esa ECOSPHERE

SPECIAL FEATURE: NEON DESIGN

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

David T. Barnett,[1,†] Paul A. Duffy,[2] David S. Schimel,[3] Rachel E. Krauss,[1] Kathryn M. Irvine,[4] Frank W. Davis,[5] John E. Gross,[6] Elena I. Azuaje,[1,9] Andrea S. Thorpe,[1,10] David Gudex-Cross,[1,11] Michael Patterson,[1,12] Jalynda M. McKay,[1,13] Joel T. McCorkel,[7] and Courtney L. Meier[1,8]

[1]*Battelle Memorial Institute, 1685 38th Street, Suite 100, Boulder, Colorado 80301 USA*
[2]*Neptune and Company, 1435 Garrison Street, Suite 100, Lakewood, Colorado 80215 USA*
[3]*NASA Jet Propulsion Lab, 4800 Grove Drive, Pasadena, California 91109 USA*
[4]*Northern Rocky Mountain Science Center, US Geological Survey, 2327 University Way, Bozeman, Montana 59715 USA*
[5]*Bren School of Environmental Science and Management, University of California, Santa Barbara, 240 Bren Hall, Santa Barbara, California 93106 USA*
[6]*Climate Change Response Program, National Park Service, Natural Resource Stewardship and Science, Fort Collins, Colorado 80525 USA*
[7]*NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland 20771 USA*
[8]*Institute of Arctic and Alpine Research, University of Colorado, Campus Box 450, Boulder, Colorado 80309 USA*

Citation: Barnett, D. T., P. A. Duffy, D. S. Schimel, R. E. Krauss, K. M. Irvine, F. W. Davis, J. E. Gross, E. I. Azuaje, A. S. Thorpe, D. Gudex-Cross, M. Patterson, J. M. McKay, J. T. McCorkel, and C. L. Meier. 2019. The terrestrial organism and biogeochemistry spatial sampling design for the National Ecological Observatory Network. Ecosphere 10(2):e02540. 10.1002/ecs2.2540

Abstract. The National Ecological Observatory Network (NEON) seeks to facilitate ecological prediction at a continental scale by measuring processes that drive change and responses at sites across the United States for thirty years. The spatial distribution of observations of terrestrial organisms and soil within NEON sites is determined according to a "design-based" sample design that relies on the randomization of sampling locations. Development of the sample design was guided by high-level NEON objectives and the multitude of data products that will be subjected to numerous analytical approaches to address the causes and consequences of ecological change. A requirement framework permeates the NEON design, ensuring traceability from each facet of the design to the high-level requirements that make the NEON mission statement actionable. Requirements were developed for the terrestrial sample design to guide the key components of the design:

1. Randomizing the sample locations ensures the unbiased collection of data, is appropriate for organisms and soil, and provides data suitable for a variety of analyses.
2. Stratification increases efficiency and allows sampling to focus on those parts of the landscape measured by other NEON observation platforms.
3. Attention to the sample size and spatial plot allocation ensures that data products will be sufficient to inform questions asked of the data and the NEON objectives.
4. Establishing a framework with the capacity for re-evaluate and design iteration allows for adaption to unexpected challenges and optimization of the sample design based on early data returns.

The utility of the NEON sampling design is highlighted by its application across terrestrial systems. The data generated from this unique design will be used to quantify patterns in: the abundance and diversity of small mammals, breeding birds, insects, and soil microbes; vegetation structure, biomass, productivity, and diversity; and soil biogeochemistry.

Key words: National Ecological Observatory Network; sampling terrestrial organisms and biogeochemistry; spatial sample design; Special Feature: NEON Design.

Received 31 October 2017; revised 5 June 2018; accepted 3 July 2018; final version received 28 November 2018. Corresponding Editor: Joseph LaManna.

[9] Present address: Mediterranean Institute for Advanced Studies, 21 Miquel Marques, Illes Balears, Spain.

[10] Present address: Washington Natural Heritage Program, Olympia, Washington 98504 USA.

[11] Present address: Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, Vermont 05405 USA.

[12] Present address: Oracle Data Cloud, 500 Eldorado Boulevard, Broomfield, Colorado 80021 USA.

[13] Present address: Civil and Architectural Engineering, University of Wyoming, 1000 East University Avemue, Laramie, Wyoming 82071 USA.

† **E-mail:** dbarnett@battelleecology.org

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 drivers of change and responses at as many as 47 terrestrial sites across the continental United States, Alaska, Hawaii, and Puerto Rico for thirty years (Vitousek 1997, Keller et al. 2008, Luo et al. 2011). The terrestrial sites encompass wildlands and a variety of gradients (e.g., land use, species invasion, nitrogen deposition) to address regional- and continental-scale questions. Within sites, measurements of atmosphere, soil, water, select organisms and disease, and airborne observations yield freely available data, enabling a new paradigm in ecological science, education, and policy.

The National Ecological Observatory Network employs automated sensors and observations to generate data regarding ecological status and trends span spatial and temporal scales. Fixed-wing aircraft census vegetation at landscape scales (~400 km^2) with high-resolution remote sensing at annual time steps; tower-based sensors capture temporally continuous fluxes over smaller spatial extents (~0.5 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 (~5–60 km^2). A complete census of organisms and biogeochemistry is biologically and financially impractical—soil microbes are ubiquitous and birds 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, a site-scale spatial sampling design is needed to direct the observation and collection of terrestrial organisms and soil to facilitate statistically rigorous inference from the scale of plots to sites and the continental Observatory.

Prescribing the number and spatial arrangement of plots for the collection of the diversity of organisms and soil observed by the NEON Terrestrial Observation System (TOS; Kao et al. 2012, Thorpe et al. 2016) in a way that informs the continental-scale Observatory presents a formidable challenge. The strategy is described herein: Guided by NEON principles and requirements, the TOS 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 overall NEON sample design for terrestrial organisms and soil, the description, justification, and study design specifics for the taxonomic groups and soil characteristics sampled are described elsewhere (Barnett et al., *in press*, Hinckley et al. 2016, Hoekman et al. 2016, 2017, Springer et al. 2016, Thorpe et al. 2016).

Design Criteria

The National Ecological Observatory Network 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 1). The sample design for organisms and soil is part of this requirement-driven hierarchical structure; high-level "upstream" requirements and "downstream" data products provide context and constraints under which specific requirements and details for the sample design were developed.

