Trends in
Ecology & Evolution

CellPress

Review

Scale gaps in landscape phenology: challenges and opportunities

Daniel S. Park,[1,5,*] Erica A. Newman,[2,3,5] and Ian K. Breckheimer[4,5]

Phenology, or the timing of life history events, can be heterogeneous across biological communities and landscapes and can vary across a wide variety of spatiotemporal scales. Here, we synthesize information from landscape phenology studies across different scales of measurement around a set of core concepts. We highlight why phenology is scale dependent and identify gaps in the spatiotemporal scales of phenological observations and inferences. We discuss the consequences of these gaps and describe opportunities to address the inherent sensitivities of phenological metrics to measurement scale. Although most studies we review and discuss are focused on plants, our work provides a broadly relevant overview of the role of observation scale in landscape phenology and a general approach for measuring and reporting scale dependence.

Highlights

Phenological responses can be heterogeneous across species, communities, and landscapes and can vary across spatiotemporal scales.

Global patterns of phenological sampling exhibit strong geographic trends and critical scale gaps.

Phenological measurements dealing with the beginning or end of a process are especially sensitive to spatial and temporal measurement scale.

Constructing scaling relationships and estimating their mathematical forms is necessary for forming null expectations about how spatiotemporal scale (resolution and extent) influences inferred phenological patterns and for resolving observed mismatches between phenological metrics measured at different scales.

Phenological scaling rules can unlock new insights into the causes and consequences of shifting phenological landscapes and into broader questions about the spatial and temporal organization of ecological communities and interactions.

Phenology in the anthropocene

Changes in **phenology** (see Glossary) are among the most dramatic biological responses to climate change and have widespread consequences for species and ecosystem processes that support human society [1–4]. Phenological responses of individual species and entire communities have shifted substantially in recent decades [5–8]. In the northeastern USA, for example, many plant species now flower 7–10 days earlier, on average, than they did 100 years ago [9,10]. Globally, phenological changes track changing climate, with spring onset arriving an average of 2.3 days earlier per decade [11]. Studies of changes in temporal community composition comprise a large and growing body of research, as such changes can affect resource availability, agriculture, trophic interactions, diversity of associated communities, and ecosystem services [12–17]. Phenological studies have traditionally focused on scales of individual organisms and plots, using field observations and experiments [18,19]. More recently, large-scale digitization of natural history collections [20], citizen science programs [21], and the advent of remotely sensed **land surface phenology** measurements via satellite [22] have facilitated research at more extensive taxonomic, spatial, and temporal scales, some of which integrates multiple methods [23–25]. Much of this new work suggests that phenologies and their responses to climate change are heterogeneous within communities and across landscapes, varying both within and among species [26–31]. Interpreting and synthesizing these diverse phenological patterns across taxonomic, temporal, and spatial scales is important for understanding their causes and making realistic predictions under climate change. This has proved to be a serious scientific challenge, exacerbated both by the lack of rigor in reporting the relevant spatiotemporal scale of studies and by the lack of a robust theoretical framework that allows us to integrate results across scales, which would thereby allow the synthesis of diverse and sometimes conflicting measurements.

The nascent field of **landscape phenology**, which studies the causes and consequences of scale-dependent spatial variation in the timing of biological events, integrates spatial patterns and temporal processes within heterogeneous environments across multiple scales to elucidate seasonal biodiversity dynamics [32,33]. Landscape phenology studies use transplant experiments, *in situ* observations, remote sensing data, and spatiotemporal analyses of phenological

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA
[2]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA
[3]Arizona Institutes for Resilience, University of Arizona, Tucson, AZ 85721, USA
[4]Rocky Mountain Biological Laboratory, Gothic, CO 81224, USA
[5]These authors contributed equally to this work

*Correspondence:
danielpark@purdue.edu (D.S. Park).

patterns to address questions related to ecosystem processes, ecological interactions, evolutionary dynamics, and ecological change. Landscape phenology studies are increasingly important in understanding how environmental heterogeneity affects species interactions [34], unpacking the relationship between climate and ecosystem processes [35], and understanding trends in economically damaging climatic events such as spring freezing [36]. To date, landscape phenology studies tend to be more commonly associated with plant systems, and thus many of the following examples pertain to such. However, many of the concepts we discuss are broadly applicable across taxa and even to abiotic patterns such as ice melt and carbon fluxes.

