both the University of Notre Dame Environmental Research Center and the Ordway-Swisher Biological Station (Figure 3.5). The deciduous forest at the Smithsonian Conservation Biological Institute supported a larger number of species than the same forest type at the University of Notre Dame Environmental Research Center and Oak Ridge National Lab which were both more species rich than the deciduous forest at Harvard Forest (Figure 3.5). Within sites, the greatest numbers of species were most frequently detected in the NLCD evergreen forest and woody wetland NLCD classes (Figure 3.6).

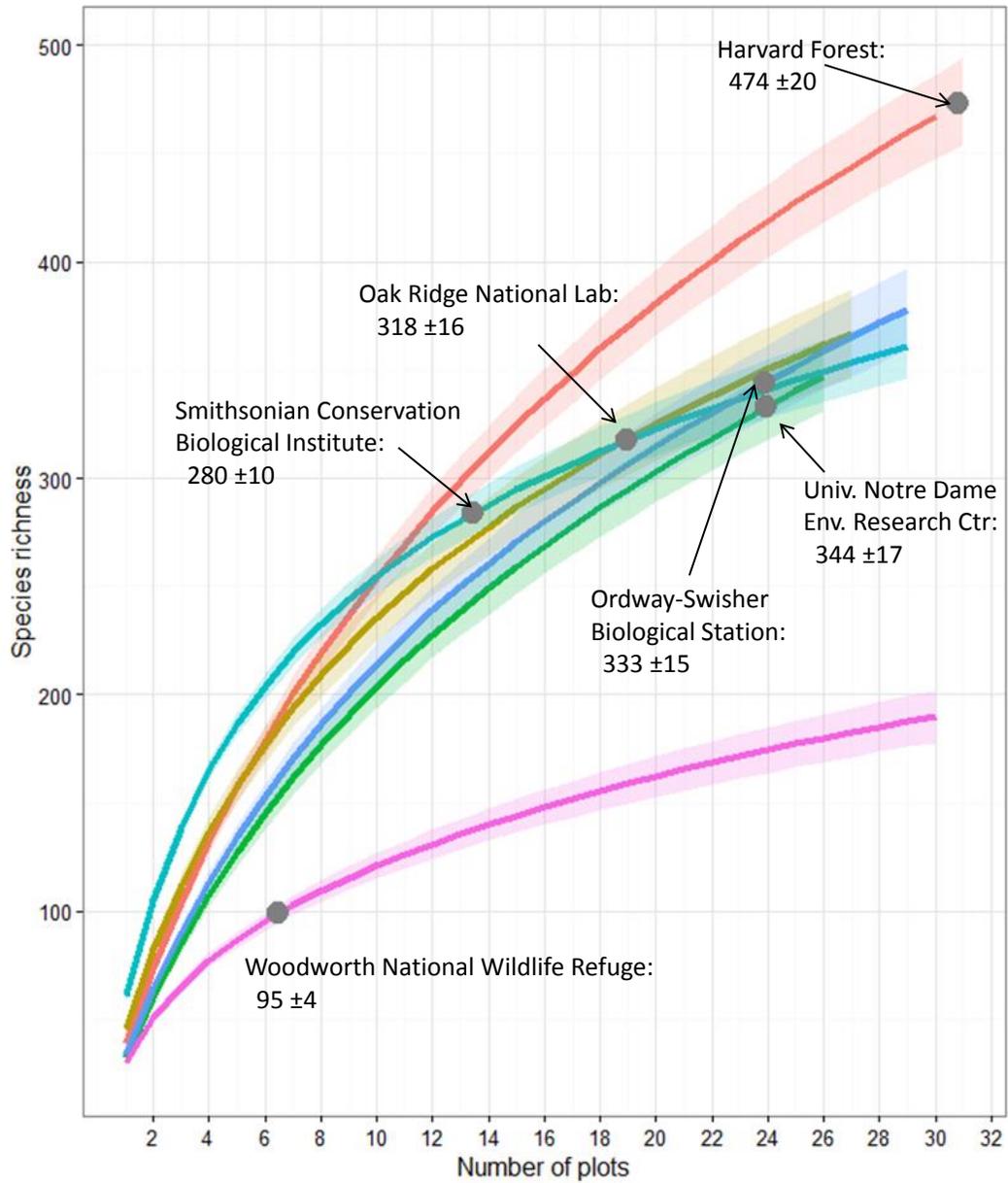


Figure 3.4. Comparisons of species richness across six sites where the slope of the species accumulation curve is equal to seven.

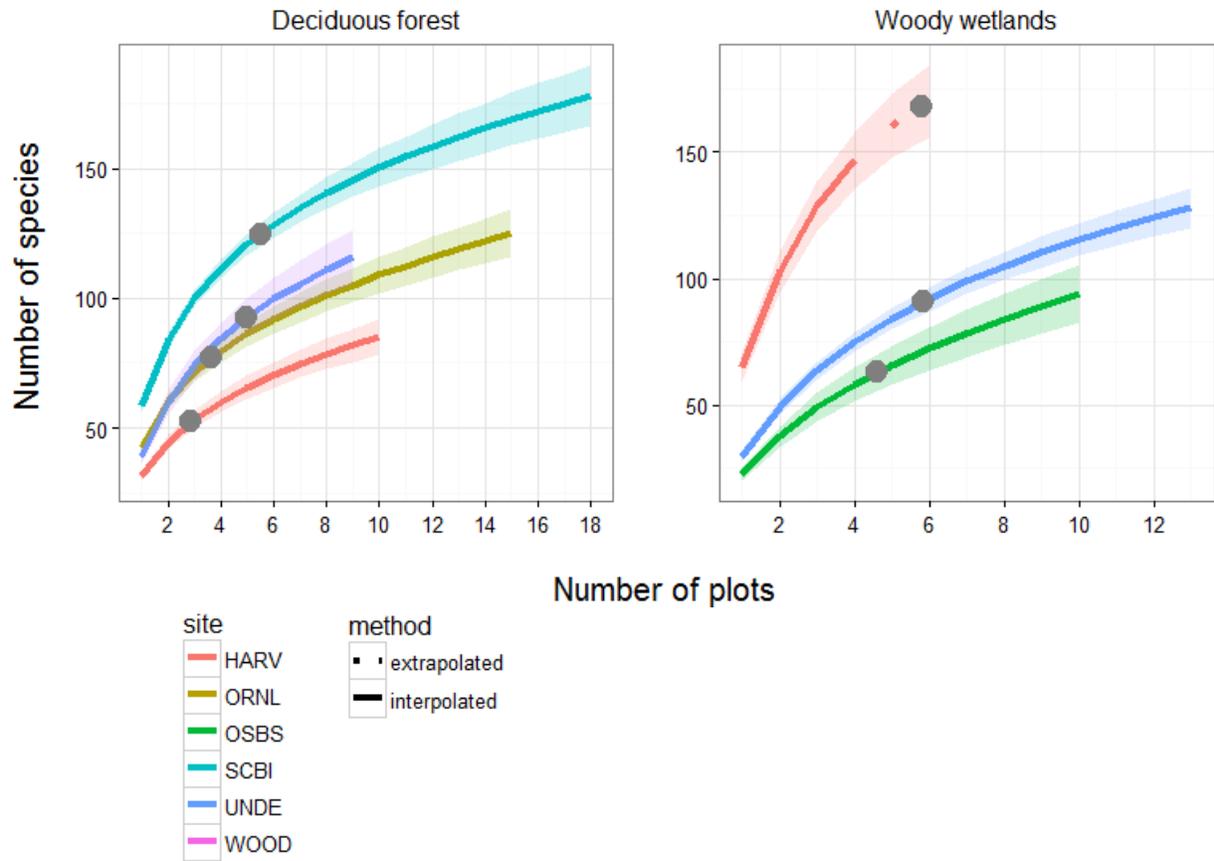


Figure 3.5. Species accumulation curves of two NLCD cover types common to three or more sites.

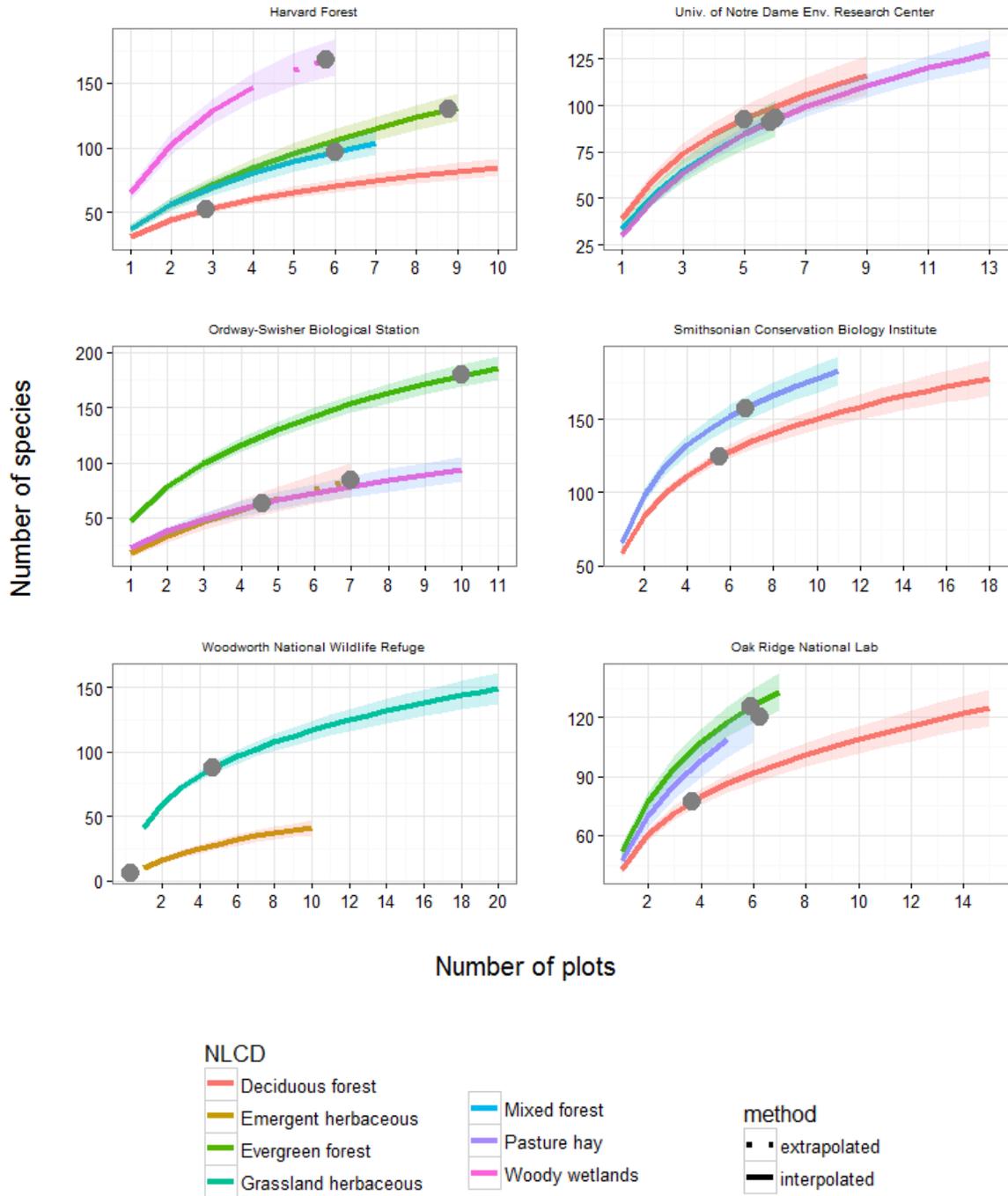


Figure 3.6. Species accumulation curves of NLCD cover classes within targeted sites.

Discussion

The design for comparing spatial and temporal patterns of plant diversity is based on assumptions informed by science and constrained by logistics. Implementing the design at multiple sites

across the United States enabled an opportunity to test assumptions, evaluate patterns of plant diversity, and acquire data capable of informing data-driven iterations for optimizing the network.

Testing assumptions of the design

The design enables a robust comparison at a slope of seven on the species accumulation curve in three ways: 1) by interpolation in cases where the slope was less than seven, 2) where the species-accumulation curve terminated at a slope not significantly different than seven, and 3) in fewer cases, where the terminal slope was greater than seven by extrapolation. That sampling to a slope greater or equal to seven required extrapolation equivalent to two or fewer plots at one site and three NLCD types suggests the prescribed sample allocation adequately characterized species accumulation curves for comparison of plant species richness at a slope of seven. More frequently (five of six sites and ten of 17 NLCD classes, Figures 3.5 and 3.7) sample sizes resulted in slopes less than seven (Table 3.2) and the number of sampled plots exceeded the number of plots that correspond to the upper and lower confidence interval at a slope of seven (Figure 3.3). Additionally, there was no positive and significant relationship between the size of the site and the number of species at a slope of seven or between the size of NLCD class and species richness at a slope of seven within sites (data not shown); there were not more species in larger sites or cover types.

Numerous explanations could account for the initial allocation at and within sites not resulting in an exact slope of seven. The allocation was motivated by general patterns. Larger areas tend to encompass more heterogeneity (Hastings 1990, Collins 1992). Expressed as micro habitats, small-scale disturbance, and gradients in resources and structural complexity, this variation tends to result in greater niche partitioning and species richness in larger areas (Hutchinson 1959, Rozenweig 1995, Stein et al. 2014). Consistent with the data presented here, the generality of the relationship is not absolute, and uncertainty regarding the interactions of heterogeneity, space, and species richness persist. Implementing comparable methods across sites in the Mountain West and Plains of the United States,

Stohlgren (2007) found that rare habitat types tended to support more plant species due to greater resource availability. Factors such as dispersal and available energy (Currie 1991) have also been shown to obscure positive heterogeneity-species richness relationships (Lundholm 2009). Furthermore, the general classifications that make the NLCD amenable to a continental stratification may complicate the assumptions and associated approach to sample allocation. The NLCD types describe broad functional groups (e.g. evergreen forest, deciduous forest, woody wetlands) that, because it is impossible to train algorithms locally due to the broad scale, suffer from registration error (Thogmartin et al. 2004) and may not capture all gradients relevant to understory and subdominant species.