The sample design for observations at local, site-specific scales must deliver data that optimally informs continental-scale ecology. Adopting the requirement framework allows traceability to elements of the continental sampling strategy and

Table 1. Connections between National Ecological Observatory Network (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
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)	Direct the collection of the raw material for continental ecology
NEON will allow extrapolation from the Observatory's local 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.	
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	
NEON's goal is to improve understanding and forecasting of ecological change at continental scales.	Efficiently capture landscape-scale pattern and trend
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	
NEON shall observe the causes and consequences of environmental change in order to establish the link between ecological cause and effect	Provide infrastructure that co-locates terrestrial measurements and links observations to other NEON data streams
NEON's measurement strategy will include coordinated and co-located measurements of drivers of environmental change and biological responses	
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	Facilitate spatial integration of NEON data with community-driven investigation
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	
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-yr lifetime	Optimize the design through iterative observation and evaluation of spatial and temporal variability

the high-level requirements that constrain the spatial observation at discrete sites across the continent (Table 1). The high-level requirements provide a starting place—a set of lower-level requirements specific to the sample design result from the integration of these requirements with input from statistically minded members of the ecological community, literature, and other long-term monitoring efforts (Table 1).

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

1. *Direct the collection of the raw data 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 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.
2. *Efficiently capture landscape-scale patterns and trends.* Organisms and soil should be measured with intensity sufficient to detect the presence of spatio-temporal trends 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. Guidance on sample location and number will be directed by the sample design (Urquhart et al. 1998, Thompson 2012). The specific sample size is ultimately determined in the science design associated with each TOS measurement (Thorpe et al. 2016); trend detection, dependent on determination of space-time covariance structures (Cressie and Wickle 2011), will depend on the diversity of analytical approaches applied to the data. Given the variety of research approaches and 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).
3. *Provide infrastructure that co-locates terrestrial measurements and links observations to other NEON data streams.* The terrestrial measurements must be co-located to enable an analytical framework for patterns and processes which vary in space and time and may effect each other (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).
4. *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).
5. *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).
6. *Optimize the design through iterative observation and evaluation of spatial and temporal trends and variability.* The number and spatio-temporal distribution of samples reflects assumptions about the variability of responses, 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 to better understand spatio-temporal patterns over the life of the Observatory. Any adjustments to the design should be made early to maximize the time series from a consistent set of plots (Lindenmayer and Likens 2009).

The high-level NEON requirements capture the essence of the NEON mission and Grand Challenges, 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.

Spatial 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 accommodates a variety of data types and questions (Olsen et al. 1999).

For terrestrial observations that span 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 spatial patterns (Carpenter 2008) and temporal trends across diverse landscapes and meet the needs of contemporary and future ecological paradigms (Cressie et al. 2009) in support of the long-term Observatory, the sample design for terrestrial organisms and biogeochemistry includes the following elements:

1. The *sample frame* is the area from which observations are made (Reynolds 2012).
2. *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.
3. *Stratification* increases efficiency (Cochran 1977) and provides a framework for describing the variability of landscape characteristics targeted by the NEON design.
4. *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).
5. *Sample allocation* allows a distribution of sampling effort appropriate to particular observations and questions.
6. *Data analysis with variance estimators* provides a solution for analysis of data with design-based inference (Stehman 2000).
7. *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 primarily focuses 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 (Thorpe et al. 2016) and taxonomic group of interest. In the case of many of the vegetation and soil observations (Thorpe et al. 2016), the frame typically corresponds to an associated management type or ownership boundary (Fig. 1). This typically includes the

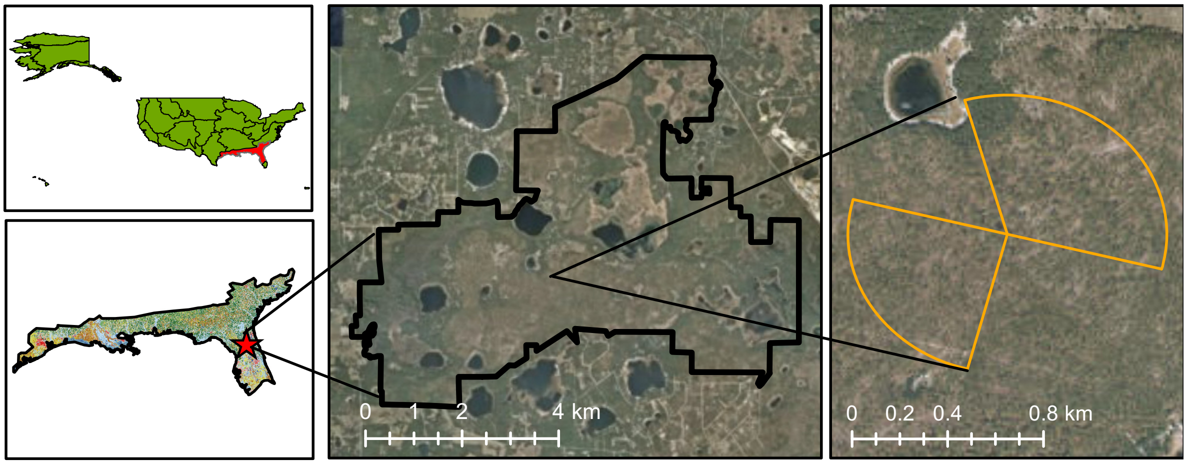

Fig. 1. National Ecological Observatory Network (NEON)'s Domain 03 is located in the southeast United States. The site at the Ordway-Swisher Biological Station in central Florida is managed as a research station by the University of Florida and includes a diversity of pine on sandy soils, broadleaf forests on wetter soils, and wet marshes. The site boundary encompasses a 34-km^2 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.

location of the tower-based sensor measurements and the aquatic measurements at some sites (Thorpe et al. 2016). 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 km^2) to larger wildland sites (e.g., part of Oak Ridge National Lab 67 km^2). At several sites, the area available for sampling is too large 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- to 80-km^2 sampling frame.

The National Ecological Observatory Network'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. 2016) 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 (Fig. 1).

