TOS SCIENCE DESIGN FOR SPATIAL SAMPLING

PREPARED BY	ORGANIZATION	DATE
David Barnett	SCI	06/07/2017
Paul Duffy	Neptune Inc.	02/28/2013
David Schimel	NASA JPL	02/28/2013
Elena Azuaje	SCI	02/28/2013
Rachel Krauss	SCI	02/28/2013
Andrea Thorpe	SCI	02/28/2013
David Gudex-Cross	SCI	02/28/2013
Joel McCorkel	NASA Goddard	02/28/2013

APPROVALS	ORGANIZATION	APPROVAL DATE
Kate Thibault	SCI	04/06/2022

RELEASED BY	ORGANIZATION	RELEASE DATE
Tanisha Waters	CM	04/06/2022

See configuration management system for approval history.

The National Ecological Observatory Network is a project solely funded by the National Science Foundation and managed under cooperative agreement by Battelle. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Change Record

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	10/26/2015	ECO-02648	Initial Release
B	08/22/2017	ECO-04861	Changes include: • Reorganization of text in Section 6.5.2 Initial minimum sample size at NEON sites • Update to Section 6.7 to include possible calculation of moments without design-based estimators when samples are allocated proportional to strata area • Other minor text edits
C	04/06/2022	ECO-06790	• Revised logo • Update to reflect change in terminology from relocatable to gradient sites

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

NEON design documents are required to define the scientific strategy leading to high-level protocols for NEON subsystem components, linking NEON Grand Challenges and science questions to specific measurements. Many NEON *in situ* measurements can be made in specific ways to enable continental-scale science rather than in ways that limit their use to more local or ecosystem-specific questions. NEON strives to make measurements in ways that enable continental-scale science to address the Grand Challenges. Design Documents flow from questions and goals defined in the NEON Science Strategy document, and inform the more detailed procedures described in Level 0 (L0; raw data) protocol and procedure documents, algorithm specifications, and Calibration/Validation (CalVal) and maintenance plans.

1.2 Scope

This document defines the rationale and requirements for Terrestrial Observation System (TOS) Science Design for Spatial Sampling in the NEON Science Design.

1.3 Acknowledgements

The authors would like to thank Frank Davis, Alan Gelfand, John Gross, Kathi Irvine, and Andy Royle for comments and contributions to this document.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Applicable Documents

AD[01]	NEON.DOC.000278	Tier 4 TOS Requirements Module
AD[02]	NEON.DOC.000001	NEON Observatory Design
AD[03]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products Catalog
AD[04]	NEON.DOC.000908	TOS Science Design for Microbial Diversity
AD[05]	NEON.DOC.000912	TOS Science Design for Plant Diversity
AD[06]	NEON.DOC.000914	TOS Science Design for Plant Biomass, Productivity, and Leaf Area Index
AD[07]	NEON.DOC.000907	TOS Science Design for Plant Phenology
AD[08]	NEON.DOC.000915	TOS Science Design for Small Mammal Abundance and Diversity
AD[09]	NEON.DOC.000911	TOS Science Design for Vectors and Pathogens
AD[10]	NEON.DOC.000906	TOS Science Design for Terrestrial Biogeochemistry
AD[11]	NEON.DOC.000916	TOS Science Design for Breeding Landbird Abundance and Diversity
AD[12]	NEON.DOC.000907	TOS Science Design for Plant Phenology
AD[13]	NEON.DOC.000909	TOS Science Design for Ground Beetle Abundance and Diversity
AD[14]	NEON.DOC.000910	TOS Science Design for Mosquito Abundance, Diversity and Phenology

2.2 Reference Documents

RD[01]	NEON.DOC.000008 NEON Acronym List
RD[02}	NEON.DOC.000243 NEON Glossary of Terms
RD[03]	National Research Council (2001) Grand Challenges in Environmental Sciences. 107 pp. The National Academies Press, Washington, D.C.
RD[04]	Schimel, D., W. Hargrove, F. Hoffman, and J. MacMahon (2007) NEON: a hierarchically designed national ecological network. Frontiers in Ecology and the Environment 5:59–59.
RD[05]	Keller, M, DS Schimel, WW Hargrove, FM Hoffman (2008) A continental strategy for the National Ecological Observatory Network. Frontiers in Ecology and the Environment 6:282-284.

3 INTRODUCTION

3.1 Overview of the Observatory

The National Ecological Observatory Network (NEON) is a continental-scale ecological observation platform for understanding and forecasting the impacts of climate change, land use change, and invasive species on ecology. NEON is designed to enable users, including scientists, planners and policy makers, educators, and the general public, to address the major areas in environmental sciences, known as the Grand Challenges (**Figure 1**). NEON infrastructure and data products are strategically aimed at those aspects of the Grand Challenges for which a coordinated national program of standardized observations and experiments is particularly effective. The open access approach to the Observatory's data and information products will enable users to explore NEON data in order to map, understand, and predict the effects of humans on the earth and understand and effectively address critical ecological questions and issues. Detailed information on the NEON design can be found in AD[01], AD[02].

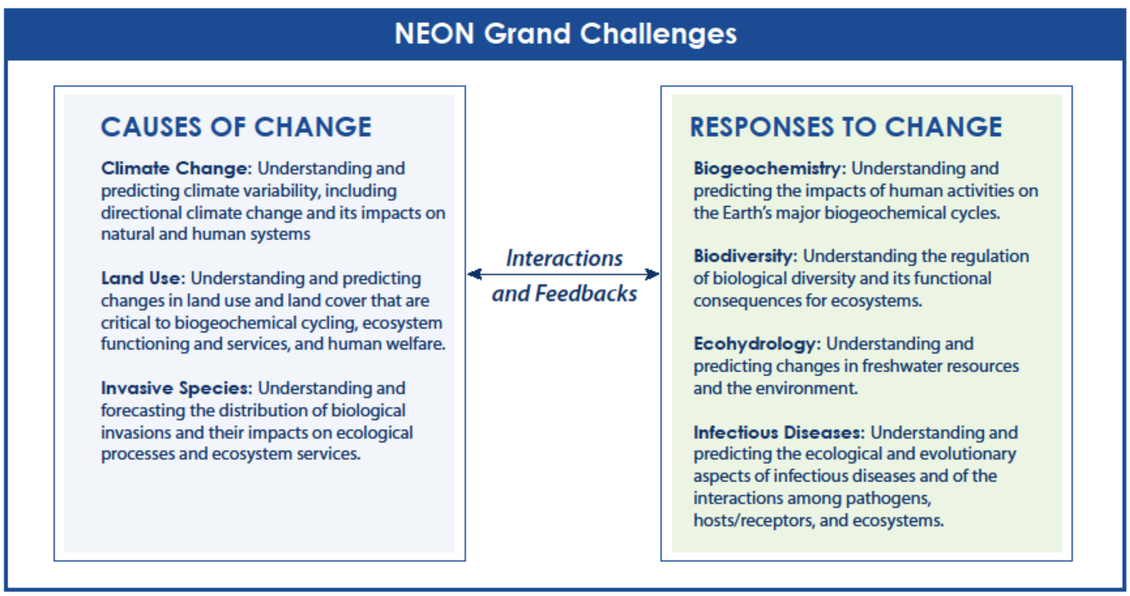

Figure 1. The seven Grand Challenges defined by the National Research Council (2001).

3.2 Components of the Observatory

There are five components of the Observatory, the Airborne Observation Platform (AOP), Terrestrial Instrument System (TIS), Aquatic Observation System (AOS), Aquatic Instrument System (AIS), and Terrestrial Observation System (TOS). Collocation of measurements associated with each of these components will allow for linkage and comparison of data products. For example, remote sensing data provided by the (AOP) will link diversity and productivity data collected on individual plants and stands by the (TOS) and flux data captured by instruments on the tower (TIS) to that of satellite-based remote sensing. For additional information on these systems, see Keller et al. 2008, Schimel et al. 2011.

3.3 The Terrestrial Observation System

The NEON TOS will quantify the impacts of climate change, land use, and biological invasions on terrestrial populations and processes by sampling key groups of organisms (sentinel taxa), infectious disease, soil, and nutrient fluxes across system interfaces (air, land, and water) (AD[01], AD[02]). The sentinel taxa were selected to include organisms with varying life spans and generation times, and wide geographic distributions to allow for standardized comparisons across the continent. Many of the biological measurements will enable inference at regional and continental scales using statistical or process-based modeling approaches. The TOS sampling design captures heterogeneity representative of each site to facilitate this inference when possible. Plot and organism-scale measurements will also be coordinated with the larger-scale airborne measurements, which provide a set of synergistic biological data products at the regional scale. Details of these design elements and algorithms can be found in individual design documents available through the NEON website (AD[04]–AD[14]).