Unsamped components of environmental heterogeneity within NLCD cover classes may compromise the ability to make comparisons at a particular slope or coverage. Coverage-based comparisons (Chao and Jost 2012) of species richness rely on a proxy for estimating the total species richness – the asymptote of the species accumulation curve - in each assemblage compared. Comparisons at a specific slope of species accumulation curves as proposed here don't explicitly rely on estimates of total species richness, but generating these curves through repeated reordering of samples and comparing species at a particular slope is an analogous procedure. The slope of the species accumulation curve represents the expected number of species that would be added with sampling another plot (Gotelli, Nicholas J.; Colwell 2011). A steep slope early in the curve trajectory suggests that the addition of each new plot is likely to result in many new species. At slope of seven the addition of a new plot is likely to add seven new species (Chao et al. 2009, Gotelli, Nicholas J.; Colwell 2011). A species accumulation curve approaching nadir is unlikely to add new species to the sample. However, when heterogeneity or area is large relative to the sample a species accumulation curve will not reach an asymptote (Chao et al. 2016). As samples cross environmental gradients and add new species to the total sampled, the slope of the curve would remain steep until the area sampled approaches the entire realm being sampled (Gotelli, Nicholas J.; Colwell 2011). An asymptotic curve can only be expected

where sample sizes are sufficient to make sampling issues more important than species turnover due to heterogeneous environments and habitat mosaics (Cam et al. 2002, Gotelli, Nicholas J.; Colwell 2011). As a result, asymptotic extrapolation in diverse systems is statistically difficult. The sample must be sufficiently large to be representative of the entire species assemblage and reliably estimate the form of the species accumulation curve, the asymptote, and to rarefy and extrapolate estimates of richness, slope and coverage (Chao and Jost 2012, Chao et al. 2014a, 2016). In the case of this investigation, the number of plots sampled, particularly in the more diverse types, may not be sufficient to adequately describe species accumulation curves associated with the NLCD cover classes given the area, heterogeneity, and species richness. Further sampling may cross unsampled within-type gradients, or encounter micro habitats that result in the detection of new species and result in modifications to the trajectory of the associated species accumulation curve. This sampling problem could be evaluated by adding more plots that target unsampled gradients (Thompson 2012) or avoided by comparing richness across types where the sample size is equal (Gotelli and Colwell 2001, Colwell et al. 2012). Rarefaction enables comparison of plant species richness at the smallest sample size collected at any site or NLCD class (Figures 3.5 and 3.7). Comparisons across sites of equal sample size ($n = 26$) resulted in different conclusions from comparisons at equal slope. Like comparisons at a slope of seven, at 26 plots Harvard Forest and Woodworth National Wildlife Refuge were the most and least species rich sites respectively, but unlike the comparison of equal slope the species accumulation curves of the other sites converged at 26 plots. There was no difference at the equal and larger sample size (Figure 3.4). Sample size is important for adequate characterization of species accumulation curves (Gotelli and Colwell 2001). Ten plots sampled at each site would have been incapable of differentiating Harvard Forest, Oak Ridge National Lab, and Smithsonian Biological Institute that were otherwise apparent at a larger sample size or at equal slope (Figure 4). Both standards for comparison suggest insufficient sampling can result in

mischaracterization of species accumulation curves and demonstrate the need for a standard at which data can be compared.

The sample allocations were put forward as an initial sampling strategy. Guided by logistical constraints and ecological theory, hypotheses about the comparable patterns of species accumulations patterns were put forward with the expectation that some curves would terminate at a slope of seven. Anticipation that some curves might miss a slope of seven accompanied the understanding that sufficient data would be collected to allow comparisons (with reasonable interpolation or extrapolation if needed) and to test and adjust the design towards sampling objectives. The rationale for sampling to an inflection point of the species accumulation curve – adequately parameterizing the species accumulation curve without oversampling - and the persistent need to efficiently acquire information suggests that optimizing the design toward the inflection point while also enabling comparison at an equal sample size with extrapolation and interpolation. In most cases, a species accumulation curve slope of seven could be obtained with fewer plots. For example, rarefaction suggests that a slope of seven would require only five plots in the grassland/herbaceous NLCD type at Woodworth National Wildlife Refuge (Figure 3.6). The sample size could be reduced at most of the cover types at Harvard Forest, but some of the effort – about two plots – would require reallocation to the woody wetland type (Figure 3.6). However, there was a disconnect across scales. The curve for all plots at Harvard Forest required extrapolation to a slope of seven and suggests that sampling should not be reduced. Data from other sites indicated that sample size could be reduced, but there was no consistency in number of plots to be eliminated across scales. This difference between patterns at different scales is prevalent in ecology (Wiens 1989, Levin 1992) and emphasizes the challenge of developing a framework for rigorous comparisons of species richness across time and space (Legg and Nagy 2006, Stohlgren 2007, Gitzen and Millspaugh 2012). Optimization and understanding will require testing and iteration over the first several years of observation (Carlson and Schmiegelow 2002, Kendall and Moore, Clinton 2012).

Comparisons across sites and strata

Species accumulation curves provide a standard for comparison of sites, of NLCD classes across and within sites. The standard – either at a particular slope of the species accumulation curve or at equal effort – handles differences in sampling effort that would otherwise complicate obfuscate conclusions about differences in species richness, and associated understanding of the factors that control patterns of diversity and differentially drive change through time. However, a complete understanding of the factors controlling plant diversity require careful experimentation, and are beyond the scope of this research.

Applications

The influx of non-native plant species (Hejda et al. 2009, Powell et al. 2011), land use change (Newbold et al. 2015), and changing environmental conditions (Tilman and Lehman 2001, Pauli et al. 2012) alter species composition, abundance, and diversity. Statistically rigorous techniques for comparing plant diversity across space and time are needed to accurately detect these changes at a variety of scales (Magurran and McGill 2011). The diversity-based framework for comparing data across space and through time presented here provides a cost-effective framework for ensuring adequate descriptions of species richness and facilitates multiple standards for comparison that result in robust insights from across disparate sites and at large spatial extents.

Ecologists increasingly seek to understand change at large spatial and temporal scales to extract more robust understanding of trends and to provide managers and policy makers with information at space-time scales that is relevant to long-term management of states, nations and continents (Keller et al. 2008, Schimel and Keller 2015). This ‘Big Data’ (Collins 2016) approach to ecology typically requires the combination of disparate datasets that originate from independent investigation, database repositories, and ecological networks (Soranno and Schimel 2014) that require manipulation and amalgamation, and – most of all – they must be comparable (LUO et al. 2011). As described previously,

comparisons of species richness are particularly difficult (Gotelli and Colwell 2001, Magurran 2011) and ecologists might do well to focus comparisons with species accumulation curves, either at equivalent slopes, coverage (Chao and Jost 2012, Chao et al. 2014a), or sampling effort (Gotelli and Colwell 2001). Integrating this comparable plant diversity data with other ecological data streams allows for an understanding of the causes and consequences of change and can enable prediction of future change (Keller et al. 2008, Oliver and Roy 2015). Understanding of relationships enables predictive models that, with the continued collection and assimilation of data can be evaluated and updated through time (Niu et al. 2014, Mouquet et al. 2015) to fine-tune understanding of trends across space and time.

Future directions

Collecting data according to a sampling design provides the opportunity to assess the performance of the design by testing assumptions and understanding extant characteristics of landscapes in question. In this case, where the intent is to repeatedly collect data to understand patterns in time there is an opportunity to further refine the collection effort and address the following questions:

What causes the differences in species accumulation curves at different scales? And, which scale should be the focus of design optimization? Subsetting data and testing different plot allocations within sites could provide insight. Regardless of the drivers of the difference, the question of the scale at which the design should focus likely depends on the question asked of the data.

How will variation through time impact the shape of species accumulation curves? Will a sample size appropriate for sampling to a slope of seven be appropriate in future years? This will simply require time and risk might be buffered by keeping more plots than the design might initially require.

How does a sample size optimized to allow for comparison impact the ability to detect trends in richness through space and time? The ability to detect trends will require depend on specific questions but would be best addressed with a data model framework focused on trend detection that

incorporates estimates of variation in space and time that can be updated with annual data collection efforts.

Some of these questions can be informed by the successive collection of data. Most will require supplemental data and testing to best understand how to collect data for trend detection at scales of the continent in addition to effectively describing changing patterns of plant species richness at time and space scales that are only just becoming possible.

REFERENCES

- Ahern, J. 2013. Urban landscape sustainability and resilience: the promise and challenges of integrating ecology with urban planning and design. *Landscape Ecology* 28:1203–1212.
- Arrhenius, O. 1921. Species and area. *Journal of Ecology* 9:95–99.
- Cam, E., J. Nichols, J. Hines, J. Sauer, R. Aplizar-Jara, and C. Flather. 2002. Disentangling sampling and ecological explanations underlying species-area relationships. *Ecology* 83:1118–1130.
- Cardinale, B. J., D. S. Srivastava, J. Emmett Duffy, J. P. Wright, A. L. Downing, M. Sankaran, and C. Jouseau. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443:989–992.
- Carlson, M., and F. Schmiegelow. 2002. Cost-effective sampling design applied to large-scale monitoring of Boreal birds. *Conservation Ecology* 6:11.
- Case, T. J. 1983. The reptiles: ecology. Pages 159–209 in T. J. Case and M. L. Cody, editors. *Island biogeography in the Sea of Cortez*. University of California Press, Berkeley.
- Chao, A., C.-H. Chiu, and L. Jost. 2014a. Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through Hill Nj numbers. *Annual Review of Ecology, Evolution, and Systematics* 45:297–324.
- Chao, A., C.-H. Chiu, and L. Jost. 2016. Statistical challenges of evaluating diversity patterns across environmental gradients in mega-diverse communities. *Journal of Vegetation Science* 27:437–438.
- Chao, A., R. K. Colwell, C. W. Lin, and N. J. Gotelli. 2009. Sufficient sampling for asymptotic minimum species richness estimators. *Ecology*.

- Chao, A., N. J. Gotelli, T. C. Hsieh, E. L. Sander, K. H. Ma, R. K. Colwell, and A. M. Ellison. 2014b. Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs* 84:45–67.
- Chao, A., T. C. Hsieh, R. L. Chazdon, R. K. Colwell, N. J. Gotelli, and B. D. Inouye. 2015. Unveiling the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology*.
- Chapin III, F. S., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper, S. Lavorel, O. E. Sala, S. E. Hobbie, M. C. Mack, and S. Diaz. 2000. Consequences of changing biodiversity. *Nature* 405:234–242.
- Cochran, W. G. 1977. *Sampling techniques*. John Wiley & Sons, New York.
- Collins, S. L. 1992. Fire frequency and community heterogeneity in tallgrass prairie vegetation. *Ecology* 73:2001–2006.
- Collins, S. L. 2016. Vegetation science in the age of big data. *Journal of Vegetation Science* 27:865–867.
- Colwell, R. K., A. Chao, N. J. Gotelli, S. Y. Lin, C. X. Mao, R. L. Chazdon, and J. T. Longino. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5:3–21.
- Colwell, R. K., and J. A. Coddington. 1994. Estimating Terrestrial Biodiversity through Extrapolation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 345:101–118.
- Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. *The American Midland Naturalist* 137:27–49.

- Damschen, E. I., S. Harrison, and J. B. Grace. 2010. Climate change effects on an endemic-rich edaphic flora: resurveying Robert H. Whittaker's Siskiyou sites (Oregon, USA). *Ecology* 91:3609–3619.
- Darwin, C. 1859. *On the origin of species by means of natural selection*. J. Murray, London.
- Diaz, S., and M. Cabido. 1997. Plant Functional Types and Ecosystem Function in Relation to Global Change. *Journal of Vegetation Science* 8:463–474.
- Fry, J. A., G. Xian, S. M. Jin, J. A. Dewitz, C. G. Homer, L. M. Yang, C. A. Barnes, N. D. Herold, and J. D. Wickham. 2011. National land cover database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing* 77:859–864.
- Gitzen, R. A., and J. J. Millspaugh. 2012. Ecological monitoring: the heart of the matter. Page *in* R. J. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge.
- Gleason, H. A. 1939. The Individualistic Concept of the Plant Association. *The American Midland Naturalist* 21:92–110.
- Gotelli, Nicholas J.; Colwell, R. K. 2011. Estimating species richness. Pages 39–54 *in* B. J. Magurran, Anne E.; McGill, editor. *Biological diversity frontiers in measurement and assessment*. Oxford University Press.
- Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4 Abstract:379–391.
- Hastings, A. 1990. Spatial heterogeneity and ecological models. *Ecology* 71:426–428.
- Hejda, M., P. Pyšek, and V. Jarosik. 2009. Impact of invasive plants on the species richness, diversity and composition of invaded communities. *Journal of Ecology* 97:393–403.
- Hooper, D. U., and P. M. Vitousek. 1997. The Effects of Plant Composition and Diversity on Ecosystem Processes. *Science* 277:1302–1305.
- Hsieh, T. C., K. H. Ma, and A. Chao. 2016a. A Quick Introduction to iNEXT via Examples.

- Hsieh, T. C., K. H. Ma, and A. Chao. 2016b. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* 7:1451–1456.
- Hubbell, S. P. 2001. *The unified theory of biodiversity and biogeography*. Princeton University Press, Princeton.
- Hurlbert, S. H. 1971. The Nonconcept of Species Diversity : A Critique and Alternative Parameters. *Ecology* 52:577–586.
- Hutchinson, G. E. 1959. Why are there so many kinds of animals. *The American Naturalist* 93:145–159.
- Johnson, D. 2012. Monitoring that matters. Page *in* R. A. Gitzen, A. B. Cooper, and D. S. Lict, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge.
- Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. *Frontiers in ecology and the environment* 6:282–284.
- Kendall, W. L., and M. Moore, Clinton. 2012. Maximizing the utility of monitoring to the adaptive management of natural resources. *in* R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Lict, editors. *Design and analysis of long-term ecological monitoring studies*. Cambridge University Press, Cambridge.
- Kraft, N. J. B., P. B. Adler, O. Godoy, E. C. James, S. Fuller, and J. M. Levine. 2015. Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* 29:592–599.
- Legg, C. J., and L. Nagy. 2006. Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management* 78:194–199.
- Levin, S. A. 1992. The problem of scale in ecology: the Robert H. MacArthur Award Lecture. *Ecology* 73:1943–1967.
- Lundholm, J. T. 2009. Plant species diversity and environmental heterogeneity: spatial scale and competing hypotheses. *Journal of Vegetation Science* 20:377–391.