Randomization

The unbiased sample generated by randomization (Cochran 1977, Thompson 2012) is the foundation of the sample design within NEON sites. 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 of these strategies support design-based inference. Sampling areas thought to be representative of a site (i.e., 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 (Lindenmayer et al. 2010). The lack of fundamental randomization results in a sample that is biased and incompatible with design-based inference to the unsampled population(s) (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 Hoef 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) are keys to understand the causes and consequences of change over the life of the Observatory.

Randomization at NEON sites.—Collectively, the design requirements provide a strong case for explicit emphasis on the characterization of spatial patterns. The NEON design addresses these constraints by sampling with a random, spatially balanced sampling framework, resulting in a probability-based study design with low to moderate variance that is both simple and flexible (Stevens and Olsen 2004). Potential sampling locations are generated with the Reversed Random Quadrat-Recursive Raster (RRQRR; Theobald et al. 2007) algorithm that is similar to the Generalized Random Tessellation Stratified (GRTS) method 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 (i.e., distributes potential random and spatially balanced sampling locations across the extent of each site) instead of a defined sample size, providing design flexibility and redundancy to assign alternative locations should a plot be unsuitable for sampling (Theobald et al. 2007). Implementation in GIS facilitates the incorporation of site boundaries, identifies barriers to sampling (e.g., roads, lakes), and enables the plot locations to be viewed in maps.

The RRQRR algorithm provides the foundation of the sample design, which consists of the specific locations for the observation and collection of terrestrial organisms and soil. Every sample unit across the sample frame receives a potential plot location that is numbered in a spatially balanced framework, addressed with a named location, randomized, and ordered in a one-dimensional list. Selecting sequential sample locations from this list provides a random, spatially balanced design (Theobald et al. 2007). Generation of the spatially balanced design is accomplished with the RRQRR function that maps 2-dimensional space into 1-dimensional space. Reversed Random Quadrat-Recursive Raster employs Morton ordering (Theobald et al. 2007), a hierarchical quadrant-recursive ordering. Morton ordering creates a recursive, space-filling address by generating 2×2 quads that are composed of upper left, lower left, upper right, and lower right cells and nested at hierarchical scales. Each cell is assigned a Morton Address that reflects the recursive order as well as a corresponding sequential Morton order. The Morton address is then reversed to achieve a systematic pattern, and the cells at each hierarchical quadrant level are randomized. The algorithm recursively generates these nested, hierarchical quads with associated Morton address, Morton order, reversed Morton address, and randomization until the cell size coincides with the specified sample unit size—a 30×30 m gridded cell in the case of the NEON design—such that a complete sample is generated for the entire sample area. Mapping the two-dimensional space into one-dimensional space as a list sorted by Morton order and sequentially selecting corresponding sample locations from this list results in a spatially balanced, random sample design (Fig. 2). The complete sample generated by the RRQRR algorithm allows design flexibility that is critical to logistical efficiency and sound science. Should a particular plot be unsuitable for sampling, the next unassigned, sequential plot on the ordered list can be included in the sample. Other reasons to include additional plot locations from the complete sample may arise. Results from initial sampling will provide data to direct iterative observations that might require different sample sizes 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.

Stratification

Stratification divides the landscape of interest into non-overlapping subareas from which sample locations are identified (Cochran 1977, Johnson 2012). This 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 ensure the design captures environmental gradients at each site. Stratification according to the National Land Cover Database (NLCD; 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.