The standardization of protocols across all sites is key to the success of NEON (and its novelty) and must be maintained at all sites through time. Thus, although specific techniques may be required at some sites (e.g., due to different vegetation types), protocols have been developed to ensure data comparability. These details can also be found in individual design documents available through the NEON website (e.g., AD[04]–AD[14]; www.NEONScience.org).

The TOS Science Designs define the scientific strategies leading to high-level sampling designs for NEON sentinel taxa, terrestrial biogeochemistry, and infectious disease, linking NEON Grand Challenges and science questions to specific measurements (AD[02]). The TOS Spatial Sampling Design document describes the sampling design that collocates observations of the components of the TOS. TOS Science Design documents were developed following input from the scientific community, including module-specific Technical Working Groups, and the National Science Foundation (AD[02]). Science Designs will be reviewed periodically to ensure that the data collected by NEON are those best suited to meet the requirements of the observatory (AD[01]), are (to the extent possible) consistent with standards used by the scientific community, and fit within the scope of NEON. Additional information on the development and review process can be found in AD[02].

4 INTRODUCTION TO THE TERRESTRIAL OBSERVATION SYSTEM SAMPLING DESIGN

4.1 Background

The National Ecological Observatory Network's (NEON) goal is to improve understanding and forecasting of ecological change at continental scales over decades (Schimel et al. 2011). The design co-locates measurements of atmosphere, soil, water, select organisms and disease, and airborne observations. Observing change by integrating measures of the drivers and ecological responses will contribute to improved understanding of ecological cause and effect (Vitousek 1997, Keller et al. 2008, Luo et al. 2011). High-level requirements derived from the NEON goal and mission guide the architecture of the design and infrastructure for the Observatory (Schimel et al. 2011). The primary requirement-driven constraint of the design is that it must assemble sufficient data collected at points and local regions (400-km^2) to enable extrapolation of these functional relationships to the scale of the continent over the course of several decades. The requirements framework permeates the NEON design, providing guidance for the design of observations and direct traceability back to the NEON mission.

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 (~400km^2) with high-resolution remote sensing at annual time steps and tower-based sensors capture temporally continuous fluxes over smaller spatial extents (~0.5km^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 (~5-60km^2). A complete census of organisms and biogeochemistry is biologically and financially impractical – 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.

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. 2016). 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 (AD[04] – AD[14]).

4.2 NEON's Contribution

This document describes a flexible, design-based sample design to direct the spatial distribution of terrestrial organism and biogeochemistry sampling at NEON sites. It is primarily designed to direct the NEON sampling efforts, but the design is also capable of accommodating auxiliary investigation by independent observers who hope to leverage NEON observations.

4.3 Purpose and Scope

This document integrates high-level science requirements and logistical constraints to provide a framework for the spatial distribution of sampling of terrestrial organisms and biogeochemistry. The justification, rationale, and study design for each of these response variables is described in other science design documents (AD[04]–AD[14]).

5 SAMPLING FRAMEWORK

5.1 Science Requirements

This science design is based on Observatory science requirements that reside in NEON's Dynamic Object-Oriented Requirements System (DOORS). Copies of approved science requirements have been exported from DOORS and are available in NEON's document repository, or upon request.

5.2 Data Products

This Science Design results in spatial metadata describing the spatial location of data collection of terrestrial organisms and biogeochemistry at NEON sites. The spatial location and other information (e.g., elevation, land cover type) are published as metadata on the NEON data portal and with Terrestrial Observation System data products.

5.3 Priorities of and Challenges for the Terrestrial Observation System Sampling Design

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 Observatory data provide a framework for the NEON design. The terrestrial observation sample design is part of this hierarchical structure. “Upstream” requirements and “downstream” data products provide context and constraints under which the design was developed. 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**.

Table 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
· 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 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.	✓ Direct the collection of the raw material for continental ecology

· 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. · 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.	✓ Efficiently capture landscape-scale pattern and trend
· 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.	✓ Provide infrastructure that co-locates terrestrial measurements and links observations to other NEON data streams
· 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.	✓ Facilitate spatial integration of NEON data with community-driven investigation
· 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

NEON developed a continental-scale design to systematically sample national variability of ecological characteristics and to allow extrapolation of local observations to the scale of the continent. The United States was partitioned into domains – twenty regions defined by a statistically rigorous clustering algorithm that grouped similar fractions of ecoclimatic variance (Hargrove and Hoffman 1999, 2004, Keller et al. 2008). The NEON sample design will characterize this continental-scale variability among domains by implementing a consistent within-domain strategy of selecting sites (e.g., Harvard Forest, Konza Prairie Biological Station, and the Northern Range of Yellowstone National Park) that are most representative of the within-domain ecoclimatic variability for long-term, intensive sampling. This high-level stratification forms a basis for the continental design, and is an integral part of the overall NEON strategy for scaling observed patterns across space and through time ((Schimel et al. 2011), **Figure 1**). These principles are carried through in the development of the local, site-scale sample design.

The sample design for observations at local, site-specific scales must deliver data that optimally informs continental-scale ecology. Adopting the requirements framework provides 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 1**; **Figure 2**). In addition to facilitating comparison across sites and at continental scales, the design must satisfy the collection of demands imposed by the unique aspects of each measurement, collocate terrestrial observations, and facilitate the integration of data with other biological and physical measurements of the observatory (Schimel et al. 2011). Maintaining generality encourages the iterative optimization of the sample effort while allowing it to remain robust to a range of questions and methods of analysis which the community may apply to NEON data products.

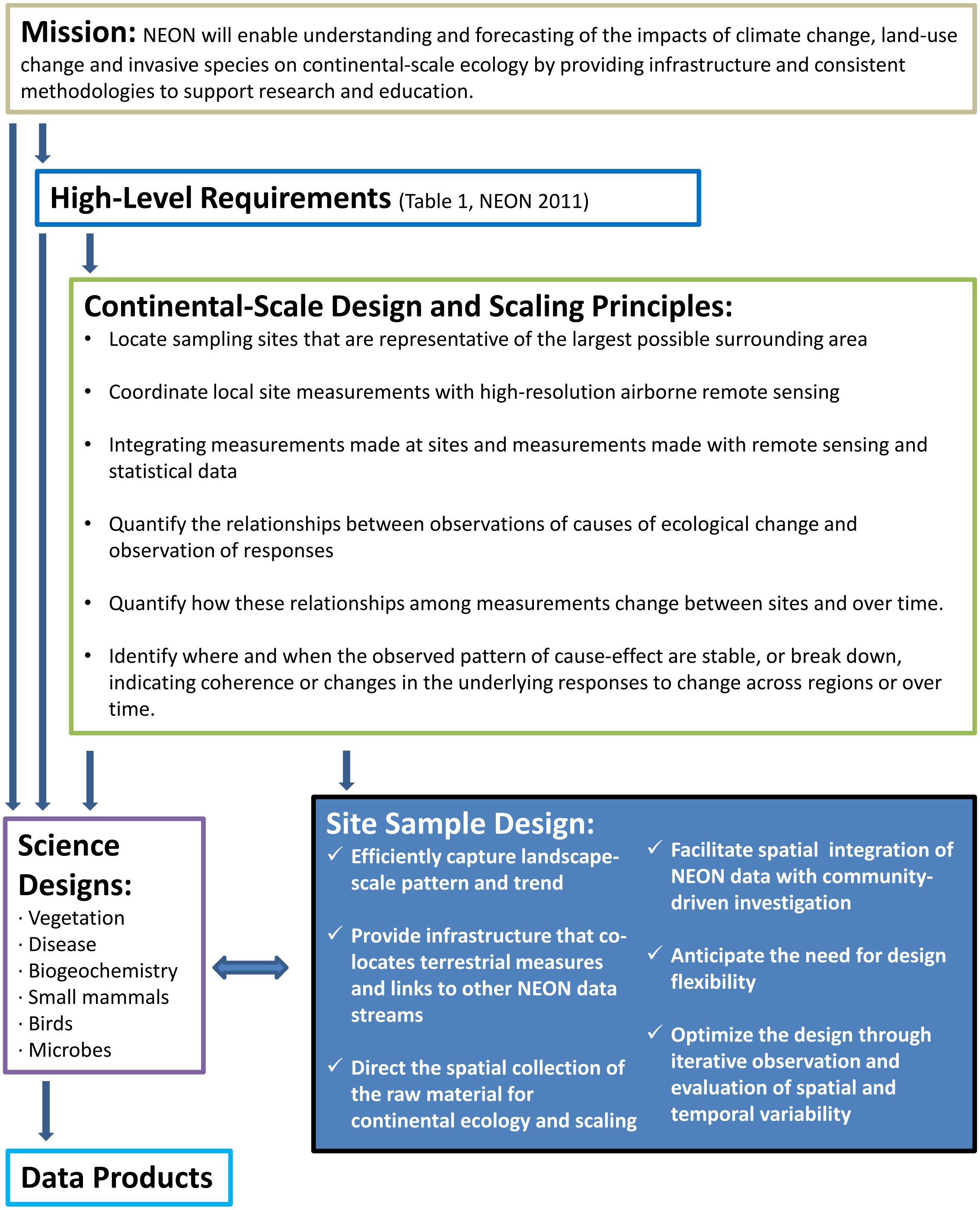

Figure 2. The NEON mission, requirements, and science designs constrain the local, site-specific sample design. The site-specific sample design scales principles and practices of the continental observatory design. Specifically, the sample design provides a framework for the spatial-temporal distribution and sample effort needed to inform continental-scale ecology. The science questions, goals, suite of observations, and data distribution considered are described in NEON (2011).