- Luo, Y., J. Melillo, S. Niu, C. Beier, J. S. Clark, A. T. Classen, E. Davidson, J. S. Dukes, R. D. Evans, C. B. Field, C. I. Czimczik, M. Keller, B. A. Kimball, L. M. Kueppers, R. J. Norby, S. L. Pelini, E. Pendall, E. Rastetter, J. Six, M. Smith, M. G. Tjoelker, and M. S. Torn. 2011. Coordinated approaches to quantify long-term ecosystem dynamics in response to global change. *Global Change Biology* 17:843–854.
- MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton University Press, Princeton.
- Maestre, F. T., J. L. Quero, N. J. Gotelli, A. Escudero, V. Ochoa, M. Delgado-Baquerizo, M. García-Gómez, M. A. Bowker, S. Soliveres, C. Escolar, P. García-Palacios, M. Berdugo, E. Valencia, B. Gozalo, A. Gallardo, L. Aguilera, T. Arredondo, J. Blones, B. Boeken, D. Bran, A. A. Conceição, O. Cabrera, M. Chaieb, M. Derak, D. J. Eldridge, C. I. Espinosa, A. Florentino, J. Gaitán, M. G. Gatica, W. Ghiloufi, S. Gómez-González, J. R. Gutiérrez, R. M. Hernández, X. Huang, E. Huber-Sannwald, M. Jankju, M. Miriti, J. Monerris, R. L. Mau, E. Morici, K. Naseri, A. Ospina, V. Polo, A. Prina, E. Pucheta, D. A. Ramírez-Collantes, R. Romão, M. Tighe, C. Torres-Díaz, J. Val, J. P. Veiga, D. Wang, and E. Zaady. 2012. Plant Species Richness and Ecosystem Multifunctionality in Global Drylands. *Science* 335:214–218.
- Magurran, A. E. 2011. Measuring biological diversity in time (and space). Pages 85–93 *in* A. E. Magurran and B. J. McGill, editors. *Biological diversity: frontiers in measurement and assessment*. Oxford University Press, Oxford.
- Magurran, A. E., and B. McGill. 2011. Challenges and opportunities in the measurement and assessment of biological diversity. *in* A. E. Magurran and B. J. McGill, editors. *Biological diversity: frontiers in measurement and assessment*. Oxford University Press, Oxford.

- Mauer, B. A., and B. J. McGill. 2011. Measurement of species diversity. Pages 55–64 *in* A. E. Magurran and B. J. McGill, editors. *Biological diversity: frontiers in measurement and assessment*. Oxford University Press, Oxford.
- May, R. M. 1988. How many species are there on Earth? *Science* 241:1441.
- McGill, B. J. 2011. Species abundance distributions. Pages 105–122 *in* *Biological diversity: frontiers in measurement and assessment*. Oxford University Press, Oxford.
- Mouquet, N., Y. Lagadeuc, V. Devictor, L. Doyen, A. Duputié, D. Eveillard, D. Faure, E. Garnier, O. Gimenez, P. Huneman, F. Jabot, P. Jarne, D. Joly, R. Julliard, S. Kéfi, G. J. Kergoat, S. Lavorel, L. Le Gall, L. Meslin, S. Morand, X. Morin, H. Morlon, G. Pinay, R. Pradel, F. M. Schurr, W. Thuiller, and M. Loreau. 2015. Review: Predictive ecology in a changing world. *Journal of Applied Ecology* 52:1293–1310.
- Newbold, T., L. N. Hudson, S. L. Hill, S. Contu, I. Lysenko, R. A. Senior, L. Borger, D. J. Bennett, A. Choimes, B. Collen, J. Day, A. De Palma, S. Diaz, S. Echeverria-Londono, M. J. Edgar, A. Feldman, M. Garon, M. L. K. Harrison, T. Alhousseini, D. J. Ingram, Y. Itescu, J. Kattge, V. Kemp, L. Kirkpatrick, M. Kleyer, D. L. P. Correia, C. D. Martin, S. Meiri, M. Novosolov, Y. Pan, H. R. P. Phillips, D. W. Purves, A. Robinson, J. Simpson, S. L. Tuck, E. Weiher, H. J. White, R. M. Ewers, G. M. Mace, J. P. W. Scharlemann, and A. Purvis. 2015. Global effects of land use on local terrestrial biodiversity. *Nature* 520:45–50.
- Niu, S., Y. Luo, M. C. Dietze, T. F. Keenan, Z. Shi, J. Li, and F. S. C. III. 2014. The role of data assimilation in predictive ecology. *Ecosphere* 5:1–16.
- Oliver, T. H., and D. B. Roy. 2015. The pitfalls of ecological forecasting. *Biological Journal of the Linnean Society* 115:767–778.
- Pauli, H., M. Gottfried, S. Dullinger, O. Abdaladze, M. Akhalkatsi, J. L. B. Alonso, G. Coldea, J. Dick, B. Erschbamer, R. F. Calzado, D. Ghosn, J. I. Holten, R. Kanka, G. Kazakis, J. Kollár, P. Larsson, P.

- Moiseev, D. Moiseev, U. Molau, J. M. Mesa, L. Nagy, G. Pelino, M. Puccinelli, G. Rossi, A. Stanisci, A. O. Syverhuset, J.-P. Theurillat, M. Tomaselli, P. Unterluggauer, L. Villar, P. Vittoz, and G. Grabherr. 2012. Recent Plant Diversity Changes on Europe's Mountain Summits. *Science* 336:353–355.
- Peet, R. K., T. R. Wentworth, and P. S. White. 1998. A Flexible, Multipurpose Method for Recording Vegetation Composition and Structure. *Castanea* 63:262–274.
- Peters, D. P. C., K. M. Havstad, J. Cushing, C. Tweedie, O. Fuentes, and N. Villanueva-Rosales. 2014. Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. *Ecosphere* 5:1–15.
- Powell, K. I., J. M. Chase, and T. M. Knight. 2011. A synthesis of plant invasion effects on biodiversity across spatial scales. *American Journal of Botany* 98:539–548.
- Preston, F. W. 1962. The canonical distribution of commonness and rarity: Part I. *Ecology* 43:185–215.
- Rozenweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge.
- Sala, O. E., F. Stuart Chapin, III, J. J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L. F. Huenneke, R. B. Jackson, A. Kinzig, R. Leemans, D. M. Lodge, H. A. Mooney, M. Oesterheld, N. L. Poff, M. T. Sykes, B. H. Walker, M. Walker, and D. H. Wall. 2000. Global Biodiversity Scenarios for the Year 2100. *Science* 287:1770-1774.
- Sanders, H. L. 1968. Marine Benthic Diversity : A Comparative Study 102:243–282.
- Schimel, D. S., and M. Keller. 2015. Big questions , big science : meeting the challenges of global ecology:925–934.
- Schimel, D. S., M. Keller, S. Berukoff, B. Kao, H. Loescher, H. Powell, T. Kampe, D. Moore, and W. Gram. 2011. 2011 science strategy: enabling continental-scale ecological forecasting. National Ecological Observatory Network, Boulder.
- Smida, A. 1984. Whittaker's plant diversity sampling method. *Israel Journal of Botany* 33:41–46.

- Soranno, P. A., and D. S. Schimel. 2014. Macrosystems ecology: Big data, big ecology. *Frontiers in Ecology and the Environment* 12:3.
- Stein, A., K. Gerstner, and H. Kreft. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* 14:866–880.
- Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association* 9:262–278.
- Stohlgren, T. J. 2007. *Measuring plant diversity, lessons from the field*. Oxford University Press, New York.
- Stohlgren, T. J., D. T. Barnett, and J. T. Kartesz. 2003. The rich get richer: patterns of plant invasions in the United States. *Frontiers in Ecology and the Environment* 1:11–14.
- Stohlgren, T. J., D. Binkley, G. W. Chong, M. A. Kalkhan, L. D. Schell, K. A. Bull, Y. Otsuki, G. Newman, M. Bashkin, and Y. Son. 1999. Exotic plant species invade hot spots of native plant diversity. *Ecological Monographs* 69:25–46.
- Stohlgren, T. J., K. A. Bull, and Y. Otsuki. 1998a. Comparison of Rangeland Vegetation Sampling Techniques in the Central Grasslands. *Journal of Range Management* 51:164–172.
- Stohlgren, T. J., K. A. Bull, Y. Otsuki, C. A. Villa, and M. Lee. 1998b. Riparian zones as havens for exotic plant species in the central grasslands. *Plant Ecology* 138:113–125.
- Stohlgren, T. J., M. B. Coughenour, G. W. Chong, D. Binkley, M. A. Kalkhan, L. D. Schell, D. J. Buckley, and J. K. Berry. 1997. Landscape analysis of plant diversity. *Landscape Ecology* 12:155–170.
- Stohlgren, T. J., M. B. Falkner, and L. D. Schell. 1995. A Modified-Whittaker nested vegetation sampling method. *Vegetatio* 117:113–121.
- Stohlgren, T. J., P. Pysek, J. Kartesz, M. Nishino, A. Pauchard, M. Winter, J. Pino, D. M. Richardson, J. R. U. Wilson, B. R. Murray, M. L. Phillips, L. Ming-yang, L. Celesti-Grapow, and X. Font. 2011.

- Widespread plant species: natives versus aliens in our changing world. *Biological Invasions* 13:1931–1944.
- Team, R. C. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Theobald, D. M., D. L. Stevens, D. White, N. S. Urquhart, A. R. Olsen, and J. B. Norman. 2007. Using GIS to generate spatially balanced random survey designs for natural resource applications. *Environmental Management* 40:134–146.
- Thogmartin, W. E., A. L. Gallant, M. G. Knutson, T. J. Fox, and M. J. Suárez. 2004. A cautionary tale regarding use of the National Land Cover Dataset. *Wildlife Society Bulletin* 32:970–978.
- Thompson, S. K. 2012. *Sampling*. Wiley, Hoboken.
- Tilman, D. G., and C. Lehman. 2001. Human-caused environmental change: impacts on plant diversity and evolution. *Proceedings of the National Academy of Science* 38:5433–5440.
- Vitousek, P. M., J. D. Aber, R. W. Howarth, G. E. Likens, P. A. Matson, D. W. Schindler, W. H. Schlesinger, and D. G. Tilman. 1997. Human alteration of the global nitrogen cycle: Sources and consequences. *Ecological Applications* 7:737–750.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- Wiens, J. A. 1989. Spatial scaling in ecology. *Functional Ecology* 3:385–397.

Introduction

Understanding global change requires an appreciation of the connectivity of ecological systems and interactions of pattern and process (Peters et al. 2008). Anthropogenic impacts to the biosphere become clearer with new observations and data, especially when integrated with predictive models (Barnosky et al. 2012). Alterations to hydrological and biogeochemical cycles impact global climate trajectories (Marshall et al. 2008). Intensive land-use disturbs systems. Trade and transportation facilitate the spread of harmful invasive species (Hobbs and Huenneke 1992) and disease (Crowl et al. 2008). When combined, these forcing factors can disrupt ecosystem function, processes, and the supported biodiversity. Understanding the causes and consequence of such change is essential for successful stewardship of the environment, and will ultimately be measured against our ability to accurately forecast and prepare for ecological changes to the systems that support life (Pejchar and Mooney 2009).

Historically, a true understanding of cause and effect relationships was limited to small-scale experiments in controlled environments (Collins et al. 1995, Gruner 2017). Experiments at large scales were impractical or overwhelmed by the many drivers of ecological variability. The coordinated collection of data on drivers and long-term observations of ecological patterns will elucidate an understanding of species-environment relationships and possibly advance theory as predictions in space and time are validated with future observations. An iterative process of predictions confronted with field validation over years and decades will iteratively improve spatio-temporal hypotheses in ecology.

The integration of species observations with predictive models can be applied at any scale. To show how specific ecological questions pertinent to local landscapes lead to hypotheses that can be

³ Additional author: Thomas Stohlgren

tested in temporal and spatial contexts, we provide an example with two invasive plant species. The spread of invasive species in Hawaii provides a stark reminder of the connectivity of even remote locations. The impact of invasion on native species and ecosystem processes is significant (Vitousek et al. 1987, Mack and D'Antonio 1998), relevant to the development of ecological theory (MacArthur and Wilson 1967, Mack and Lonsdale 2002), and provides a suite of management and policy challenges.

Local questions and baseline data relevant to invasive plant species management

Management of invasive plant species at local scales, such as the Hakalau Forest National Wildlife Refuge and the Lapahoehoe Experimental Forest Unit on the island of Hawaii, is fraught with questions that could benefit from projections of the influence of environmental change:

- Which invasive species pose the greatest risks to the environment, economy, and human health?
- Where are these high priority species now?
- Where are they likely to spread given the future change in climate, land use, and other anthropogenic drivers?
- Where should resources for monitoring, mitigation, and containment take place (i.e., which species to manage in which areas) in space and time?

These sites are part of the largest remaining native-plant dominated forest in Hawaii, but non-native plant species exist, are a topic of research (DeWalt 2006), and are the subject of long-term monitoring as part of the Hawaii Permanent Plot Network (www.hippnet.hawaii.edu) and efforts to restore native species at Hakalau Forest National Wildlife Refuge. This analysis was informed by an assessment of native and non-native vegetation collected at the Hakalau Forest National Wildlife Refuge in 2007.

Stratified random sampling that integrates vegetation and environmental gradients captures biotic responses to climate change, land use change, and invasive plant species (Stohlgren 2007, Crowl et al. 2008). A vegetation map of Hawaii (Jacobi 1989) provided the strata for a stratified-random design.

We measured vegetation in circular, multi-scale vegetation plots (Barnett et al. 2007) to capture gradients of elevation and soil within each vegetation type. Plant species incidence and cover was observed in three 1-m² subplots and the presence of all species was recorded in the 168-m² area in the 76 plots sampled (Figure 4.1). The plot is comparable to multiple-scale plot designs used in many monitoring efforts (Frayer and Furnival 1999, Stohlgren 2007). The documentation of both native and non-native plant species provides a system to track invasion and interactions between native and non-native species (Stohlgren 2007).

