Working across space and time: nonstationarity in ecological research and application

Christine R Rollinson1*, Andrew O Finley2,3, M Ross Alexander4†, Sudipto Banerjee5, Kelly-Ann Dixon Hamil6, Lauren E Koenig7, Dexter Henry Locke8‡, Megan Peterson9, Morgan W Tingley10,11, Kathryn Wheeler12, Casey Youngflesh10,11, and Elise F Zipkin2,13

Ecological research increasingly considers integrative relationships among phenomena at broad spatial and temporal domains. However, such large-scale inferences are commonly confounded by changing properties in the processes that govern phenomena (termed nonstationarity), which can violate assumptions underlying standard analytical methods. Changing conditions are fundamental and pervasive features in ecology, but their influence on ecological inference and prediction increases with larger spatial and temporal domains for a host of factors. Fortunately, tools for identifying and accommodating potentially confounding spatial or temporal trends are available, and new methods are being rapidly developed. Here, we provide guidance for gaining a better understanding of nonstationarity, its causes, and how it can be addressed. Acknowledging and addressing non-constant trends in ecological patterns and processes is key to conducting large-scale research and effectively translating findings to local policies and practices.

In a nutshell:

- Ecological systems are under constant change and frequently exhibit spatially and temporally varying trends, which can cause substantial challenges for analysis and application of ecological data
- Most biological systems display potentially confounding spatial or temporal trends (nonstationarity) at some scale, presenting a key challenge for macrosystems research
- Accounting for nonstationarity in ecological processes can improve both inference and prediction
- Ecological research needs to accommodate spatial and temporal variability in ecological patterns and processes to be better translated into actionable information for stakeholders

(continued on last page)
in terms of properties of ecological processes and how those processes are represented in statistical models or analyses, has implications not only for how ecological systems are studied, but also for subsequent inferences and predictions (Miller and Hanham 2011).

Acknowledging nonstationarity in ecological processes is key to bridging large-scale research to local policies and practices. Accommodating spatial and temporal variability in the factors and ecological processes that determine habitat suitability in species distribution modeling has been a long-standing challenge for both the research and conservation communities (e.g., Osborne et al. 2007; Thuiller et al. 2008). In restoration, the use of historical reference states as targets is complex due to temporal changes in conditions, such as climate and land-cover patterns (Swetnam et al. 1999; Higgs et al. 2014). The use of specific management practices (e.g., fire) is known to have mixed results among locations due to spatial variability in site and environmental conditions (McEwan et al. 2011). The rise of macrosystems biology via “big data,” new technologies in fields like remote sensing, and advances in data integration have led to novel discoveries in global-scale and emergent phenomena that are not possible from smaller scale studies based on a few sites or short temporal extent alone (Zipkin et al. 2021). Translating such advances into reliable and actionable information for practitioners and policy makers requires that spatial and temporal variability in ecological patterns and processes be accounted for in both inference and prediction (Rodo et al. 2002; Gouveia et al. 2013).

Nonstationarity as a property of ecological systems

Assuming that an ecological system is stationary is akin to viewing a system as being in equilibrium, in that both states are often highly dependent on the scale and/or the question being posed. For example, the landscape mosaic concept from geography postulates that a landscape may have consistent properties such as median age or composition through time when viewed at a coarse scale, but this pattern is maintained by periodic disturbances with localized sites undergoing constant change (Turner 2005). Conversely, genetic change leading to adaptation or speciation is widely observed, but only over multiple generations, and therefore stationary population genetics may be a valid assumption for research on finer scales (Jost et al. 2018). Arguments have been made that all systems are nonstationary at some scale, and that scale often corresponds with macrosystems research that spans broad spatial and temporal scales (Heffernan et al. 2014; Wolkovich et al. 2014; Collins et al. 2018).