A set of lower-level requirements specific to the sample design link directly to the high-level NEON requirements (**Table 1**). These provide context for the terrestrial observation sample design that specifies the site-specific location and number of observations needed to adequately describe local patterns and trends. The sample design requirements include:

- ***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 goal of the terrestrial sampling design is to direct the observation of terrestrial organisms and biogeochemistry endpoints for long-term trend detection within specific NEON sites, facilitate comparability across sites as well as with other ecological investigations, and contribute to the understanding of the cause and consequence of ecological change.

6 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 facilites 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 accomodating 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 sptial 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 terresterial 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.

6.1 Overview of the Terresstrial Observation System Sampling Strategies

A site-level spatial sampling design that can be applied consistently across NEON domains has been developed for the Terrestrial Observation System. Many of the specific locations of the Terrestrial Observatory System sampling elements are collocated with each other and with environmental sensors (i.e., within the flux zone of the tower) to allow comparison of the data streams. Within a site, organismal and soil sampling for the TOS has been collocated to the extent possible to optimize linkages between data products. TOS sampling occurs with three different plot types:

- **Tower Plots** provide a direct link between NEON's Terrestrial Observation System and Terrestrial Instrument System platforms. Measurements in these plots include above and below-ground plant productivity and biomass, plant diversity, soil biogeochemistry, and soil microbial community diversity and abundance. In addition~~,~~ individual plants are marked for phenological observation along a square 'loop' transect. Tower Plots are placed in the 90% flux area of the primary and secondary (if applicable) airsheds of each NEON tower. If the requisite number of plots cannot be established in the airshed(s), Tower Plots can also be placed outside the airsheds not further from the tower than the length of the vector defining the extent of the primary airshed (e.g., vector length is 220m at Abby Road and vector length is 1380m at Wind River Experimental Forest). Tower Plot locations for observations of plant biomass, productivity, diversity and foliar biogeochemistry, and soil biogeochemistry and microbial diversity and abundance are selected from this realm (available for each site on the NEON data portal) according to a spatially balanced, random design (see section 6.3). Phenology Plots were subjectively located. More information about Tower Plots can be found in Appendix 1.

- **Distributed Plots** are established in an effort to describe organisms and processes throughout NEON sites. Observations of vegetation, soil and beetles are collocated to maximize the value of data streams. In an effort to maintain collocation but minimize disturbance to target taxa, sampling for ticks, breeding birds, and small mammals often occur at or adjacent to these observations of vegetation, soil and beetles. Mosquito traps, due to design and logistical constraints, are located in close proximity to accessible roads. Distributed Plot sampling locations within NEON sites (see section 6.2) are selected according to a stratified-random and spatially balanced design (sections 6.3 and 6.4), but a subset of these locations can be treated as a purely random sample (section 6.3.1). More information about Distributed Plots can be found in Appendix 1.

- **Gradient Plots** are established to better capture site-level gradients in vegetation structure, leaf area index, or plant canopy chemistry. For these sites, Gradient Plots will be established using a targeted, non-random approach informed by NEON's Aerial Observation Platform remote-sensing data. These plots may include subplots for sampling plant biodiversity, soil and plant biogeochemistry, and soil microbiota.

6.2 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 and taxonomic group of interest. In the case of vegetation and soil observations at Distributed Plots, the frame corresponds to an associated management or ownership boundary (**Figure 3**). This typically includes the location of the tower-based sensor measurements and the aquatic measurements at some sites. 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 at NEON is variable, from small landscapes (e.g., an agricultural site near 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 to be sampled given budget and travel constraints. 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 km^2 sampling frame.

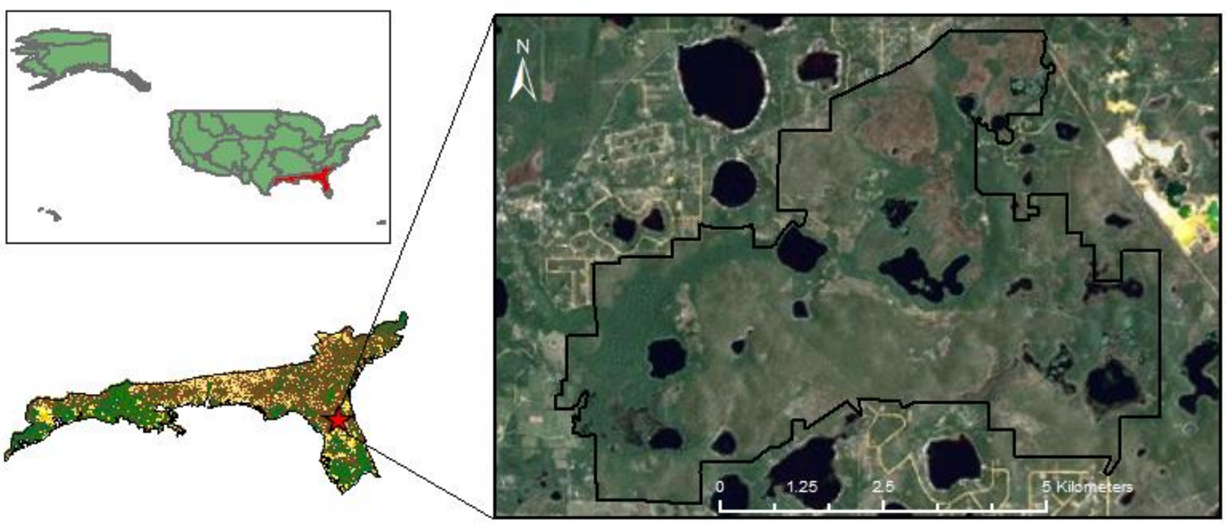

Figure 3. The Ordway-Swisher Biological Station in central Florida is managed as a research station by the University of Florida and includes a diversity of pine on sandy soils, broadleaf forests on wetter soils, and wet marshes. The site boundary encompasses a 34 km^2 area.

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. 2016, Appendix 1) 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 4**).

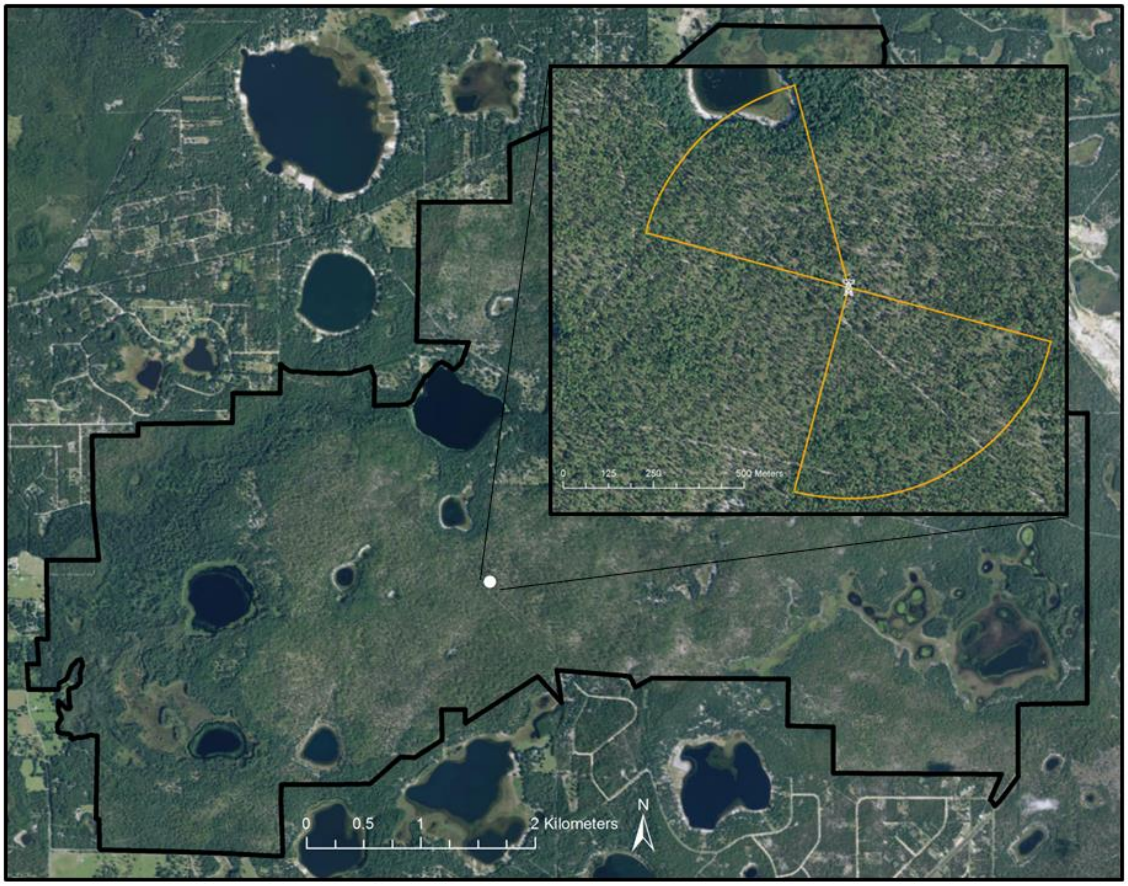

Figure 4. The Ordway-Swisher tower (in white) mostly observes two areas of the landscape, primary and secondary airsheds (in orange) that comprise the sample frame for Tower Plots at the site.