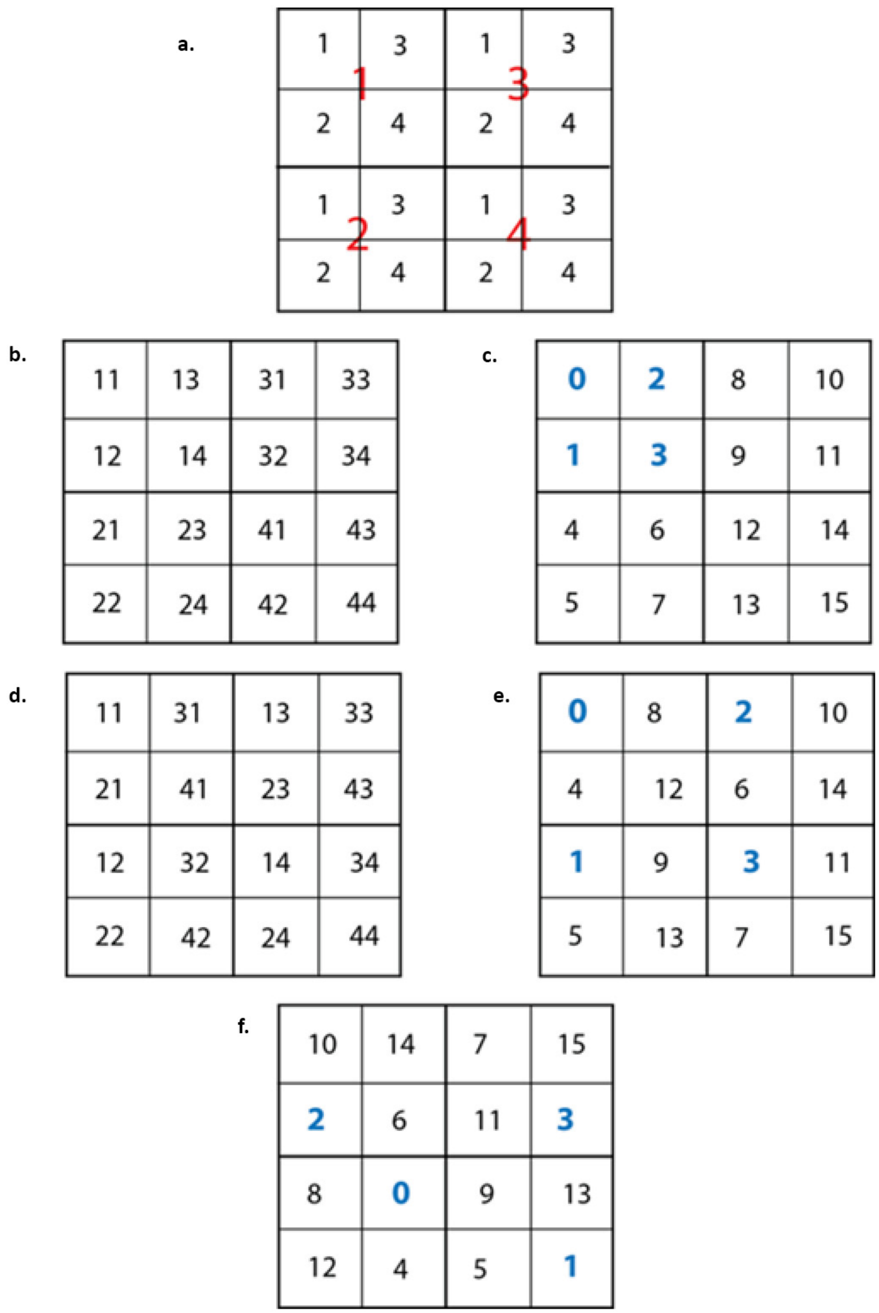

Fig. 2. The spatially balanced Reversed Random Quadrat-Recursive Raster (RRQRR) design for locating sample plots across National Ecological Observatory Network (NEON) sites. RRQRR assigned Morton addresses to a very large number of cells in a raster. The steps to create a spatially balanced list based 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), each cell is assigned a Morton addresses that represents the recursive order; (c) each cell is also assigned a sequential Morton order; (d) the Morton Address is reversed to create a uniform systematic pattern; (e) the systematic pattern is also reflected in the sequential Morton order; (f) and cells are randomized at each quadrant level (Theobald et al. 2007).

First, stratification is an integral part of the NEON design at multiple scales, and when applied to the terrestrial sample design, it provides consistency and ensures observations describe local landscape characteristics essential to the continental-scale Observatory. The National Ecological Observatory Network 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 (Fig. 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 and other ecosystem properties that drive ecological response (Chapin et al. 2012, Clark et al. 2012, Sala et al. 2012). Observing terrestrial biogeochemistry and organisms in dominant vegetation types at each 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.

Second, stratification by land cover allows differential allocation of resources and sampling effort across cover types. 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 facilitate comparison of ecological response across vegetation types within a site and, crucial to the success of the continental Observatory, comparison among NEON sites as well.

Caveats associated with stratification by cover type merit recognition (e.g., vegetation will change over time; Scott 1998). The National Ecological Observatory Network hopes to capture this change, but the choice of dynamic strata will complicate design-based inference (Fancy et al. 2009). As such, NEON will track plot-specific changes in strata and develop statistical methods to deal with dynamic strata adjustments to design-based estimators and the inclusion probability of each sampling stratum (Wikle and Royle 1999, Stevens and Olsen 2004, Luo et al. 2011) that will be available to users. Many other long-term monitoring efforts either do not stratify or

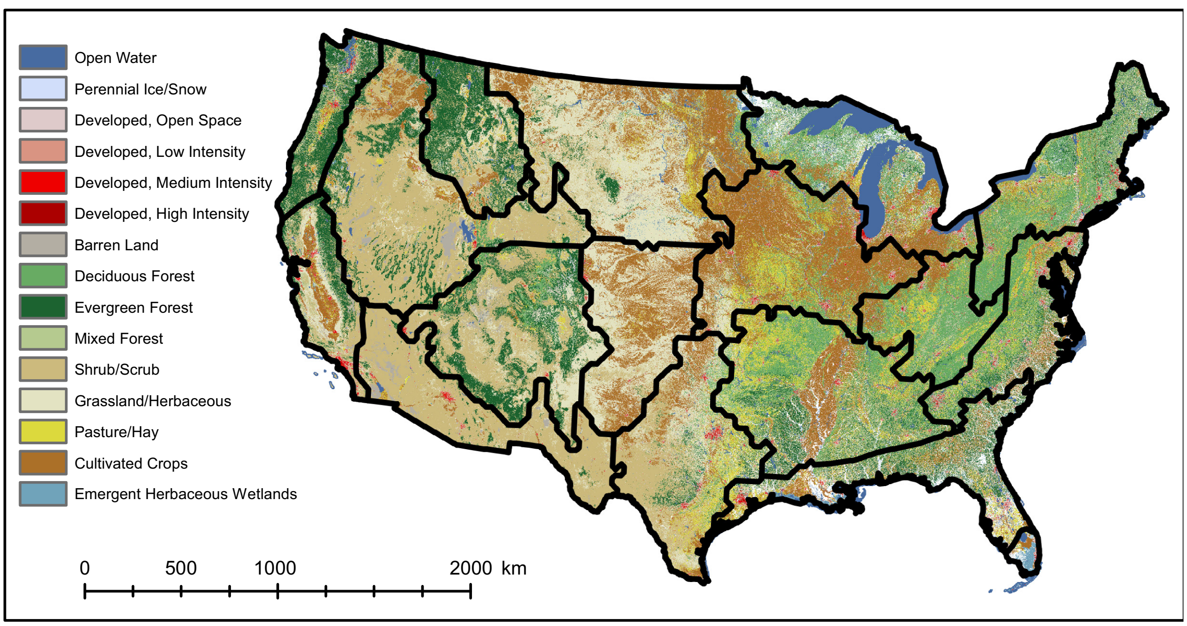

Fig. 3. A subset of NEON domains layered on top of land cover types as described by the National Land Cover Database.

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 is mapped 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 locations was based on vegetation patterns inferred from 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 across all domains.

*Stratification at NEON sites.—*The land cover vegetation strata were based on the NLCD (Fry et al. 2011), which was developed 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 land cover (Fig. 4). In the context of the RRQRR sample design, stratification is achieved by intersecting points from the ordered sample list with each land cover type by assigning an inclusion probability of one for areas associated with the target vegetation type and zero for non-target types. In other words, the ordered one-dimensional list developed by the RRQRR remains unchanged; selecting points within a particular land cover type filters that list such that plots are skipped to distribute plots across target strata, but the ordered list is maintained within each strata. The result is a random, spatially balanced sample design that is stratified by land cover (Fig. 4).