Change in the properties of ecological systems (i.e., nonstationarity) can arise from multiple sources, including variation in underlying environmental conditions (e.g., soils, topography, climate) or as a result of abrupt events (e.g., disturbance) (Figure 2). These and other causes of nonstationarity can result in spatial or temporal differences in properties like the mean or variance of ecosystem states or traits (Figure 1). Nonstationary properties of ecological processes frequently result from multiple, interacting factors that vary across space or time (Schmidt et al. 2014). The relationships among these factors are complex and often unknown. Nonstationarity is not unique to macrosystems research, but working at large scales increases the likelihood that the ecological processes of interest may vary within the spatial or temporal study domain, potentially leading
to complications for cross-scale inference and prediction (Cornulier and Bretagnolle 2006; Finley 2011; Miller and Hanham 2011).

Spatial or temporal variability of ecological factors in a dataset does not necessarily qualify as nonstationarity and is often a strength rather than an intrinsic obstacle for describing ecological relationships that underlie an observed phenomenon. For example, prediction of species distributions typically relies on observations of species occurrence or abundance over a wide range of climatic conditions to describe potential suitable habitat (WebPanel 1). In this instance, spatial gradients in a factor like mean annual temperature are not inherently problematic if they have a consistent relationship with the observed response of abundance. Predictions and inference become more complicated if the spatial or temporal relationships between the predictor variables and the observed response are not constant. For example, if habitat characteristics or local adaptations lead a population to have unique relationships with a key environmental factor, local abundance estimates may be inaccurate unless spatial nonstationarity is taken into account (WebPanel 1). In turn, accurate distribution and abundance estimates are important for identifying conservation priorities and allocation of financial and human resources (Jetz et al. 2008; Johnston et al. 2015).

When is nonstationarity important for research and application?

As with other scientific disciplines, ecological research is tasked with making new discoveries. Doing so often requires reliance on carefully constructed statistical models, which are simplifications of reality, to test hypotheses that explain observed phenomena across space and time. Because it is not possible to measure or fully quantify all of the causal processes that generate an observed phenomenon, there will always be some level of undescribed variability or model uncertainty. In cases where nonstationarity – and not just general uncertainty – is a consideration, conclusions drawn from a single location or point in time are insufficient for explaining large-scale patterns because they are based on only a subset of the conditions over which broad ecological processes occur (Symstad et al. 2003; Soranno et al. 2014).

Most relationships in ecological research are not fully understood, and therefore determining whether consideration of nonstationarity in study design and analysis has improved causal inference is difficult, particularly when results contradict those of past studies. However, studies using simulated data, where the true data-generating process is known, have demonstrated that failure to take into account spatial heterogeneity through additional processes or random effects can lead to results that contradict reality (Dixon Hamil et al. 2016). Analyses using methods that account for spatial nonstationarity have been used to challenge past research that has wide-ranging policy implications. For example, accounting for spatial nonstationarity in analyses of tree cover in Philadelphia, Pennsylvania, indicated that relationships between tree cover and demography are more complex than previously portrayed (Locke et al. 2016). Accounting for the complex and spatially nonstationary relationships between demography and tree cover can lead to more effective urban planning policies.

Failure to account for nonstationarity in models of ecological processes can cause agents of change to be misattributed to other, unrelated factors, resulting in flawed inference or prediction. The challenge of accommodating nonstationarity for ecological inference is common in many ecological fields, and particularly those focused on identifying drivers of change in systems undergoing multiple alterations (Wolkovich et al. 2014). For example, tree-ring-based studies that investigate the effects of 20th-century climate change on tree growth must often contend not only with a nonstationary climate, but also...
changes in precipitation, nitrogen deposition, and disturbance dynamics (Figure 3; McEwan et al. 2011; Wolkovich et al. 2014). These additional effects do not necessarily need to be explicitly included in a model focused on the effects of temperature, but accounting for temporal trends in individual growth rates through a priori detrending or a single additional covariate improves stationarity of estimated effects and reduces the inferred effect of temperature on tree growth (Figure 4; WebPanel 2). Consequently, detrending or otherwise accommodating age- or size-related growth trends in tree-ring width is a fundamental aspect of dendrochronological research for both inference and prediction (Peters et al. 2015). However, partitioning the partial effects and attributing change to a single factor or variable is challenging when nonstationarity is present, especially in multiple, collinear predictors. These covarying trends and complex interactions in both ecological processes and statistical models complicate identifying the “true” effects of predictors on a response (WebPanel 2).