6.3 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 landsape, 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.

6.3.1 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. 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 (**Figure 5**) 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 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.

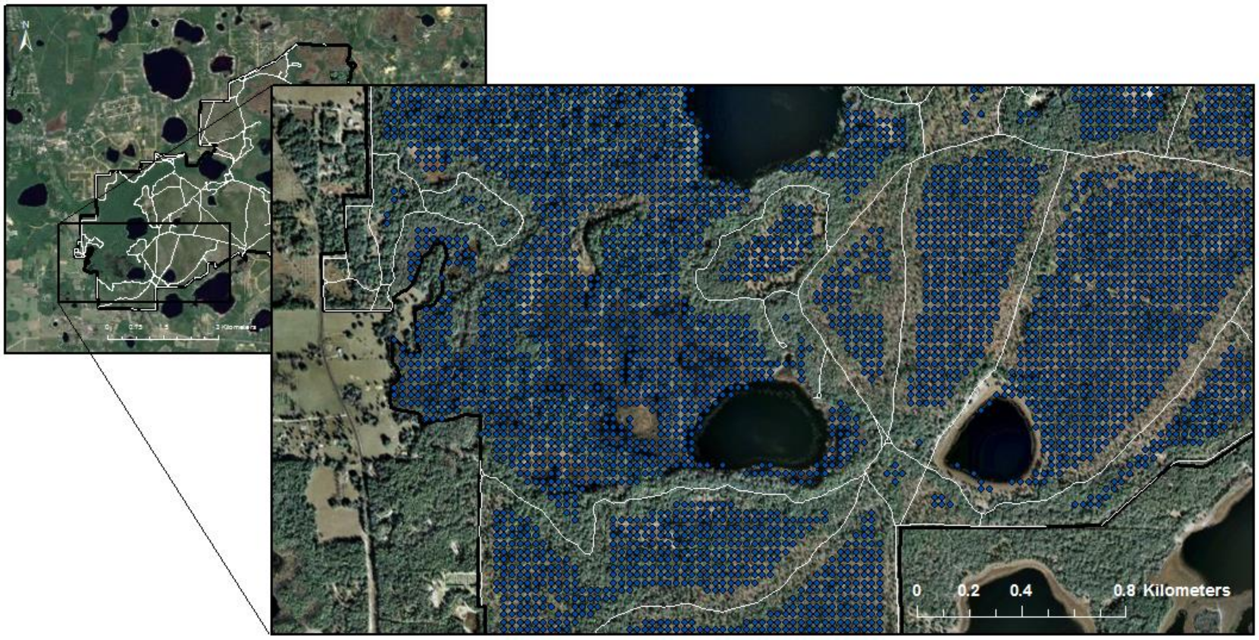

Figure 5. The complete Reversed Random Quadrat-Recursive Raster (RRQRR) sample displayed in a portion of the Ordway-Swisher Biological Reserve in Florida displays a potential sample location (blue) in each 30-m pixel. Areas unsuitable to the sampling of terrestrial organisms and soil were removed from the target population. Examples of 'exclusion' areas include the roads (buffered by 50 m to prevent plots from intersecting roads), power lines (50 m), and standing water (10 m). While not a target of the NEON sampling effort, sampling locations exist across the entire sample frame should these areas be of interest to complimentary sampling efforts.

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 2x2 quads that are composed of lower-left, upper-left, lower-right, and upper-right cells numbered and nested at hierarchical scales (**Figure 6**). 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).

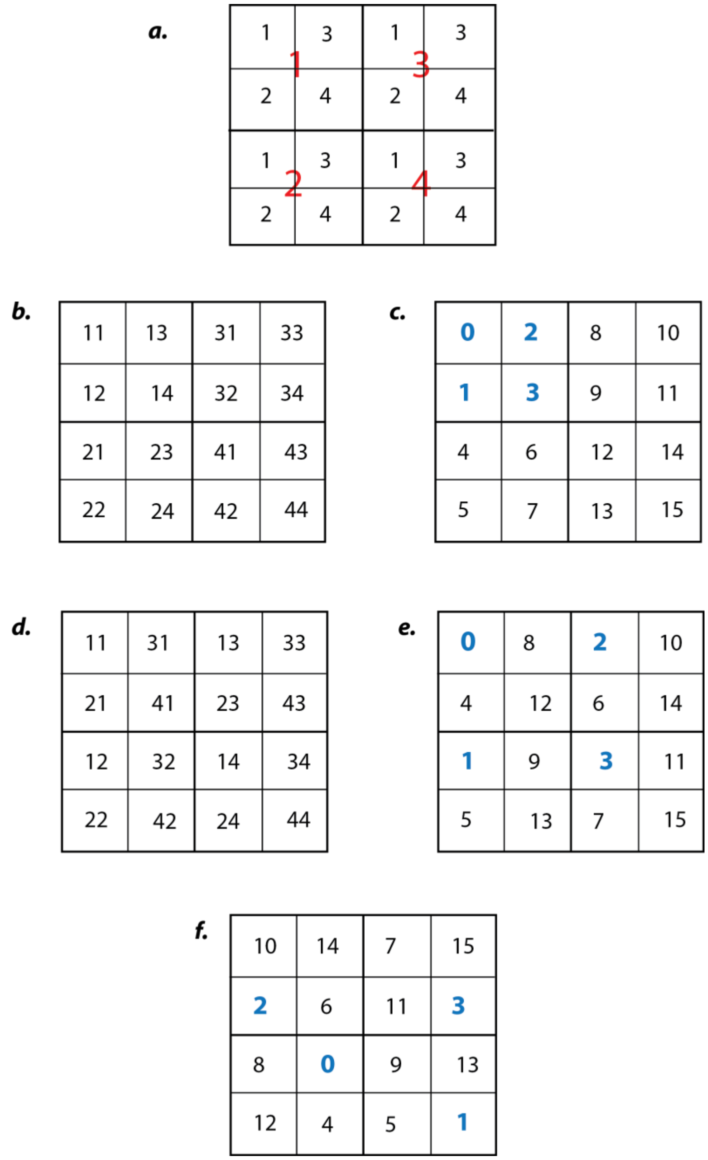

Figure 6. The spatially balanced RRQRR design for locating sample plots across NEON sites. RRQRR assigned Morton addresses to a very large number of cells in a raster. The steps to create a spatially balanced list based on the RRQRR design include (a) the recursive order formation of the Morton Address 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).

6.4 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 stratificaiton 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 7**). 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 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. 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 variablility associated with an ecological response across the landscape allows for a distribution of observations that will stabilize variance of estimators among strata. Appoximately 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, comparision 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 track plot-specific changes in strata and develop statistical methods to deal with dynamic strata that will be available to data users through the NEON document library (Wikle and Royle 1999, Stevens and Olsen 2004, Luo et al 2011). Many 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.