Using a subset of the NEON sampling design and plots as a simple random sample

The spatially balanced, random sampling locations generated by the RRQRR algorithm provide the sample design with flexibility. Users of NEON data in need of a strictly random sample, not the stratified-random sample design, may conduct analyses with a subset—or all in some cases—of the plots sampled by including plots in ordered RRQRR list that are not interrupted for the NLCD stratification. The initial steps of the sample generation (Fig. 2), prior to the filtering of potential plot locations by the NLCD strata, result in a design that conforms to assumptions of a random sample (Theobald et al. 2007). At sites characterized by a single NLCD type, the NEON design is analogous to a simple random design (Table 2). With multiple strata, potential viable sample locations (non-viable plots are skipped for safety and logistical challenges, etc.) from the initial one-dimensional ordered list are only skipped to allow the ordered allocation of target sample sizes (see *Minimum sample size* and *Sample allocation*) across each NLCD type. Those plots that adhere to the one-dimensional RRQRR list without interruption for stratification purposes can be treated as a simple random sample (Theobald et al. 2007). The number of sample locations and the fraction of the total sampling effort that can be considered random depend on site size, heterogeneity, and the evenness of target strata. Every sample location can be considered random at homogeneous sites, while those sites characterized by numerous strata result in a relatively smaller sample size available to any analysis dependent on a random sample (Table 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. This design flexibility makes the data more broadly available to a variety of NEON data consumers, ecological questions, and statistical applications.

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 organisms, soil, 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 these data are sufficient to describe 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

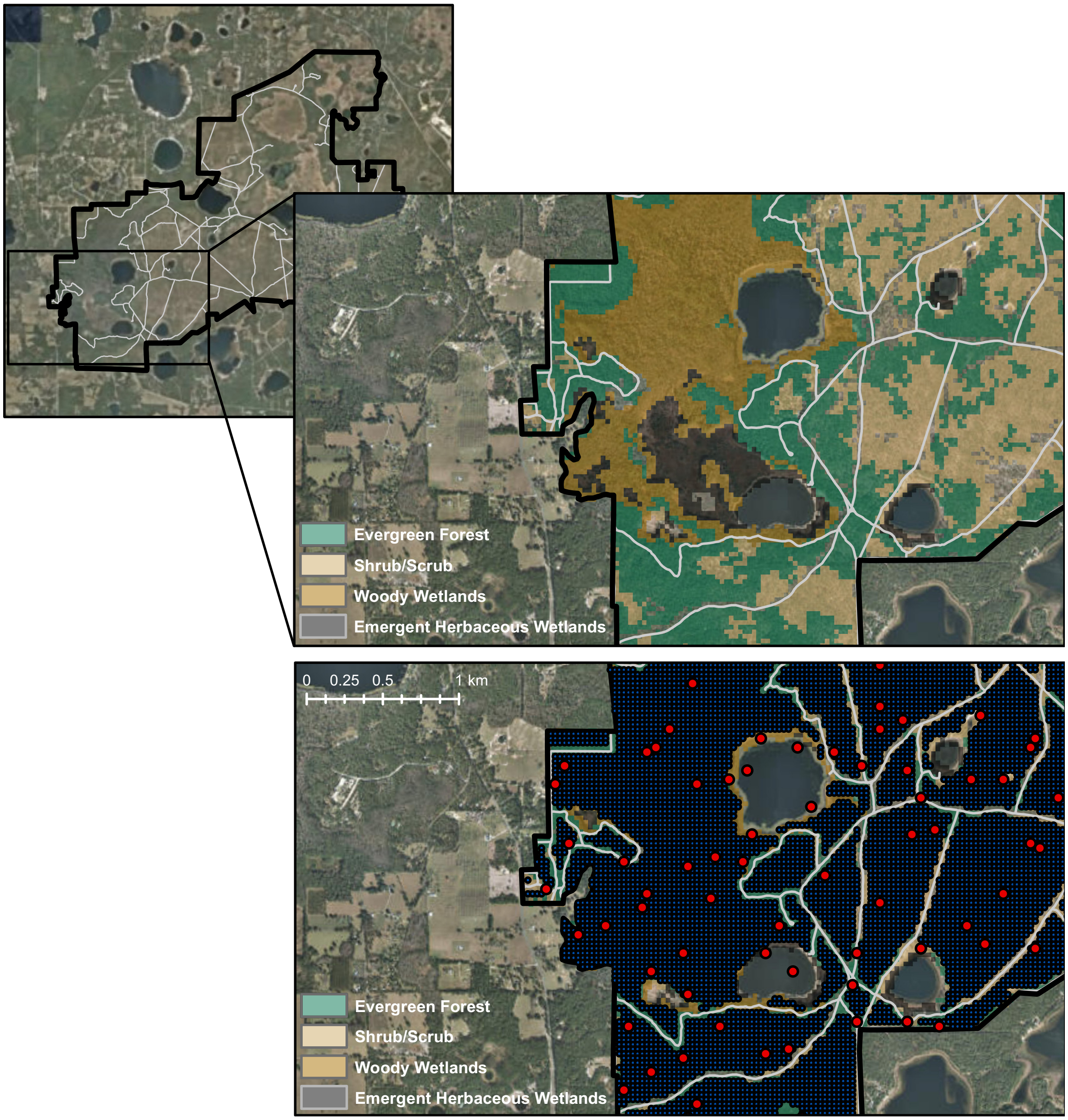

Fig. 4. Stratification by the National Land Cover Database at the Ordway-Swisher Biological Station (a). Blue dots represent potential sampling locations from the spatially balanced and randomized sample, and red points indicate hypothetical sample locations selected from the complete sample (b).

Millspaugh 2012) across 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, and 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

Table 2. The sample design for Distributed Plots sampling within National Ecological Observatory Network sites follows a stratified-random design.

Site, plot subtype, and NLCD cover types	Stratified-random plots: Area (km^2)	Stratified-random plots: No. plots	No. of random plots
KONZ			
Base plot		30	19
Grassland/herbaceous	29.8	23	
Deciduous forest	3.3	7	
Mosquito point		10	10
Grassland/herbaceous	4.9	9	
Deciduous forest	0.3	1	
Mammal grid		6	5
Grassland/herbaceous	28.2	4	
Deciduous forest	3.1	2	
Tick plot		6	3
Grassland/herbaceous	29.8	4	
Deciduous forest	3.3	2	
Bird grid		12	7
Grassland/herbaceous	29.8	9	
Deciduous forest	3.3	3	
TALL			
Base plot		30	10
Deciduous forest	16.6	10	
Evergreen forest	18.2	11	
Mixed forest	13.8	9	
Mosquito point		10	1
Deciduous forest	1.8	3	
Evergreen forest	3.1	4	
Mixed forest	1.6	3	
Mammal grid		8	3
Deciduous forest	15.4	3	
Evergreen forest	15.9	3	
Mixed forest	12.4	2	
Tick plot		6	5
Deciduous forest	16.6	2	
Evergreen forest	18.2	2	
Mixed forest	13.8	2	
Bird grid		15	4
Deciduous forest	16.6	5	
Evergreen forest	18.2	5	
Mixed forest	13.8	5	
JORN			
Base plot		30	30
Shrub/scrub	45.7	30	
Mosquito point		10	10
Shrub/scrub	45.7	10	
Mammal grid		6	6
Shrub/scrub	45.7	6	
Tick plot		6	6
Shrub/scrub	45.7	6	
Bird grid		7	7
Shrub/scrub	45.7	10	