The enormity of the plant species invasion in Hawaii paired with fiscal constraints mandates a triage approach to monitoring and control (Parker et al. 1999, Lawler 2009). Records of invasive species in the plot-based assessment were integrated with models that describe single species distributions. Modeling individual species allows the species-specific response to environmental drivers (Higgins et al. 2001, McMahon et al. 2009). Models provide a chance to better understand the processes associated with observed patterns of vegetation (Kerr et al. 2007, McMahon et al. 2009). Modeling tools and regional coordination can guide priorities for the prevention, early detection, and control of invasive species (Hulme 2003). Organizations like the Hawaii Ecosystem at Risk (HEAR; www.hear.org) and weed management association rank species virulence for specific areas. HEAR ranks both kikuyugrass (*Pennisetum clandestinum*) and common velvetgrass (*Holcus lanatus*) as 'high risk' species (Daehler et al. 2004), and both are of special concern at the Hakalau Forest National Wildlife Refuge. The combination of regional and local concern qualified these species for management and long-term monitoring at both sites and makes these species suitable targets for understanding distributions through spatio-temporal modeling.

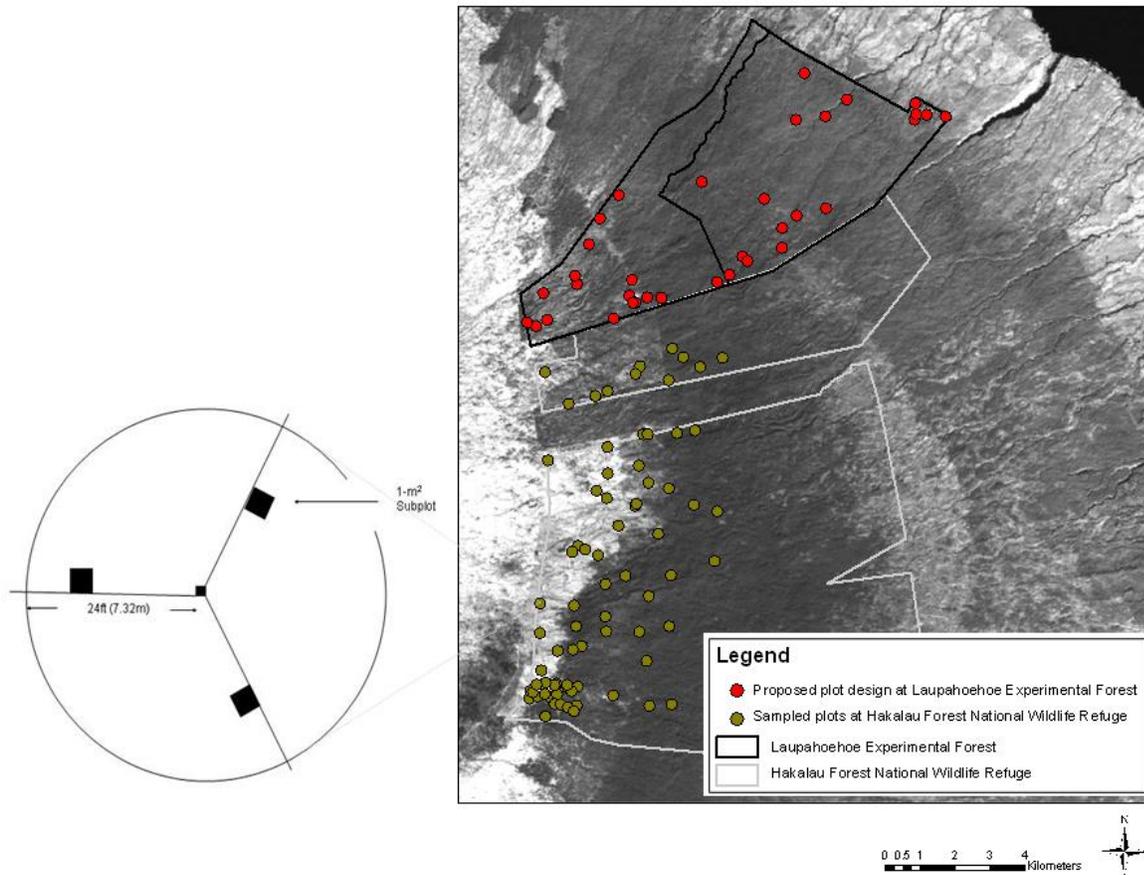


Figure 4.1. Multi-scale vegetation plots were sampled at the Hakalau Forest National Wildlife Refuge. A proposed long-term monitoring design where plant species could be monitored through time was proposed for the Laupahoehoe Experimental Forest, but these plots were not sampled as part of the described study.

An iterative framework for evaluating spatial and temporal hypotheses

Treating spatial and temporal model results of species distributions hypotheses creates a framework for understanding systems and interactions to be tested, verified, and refined. Like small-scale experiments, these hypotheses are best evaluated when compared to independent data (Hijmans 2012). Obtaining that data in this case require time, even decades. Repeated forecasts and data collection that validates models and targets new gradients will refine both the models and sample design to iteratively improve understanding through time. We propose a simple conceptual framework that will iteratively evaluate predictions through space and time to allow proactive management of

species, and furthers the understanding of the drivers of invasion (Figure 4.2). The research questions above can be combined with the data from the Hakalau Forest National Wildlife Refuge and the Laupahoehoe study design to formulate hypotheses.

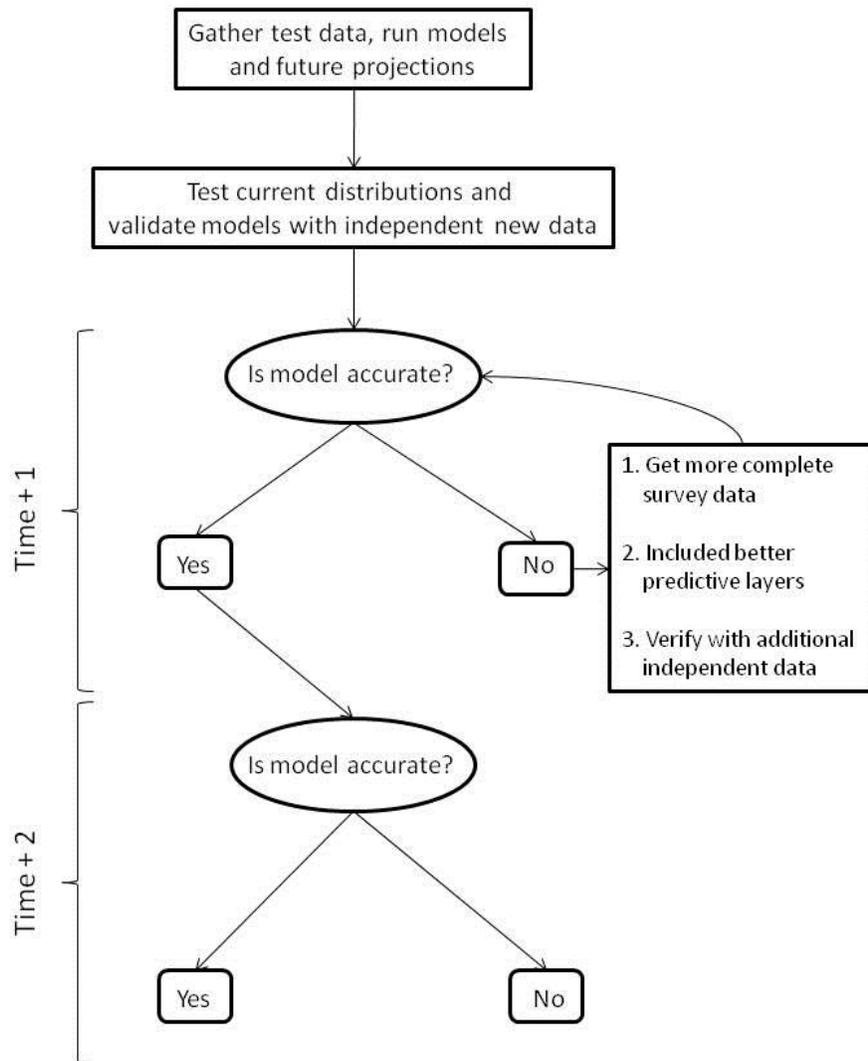


Figure 4.2. An iterative process of testing spatial hypotheses will identify areas of uncertainty for further study thereby refining models that define the interactions of pattern and process over extended time periods. In many cases, the preliminary habitat suitability model may not be as accurate as needed for specific management purposes. Some of the steps needed to address uncertainty are shown. Because predictive layers change over time (e.g., climate, vegetation, disturbance layers, changes in threats), and because species-environment relations may change due to species adaptations and changing interactions among species (e.g., disease, predation), habitat suitability modeling must be an iterative process. New data will always contribute to a better understanding of the system.

Hypothesis 1: Species distributions and potential habitat suitability change predictably in space

Habitat suitability models attempt to determine the environmental conditions in which a species can persist and can be applied to project that relationship across a defined space (Elith et al. 2006, Elith et al. 2009). The accuracy of habitat suitability models depends on the number and distribution of observed species locations, the completeness of the species surveys, and the resolution and relatedness of predictor layers in the models (e.g., climate, geology and soils, vegetation, and remote sensing layers; Elith and Leathwick 2009). Once a habitat suitability map is generated, it can be validated by withholding some of the occurrence data for testing, or with new surveys (Hijmans 2012). These models do not explicitly reflect intricacies of species interactions, dispersal, and other community assembly filters, but they do provide a baseline approximation of the potential range and distribution of the species (Lawler 2009). In this example, we modeled the two plant species' (*Pennisetum clandestinum* observed in 50 plots, and *Holcus lanatus* detected in 28 plot) distributions with Maxent (Phillips et al. 2006) which compared favorably to many other techniques (Elith et al. 2006, Ortega-Huerta and Peterson 2008). Maxent employs a machine learning method based on principals of maximum entropy to probabilistically describe locations (cells) with conditions conducive to species occurrence based on individual environmental variables. Environmental variables included 19 bioclimatic variables (Hijmans et al. 2005) and five topographic characterizations and (Table 4.1). We acquired climate data for the 2007 model from Daymet (www.daymet.org) and future climate predictions from the Canadian Centre for Climate Modeling and Analysis (CCCMA) A2a predictions (www.worldclim.org). Both of the climate scenarios were scaled from 1-km resolution to 30 meters (Wood et al. 2004) using a bilinear transformation (Environmental Systems Research Institute (ESRI), Redlands, CA). Despite the coarse resolution of the climate projections, the species information intersected a diversity of cells (kikuyugrass with 49 locations in 28 different 1-km² grid cells, and velvet grass with 28 locations in 18 different 1-km² grid cells). Cross-correlated environmental variables were removed ($r > 0.8$) and those layers retained

were selected based on those likely to have biological relevance to the controlling plant species (Table 1). We report the Maxent test area under the curve (AUC) values generated from 25 iterations with 15% the data withheld for training and testing purposes.

Measures of climate provide a reasonable starting point for evaluating current and future patterns of plant species distribution because climate provides the context within which other drivers exist (Carpenter 2008). These models suggest that climate and physical landscape characteristics may control current and future species distributions (Table 4.1). Climate and other broad-scale drivers (e.g. topography) interact with processes and patterns at finer scales that tend to influence pattern at finer scales than that of this study (Peters et al. 2008). Measures of topography seemed to have minimal influence in this study (Table 4.1), which may be a result of the high degree of correlation (or a direct mathematical dependency depending on how the climate data were interpolated) with other predictor variables such as temperature and precipitation. Furthermore, model predictions can be highly dependent on the variables selected (Dormann et al. 2008). Evaluation of the models with existing data indicates that they are not particularly robust (Figures 4.2 and 4.3); Swets (1988) suggests that AUC values ranging from 0.7-0.9 are characteristic of models useful for some purposes. Generalist species (Evangelista et al. 2008) and the inclusion of coarse-scaled environmental data (Wood et al. 2004) likely challenged the accuracy of both models (Evangelista et al. 2008) and result in a tentative acceptance of the hypothesis on how the predictor layers relate to the species distributions in time and space. These models should be viewed as first approximations (Stohlgren and Schnase 2006) that could be improved with the integration of fine-scale remote sensing and climate data. True model validation will be possible with the implementation of the collection of NEON organism data at the Laupahoehoe Experimental Forest.

Validation evaluates models, and, especially as part of an iterative process, improves anticipatory capabilities (Currie et al. 2004, Morin and Thuiller 2009). Many models incorporate testing

and validation, but true testing of generality requires evaluation against new data, at new places, and at new time steps (Thuiller et al. 2005, Heikkinen et al. 2006). The importance of unique site characteristics (land-use, management, disturbance, other invasive species, different environmental extremes) challenges the applicability of such efforts (Araujo et al. 2005), and comparison of our projections to independent observation from new studies or from ongoing monitoring by at Hakalau Forest National Wildlife Refuge and the Laupahoehoe Experimental Forest. Models become a learning tool in the context of long-term observation (Clark 2001, McMahon et al. 2009). The disparity between prediction and observation will highlight uncertainty that can be addressed with alterations to study design and targeted observations, the acquisition of improved environmental metrics, and experiments aimed at a better understanding of the underlying ecological processes. Integrating this new information into subsequent model runs and iteratively comparing drivers and spatial description to the next observation provides opportunities to better understand interactions of invasion, environment, and climate.

Hypothesis 2: Species distributions and potential habitat suitability change predictably in time

Species distributions respond to climate, land-use, disturbance, soil, and landcover and are further mediated by interactions with other species (Chapin et al. 2011). As these forcing factors undergo change, there is increasing evidence species distributions and diversity also changing in response (Parmesan 2006, Lenoir et al. 2008, Bremer and Farley 2010, Chen et al. 2011). Suitable habitat shifts as climate changes. Climate model projections suggest that climate in the part of Hawaii occupied by these study sites is likely to change in future decades, and the distribution of suitable habitat for kikuyugrass and common velvetgrass will likely change as well. This temporal hypothesis – based on extant species-environment relationships - suggests that the prevalence of kikuyugrass across this landscape is likely decrease with reductions in suitable habitat (Figure 4.3). With time the leading edge of suitable habitat for common velvetgrass, however, will extend through the area with time, increasing the probability or increases in velvetgrass throughout the study site (Figure 4.4).