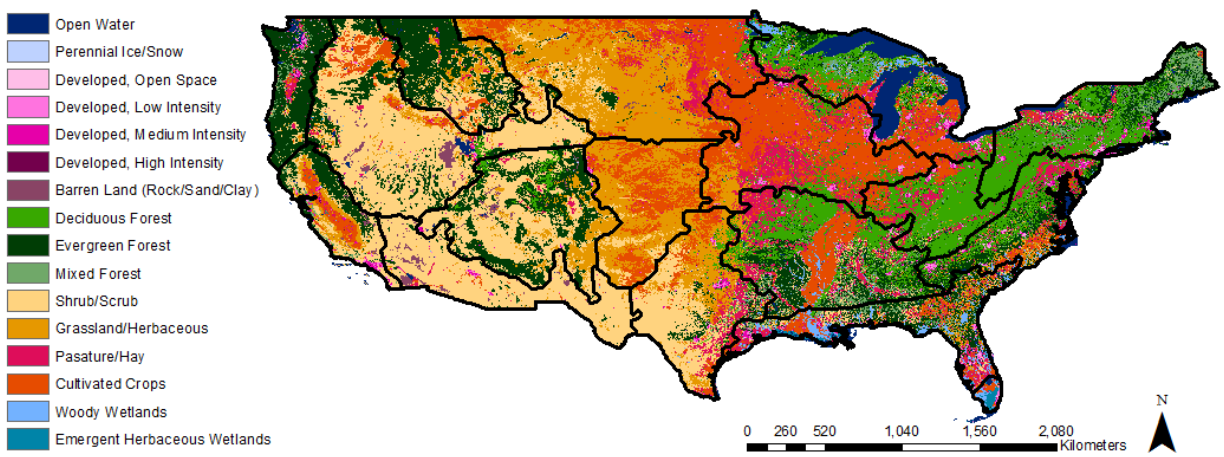

Figure 7. A subset of NEON domains layered on top of land cover types as described by the National Land Cover Database.

6.4.1 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 7** and **Figure 8**). 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 to 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 vegetation (**Figure 9**).

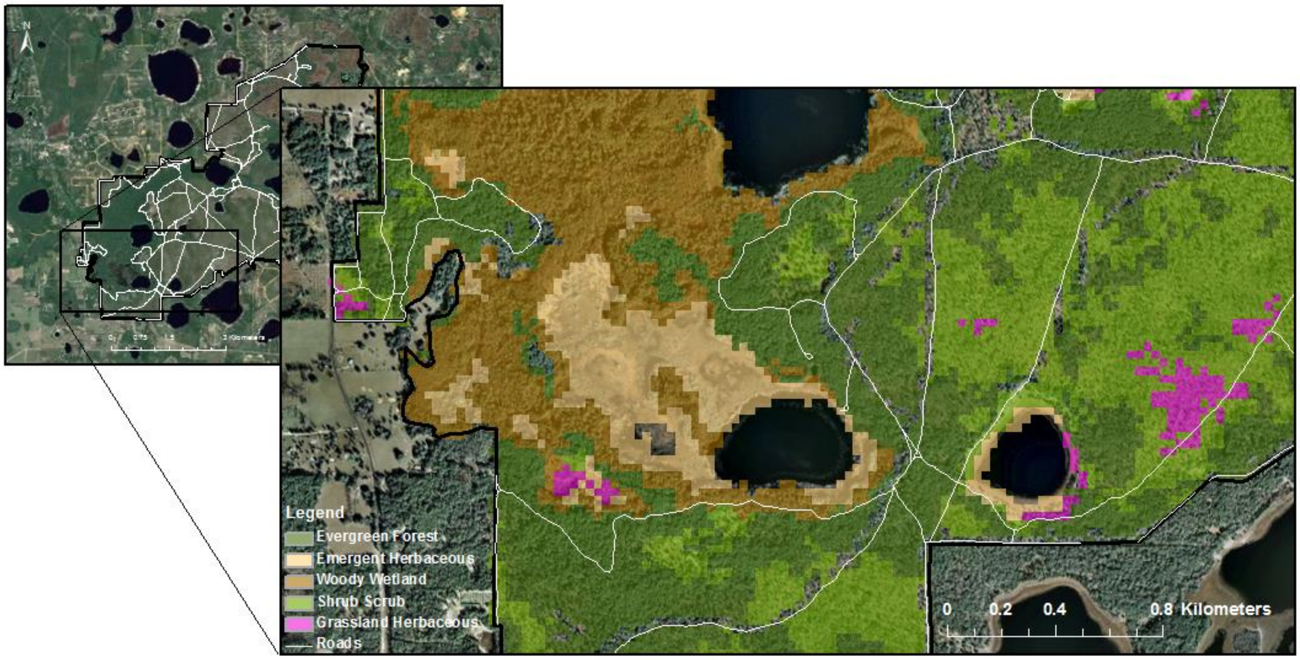

Figure 8. Stratification by the National Land Cover Database at the Ordway-Swisher Biological Station.

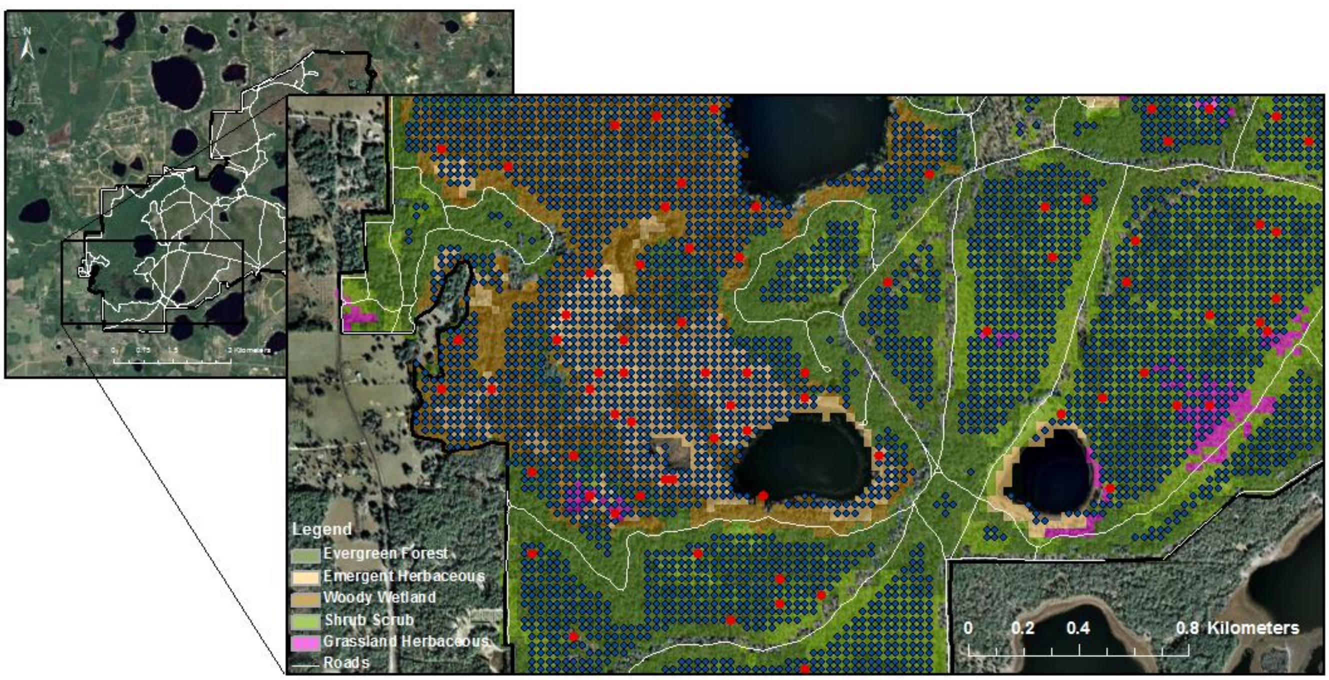

Figure 9. The Spatially balanced and randomized sample at the Ordway-Swisher Biological Station. Red points indicate plots selected from the complete sample. Plots are selected from a spatially-balanced, one dimensional list that is filtered by vegetation type.

6.5 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. The initial steps of the sample generation (Figure 6), 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 analagous to a simple random design (**Table 2**). With multiple strata, potential viable (non-viable plots are skipped for safety and logistical challeges etc.) sample locations from the initial one-dimensional ordered list are only skipped to allow the ordered allocation of target sample sizes (see 6.6 Minimum sample size and 6.7 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. 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 homogeneous sites, while those sites represented 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 flexibilty makes the data more broadly available to a variety of NEON data consumers, ecological questions, and statistical applications.

Table 2. The sample design for Distributed Plots sampling within NEON sites follows a stratified-random design. 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.

Site	**Subtype**	**Stratified-random plots**			**Number of random plots**
		NLCD cover type	**Area (km^2)**	**Number of plots**	
KONZ	Base plot	Grassland/herbaceous Deciduous forest	29.8 3.3	23 7 Total: 30	19
KONZ	Mosquito point	Grassland/herbaceous Deciduous forest	4.9 0.3	9 1 Total: 10	10
KONZ	Mammal grid	Grassland/herbaceous Deciduous forest	28.2 3.1	6 2 Total: 8	5
KONZ	Tick plot	Grassland/herbaceous Deciduous forest	29.8 3.3	4 2 Total: 6	3
KONZ	Bird grid	Grassland/herbaceous Deciduous forest	29.8 3.3	9 3 Total: 12	7

TALL	Base plot	Deciduous forest Evergreen forest Mixed forest	16.6 18.2 13.8	10 11 9 Total: 30	10
TALL	Mosquito point	Deciduous forest Evergreen forest Mixed forest	1.8 3.1 1.6	3 4 3 Total: 10	1
TALL	Mammal grid	Deciduous forest Evergreen forest Mixed forest	15.4 15.9 12.4	3 3 2 Total: 8	3
TALL	Tick plot	Deciduous forest Evergreen forest Mixed forest	16.6 18.2 13.8	2 2 2 Total: 6	5
TALL	Bird grid	Deciduous forest Evergreen forest Mixed forest	16.6 18.2 13.8	5 5 5 Total: 6	4
JORN	Base plot	Shrub/scrub	45.7	30	30
JORN	Mosquito point	Shrub/scrub	45.7	10	10
JORN	Mammal grid	Shrub/scrub	45.7	6	6
JORN	Tick plot	Shrub/scrub	45.7	6	6
JORN	Bird grid	Shrub/scrub	45.7	10	7

6.6 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 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 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 patterns 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 are temporal patterns of bud burst different between high and low elevation sites 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 within the RRQRR grid at each site. Values for the parameters in the statistical model that have relevance to these calculations must be specified based on 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.