Notes: However, an inherent flexibility in the generation of these sample location allows a subset of Distributed Plots to 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.

questions 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 the temporal trends of a given response of interest between two populations of interest? Examples of specific questions enabled by NEON data might include:

1. 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?
2. How do trends in invasive plant species richness differ between a wildland site and a site managed for cattle grazing in Domain 12?
3. How do temporal patterns of plant phenology (e.g., bud burst) differ between high and low elevation sites in Domain 17?

The 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 disparate populations of responses) in specific parameters to constrain the sample size. This approach does not provide a unique solution; rather it provides a range of minimum sample sizes that correspond to the range of parameter values considered. In this way, the differences in minimum sample size as a function of the populations considered can be accounted for. 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., 30 m × 30 m 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 (1971) and utilizes the approach of 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:

$$\begin{aligned} Y_i = {} & \mu_0 + \mu_{0i} + \alpha_1 \times \text{time} + \beta_{1i} \times \text{time} \\ & + \alpha_2 \times \text{site} + \beta_{\text{int}} * (\text{site} \times \text{time}) + \varepsilon_i \end{aligned} \tag{1}$$

where Y_i is a vector representing observations through time t (i.e., the number of repeat measurements) at the ith sampling location; with respect to measurement i, μ_{0i} is a random intercept, β_{1i} is a random slope of time for the ith sampling location; μ_0 is a fixed intercept; 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); and $\boldsymbol{\varepsilon}_i$ is a vector representing errors through time t (i.e., the number of repeat measurements) at the ith sampling location.

The parameters (Eq. 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 as $\tau = (\mu_0, \alpha_1, \alpha_2, \beta_{\text{int}})$. Using this grouping of the parameters, the Eq. 1 can be re-written as

$$Y_i = \mathbf{X}_i\tau + \mathbf{M}_i\lambda_i + \boldsymbol{\varepsilon}_i \tag{2}$$

where $\mathbf{X}_i$ is a design matrix with t rows and p columns, and $\mathbf{M}_i$ is a matrix with t rows and q columns. Here $q \leq p$ and the columns of $\mathbf{M}_i$ are also columns of $\mathbf{X}_i$.

This formulation (Eq. 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).

$$\frac{\hat{\beta}_{\text{int}}}{\sqrt{\text{Var}\left(\hat{\beta}_{\text{int}}\right)}} \sim \text{N}(0,1) \tag{3}$$

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

$$n = \frac{\left(Z_{(1-\alpha/2)} + Z_\beta\right) \times \left(\mathbf{X}_1^T\mathbf{V}^{-1}\mathbf{X}_1 + \mathbf{X}_2^T\mathbf{V}^{-1}\mathbf{X}_2\right)^{-1}}{{\beta_{\text{int}}}^2} \tag{4}$$

where Z represents the quantile from the standard normal distribution corresponding to the desired error rate for the type I and type II errors; $\mathbf{X}_1$ is the design matrix corresponding to samples of one population of interest; $\mathbf{X}_2$ is the design matrix corresponding to the samples of the other population of interest; $\mathbf{V}$ 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 initial data collection at NEON sites. Ultimately, four levels of population variance were assessed (Table 3).

In the absence of time series data, temporal parameters were estimated with ten years of MODIS-derived Normalized Difference Vegetation Index (NDVI) data 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 (Fig. 5). Correlations of these NDVI data informed the range of temporal correlations initially specified in the sample size calculations (Fig. 5, Table 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 display trends more closely aligned with an autoregressive framework.

Type I error tolerance was assessed at levels of 0.05 and 0.10. 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 >20% different from one another.

In the case of the compound symmetric temporal specification, there was a monotonic, yet nonlinear relationship between the number of

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

samples, the temporal correlation, the population variance, and collection of data through time (Fig. 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. 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 (when type I error rate = 0.01), 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 3). The magnitude of the correlation associated with the autoregressive process demonstrated a lack of monotonicity between the number of samples and the number of years data is collected (Fig. 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 if 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 >20% different from one another.

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 (>5% spatial 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