Table 4.1. Environmental and topographic variables and the percent contribution to the model of each species on the island of Hawaii, HI.

Environmental variable	Percent contribution	
	Kikuyugrass	Common velvetgrass
Minimum temperature of coldest month	39.9	a
Precipitation of wettest month	20.5	a
Maximum temperature of warmest month	15.3	a
Mean annual precipitation	5.5	a
Precipitation seasonality	5.5	a
Distance from stream	4.7	3.5
Aspect, degrees from north	4.6	1.9
Aspect, degrees from east	2.0	3.6
Isothermality	1.5	
Annual mean temperature	0.2	2.0
Slope	0.2	1.3
Precipitation of warmest quarter	a	46.9
Precipitation of driest month	a	29.3
Mean diurnal range in temperature	a	8.8
Temperature annual range	a	2.9
Mean temperature of warmest quarter	a	a
Mean temperature of coldest quarter	a	a
Precipitation of wettest quarter	a	a
Precipitation of driest quarter	a	a
Precipitation of coldest quarter	a	a
Temperature seasonality	a	a
Mean temperature of wettest quarter	a	a
Mean temperature of driest quarter	a	a
Elevation	a	a

^a Cross-correlated variable removed from model.

As temporal hypotheses these models cannot be formally accepted and should be continually challenged, but they do imply we may anticipate different futures for each of these species. The iterative hypothesis process supports real-world decisions on how to manage for change. In the context of limited budgets and numerous priorities, anticipating invasion with a data-driven approach allows a cost-effective for prioritizing monitoring and control (McGeoch et al. 2015). Priority should focus on those species poised to impact native species and challenge management goals (Blossey 1999, Ficetola

et al. 2007) such as restoring native forest cover and increasing habitat for endangered fauna (Camp et al. 2010). Managers could prioritize strategies to control and monitor common velvetgrass, and reduce efforts to contain kikuyugrass that is likely to be locally challenged by reductions in suitable habitat.

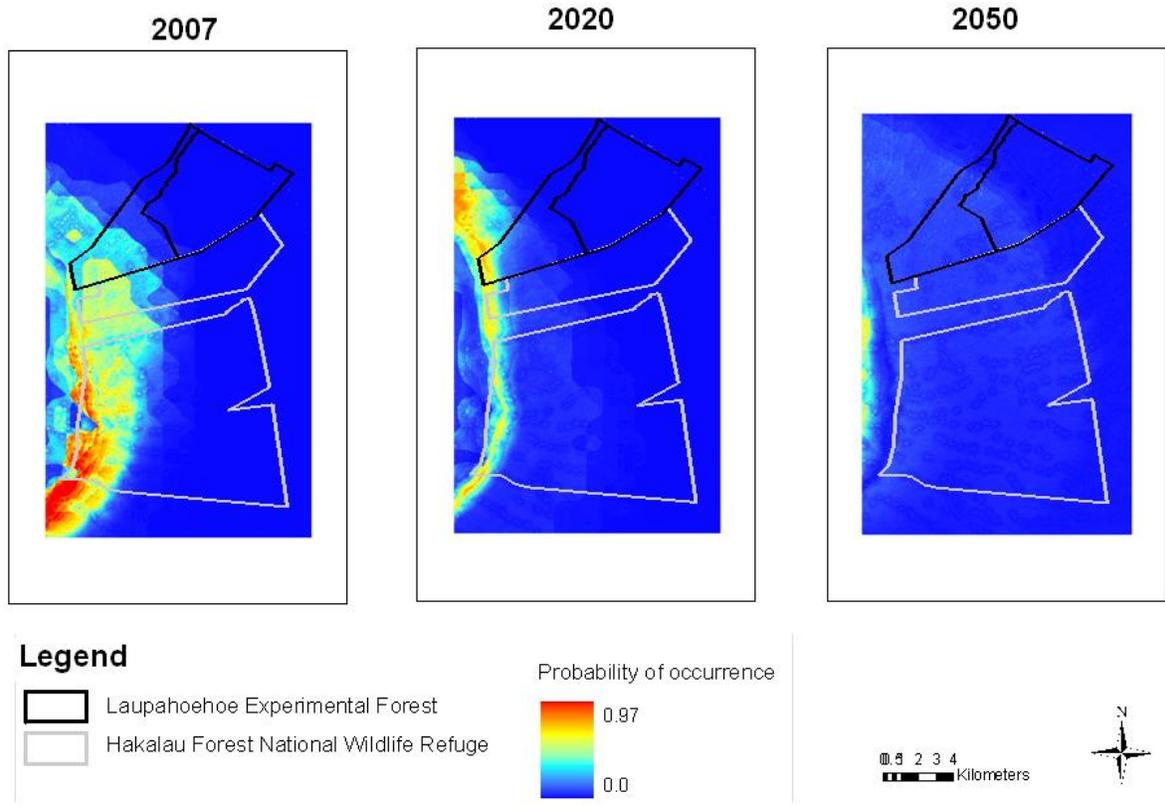


Figure 4.3. The estimated current and anticipated future distribution of kikuyugrass at Hakalau Forest National Wildlife Refuge and the Laupahoehoe Experimental Forest.

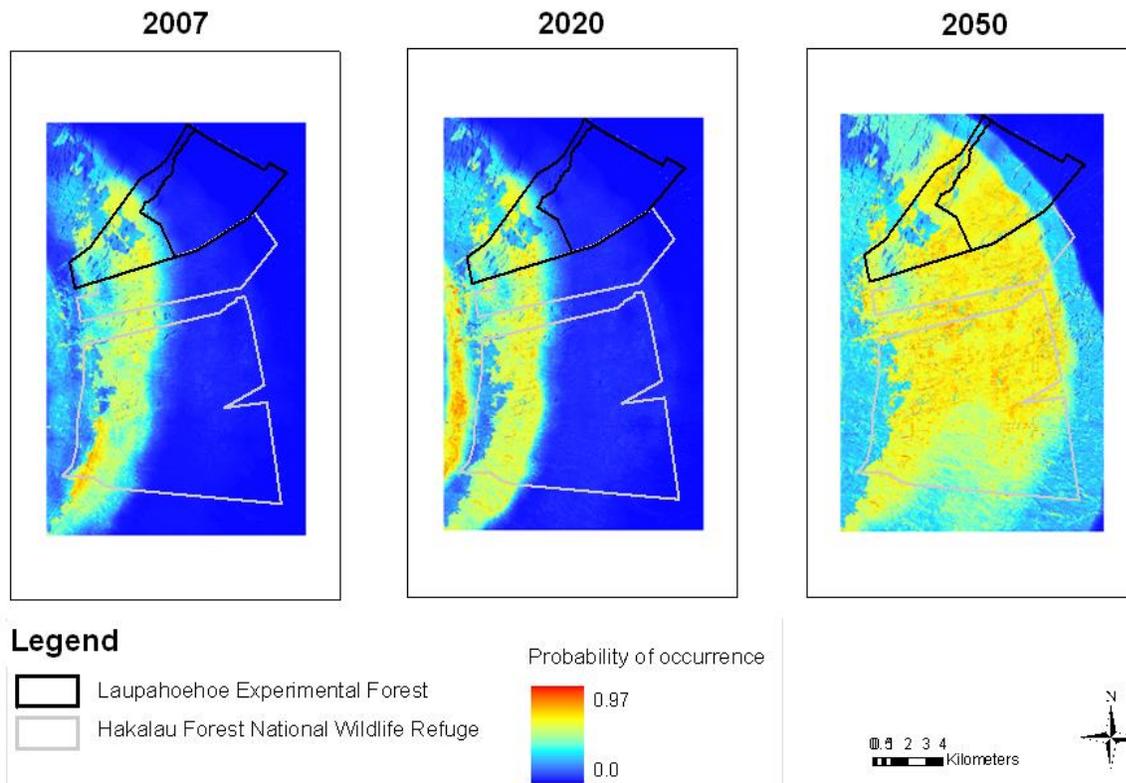


Figure 4.4. The estimated current and anticipated future distribution of common velvetgrass at Hakalau Forest National Wildlife Refuge and at the Laupahoehoe Experimental Forest.

Hypothesis 3: Spatial and temporal trends in invasion are best measured with a sampling design that captures biotic and abiotic gradients

The study design is responsible for directing observation that results in an unbiased description of current and future conditions. A definitive evaluation of this hypothesis depends on future collection of the complete suite of abiotic and biotic data elements. However, these projections of species distributions provide a chance to test the ability of a proposed long-term plot design to capture current and future biological heterogeneity at the Laupahoehoe Experimental Forest (Figure 4.1). Sample sites must be positioned to capture the diversity of distribution characteristics to accurately calibrate the species-environment relationships if drivers of invasion are to be understood (Ibanez et al. 2009). The study site is dominated by habitat of medium or marginal suitability for common velvetgrass, and these

areas should be sufficiently represented by the study design (Table 4.2). Even those small parts of the site that are projected to have a greater likelihood of supporting the species seem to be well represented by the study design (Table 4.2). While more data and further analyses are needed, at least for common velvetgrass, we can currently accept the hypothesis that this design captures the abiotic and biotic gradients responsible for the distribution of this species across space and time.

Table 4.2. The number of plots in a proposed long-term monitoring design in classes of probabilistically described habitat suitability for common velvetgrass (*Holcus lanatus*) at the Laupahoehoe Experimental Forest.

Probability of occurrence	2007		2020		2050	
	% area	Number of sample sites	% area	Number of sample sites	% area	Number of sample sites
0 - 0.2	88	27	72	19	41	10
0.2 - 0.4	9	5	13	4	27	13
0.4 - 0.6	2	3	13	9	32	11
0.6 - 0.8	< 0.01	0	2	3	0.41	1
0.8 - 1.0	0	0	< 0.01	0	0	0

If projections of invasion change, new sample sites can be added to provide monitoring infrastructure that will capture the front edge of invasion (Regniere et al. 2009). Preemptive sampling in projected hotspots of invasion provides early detection and catalogues conditions prior to invasion to better describe interspecific interactions after the arrival of target species (Lee et al. 2008). Budget constraints will limit the number of sample sites dedicated to capturing the leading edge of common velvetgrass dispersal. In addition to plots that measure all species, further documentation and validation of common velvetgrass can be achieved with rapid assessment mapping techniques that provide a cost-effective and efficient way to increase sample size and improve model predictions (Barnett et al. 2007).

Caveats

Habitat suitability models have drawbacks (Jarnevich et al. 2015). Unlike mechanistic models they do not account for specific factors that likely contribute to patterns of diversity such as interspecific interactions, dispersal, and propagule pressure (Davis et al. 1998, Kearney and Porter 2009). Process

models that include these variables may more quickly provide information about the impact drivers have on patterns of diversity, but they are difficult to parameterize, especially at large scales with many interacting species and processes (Kerr et al. 2007, Lawler et al. 2009, Morin and Thuiller 2009). Simple models can often capture the essence of a system in ways that proves to be elusive to models that rely heavily on assumptions and estimations (McMahon et al. 2009).

Unforeseen changes and deficiencies in data may compromise predictions of future states:

- Climate may not change as predicted in the climate models used in this analysis (Millar et al. 2007).
- Coarse-scale climate models, even when downscaled with precision, may not be sufficient to model fine-scale species distributions (Jarnevich and Stohlgren 2008, Ashcroft et al. 2009, Jarnevich and Stohlgren 2009).
- Disparate sources of climate data might add uncertainty to predicted distributions. The models were parameterized, in part, with historic climate conditions from Daymet (daymet.ornl.gov) but forecasts were derived with climate projections from the Canadian Center of Climate. The different methods, assumptions and source in the climate data could explain some of the observed patterns or uncertainty in the model that was not calculated (Mika et al. 2008).
- Species assemblages may decouple. Groups of species that typically interact may react differently to a changing climate, and to the degree that interspecific interactions control diversity, future patterns of diversity may also be more different than a correlative habitat suitability model can anticipate (Pearman et al. 2008).
- Species may encounter a variety of thresholds that change dispersal and establishment at future suitable sites or affinity for habitat where they currently exist (Pearman et al. 2008).

We may not have captured the existing environmental envelope. The Hakalau Forest National Wildlife Refuge study may have failed to capture individuals that would have changed the perceived

species-environment relationship which would compromise our capacity to predict the extent of future distributions (Thuiller et al. 2006, Ibanez et al. 2009). Our inventory captured a snapshot in time of two invasive species likely in the midst of invading all suitable sites or niches.

Additionally, there are a variety of reasons kikuyugrass may not recede as rapidly nor as completely as described by the temporal hypothesis (Figure 4.4). Areas that currently support the species but are probabilistically described by the this hypothesis as less likely to be of suitable habitat in the future, are referred to as the receding edge, or extinction zones (Thuiller et al. 2008). However, receding edges are frequently overestimated with habitat suitability models (Lawler et al. 2009, Randin et al. 2009). Large-scale data, and climate data downscaled from coarse grain, may miss relict microclimates that might allow isolated pockets of suitable habitat and associated species to survive (Thomas et al. 2006, Ashcroft et al. 2009). Even when exposed to complete spatial shifts of suitable habitat many species demonstrate plasticity that allows them to persist or even adapt (Thuiller et al. 2008) to conditions outside bounds of typical environmental envelopes (Davis et al. 2005). They may not thrive at these locations, but they may persist, even if only as sink populations reliant on continued invasion to sustain the population (Davis et al. 1998).

Improvements

This exercise demonstrates the benefits of testing spatial and temporal hypotheses iteratively with long-term observations, but the approach is portable to landscape, regional, and continental scale monitoring programs. The modeling techniques applied can be improved and augmented. Maxent provides a good description of anticipated ecological states subjected to a changing climate, but others have demonstrated that ensemble or model averaging techniques that combine output of competing habitat suitability models might improve predictions (Stohlgren et al. 2010, Morisette et al. 2013). Several methods have also demonstrated the power of integrating mechanistic models to further constrain habitat suitability distributions and explicitly define relationships between pattern and process

(Broennimann et al. 2006, Kearney and Porter 2009). Regardless of the technique, iterative comparison of prediction against observation will pinpoint uncertainty that can be assessed with observation or targeted experimental investigation. This process repeated across time scales rarely seen in ecological investigation will provide new insight and predictive power.