6.6.1 Initial Sample Size Calculation

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 following: specification of the number of repeat measurements within the course of the study, the correlation structure, and the magnitude of the correlation associated with the repeated measures. 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 ith sampling location
- with respect to measurement *i*, μ_{0i} is a random intercept, β_{1i} is a random slope of time for the
- μ_0 is a fixed intercept
- ith sampling location
- $\alpha_1$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 ith 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,

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

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 ≤ 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).

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

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,

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

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

6.6.2 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 (⍰^{2}) was estimated across the groups of organisms to be sampled 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). Ultimately, four levels of population variance were assessed (**Table 3**).

The significant difference at which the power constraint is imposed also required specification. 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.

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 10**). Correlations of these NDVI data informed the range of temporal correlations (?) initially specified in the sample size calculations (**Figure 10**, **Table 3**).

The form of the temporal correlation structure - compound symmetric or first order autoregressive correlation structures were considered (Yi and Panzarella, 2002) - 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.

In the case of the compound-symmetric temporal specification there was a monotonic, yet non-linear relationship between the number of samples, the correlation, the population variance, and collection of data through time (**Figure 11**, **Table 3**). 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 5-22, 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 are collected (**Figure 12**). Although there was little evidence for the use of the autoregressive correlation structure, it is likely some of the sites will actually display trends more closely aligned with an autoregressive framework (**Figure 10**). Specifically, the monotonic behavior associated with the compound symmetric correlation structure did not translate to the results obtained by using the first order autoregressive model. This finding is similar to results from Yi and Panzarella (2002).

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.

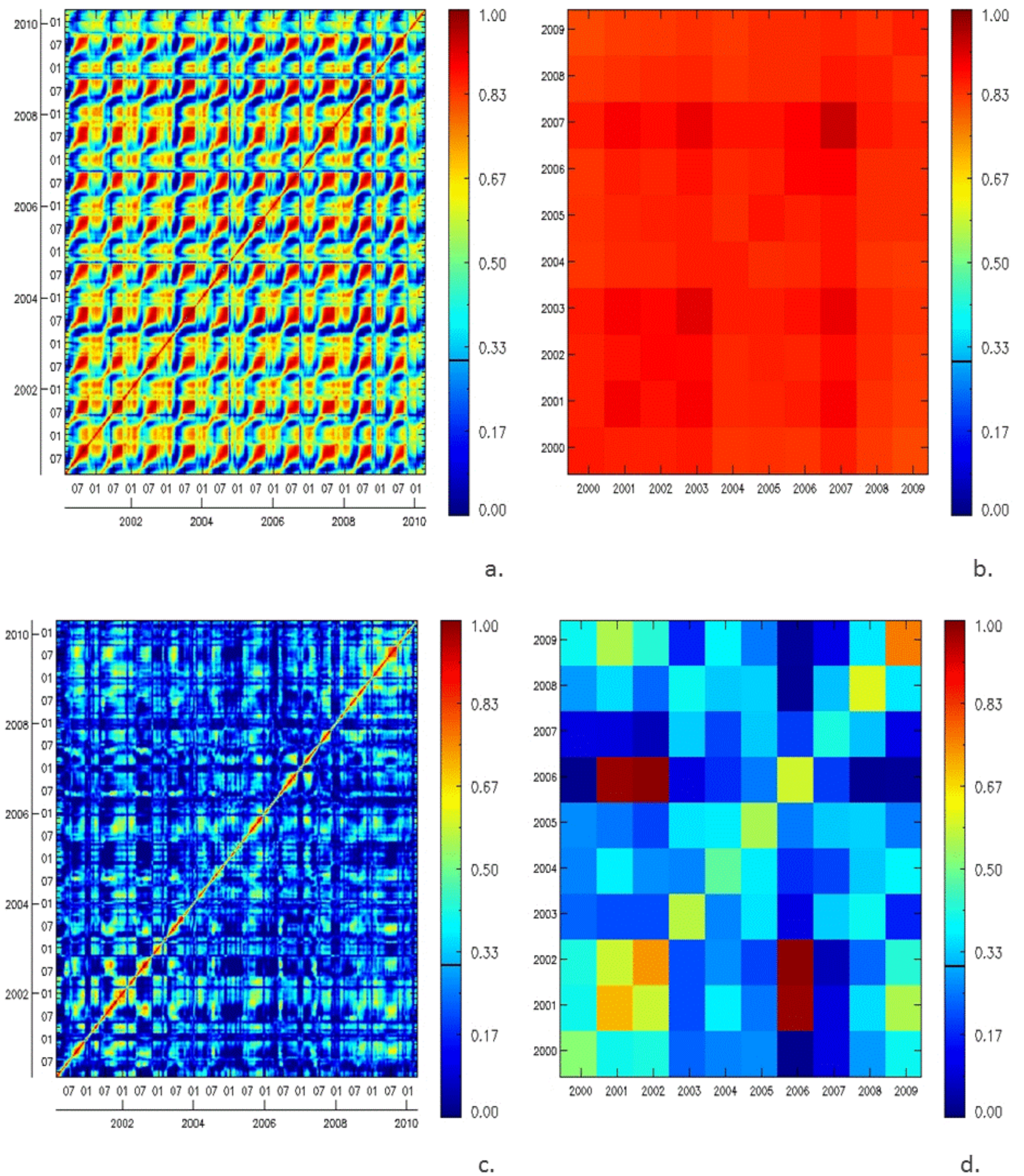

Figure 10. Annual temporal correlations from 2000 - 2010 of normalized difference vegetation index (NDVI) at Yellowstone National Park (a) and the Central Plains Experimental Range (c) and site-wide correlations averaged over peak greenness interval of ninety-five days (Julian Day 165 – 260) at Yellowstone National Park (b) and over a ninety-one day interval (Julian Day 166 – 257) at the Central Plains Experimental Range (c). 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.