Addressing nonstationarity in models and analyses

Nonstationarity is a pervasive and persistent challenge for ecological research and its application to solving real-world problems. Although there is no simple, universal approach for identifying and addressing spatial or temporal nonstationarity in ecology, approaches to identifying and addressing nonstationarity do exist. Carefully considering the objective or hypothesis connecting measured factors and observed phenomena is the first and most important step in any analytical work. In addition to determining what the key predictors and response variables are, the research or application objective determines what steps are most appropriate for addressing any evidence of nonstationarity that arise (Figure 5).

From a technical perspective, nonstationarity is defined as differences in the statistical characteristics (eg mean, variance, covariance) of a variable, or in statistical relationships across space or time (Figure 1; Schabenberger and Gotway 2005; Banerjee et al. 2014). Changes in a variable’s mean are referred to as first-order and are often relatively easy to detect and address through methods described below (Figure 1b). Changes in variance and covariance are second-order and may be more challenging to diagnose, requiring collaboration with statistical experts (Figure 1c). When discussing nonstationarity statistically, we often start with the assumption that ecological variables are regulated by highly complex and largely unobservable mechanisms (“data-generating processes”) that will never be fully represented in analytical models. Various approaches for addressing nonstationarity in statistical analysis exist, but many approaches are beyond what most ecologists are traditionally exposed to during their academic training.

Diagnosing nonstationarity

The first step in designing a model that can accommodate nonstationarity is to identify the spatial and temporal scales of the data and desired inference, as well as the factors or processes that are likely to be of greatest importance. Following that, exploratory data analysis can help determine if modeling efforts that explicitly accommodate nonstationarity are warranted (Figure 5). Visual inspection of raw data and model residuals is often the best starting point for identifying nonstationarity in initial statistical model development (eg Figure 1; Figure 4a). If predictor variables, observed responses, and the distribution of residuals from a model are evenly distributed across the study’s spatial and/or temporal domain, then stationarity can be assumed and many common analytical approaches, such as simple linear regression, are appropriate. However, if the statistical properties (eg mean and variance) among model residuals change across the domain, then the initial model may fail to adequately capture the ecological processes generating the response variable. Once nonstationarity is detected, several methods are available and widely used in ecology for addressing spatial or temporal trends affecting ecological data. In general, approaches can be sorted into two categories (Figure 5): (1) those that describe the source of nonstationarity by modifying the inferential scope of the model so that ecological relationships can vary across a study domain, or (2) those that accommodate nonstationarity through latent or “hidden” processes that retain the initial inferential model structure (eg fixed effects) but add additional complexity that accounts for previously undescribed spatial or temporal trends.

Describing nonstationarity

The appearance of nonstationarity in ecological models or analyses indicates additional factors or ecological processes that strongly influence the observed phenomenon of interest.
In many cases, particularly in theoretical or research ecology, this may be the main objective of the study; in such instances, the hypothesis and model structure may need to be revisited and additional covariates included to better represent the phenomenon-generating ecological processes (Figure 5; eg Schmidt et al. 2014; Abbott et al. 2017). In other cases, there may not be enough information or data available to add additional factors to a model to explain the spatial or temporal variability in ecological processes, and therefore the model structure may need to be altered so that processes vary spatially, which may prompt new hypotheses and future research into the sources of nonstationarity. One common approach to this is to first divide the analytical domain into overlapping subregions where stationarity can be assumed, and then the multiple model parameters and processes can be interpolated among those subregions (eg geographically weighted regression; WebPanel 1; Brunsdon et al. 1996; Mellin et al. 2014). However, more robust methods exist that allow estimation of the changing relationships between response and predictor variables across spatiotemporal domains using a single modeling approach (eg spatially varying coefficient models; Finley 2011; Jarzyna et al. 2014; Risser and Turek 2019).