Measures of land use and land cover often have significant predictive power of species distributions at small scales (Broennimann et al. 2006, Ficetola et al. 2007, Ashcroft et al. 2009), but they can be difficult to quantify. Historic land-use records often exist, but rarely in a spatial form. Where historic landcover maps exist, they can be compared and digitized, but disparity in techniques and classification often requires fine-scale resolution data. Most importantly, it is difficult to predict how these features will change in the future and what impact they will have on species distributions (Heikkinen et al. 2006).

Data on species distributions and predictive layers are expected to improve with time. The repeated natural experiments will direct new observation and experiments to areas and process that reduce model uncertainty. Projections of climate change will likely improve with new techniques and understanding and likely become available at smaller resolution (Ashcroft et al. 2009). Remote sensing data from satellites is improving, and fixed-wing packages with high-resolution hyperspectral and LiDAR capabilities are becoming increasingly available (Kampe 2010, Asner et al. 2012). These data will provide a better understanding of the causes and consequences of change to enable proactive management and improved preparation for resulting impacts to ecosystems and the services they provide.

REFERENCES

- Araujo, M. B., R. G. Pearson, W. Thuiller, and M. Erhard. 2005. Validation of species-climate impact models under climate change. *Global Change Biology* 11:1504-1513.
- Ashcroft, M. B., L. A. Chisholm, and K. O. French. 2009. Climate change at the landscape scale: predicting fine-grained spatial heterogeneity in warming and potential refugia for vegetation. *Global Change Biology* 15:656-667.
- Asner, G. P., D. E. Knapp, J. Boardman, R. O. Green, T. Kennedy-Bowdoin, M. Eastwood, R. E. Martin, C. Anderson, and C. B. Field. 2012. Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sensing of Environment* 124:454-465.
- Barnett, D. T., T. J. Stohlgren, C. Jarnevich, S., G. W. Chong, J. Ericson, A., T. Davern, R., and S. Simonson, E. 2007. The art and science of weed mapping. *Environmental Monitoring and Assessment* 132:235-252.
- Barnosky, A. D., E. A. Hadly, J. Bascompte, E. L. Berlow, J. H. Brown, M. Fortelium, W. M. Getx, J. Harte, A. Hastings, P. A. Marquet, N. D. Matinez, A. Mooers, P. Roopnarine, G. Vermeij, J. W. Williams, R. Gillespie, J. Kitzes, C. Marshall, N. Matzke, D. P. Mindell, E. Revilla, and A. B. Smith. 2012. Approaching a state shift in Earth's biosphere. *Nature* 486:52-58.
- Blossey, B. 1999. Before, during and after: the need for long-term monitoring in invasive plant species management. *Biological Invasions* 1:301-311.
- Botkin, D. B., H. Saxe, M. B. Araujo, R. Betts, R. H. W. Bradshaw, T. Cedhagen, P. Chesson, T. P. Dawson, J. R. Etterson, D. P. Faith, S. Ferrier, A. Guisan, A. S. Hansen, D. W. Hilbert, C. Loehle, C. Margules, M. New, M. J. Sobel, and D. R. B. Stockwell. 2007. Forecasting the effects of global warming on biodiversity. *Bioscience* 57:227-236.

- Broennimann, O., W. Thuiller, G. Hughes, G. F. Midgley, J. M. R. Alkemade, and A. Guisan. 2006. Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology* 12:1079-1093.
- Bremer, L. L., and K. A. Farley. 2010. Does plantation forestry restore biodiversity or create green deserts? A synthesis of the effects of land-use transitions on plant species richness. *Biodiversity and Conservation* 17:3893-3915.
- Camp, R. J., T. K. Pratt, P. M. Gorresen, J.J. Jeffrey, and B. L. Woodworth. 2010. Population trends of forest birds at Hakalau Forest National Wildlife Refuge, Hawai'i. *Condor* 112:196-212.
- Carpenter, S. 2008. Emergence of ecological networks. *Frontiers in Ecology and the Environment* 6:228.
- Chapin, F. S., P. A. Matson, and P. Vitousek. 2011. *Principles of terrestrial ecology*. Springer, New York, NY.
- Chen, I., J. K. Hill, R. Ohlemuller, D. B. Roy, and C. D. Thomas. 2011. Rapid range shifts of species associated with high levels of climate warming. *Science* 333:1024-1026.
- Clark, J. S. 2001. Ecological forecasts: an emerging imperative. *Science* 293:657-660.
- Collins, S. L., S. M. Glenn, and D. J. Wood. 1995. Experimental analysis of intermediate disturbance and initial floristic composition: decoupling cause and effect. *Ecology* 76:486-492.
- Crowl, T. A., T. O. Crist, R. R. Parmenter, G. Belovsky, and A. E. Lugo. 2008. The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment* 6:238-254.
- Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J. F. Guegan, B. A. Hawkins, D. M. Kaufman, J. T. Kerr, T. Oberdorff, E. O'Brien, and J. R. G. Turner. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters* 7:1121-1134.
- Daehler, C. C., J. S. Denslow, S. Ansari, and H. Kuo. 2004. A risk assessment system for screening out invasive pest plants from Hawai'i and other Pacific Islands. *Conservation Biology* 18:360-368.

- Davis, A. J., L. S. Jenkinson, J. H. Lawton, B. Shorrocks, and S. Wood. 1998. Making mistakes when predicting shifts in species range in response to global warming. *Nature* 391:783-786.
- Davis, M. B., R. G. Shaw, and J. R. Etterson. 2005. Evolutionary responses to changing climate. *Ecology* 86:1704-1714.
- DeWalt, S. J. 2006. Population dynamics and potential for biological control of an exotic invasive shrub in Hawaiian rainforests. *Biological Invasions* 8:1145-1158.
- Dormann, C. F., O. Schweiger, P. Arens, I. Augenstein, S. Aviron, D. Bailey, J. Baudry, R. Billeter, R. Bugter, R. Bukacek, F. Burel, M. Cerny, R. De Cock, G. De Blust, R. DeFilippi, T. Diekotter, J. Dirksen, W. Durka, P. J. Edwards, M. Frenzel, R. Hamersky, F. Hendrickx, F. Herzog, S. Klotz, B. Koolstra, A. Lausch, D. Le Coeur, J. Liira, J. P. Maelfait, P. Opdam, M. Roubalova, A. Schermann-Legionnet, N. Schermann, T. Schmidt, M. J. M. Smulders, M. Speelmans, P. Simova, J. Verboom, W. van Wingerden, and M. Zobel. 2008. Prediction uncertainty of environmental change effects on temperate European biodiversity. *Ecology Letters* 11:235-244.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Marion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmerman. 2006. Novel methods improve prediction of species' distribution from occurrence data. *Ecography* 29:129-151.
- Elith, J. and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40:677-697.
- Ficetola, G. F., W. Thuiller, and C. Miaud. 2007. Prediction and validation of the potential global distribution of a problematic alien invasive species - the American bullfrog. *Diversity and Distributions* 13:476-485.

- Frayer, W. E., and G. M. Furnival. 1999. Forest survey sampling designs: a history. *Journal of Forestry* 97:4-10.
- Gruner, D. S., M. E. S. Bracken, S. A. Berger, B. K. Eriksson, L. Gamfeldt, B. Matthiessen, S. Moorthi, U. Sommer, and H. Hillebrand. 2017. Effects of experimental warming on biodiversity depend on ecosystem type and local species composition. *Oikos* 126: 8–17.
- Heikkinen, R. K., M. Luoto, M. B. Araujo, R. Virkkala, W. Thuiller, and M. T. Sykes. 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography* 30:751-777.
- Higgins, S. I., D. M. Richardson, and R. M. Cowling. 2001. Validation of a spatial simulation model of a spreading alien plant population. *Journal of Applied Ecology* 38:571-584.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965-1978.
- Hijmans, R. J. 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93: 679–688.
- Hobbs, R. J., and L. F. Huenneke. 1992. Disturbance, diversity, and invasion: implications for conservation. *Conservation Biology* 6:324-337.
- Hulme, P. E. 2003. Biological invasions: winning the science battles but loosing the conservation war? *Oryx* 37:178-193.
- Ibanez, I., J. A. Silander, A. M. Wilson, N. Lafleur, N. Tanaka, and I. Tsuyama. 2009. Multivariate forecasts of potential distributions of invasive plant species. *Ecological Applications* 19:359-375.
- Jacobi, J. D. 1989. Vegetation maps of the upland plant communities on the islands of Hawai'i, Maui, Moloka'i, and Lana'i. *in*.

- Jarnevich, C. S., and T. J. Stohlgren. 2009. Near term climate projections for invasive species distributions. *Biological Invasions* 11:1373-1379.
- Jarnevich, C. S., T. J. Stohlgren, S. Kumar, J. T. Morissette, and T. R. Holcombe. 2015. Caveats for correlative species distribution modeling. *Ecological Informatics* 29:6-15.
- Kampe, T. U., B. R. Johnson, M. Kuester, and M. Keller. 2010. NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing* 4:043510.
- Kearney, M., and W. Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters* 12:334-350.
- Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. *Frontiers in Ecology and the Environment* 6:282-284.
- Kerr, J. T., H. M. Kharouba, and D. J. Currie. 2007. The macroecological contribution to global change solutions. *Science* 316:1581-1584.
- Lawler, J. J. 2009. Climate Change Adaptation Strategies for Resource Management and Conservation Planning. Pages 79-98 *in* Year in Ecology and Conservation Biology 2009. Blackwell Publishing, Oxford.
- Lawler, J. J., S. L. Shafer, D. White, P. Kareiva, E. P. Maurer, A. R. Blaustein, and P. J. Bartlein. 2009. Projected climate-induced faunal change in the Western Hemisphere. *Ecology* 90:588-597.
- Lee, H., D. A. Reusser, J. D. Olden, S. S. Smith, J. Graham, V. Burkett, J. S. Dukes, R. J. Piorkowski, and J. McPhedran. 2008. Integrated monitoring and information systems for managing aquatic invasive species in a changing climate. *Conservation Biology* 22:575-584.
- Lenoir, J., J.C. Gegout, P. A. Marquet, P. Ruffray, and H. Brisse. 2008. A significant upward shift in plant species optimum elevation during the 20th century. *Science* 320:1768-1771.

- Lodge, D. M., S. Williams, H. H. MacIassac, K. R. Hayes, B. Leung, S. Reichard, R. N. Mack, P. B. Moyle, M. Smith, D. A. Andow, T. J. Carleton, and A. McMichael. 2006. Biological invasions: recommendations for U.S. policy and management. *Ecological Applications* 16:2035-2054.
- MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton University Press, Princeton, NJ.
- Mack, M. C., and C. M. D'Antonio. 1998. Impacts of biological invasions on disturbance regimes. *Trends in Ecology and Evolution* 13:195-198.
- Mack, R., N., and M. Lonsdale, W., editors. 2002. *Erradicating invasive plants: Hard-won lessons for islands*. IUCN, Gland.
- Marshall, J. D., J. M. Blair, D. P. C. Peters, G. Okin, A. Rango, and M. Williams. 2008. Predicting and understanding ecosystem responses to climate change at continental scales. *Frontiers in Ecology and the Environment* 5:273-280.
- McGeoch, M. A., P. Genovesi, P. J. Bellingham, M. J. Costello, C. M. McGrannachan, and A. Sheppard. 2016. *Biological Invasions* 18:299-314.
- McMahon, S. M., M. C. Dietze, M. H. Hersh, E. V. Moran, and J. S. Clark. 2009. A Predictive Framework to Understand Forest Responses to Global Change. Pages 221-236 *in* Year in Ecology and Conservation Biology 2009. Blackwell Publishing, Oxford.
- Mika, A. M., R. M. Weiss, O. Olfert, R. A. Hallett, and J. A. Newman. 2008. Will climate be beneficial or detrimental to the invasive swede midge in North America? Contrasting predictions using climate projections from different general circulation models. *Global Change Biology* 14(8):1721-1733.
- Millar, C. I., N. L. Stephenson, and S. L. Stephens. 2007. Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* 17:2145-2151.

- Morin, X., and W. Thuiller. 2009. Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology* 90:1301-1313.
- Morisette, J. T., C. S. Jarnevich, T. R. Holcombe, C. B. Talbert, D. Ignizio, M. K. Talbert, C. Silva, D. Koop, A. Swanson, and N. E. Young. 2013. VisTrails SAHM: visualization and workflow management for species habitat modeling. *Ecography* 36: 129–135.
- Ortega-Huerta, M. A., and A. T. Peterson. 2008. Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. *Revista Mexicana De Biodiversidad* 79:205 - 216.
- Parker, I. M., D. Simberloff, W. M. Lonsdale, K. Goodell, P. M. Wonham, M. H. Kareiva, B. Von Holle, P. B. Moyle, J. E. Byers, and L. Goldwasser. 1999. Impact: toward a framework for understanding the ecological effects of invaders. *Biological Invasions* 1:3-19.
- Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics* 37:637-669.
- Pearman, P. B., A. Guisan, O. Broennimann, and C. F. Randin. 2008. Niche dynamics in space and time. *Trends in Ecology & Evolution* 23:149-158.
- Pejchar, L., and H. A. Mooney. 2009. Invasive species, ecosystem services and human well-being. *Trends in Ecology and Evolution* 24:497-504.
- Perry, A. L., P. J. Low, J. R. Ellis, and J. D. Reynolds. 2005. Climate change and distribution shifts in marine fishes. *Science* 308:1912-1915.
- Peters, D. P. C., P. M. Groffman, K. J. Nadelhoffer, N. B. Grimm, S. L. Collins, W. K. Michener, and M. A. Huston. 2008. Living in an increasingly connected world: a framework for continental-scale environmental science. *Frontiers in Ecology and the Environment* 6:229-237.
- Phillips, S., J., R. Anderson, P., and R. Schapire, E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190:231-259.