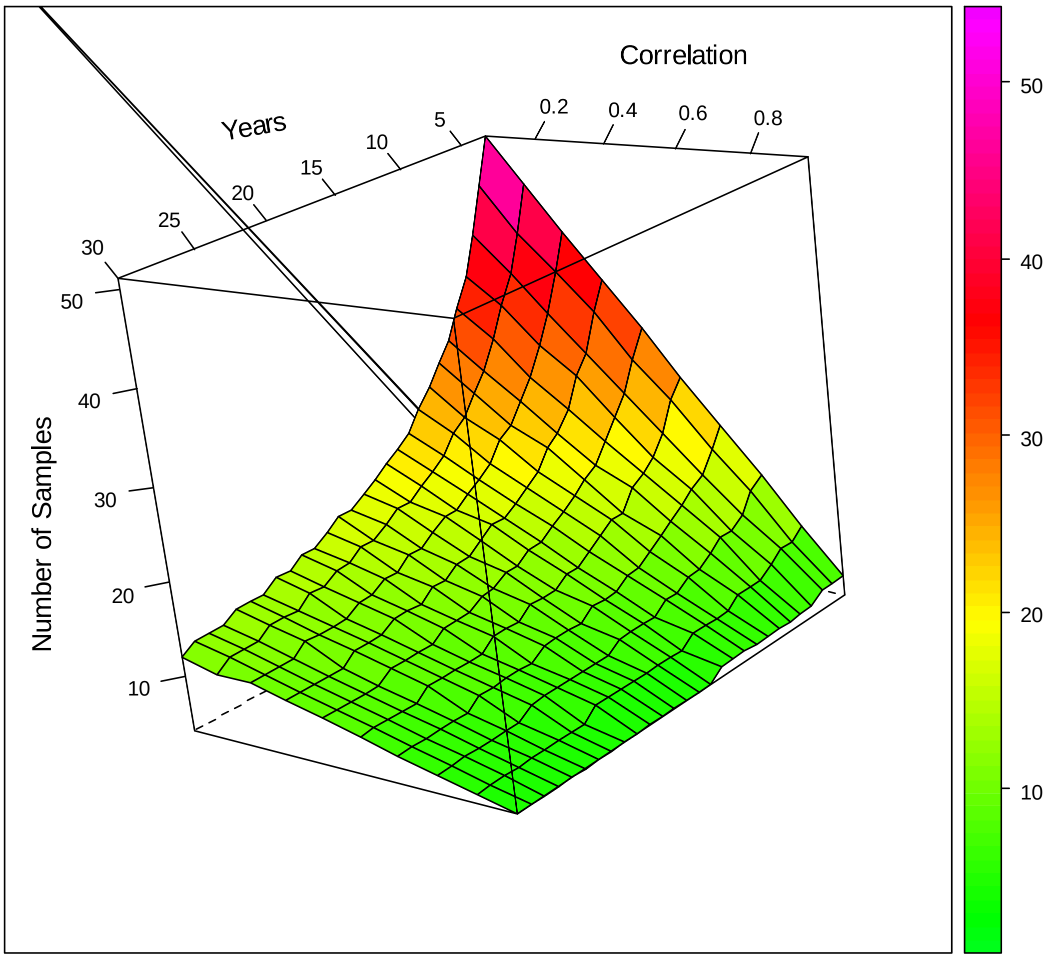

Figure 11. Minimum sample size as a function of years and temporal correlation. Type I error is set at 0.1 and compound symmetric correlation structure is assumed.

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**			
σ^2 = 0.25					**σ^2 = 0.25**			
Year	**ρ=0.25**	**ρ=0.50**	**ρ=0.75**		**Year**	**ρ=0.25**	**ρ=0.50**	**ρ=0.75**
10	13	10	7		10	16	12	8
20	9	7	5		20	11	9	7
30	7	6	5		30	9	7	6
σ^2 = 0.50					**σ^2 = 0.50**			
10	22	16	10		10	28	20	12
20	14	10	7		20	17	13	9
30	11	8	6		30	13	10	7
σ^2 = 0.75					**σ^2 = 0.75**			
10	31	22	13		10	39	28	16
20	19	14	9		20	24	17	11
30	14	11	7		30	18	13	9
σ^2 = 1.00					**σ^2 = 1.00**			
10	40	28	16		10	51	35	20
20	24	17	10		20	30	21	13
30	17	13	8		30	22	16	10

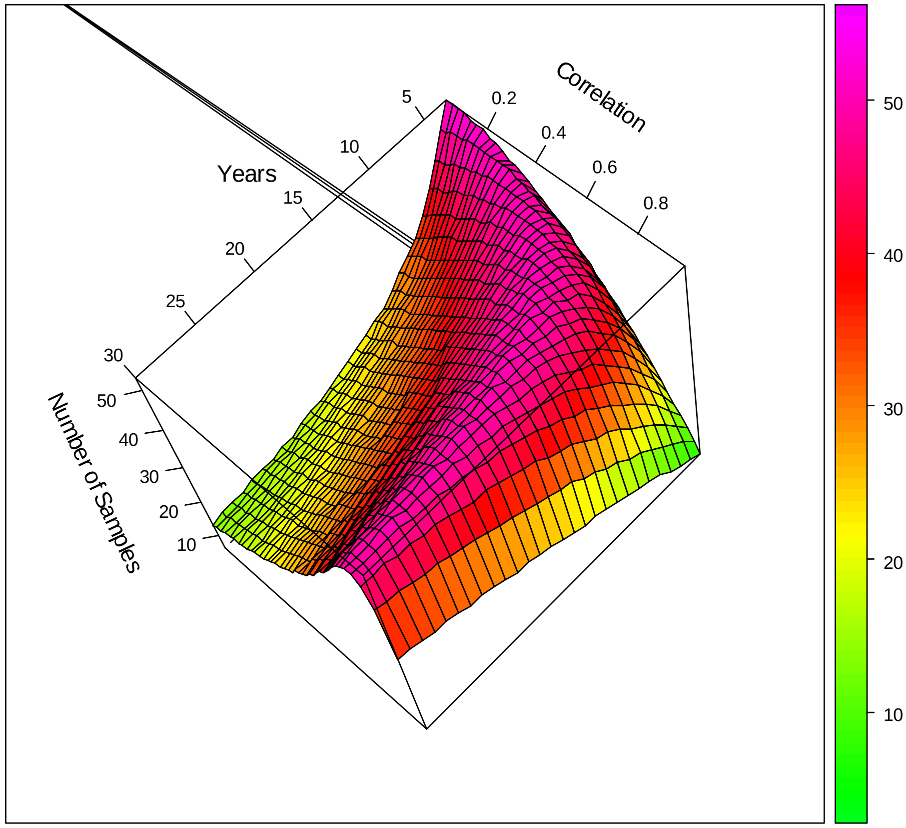

Figure 12. Minimum sample size as a function of years and temporal correlation for the autoregressive structure; type I error is set at 0.1.

6.7 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 initially 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. 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.

6.8 Data Analysis

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 any approach might benefit from a particular 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 the area of each stratum or with a design based estimator.

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). It should be noted that the small mammal and mosquito diversity and abundance protocols that are limited to a subset of the sample frame are allocated according to proportions of the entire sample frame.

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,

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

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,

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

Where,

s_i^2 = the sample variance from the i^th stratum.
n_i = the number of units sampled from the i^th stratum.

The number of pixels 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

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

Where,

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

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

and,

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

An additional consideration that may be of use for subsequent design optimization is the relative amount of resources that are needed for sampling among sites. If the total cost of sampling at a given site is broken down into a linear combination of fixed cost and the variable cost for each stratum, it can be expressed as follows

$$c = c_0 + \sum_{i=1}^{S} c_i\, n_i \qquad [10]$$

Where,

c_0 = the fixed cost per stratum.

c_i = the cost per sample within the ith stratum.

Using this representation of total cost, the optimal allocation among strata taking into account both the cost and standard deviation among strata is given by

$$n_i = \frac{(c-c_0)N_i\sigma_i/\sqrt{c_i}}{\sum_{i=1}^{S} N_i\sigma_i\sqrt{c_i}} \qquad [11]$$

This information can be used to consider subsequent optimization of the design once several years of data have been collected. This approach will be especially useful once better estimates of cost and variability have been obtained from the first several years of sampling.

The simplest approach to dealing with the estimated variance for mean estimates in a spatially-balanced, stratified-random design was demonstrated. More complex approaches yield lower variance estimators in certain situations; however, the level of applicability for a given approach is inversely proportional to the number of assumptions that need to be made for implementation. Since this sample design will accommodate the co-location of multiple observations across dozens of sites spanning numerous biomes, the most general and applicable approach was developed.

6.9 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 contraints, 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 unavilable 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 classification 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.

7 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.

8 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.

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 2edition. Springer.

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, NY.

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. Wiley.

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.

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.

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. *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.

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. *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.

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.

Lohr, S. L. 2010. Sampling design and analysis. Brooks/Cole Cengage Learning, Boston.

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;.

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 compoent estimation. Biometrics **27**:1-&.

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.

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 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. 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 spampling desings of the National Ecological Observatory Network Terrestrial Observation System. Ecosphere 7(12) 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 (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.

9 APPENDICES

9.1 Appendix 1. Implementation of the Study Design

The NEON Terrestrial Observation System (TOS) will describe spatio-temporal patterns of organisms and biogeochemistry within sites. In many cases, the site boundary corresponds to a legal ownership or administrative boundary. In some cases, a more restrictive boundary has been defined, because the sites (e.g., Yellowstone National Park) are too large to be sampled with available resources. In other cases, a site is composed of two or more areas with one or more land owners located in close proximity (e.g., the Harvard Forest site; for a map and table of all sites, see http://neonscience.org).

A site-level spatial sampling design that can be applied consistently across NEON domains has been developed for the Terrestrial Observation System. Many of the specific locations of the TOS sampling elements are collocated with each other and with environmental sensors (i.e., within the flux zone of the tower) to allow comparison of the data streams. Sampling has been optimized to ensure efficient and effective sampling within the budgeted scope of NEON.

Sampling is distributed throughout each site according to three different **plot types** (**Table 4**)Tower Plots, Distributed Plots and Gradient Plots (see Section 6.1):

- **Tower Plots** provide a direct link between NEON's Terrestrial Observation System and Terrestrial Instrument System platforms. Measurements in these plots include above and below-ground plant productivity, soil biogeochemistry, and soil microbial community diversity and function. In addition~~,~~ individual plants are marked for phenology observation along a square 'loop' or plot perimeter. Tower Plots were placed in the 90% flux area of the primary and secondary (if applicable) airsheds of each NEON tower. If the requisite number of plots could not be established in the airshed(s), Tower Plots were also placed outside the airsheds not further from the tower than the length of the vector defining the extent of the primary airshed. Tower Plot locations for observations of plant biomass, productivity, diversity and foliar biogeochemistry, and soil biogeochemistry and microbial diversity and abundance are selected from this realm (available for each site on the NEON data portal) according to a spatially balanced, random design (see section 6.3). Phenology plots were subjectively located.