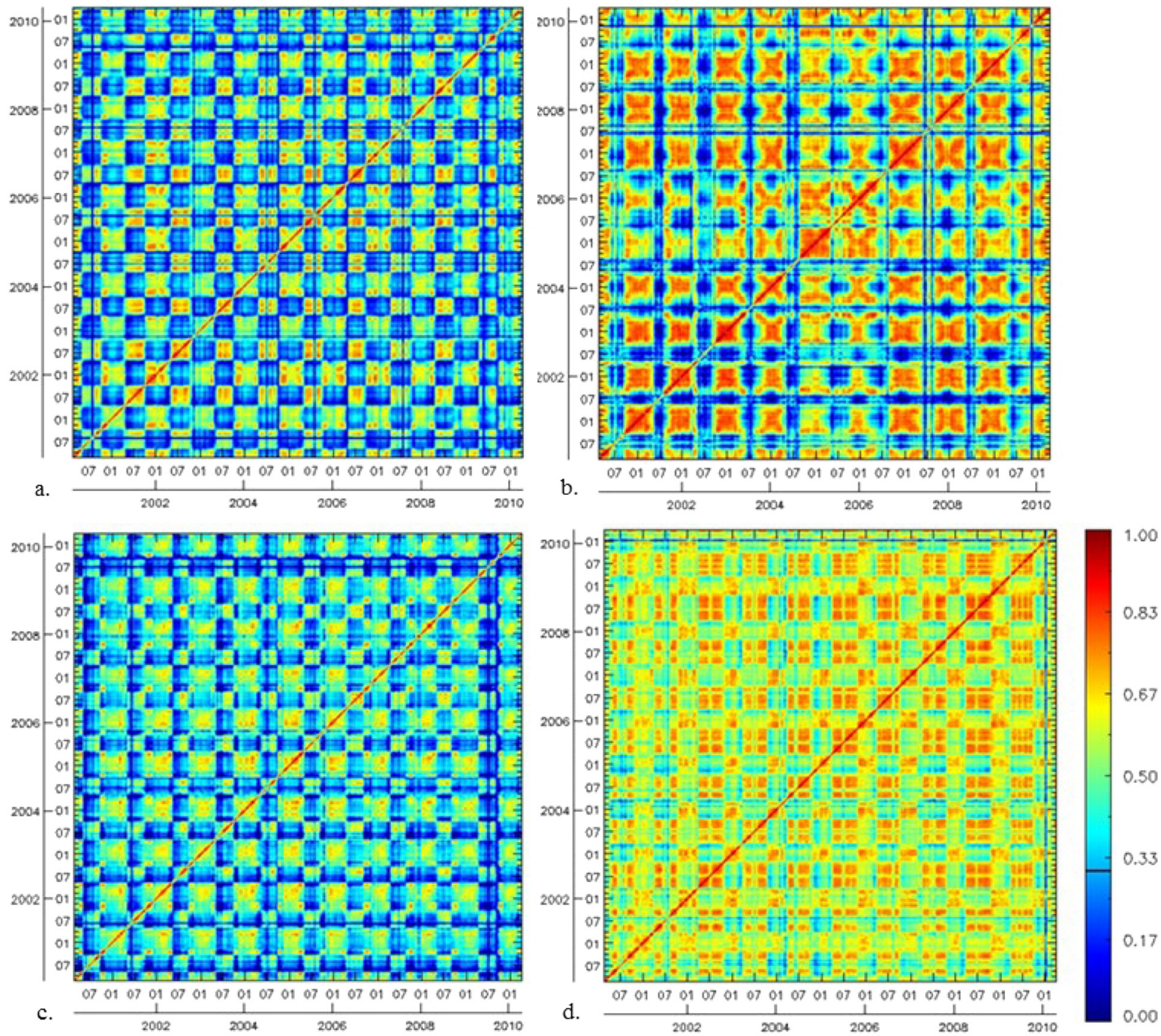

Fig. 5. Annual temporal correlations from 2000 to 2010 of normalized difference vegetation index (NDVI) at Harvard Forest (a), Ordway-Swisher Biological Station (b), University of Notre Dame Environmental Research Station (c), and Oak Ridge National Lab (d). The lack of a consistent decay in temporal 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.

selected to represent dominant vegetation types across the NEON domains. The co-located terrestrial measurements will focus on quantifying variability within 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 5% and 10% of aerial coverage. Given a fixed sampling effort, there is a trade-off in selecting the level for inclusion of vegetation classes between 5% and 10%; sampling vegetation types <10% (but >5%) pulls samples away from the more representative vegetation classes.

Excluding rare vegetation is not without trade-offs. Disproportionate numbers of species may be endemic to rare vegetation types (Stohlgren et al. 1998), and rare vegetation types might be

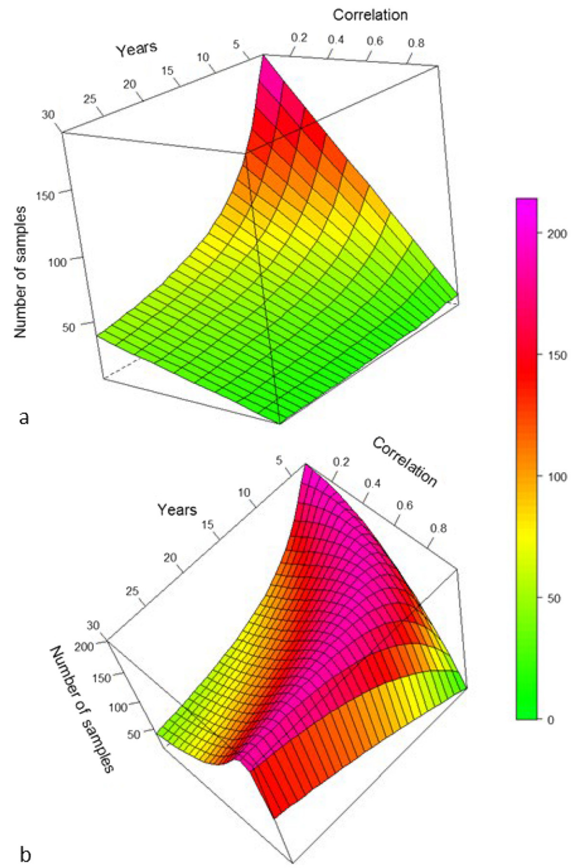

Fig. 6. Minimum sample size as a function of years and temporal correlation for the compound symmetric correlation structure (a) and the autoregressive structure (b) with the type I error set at 0.1 (Data S1).

differentially susceptible to environmental change (Stohlgren 2007, Suding et al. 2008). These rare types (e.g., riparian corridors or ecotones) may be targeted in iterative sampling efforts or by efforts organized by members of the ecological community.

Data analysis with variance estimators.—Data collected according to the spatially balanced and stratified-random design is robust to a variety of design estimation and modeling techniques (Sarndal 1978, Cressie et al. 2009). While a particular approach might benefit from a model-based sample design or stratification conducive to a specific question, most analytical and data assimilation approaches can accommodate data based on principles of randomization. Perhaps the most simple approach to inference leverages the probabilistic nature of random design with design-based inference (Reynolds 2012). In the context of the NEON data, design-based inference can be handled by simply treating the data as a simple random sample when samples are allocated proportional to strata area. The samples of all TOS protocols from Distributed Plots, Grids, and Points except plant diversity are allocated in proportion to the area of the strata. In these cases, observations carry equal information content to the larger population. This self-weighting sample—the sample weights are equal—allows the resulting data to be treated as a simple random sample when calculating statistical moments (Cochran 1977, Lohr 2010). However, in cases when there is lower variance within strata, these simple random sample estimators will have higher variance than estimators specifically designed to handle data from a stratified-random design.

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

$$\bar{y}_{\mathrm{strat}} = \frac{1}{\mathrm{N}} \sum_{i=1}^{S} \mathrm{N}_i \bar{y}_i \tag{5}$$

where $\bar{y}_i$ is the sample mean from the ith stratum; N_i is the number of units in the ith stratum; N is the number of units across all strata; and S is the number of strata.