Accommodating nonstationarity

Although ecologists often seek to explicitly describe the reasons ecological processes vary over space and time, occasionally the objective of a study needs to merely accommodate nonstationarity to permit valid predictions and inferences about specific processes or ecological behaviors (Figure 5). One method for accommodating spatial or temporal variability that is becoming increasingly standard in ecological analyses, particularly in first-order nonstationarity with differences in mean effects among observational units, is the use of hierarchical random effects (Cressie et al. 2009). Other methods, such as data assimilation techniques, where local states or parameters are used to probabilistically update model predictions or parameters at specified spatiotemporal intervals, can improve predictions in nonstationary environments (Luo et al. 2011; Niu et al. 2014). For example, the complex relationships among population dynamics, climate, and oceanic circulation make jellyfish outbreaks difficult to predict; data assimilation allows integration of multiple sources of observations (eg satellite imagery, social media reports) to update predictions and produce more accurate forecasts in (for this example) the Gulf of Maine when the drivers of spatiotemporal dynamics of outbreaks and subsequent public health risks are poorly constrained (Record et al. 2018). Formal forecasting techniques that assimilate site-specific parameters and observations have also been proposed as a useful approach for improving restoration decisions in the face of a changing (nonstationary) climate (Hardegree et al. 2018). Critical evaluation of the spatiotemporal trends in estimated parameters or assimilation weights can provide critical insights into the key ecological processes generating statistical nonstationarity to be the focus of future inference-focused efforts (WebPanel 1).

Conclusions

Explaining patterns and processes across space and time is a fundamental aspect of ecological research and is essential for providing information to guide natural resource and policy decision making. The development of new technologies and data streams that facilitate large- and cross-scale research on ecological phenomena has increased our ability to detect, study, and account for how relationships among factors and processes vary across spatiotemporal domains (Heffernan et al. 2014; Soranno et al. 2014). The role of nonstationarity and its impacts on ecological inference and prediction is often dependent on the scale and scope of research hypotheses or project objectives (Figure 5). Consequently, many methods exist and have been applied to ecology to account for nonstationarity in the statistical models we use to formally test research hypotheses and make predictions about ecological phenomena. These models will never fully represent the complexity of ecological systems, but careful consideration of how and why processes vary across space or time is an important step toward improving ecological research and its application to solving real-world challenges.
Figure 5. Conceptual workflow for detecting and addressing nonstationarity when building ecological models. First, a formal model describing the relationships between measured predictors and observed responses is developed; second, diagnosing spatial and/or temporal trends in model residuals can help determine whether nonstationarity needs to be formally addressed in the model. Approaches for addressing nonstationarity depend on the ecological hypothesis or objective and may require modifying model structure to include additional covariates or latent processes that implicitly account for non-focal spatiotemporal trends. Once statistical assumptions of stationarity have been met, robust inference and predictions can be made using the developed model.

Acknowledgements

Publication of this Special Issue was funded by the US National Science Foundation (NSF award number DEB 1928375). All data and code used in figures and examples can be found at https://github.com/MortonArb-ForestEcology/MSB_Non-Stationarity. Special thanks to A Hipp and L Fitzpatrick for providing valuable feedback on this paper. We thank the thousands of volunteers who annually conduct the North American Breeding Bird Survey for the observations and data used in WebPanel 1. We also thank N Pederson, D Bishop, Harvard Forest, and NSF EF 1241930 for the data used in the tree-ring analysis in WebPanel 2. This work was funded by numerous organizations, including NSF EF 1702996 (CRR), 1638648 (DHL), EF 1703048 (MWT), EF 1442439 (LEK), 1638577 (KW), and 1702635 (EFZ). Additional support was provided by NSF GRFP 1247312 (KW) and the National Socio-Environmental Synthesis Center under funding received from NSF DBI-1639145 (DHL). The findings and conclusions in this publication are those of the authors and should not be construed to represent any official US Department of Agriculture or US Government determination or policy.

References


This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

### Supporting Information

Additional, web-only material may be found in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/fee.2298/supinfo

7Natural Resources and the Environment, University of Connecticut, Storrs, CT; 8National Socio-Environmental Synthesis Center, Annapolis, MD; ‡current address: US Department of Agriculture, Forest Service Baltimore Field Station, Baltimore, MD; 9Department of Plant Biology, University of Georgia, Athens, GA; 10Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT; 11Ecology and Evolutionary Biology, University of California–Los Angeles, Los Angeles, CA; 12Earth and Environment, Boston University, Boston, MA; 13Department of Integrative Biology, Michigan State University, East Lansing, MI