- **Distributed Plots** were established according to a stratified-random and spatially balanced design (section 6.3 and 6.4) in an effort to describe organisms and process with plot, point, and grid sampling distributed throughout NEON sites. At some plots, collocation of plant and soil sampling will occur to maximize the value of data streams. Sampling for beetles, mosquitoes, ticks, birds, and small mammals will also occur at or adjacent to a subset of these plots, with additional constraints on plot locations for mosquitoes and small mammals, which must be relatively close to a road.

- **Gradient Plots** At some sites, Distributed Plots may fail to fully capture site-level gradients in vegetation structure, leaf area index, or plant canopy chemistry. For these sites, Gradient Plots will be established using a targeted, non-random approach informed by NEON's Aerial Observation Platform remote-sensing data. These plots may include subplots for sampling plant diversity, soil and plant biogeochemistry, and soil microbiota.

NEON adopted and created a variety of sampling protocols to facilitate the observations of organisms and soil. The plots, points, and grids, termed a plot **subtype** in the context of the NEON design, were placed according to parameters and requirements described in corresponding science design documents. An abridged description of each collection method and design details follows:

- **Base Plots** are square multiscale plots designed for Distributed (10–30/site), Tower (4–30/site), and Gradient Plot sampling. The collection of soil microbe (at 4 Tower Plots and 6 Distributed Plots), soil biogeochemistry (at 4 Tower Plots and 6 Distributed Plots) and vegetation biomass, productivity, biogeochemistry, and diversity (at 3-30 Tower Plots and up to 30 Distributed Plots). The Base Plot is centered on the point selected for sampling from the stratified-random and spatially balanced study design. While most frequently established as 20m x 20m or 40m x 40m Tower Plots and 40m x 40m as Distributed Plots, Base Plots are designed to be up to 80m x 80m in size. At most sites, the south-west corner and plot center are marked with permanent 'primary' aluminum monument markers and, when possible, at other corners with a 'secondary' marker appropriate to the conditions and regulations of each site. Other Base Plot notes and criteria:
    - Distributed Base Plots are allocated proportional to the NLCD cover types within the sampling frame for all sampling except for plant diversity which is allocated in proportion to the square root of stratum size (AD[05])
    - Distributed Base Plot edges must be separated by a minimum of 55m
    - Tower Base Plots are placed in the primary and secondary airsheds, and outside these areas but in close proximity (bounded by airshed vector length) to the tower when necessary
    - Tower Base Plot edges must be separated by a distance 150% of one edge of the plot (e.g., 40m x 40m Tower Base Plots must be 60m apart)
    - Base Plot centers must be greater than 50m from large paved roads and plot edges must be at least 10m from two-track dirt roads
    - Base Plot centers must be at least 50m from buildings and other non-NEON infrastructure
    - Streams larger than 1m must not intersect Base Plots
    - Minimum accuracy requirement of GPS-derived coordinates (horizontal precision) of the Base Plot corners and center: 0.3m

- **Tick Plots** (6/site) are 40m x 40m plots that are collocated with Distributed Base Plots. To reduce the probability that the sampling activities associated with Base Plots impact tick diversity and distribution (e.g., technicians inadvertently attracting or redistributing ticks), the Tick Plot center is offset from Base Plot center according to a specified distance (150m +/- 15m) and a randomly chosen direction established prior to establishment of plots in the field. If the location of the plot is not suitable for sampling (e.g., too close to a road or other infrastructure) a different random direction is selected. In the few instances that a tick plot could not be collocated with a Distributed Base Plot, the next available Base Plot is selected from the Morton Ordered sampling list (section 6.3). Other Tick Plot notes and criteria:

- Tick Plots are allocated proportional to the NLCD cover types within the sampling frame
- The edge of the Tick Plot is greater than 150m from the edge of other NEON plots and infrastructure
- The centers of Tick Plots must be separated by a minimum of 500m
- Streams must not bisect the edge of the Tick Plot
- Tick Plots are marked with permanent markers at the center (point 41) and the south-west corner (point 21)
- Minimum accuracy requirement of GPS-derived coordinates (horizontal precision) of the Tick Plot corners and center: 2m

- **Mammal Grids** (3–8/site), consisting of 100 trapping locations at 10m spacing, are 90m x 90m. Due to the equipment and time required to complete sampling, the center (trap location E5) of these grids is not more than 300m from roads that can be accessed by NEON technicians. Where possible, these grids are collocated with Distributed Base Plots by placing them a specified distance (150m +/- 50m) and random direction from the center of the Base Plot. When fewer than 6 Distributed Base Plots are within 300m of roads, the Mammal Grid centers are placed at a random azimuth and specified distance (150m +/- 50m) from the next available sample location from the Morton Order sampling list that are within 300m of roads. Other Mammal Grid notes and criteria:

  - Mammal Grids are allocated proportional to the NLCD cover types within the sampling frame
  - More than 50% of the Mammal Gird must fall within the target NLCD cover type
  - Trapping points must not fall within streams, lakes, or ponds
  - The edges of any two Mammal Grids must be separated by a minimum of 135m between the edges of grids
  - The edge of Mammal Grids must be at least 25m from paved roads
  - Dirt roads less than 10m in width may bisect the Mammal Grid
  - Grids are marked with permanent markers at the center (E5) and the south-west corner (A10)
  - Minimum accuracy requirement of GPS-derived coordinates (horizontal precision) of the Mammal Grid center (E5): 2m

- **Bird Grids** (5-15 Grids/site), consisting of 9 points at 250m spacing, are 500m x 500m. Where possible, Bird Grids are collocated with Distributed Base Plots by placing the Bird Grid center (B3) in close proximity to the center of the Base Plot. At smaller sites, a single point count is done at the south-west corner (point 21) of the Distributed Base Plot. Other Bird Grid notes and criteria:

  - Bird Grid centers are allocated proportional to the NLCD cover types within the sampling frame
  - More than 50% of the points on a Bird Gird must fall within the target NLCD cover type
  - The edges of any two Bird Grids must be separated by a minimum of 250m; this distance is typically greater than 500m
  - Bird Grids are marked with permanent markers at the center (B2)
  - Minimum accuracy requirement of GPS-derived coordinates (horizontal precision) of the Bird Grid center: 2m

- **Mosquito Points** (10/site) are the points at which CO_2 traps are established. Due to the frequency of sampling and temporal sampling constraints, Mosquito Points are located within 45m of roads accessible to sampling by NEON technicians. Other Mosquito Point notes and criteria:

  - Mosquito Points are allocated proportional to the NLCD cover types within the sampling frame
  - Mosquito Points must be greater than 5m from roads
  - Any two Mosquito Points must be separated by a minimum of 310m
  - Points must be at least 10m from the edge of other NEON sample locations
  - Mosquito Points are marked with permanent markers
  - Minimum accuracy requirement of GPS-derived coordinates (horizontal precision) of the Mosquito Points: 2m

- **Phenology** observations are made on individuals located on the perimeter of a 200m x 200m loop or transect (2/site). The majority of the observations are made on individuals on a loop (subtype specification = 'primary') located in the primary airshed of the NEON tower. If there is insufficient space in the primary airshed, the loop is placed in the secondary airshed, or in close proximity to the tower. If the Phenology plot is not north of the tower, additional observations (subtype specification = 'phenoCam') are made on a second set of individuals directly north of the NEON tower to help calibrate phenology camera images captured from sensors on the tower. Other Phenology notes and criteria:

  - Points on the perimeter of the Phenology sampling plot must be greater than 60m from roads and other infrastructure
  - The southwest corner of the Phenology plot is marked with a permanent marker

- Minimum accuracy requirement of GPS-derived coordinates (horizontal precision) of the points on the Phenology Plot: 0.3

Table 4. An overview of the plot types and subtypes deployed to measure the taxonomic groups and soil targeted by the NEON Terrestrial Observation System.

Plot type	Plot subtype	Taxonomic group/module
Distributed Plots (Plots, Grids, Points)	Base Plot	Plant Productivity and Biomass
		Plant Diversity
		Microbes
		Biogeochemistry
		Beetles
	Mammal Grid	Small Mammals
	Bird Grid	Breeding Birds
	Tick Plot	Ticks
	Mosquito Point	Mosquitos
Tower Plots	Base Plot	Plant Productivity and Biomass
		Plant Diversity
		Microbes
		Biogeochemistry
	Phenology Plot	Plant Phenology
Gradient Plots (Plots, Grids, Points)	Base Plot	Plant Productivity and Biomass
		Biogeochemistry
		Plant Diversity

9.2 Appendix 2. 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<-.5/(seq(1: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 gradient 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 gradient) 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)
sigsqs<- 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 = sigsqs[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 = sigsqs[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)
```