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

$$\widehat{\mathrm{Var}}(\bar{y}_{\mathrm{strat}}) = \sum_{i=1}^{S} \left(\frac{\mathrm{N}_i}{\mathrm{N}}\right)^2 \left(\frac{\mathrm{N}_i - n_i}{\mathrm{N}_i}\right) \frac{s_i^2}{n_i} \tag{6}$$

where s_i^2 is the sample variance from the ith stratum and n_i is the number of units in the sample from the ith stratum.

The estimators adjust for differences in sampling intensity within each strata by, in the case of the mean, dividing the number of units in each stratum by the number of units across all strata (Eq. 5), and sample variance of each stratum by the number of units in the sample from each stratum in the case of the variance estimator (Eq. 6). The area is computed using the 30-m^2 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

$$\bar{y}_{\mathrm{strat}} \pm Z_{(\alpha/2)} \times \left(\widehat{\mathrm{Var}}(\bar{y}_{\mathrm{strat}})\right)^{1/2} \tag{7}$$

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

Few of the sites in the initial implementation will have strata with sufficiently large samples that allow this approximation (Eq. 7). For strata with sample sizes smaller than 30, Thompson (2012) suggests using a t-distribution with degrees of freedom approximated using Satterthwaite's method

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

where d is the Satterthwaite approximation for the degrees of freedom and

$$a_i = \mathrm{N}_i(\mathrm{N}_i - n_i)/n_i \tag{9}$$

where a_i = the variance coefficients.

Adaptation

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 due to:

1. 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).
2. Plots may intersect buildings, roads, or other developments or natural features such as rock formations that are not suitable for NEON sampling.
3. The location may be inaccessible due to steep slopes or other natural features that pose danger to field technicians.
4. The time to travel to remote locations may make the observation too costly. The National Ecological Observatory Network 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.

5. NLCD classification error will result in plot locations that do not land in the target vegetation type. If the misclassification is limited in extent, the next appropriate plot in the target NLCD class can be selected. If NLCD class does not exist at a site, erroneous characterizations may need to be corrected with the NEON remote sensing platform.

Linking continuous surfaces with ground-based point measurements will provide new ways to measure ecological patterns and trends (Ollinger et al. 2008). Where remote sensing proxies for ground measurements are robust, or in the rare case that there is a 1:1 comparison between a ground measurement and a remotely sensed measurement, the airborne data approximate 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.

For many processes, NEON will not be able to determine whether 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 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 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 be capable of sampling such that cohesive integration of drivers and response can be achieved. Detecting these relationships and temporal trends across multiple taxonomic groups is a challenge, and time and rigorous analyses are required to determine the efficacy of NEON data in this context. The National Ecological Observatory Network 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 efforts will provide unbiased data products that can be assimilated into both design- and model-based approaches to statistical inference for the efficient detection of trends scalable within the context of the NEON design.

Acknowledgments

The National Ecological Observatory Network is a project sponsored by the NSF. The following grants supported this work: EF-1029808, EF-1138160, and

DBI-0752017. The authors would like to thank Alan Gelfand, Dave Theobald, Andy Royle, and Scott Urquhart for comments and contributions to the design and/or this manuscript, as well as NEON staff and the scientific community that provided guidance, expertise, and critical review of the NEON design. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Literature Cited

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 USA 105:4519–4523.

Barnett, D., et al. *In press*. The plant diversity sampling design for the National Ecology Observatory Network. Ecosphere.

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.

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. Second edition. Springer, New York, New York, USA.

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 367:236–246.

Cochran, W. G. 1977. Sampling techniques. Third edition. John Wiley and Sons, New York, New York, USA.

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. Wickle. 2011. Statistics for spatio-temporal data. John Wiley and Sons, New York, New York, USA.

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.

Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. Methods in Ecology and Evolution 1:330–342.

Elmendorf, S. C., et al. 2016. The plant phenology monitoring design for the National Ecological Observatory Network. Ecosphere 7:e01303.

Fancy, S. G., and R. E. Bennetts. 2012. Design and analysis of long-term ecological monitoring studies. Pages 481–497 *in* R. A. Gitzen, J. J. Millspaugh, C. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies. 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. Millspaugh. 2012. Ecological monitoring: the heart of the matter. Pages 3–22 *in* R. A. 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, UK.

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., et al. 2016. The soil and plant biogeochemistry sampling design for The National Ecological Observatory Network. Ecosphere 7: e01234.
Hoekman, D., et al. 2016. Design for mosquito abundance, diversity, and phenology sampling within the National Ecological Observatory Network. Ecosphere 7:e01320.
Hoekman, D., et al. 2017. Design for ground beetle abundance and diversity sampling within the National Ecological Observatory Network. Ecosphere 8:e01744.
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. Wikle, 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. Pages 54–73 *in* R. A. 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, UK.
Kao, R. H., et al. 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3:115.
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 USA 107:21957–21962.
Lohr, S. L. 2010. Sampling: design and analysis. Brooks/Cole, Boston, Massachusetts, USA.
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. Pages 101–125 *in* R. A. 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, UK.
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., et al. 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 USA 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. Pages 23–53 *in* R. A. 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, UK.
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 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., et al. 2011. 2011 science strategy: enabling continental-scale ecological forecasting. The National Ecological Observatory Network, Boulder, Colorado, USA.

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 27:1–76.

Smith, D. R., Y. Lei, H. A. Walter, and J. A. Young. 2012. Incorporating predicted species distribution in adaptive and conventional sampling designs. Pages 381–396 *in* R. A. 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, UK.

Springer, Y. P., et al. 2016. Tick-, mosquito-, and rodent-borne parasite sampling designs for the National Ecological Observatory Network. Ecosphere 7:e01271.

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, New York, USA.

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., et al. 2010. A multiscale and multidisciplinary investigation of ecosystem-atmosphere CO_2 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.

Thompson, S. K. 2012. Sampling. John Wiley & Sons, Hoboken, New Jersey, USA.

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:e01627.

Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. Ecological Applications 8:246–257.

Ver Hoef, J. 2002. Sampling and geostatistics for spatial data. Ecoscience 9:152–161.

Vitousek, P. M. 1997. Human domination of Earth's ecosystems (vol 277, pg 494, 1997). 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.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2540/full