- Randin, C. F., R. Engler, S. Normand, M. Zappa, N. E. Zimmermann, P. B. Pearman, P. Vittoz, W. Thuiller, and A. Guisan. 2009. Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology* 15:1557-1569.
- Regniere, J., V. Nealis, and K. Porter. 2009. Climate suitability and management of the gypsy moth invasion into Canada. *Biological Invasions* 11:135-148.
- Stohlgren, T. J. 2007. *Measuring plant diversity, lessons from the field*. Oxford University Press, New York.
- Stohlgren, T. J., and J. L. Schnase. 2006. Risk analysis for biological hazards: what we need to know about invasive species. *Risk Analysis* 26:163-173. Note 2006 and full citation for references.
- Stohlgren, T. J., P. Ma, S. Kumar, M. Rocca, J. T. Morissette, C. S. Jarnevich, and N. Benson. 2010. Ensemble Habitat Mapping of Invasive Plant Species. *Risk Analysis* 30: 224–235.
- Thomas, C. D., A. M. A. Franco, and J. K. Hill. 2006. Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution* 21:415-416.
- Thuiller, W., C. Albert, M. B. Araujo, P. M. Berry, M. Cabeza, A. Guisan, T. Hickler, G. F. Midgely, J. Paterson, F. M. Schurr, M. T. Sykes, and N. E. Zimmermann. 2008. Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology Evolution and Systematics* 9:137-152.
- Thuiller, W., D. M. Richardson, P. Pysek, G. F. Midgley, G. O. Hughes, and M. Rouget. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11:2234-2250.
- Thuiller, W., D. M. Richardson, M. Rouget, S. Proches, and J. R. U. Wilson. 2006. Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* 87:1755-1769.

Vitousek, P. M., L. R. Walker, L. D. Whiteaker, D. Mueller-Dombois, and P. A. Matson. 1987. Biological invasion by *Myrica faya* alters ecosystem development in Hawaii. *Science* 238:802-804.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier. 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climate Change* 62:189 - 216.

CONCLUSIONS ABOUT THE NEON SAMPLE DESIGN FOR PLANT DIVERSITY

Introduction

The goal of this effort was to design and test methods that will enable the National Ecological Observatory Network to understanding the causes of and consequences of change on plant diversity. What follows is a summary of each chapter and a summary of some of the challenges and future directions.

The terrestrial organism and biogeochemistry spatial sampling design for the National Ecological Observatory Network

The need to support the high-level objectives of NEON constrained the development of the sampling design for plant diversity – all terrestrial organisms and soil – at sites across the United States and Puerto Rico. The foundation of the design is the random sample that provides unbiased descriptions and functions with many analytical approaches. Stratification aims to increase efficiency and characterization of dominant cover types measured by the tower-based sensor data streams. Design redundancy and flexibility allows distinct sample sizes and allocation across strata within a site, while allowing collocation of taxonomic groups. To ensure that plant diversity design was capable of detecting and differentiating trends, power analyses were tested with real NEON data. Results suggested that the proposed sample size of 30 was not generally sufficient to detect and differentiate site trends after decades. If observed patterns of spatial variability persist and estimates of temporal variation (made by NDVI in the absence of time series data) are accurate, more sites might be necessary to detect trends, particularly if conclusive and actionable data are needed before 20-30 years.

These findings present both challenges and opportunities. It is unlikely that the National Science Foundation will provide additional funding to increase the sample size for plant diversity sampling. Given

this zero-sum game, if it is determined that this relatively stringent test (Yi and Panzerelli 2002, Thompson 2012) should still guide data collection efforts, several options exist:

- Reallocate sampling effort across sites. This analysis demonstrated variability in patterns of spatial and temporal variability across sites. Further analysis might demonstrate that smaller sites or homogenous sites (e.g., agricultural sites) are characterized by less variation in space, time, or both; sampling effort could be moved from these sites to more complex sites. The assumption of equal sample size would need to be relaxed and that could change analysis projections (Yi and Panzarella 2002).
- Time and funding could be pulled from a different NEON sampling activity. Analysis might show that some taxonomic groups are oversampled given objectives, or that sample that is determined to be oversampled given objectives.
- Filling in gaps could be left to the ecological community who could sample with comparable methods from additional plots from the NEON sampling design. This option might be less popular given the demands of the NEON Operations funding on resources available to the entire ecological community.
- The diversity of NLCD cover types within a site could be reduced. By focusing on dominant types at each site that are also measured by tower-based sensors at each site, variability would effectively be reduced, likely reducing the number of plots required to sample each site with sufficient intensity as described by the test.

There are other challenges associated with the design. The stratified-random approach adds complexity to the end users' efforts with NEON data given the need to determine the first couple of moments with the design-based estimators. A simple random sample would obviate this need and avoid data processing errors. The subset of plots that can be treated as a random sample will be available through the NEON portal for those that would prefer the simplicity of a random design. A slightly

different approach would be to allocate plots directly proportional to the strata sampled. In this special case, the data become a 'self-weighting sample' that allows the resulting data to be treated as a simple random sample (Thompson 2012). Slight modifications could be made to the allocation to achieve this design for the sake of simplicity, particularly if the hypothesis associated with the allocation proportional to the square-root of strata area are not validated. Ongoing data collection and analysis and engagement with the community of users will provide direction for future design iterations.

Strategies for comparing plant diversity in a national network of sites

The data were able to generate meaningful comparisons of sampled plant species richness across sites and strata within sites. Assuming the sample is sufficient relative to the heterogeneity of the site in question, comparison appears to be a robust and possibly more descriptive and insightful standard for comparison of richness. The hypothesis that sampling according to the proportion of the square-root of the area to reach the inflection point of the species-accumulation curve did not hold in most strata (30%), nor at most sites (33%). In most cases, the sample size of 30, set by the constraints of the sample design manuscript, exceeded the inflection point. The cause does not appear to be purely related to sample size. Assuming the relative difference that the strata-specific sampling exceeds the inflection point were to hold with a reduced sample size, allocating samples proportional to the square-root of the area – at least at the sites tested, does not effectively capture the proposed efficiency and information content proposed.

Challenges to the design present opportunities. The design does provide information – particularly with insights from interpolation and extrapolation - about how to optimize the design. Insights from interpolation and extrapolation suggest that sampling to the inflection point would be most effective by simply adjusting the strata- or site-specific sample size by 'sliding' along the species accumulation curve to the inflection point. Similarly, the degree to which the inflection point comparison is robust could be evaluated at each site by simply adding test plots in targeted strata.

These could be random or they could be targeted to environmental gradients that seem to be missed by the sample as described by NEON's remote sensing platform (Kampe et al. 2010, Stohlgren et al. 2010). Alternatively, the principal components of the hyperspectral data can be treated as a surrogate of plant diversity data. Simulations that compare 'species' accumulation curves of these reflectance values based on NEON plot locations could be compared to the impact of additional samples from the realm, unsampled by the sample design for plot sampling. Asking, and answering, 'what did we miss?' further evaluates assumptions and hypotheses and provides direction for cost-effective design optimization.

Planning for climate change when designing invasive plant species studies

The results suggest that species vary predictably in space and possibly through time. In the case of this investigation, the range one invasive grass species is expected to expand while the other will likely undergo a reduction in suitable habitat in the study area. These patterns enable action. Plots could be placed strategically for model validation and for a monitoring system designed to detect change. This preemptive sampling will have a higher chance of detecting change with a network of plots capable of sampling a very small part of each site. Managers focused on control of invasive species and habitat restoration might also leverage these models to focus control efforts on the species likely to expand, and continuing to evaluate the expected trend of the species expected to undergo a reduction in suitable habitat.

This model-informed sampling approach is not currently in the scope of the NEON design for plant-diversity. The funding is not available for plot-based sampling, but perhaps managers or owners of the site or members of the ecological community would pursue such an approach. NEON has set aside some funding for 'gradient plots' designed to record unsampled landscape features (Thorpe et al. 2016), but efforts that validate the airborne platform (e.g., vegetation structure, leaf area index, foliar chemistry) or to better link the flow of energy and matter from the atmosphere to aquatic systems with targeted biogeochemistry observations (Hinkley et al. 2015).

Notes and recommendations for the next 30 years of NEON monitoring

A tension between designs

The results of the three manuscripts that make the body of this work describe the logic and implementation of the NEON plant diversity design and provide insights and test assumptions of the design. Conspicuously, the testing results in conclusions that seem to provide contrasting paths for optimizing the design to support NEON goals.

The first manuscript that lays out the study design, sample size and allocation implemented a simplistic constraint – both for logistics and sample size calculation – that all sites are sampled with the same intensity. However, the paper focused on generating valid comparisons by sampling to the inflection point of the species accumulation curve suggests that a design prioritizing comparison would reduce sample size at some sites, possibly reallocating effort to more diverse strata and sites. Logistical simplicity, the capacity to compare richness by rarefaction – moving down a species accumulation curve – and the need to ensure capacity to detect trends means that initially NEON will continue to sample the same sample size across sites.

Also related to sample size, the single season of plant diversity sampling that informs spatial sample variance parameters and the NDVI estimate of temporal variation suggests that a sample size of 30 plots per site is marginally sufficient and even insufficient for detecting and differentiating trends in plant species richness per plot across sites. However, as indicated, results from the comparison paper might suggest a reduction in sample size would be feasible for many strata and sites. The need to ensure the capacity to detect trends means that there will not be a reduction in the sample size as NEON as part of the initial sample design.

The notion of preemptive sampling to increase the probability of detecting change would result in the placement and inclusion of plant diversity plots placed subjectively or according to a model-based design. However, the random component is central to the design. Some analytical techniques are

agnostic with respect to sample design, capable of assimilating data collected in numerous ways (Evangelista 2008). However, design-based inference (Thompson 2012) and the calculation of basic statistics from plots collected from a stratified-random design must be calculated with specific estimators (Thompson 2012). Serving plant diversity data collected from two different designs to the ecological community by the NEON data portal is likely to be confusing and prone to errors. With this complexity and the demand from other protocols on gradient plots, the initial plant diversity design will, at least initially, rely on plots sampled according to the stratified-random design.

These contra indicators suggest divergent paths. The struggle they engender and the compromise they mandate occur in the context of the plant diversity design is a microcosm of similar challenges across the many protocols associated with terrestrial organisms and soil (the Terrestrial Observation System), the aquatic observations (Aquatic Observation System) and sensors (Aquatic Instrument System) and the system for sensing fluxes of soil and atmosphere and measuring climate (Terrestrial Instrument System). Resolution and a path forward in the form of a design is found in reliance on hierarchical requirements that connect decisions to the high-level NEON mission and in consultation with the ecological community who will work with the data to derive understanding of the impacts of change.

Additional considerations

The design and launch of the NEON project was riddled with difficult decisions and tradeoffs. In many cases there was no single, correct answer. Outcomes reflected previous efforts, data, and testing, but also input from ecologists from outside NEON that generously donated countless hours to shape the Observatory, as well as bias and history of the NEON staff. Ultimately the persistent and driving goal was to adhere to a requirements framework that ensured each decision reflected guiding constraints that link to the NEON mission (Keller et al. 2008, Schimel and Keller 2015). Some of the considerations theoretical gymnastics include the following:

- The data and the science NEON enables must justify the cost. The design and creation of NEON was funded by the NSF Major Research Equipment and Facilities Construction (MREFC) which is responsible for scientific infrastructure. However, operating NEON is likely to be funded by the NSF Biology program. That NEON will pull substantial capital from the same budgets that fund principal investigator research means two things:
 1. The data must be useful. NEON data must provide plant diversity data that the ecology community needs to answer questions and solve problems at the scale of the continent and over decades. It must serve as a backbone for ecological understanding developed by the next generation of ecologists. If records indicate that the plant diversity data is not popular with those downloading data, changes to the design and collection strategy should be considered.
 2. The data must be informative. After 30 years of data collection it would be unacceptable to indicate that, “we need more data” to provide information about causes of change on ecology.
- There needs to be quality control and assurance of the data. The best design and copious funding could be wasted with the collection of data that inflates uncertainty. NEON is currently developing QA/QC plan, but it has been estimated that approximately 20% of a budget should be dedicated to evaluating data quality (Stohlgren 2007). In addition to testing the accuracy of species identification, rates of species detection and estimates of cover should be evaluated for inclusion in statistical models (Clark 2003, Cressie et al. 2009). Quality-assured data will enable the NEON mission. Erroneous data can obscure any signal, but also point to the need for adjustments such as better training or the need to alter timing of sampling such that more species possess diagnostic parts.

- The design must be tested and optimized. Data will provide information about the design. While this information can guide efforts to optimize the design, these iterations should not be prolonged over many years. Temporal continuity will be critical to describing trends and incessant design alterations would challenge NEON objectives.
- Despite a significant dedication of funds, it is simply impossible to get everything desired. It would have been optimal to expand the definition of plant diversity to include a suite of functional traits such or an expanded focus on mapping invasive plant species. Field work and the maintenance of systems to evaluate and serve the data are expensive and have left little in the way of money to expand the NEON scope.

Ultimately the goal of the Observatory is to enable transformative science. Perhaps the best analogy is high-level or pooled funding for super collider science. The physics and astronomy community placed a premium on scientific infrastructure no independent investigator could fund or maintain that had the potential to push the boundaries of scientific knowledge (Eichten et al. 1984). It is hoped that NEON, combined with technology and evolving statistical approaches capable of handling big data, will do the same for ecology and understanding environmental change during the Anthropocene.

REFERENCES

- Chapin, F. S. and E, Fernandez. 2013. Proactive ecology for the Anthropocene. *Elemental Science of the Anthropocene* 1:3.
- Clark, J. S. 2003. Uncertainty and variability in demography and population growth: a hierarchical approach. *Ecology* 84(6):1370-1381.
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters* 8(1):2-14.
- Cressie, N., C. A. Calder, J. S. Clark, J. M. ver Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* 19(3):553-570.
- Evangelista, P. H., S. Kumar, T. J. Stohlgren, C. S. Jarnevich, A. W. Crass, J. H. Norman, and D. T. Barnett. 2008. Modelling invasion for a habitat generalist and specialist. *Diversity and Distributions* 14(5):808-817.
- Eichten, E., I. Hinchliffe, K. Lane, and C. Quigg. 1984. Supercollider physics. *Reviews of Modern Physics* 58:1056.
- Hinckley, E. S., G. B. Bonan, G. J. Bowen, B. P. Colman, P. A. Duffy, C. L. Goodale, B. Z. Houlton, E. Marin-Spiotta, K. Ogle, S. V. Ollinger, E. A. Paul, P. M. Vitousek, K. C. Weathers, D. G. Williams. 2016. *Ecosphere* 7(3): article e01234. DOI: 10.1002/ecs2.1234.
- Schimel, D. S. and M. Keller. 2015. Big questions, big science: meeting the challenges of global ecology. *Oecologia* 177:925-934.
- Stohlgren, T.J., 2007. *Measuring plant diversity, lessons from the field*. Oxford University Press, New York.
- Thompson, S. K. 2012. *Sampling*. Wiley, New Jersey.

Thorpe, A. T., Barnett, D. T., Elmendorf, S. C., Eve-Lyn Hinckley, Hoekman, D., Jones, K. D., LeVan, K. E., Meier, C. L., Lee Stanish, Thibault, K. M., 2016. Introduction to the sampling designs of the National Ecological Observatory Network Terrestrial Observation System. *Ecosphere* 7(12): article e01627. DOI: 10.1002/ecs2.1627.

Yi, Q. L. and T. Panzarella. 2002. Estimating sample size for tests on trends across repeated measurements with missing data based on the interaction term in a mixed model. *Controlled Clinical Trials* 23:481-496.

APPENDIX 1. SAMPLE SIZE CALCULATIONS

The following is R code developed for the initial estimation of sample size calculations:

```
#### Sample size calculation for the test of differences in the slope
#### between two independent sets of samples in a repeated measures model
#### with both fixed and random effects
#### see [Q. Yi, T. Panzarella/Controlled Clinical Trials 23 (2002) 481–496]

# t is the number of repeated measurements, not necessarily the number of years
# samp.freq is the number of samples per year
# sigsq is the estimate of the common population variance
# corr is the parameter for correlation in either compound symmetric or
# first order autoregressive model
# AR is a flag to determine whether CS or AR correlation structure should be used
# alpha is the acceptable type I error level
# beta is the acceptable type II error level specified as defined below
# slopes.random is a logical indicating whether slopes should be considered random

rep.meas.lmm<-function(t = 5, sigsq = 1, corr = 0.5, AR = F, alpha = 0.05, beta = 0.8, slopes.random=T,
samp.freq = 1)
{

require(ramps)
require(MASS)
# beta.int is related to a one-unit change of time and the length of one unit
# of time varies with the number of measurements, it requires a corresponding adjustment for
# the number of repeated measurements within the fixed duration. This is also the case for the
# variance of random slopes. hence beta.int=0.5/(t-1), and var(beta.int) = 0.05*4/(t-1)^2
# This keeps the magnitude of the difference in slopes between the two groups and random variation
# constant within a fixed duration. (Yi and Panzarella 2002)
s<-samp.freq
b.int<-0.5/(s*(t-1))
# fix t=0 in the denominator
b.int[1]<-0

d.mat <- data.frame(time=c(0:(t-1)))
X<-model.matrix(~time,d.mat)
# main effects design matrix for core site
X1<-cbind(X,X)
# main effects design matrix for relocatable to be compared to core site
X2<-cbind(X,matrix(rep(0,t*2),nrow=t))
# random effects matrix
Z<-X[,c(1,2)]

# comp symm correlation matrix
```

```

cor.CompSymm<-corCompSymm(corr)
cor.Symm.init<-Initialize(cor.CompSymm,data=X1)
R<-corMatrix(cor.Symm.init)

# R for AR(1)
if(AR==T){
coef<-seq(1:t)
for(i in 2:t){
coef<-rbind(coef,c(rev(seq(1,i)),seq(2,t))[1:t])
}
coef <- (coef - 1)/s
R <- matrix(corr, nrow = t, ncol=t)
diag(R) <- 1
R <- R^coef
}

# Not considering the variance of b.int as nonzero
# specifying the variance of b.int
var.b1.i<-0

# Power constraint from Yi and Panzarella (2002)
# page 458, results 1 paragraph, last sentence.
# Their constraint corresponds to a power of 80% at a difference
# between slopes (at the core site and relocatable) of roughly 11%
# run (0.05/(5-1))/sqrt(0.05*4/(5-1)^2) to check this
# if(slopes.random==T){var.b1.i<-0.05*(4/((t-1)^2))}

# The next line specifies the power at a difference of
# slopes of roughly 20%
# run (0.05/(5-1))/sqrt(0.05*1.25/(5-1)^2) to check this
if(slopes.random==T){var.b1.i<-0.05*(1.25/((t-1)^2))}

D<-matrix(c(0,0,0,var.b1.i),ncol=2, byrow=T)

V<-Z%%D%%t(Z)+sigsq*R
v.inv<-solve(V)
z.alp<-qnorm(1-(alpha/2))
z.bet<-qnorm(beta)

t1<- (z.alp+z.bet)^2
t2<- solve(t(X1)%%v.inv %% X1 + t(X2)%%v.inv %% X2)
t3<- 0.5/(t-1)
return(ceiling(((t1*t2)/(t3^2))[4,4]))
}

#####
#####

```

```
#####
#####
#####
#####
```

```
# specifying the parameters of interest for the generation of tables
corrs<- c(0.25 ,0.50, 0.75)
sigqs<-c(0.25 ,0.50, 0.75, 1.00)
years<- c(10, 20, 30)
```

```
# code for table
samp.vec<-NA
for(i in 1:3){
  for(j in 1:4){
    for(k in 1:3){
      samp.vec<-c(samp.vec,rep.meas.lmm(t = years[k], sigsq = sigqs[j], corr = corrs[i], AR = F, alpha = 0.1,
      beta = 0.8))
    }
  }
}
matrix(samp.vec[-1],ncol=3)
```

```
# code for table
samp.vec<-NA
for(i in 1:3){
  for(j in 1:4){
    for(k in 1:3){
      samp.vec<-c(samp.vec,rep.meas.lmm(t = years[k], sigsq = sigqs[j], corr = corrs[i], AR = F, alpha = 0.05,
      beta = 0.8))
    }
  }
}
matrix(samp.vec[-1],ncol=3)
```

```
# code for figure
corrs<- seq(0.05,0.95,0.1)
years<- c(3:30)
```

```
samp.vec<-NA
for(i in 1:length(corrs)){
  for(k in 1:length(years)){
    samp.vec<-c(samp.vec,rep.meas.lmm(t = years[k], sigsq = 0.50, corr = corrs[i], AR = F, alpha = 0.1, beta =
    0.8))
  }
}
```

```
fig1.df<-data.frame(z = samp.vec[-1])
```

```

fig1.df$x<-rep(years,length(corrs))
fig1.df$y<-rep(corrs,each=length(years))
require(lattice)
wireframe(z ~ x * y, fig1.df,
drape = TRUE, zoom=0.875,
          xlab=list(c("Years"),rot=10,cex=1.1),
          ylab=list(c("Correlation"),rot=0,cex=1.1),
          zlab=list(c("Number of Samples"),rot=90,cex=1.1),
aspect = c(0.75, .85),
light.source = c(10,10,10),
col.regions = rev(rainbow(length(corrs)*length(years),start=0.825,end=0.35)),
add.legend=T,
screen = list(z = -110, x = -70, y = -20),
          scales = list(arrows = F)
)

# code for figure
corrs<- seq(0.025,0.975,0.05)
years<- c(3:30)

samp.vec<-NA
for(i in 1:length(corrs)){
for(k in 1:length(years)){
samp.vec<-c(samp.vec,rep.meas.lmm(t = years[k], sigsq = 0.50, corr = corrs[i], AR = T, alpha = 0.1, beta =
0.8))
}
}

fig1.df<-data.frame(z = samp.vec[-1])
fig1.df$x<-rep(years,length(corrs))
fig1.df$y<-rep(corrs,each=length(years))

wireframe(z ~ x * y, fig1.df,
drape = TRUE, zoom=0.875,
          xlab=list(c("Years"),rot=0,cex=1.1),
          ylab=list(c("Correlation"),rot=-35,cex=1.1),
          zlab=list(c("Number of Samples"),rot=-65,cex=1.1),
aspect = c(0.75, .85),
light.source = c(10,10,10),
col.regions = rev(rainbow(length(corrs)*length(years),start=0.825,end=0.35)),
add.legend=T,
screen = list(z = -130, x = -30, y = -10),
          scales = list(arrows = F)
)

#####
#####

```

```
#####
#####
#####
#####
# Sample code to confirm the bottom half of table 1 in Yi and Panzaralla (2002) p. 485
# before running this, reset the power constraint to that which they used by
# uncommenting the following line
# if(slopes.random==T){var.b1.i<-0.05*(4/((t-1)^2))}

rep.meas.lmm(t = 5, sigsq = 1, corr = 0.2, AR = F, alpha = 0.05, beta = 0.8)
rep.meas.lmm(t = 5, sigsq = 1, corr = 0.5, AR = F, alpha = 0.05, beta = 0.8)
rep.meas.lmm(t = 5, sigsq = 1, corr = 0.8, AR = F, alpha = 0.05, beta = 0.8)

rep.meas.lmm(t = 5, sigsq = 1, corr = 0.2, AR = T, alpha = 0.05, beta = 0.8)
rep.meas.lmm(t = 5, sigsq = 1, corr = 0.5, AR = T, alpha = 0.05, beta = 0.8)
rep.meas.lmm(t = 5, sigsq = 1, corr = 0.8, AR = T, alpha = 0.05, beta = 0.8)

rep.meas.lmm(t = 9, sigsq = 1, corr = 0.2, AR = F, alpha = 0.05, beta = 0.8,samp.freq = 2)
rep.meas.lmm(t = 9, sigsq = 1, corr = 0.5, AR = F, alpha = 0.05, beta = 0.8,samp.freq = 2)
rep.meas.lmm(t = 9, sigsq = 1, corr = 0.8, AR = F, alpha = 0.05, beta = 0.8,samp.freq = 2)

rep.meas.lmm(t = 9, sigsq = 1, corr = 0.2, AR = T, alpha = 0.05, beta = 0.8,samp.freq = 2)
rep.meas.lmm(t = 9, sigsq = 1, corr = 0.5, AR = T, alpha = 0.05, beta = 0.8,samp.freq = 2)
rep.meas.lmm(t = 9, sigsq = 1, corr = 0.8, AR = T, alpha = 0.05, beta = 0.8,samp.freq = 2)
```

APPENDIX 2: SAMPLE MEAN AND SAMPLE VARIANCE WITH DESIGN-BASED ESTIMATORS

```
---
title: "Mean and Variance"
author: "dBarnett"
date: "June 8, 2016"
output: word_document
---

``{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)

rm(list=ls())
library(dplyr)
library(stringr)

#Set working directory
setwd('C:/evalSampleSize')
```

``{import files and organize: sample size}

#pull sample size and area by plot type and NLCD class
sampleSize3<-read.csv(file = "data/sampleSize3.csv",
 head=TRUE, sep=",", stringsAsFactors = FALSE)
sampleSize3$siteID<-sampleSize3$site
sampleSize3$nlcdClass<-sampleSize3$NLCD
sampleSize3$Ni<-sampleSize3$areakm2

limitSampSize<-select(sampleSize3, siteID, nlcdClass, Ni)
```

``{import files: mass data}

#bring in mass data (available on NEON portal)

osbs<-read.csv(file = "data/osbs.csv",
  head=TRUE, sep=",", stringsAsFactors = FALSE)

ornl<-read.csv(file = "data/ornl.csv",
  head=TRUE, sep=",", stringsAsFactors = FALSE)

onaq<-read.csv(file = "data/onaq.csv",
  head=TRUE, sep=",", stringsAsFactors = FALSE)

#combine data
```

```

rawData<-rbind(harv,osbs,ornl,tall,wood,onaq)

#distributed plots (tower plots not distributed according to sample design)
rawDataDis<-rawData[which(rawData$plotType == "distributed"),]

#reduce number of fields
rawDataDisLimit<-select(rawDataDis, siteID, plotID, dryMass)

#sum across samples within a plot
massPlot<-ddply(rawDataDisLimit, 'plotID', function(x) c(count=nrow(x), sumPlot=sum(x$dryMass)))
```

```${import files: nlcd data}```

#bring in boutData for nlcdClass (available from NEON portal)

osbsBout<-read.csv(file = "data/osbsBout.csv",
  head=TRUE, sep=";", stringsAsFactors = FALSE)

ornlBout<-read.csv(file = "data/ornlBout.csv",
  head=TRUE, sep=";", stringsAsFactors = FALSE)

onaqBout<-read.csv(file = "data/onaqBout.csv",
  head=TRUE, sep=";", stringsAsFactors = FALSE)

#combine bout data
rawDataBout<-rbind(harvBout,osbsBout,ornlBout,tallBout,woodBout,onaqBout)

#select variables want from bout
varsDataBout<-select(rawDataBout, siteID, plotID, nlcdClass)

#remove duplicates
limitVarsDataBout<-unique(varsDataBout)
```

```${import files: combine and calculate}```

#combine mass data with bout data to add nlcdClass
massNlcd<-full_join(massPlot, limitVarsDataBout)

#calculate mean and sample variance by nlcdClass
meanStrat<-ddply(massNlcd, c('siteID', 'nlcdClass'), function(x) c(ni=nrow(x), yi=mean(x$sumPlot),
si2=var(x$sumPlot)))

#add strat to frame
meanStratArea<-merge(meanStrat, limitSampSize, by = c("siteID", "nlcdClass"), type = "left", match =
"all" )

#limit to records of strata that were sampled

```

```

meanStratAreaValues<-filter(meanStratArea, yi != 'NA')

#remove duplicates
meanStratAreaNoDups<-unique(meanStratAreaValues)

#calculateMean
meanStratAreaNoDups$yiNi<-meanStratAreaNoDups$yi*meanStratAreaNoDups$Ni

mean<-ddply(meanStratAreaNoDups, 'siteID', function(x) c(count=nrow(x), Es=sum(x$yiNi),
N=sum(x$Ni)))

mean$Ybar<-mean$Es/mean$N

#calculateVariance
addArea<-merge(meanStratAreaNoDups, mean, by = c("siteID"), type = "left", match = "all")
addArea$byStrat<-((addArea$Ni/addArea$N)^2)*((addArea$Ni-
addArea$ni)/addArea$Ni)*(addArea$si2/addArea$ni)

varHat<-ddply(addArea, 'siteID', function(x) c(count=nrow(x), varHat=sum(x$byStrat)))
'''

```