Why is phenology scale dependent?

It has been long recognized that ecological patterns are scale dependent, from matters of species diversity and abundance [37,38] to relationships among native and exotic taxa [39,40]. The ecology of temporal events is no exception. The conclusions we draw about phenological transitions such as flowering onset [41], leaf unfolding [42], or insect emergence [43] are inherently sensitive to how the original measurements were aggregated over space, time, and taxonomy. This '**scale dependence**' arises for three major reasons. First, with scale defined as the grain (or resolution) and extent of a set of measures, the spatiotemporal scale at which measurements are made will change the mean and variance of a metric, even in completely homogeneous environments, purely due to statistical aggregation [44] (Figure 1). These biases are additional to other effects, such as those based on measurement error and lack of precision due to small sample sizes, and therefore need to be accounted for when comparing metrics across scales. Second, many environmental cues that trigger life-history transitions (e.g., temperature) can vary dramatically across space. Some common environmental cues, such as the timing of changes in day length across the season, vary smoothly across latitude, while others, such as the departure of seasonal snow in temperate mountains, vary greatly over much shorter distances [45]. Third, individuals, populations, and ecological communities often respond differently to a given set of environmental cues. This can occur because of differences in the plastic responses of individuals or species (phenological sensitivities or norms of reaction [46]) or because of ecological interactions that result in different distributions of phenological sensitivities within communities [47]. These processes often occur simultaneously, leading to complex patterns of phenological variation across space, and scale-dependent relationships between measures of phenology and their underlying drivers (Figure 2).

Phenological events, such as the beginning and end of a growing season, dates of migration or hibernation, or dates of first and last frost, can vary in how sensitive they are to issues of scaling and sampling (Figure 1). The degree of scale dependence of any landscape metric ultimately depends on the 'window size,' or spatial or temporal grain, to which a metric is being aggregated [44], the spatial and temporal extent studied [48], endogenous factors such as landscape heterogeneity and spatial autocorrelation of the locations of species' individuals [49], and the degree of aggregation across different species or functional types [50]. Because many current studies of phenology use data from a single spatial or temporal scale, conclusions from those studies are also limited to those scales [44,51,52], a 'lurking' limitation that is not often recognized [53]. Most attempts to directly link observations made at different scales in landscape phenology (e.g., ground-based observations of individuals vs. satellite-derived landscape observations) have relied on direct correlations between measurements made at different spatial or temporal grains, which often yield relationships with significant unexplained variability or even direct contradictions [54–58]. Direct integration of data at different scales and resolutions in a remote sensing context can result in loss of information, introduction of mathematical artifacts, and conclusions that are dependent on how the data were aggregated [50,59]. Measurements taken using remote sensing tools often aggregate over taxonomy as well as space, leading to ambiguity in the best way to link species-level field measurements with remote sensing [50,60]. For these reasons, it is not

Glossary

Environmental forcing: an exogenous perturbation to a baseline trend of environmental patterns or factors; for example, climate change may force start of season phenology to occur earlier in the year than historical baselines.

Exogenous: arising from factors external to the narrowly defined system under consideration; for example, climate change is exogenous to the effect of landscape heterogeneity on phenology (where factors of landscape heterogeneity can be quantified in the absence of climate change).

Fully nested data: data with sufficient spatial and/or temporal precision that individual records can be resolved at a given small scale and then aggregated to larger or 'coarser' scales that fully contain the smaller units and all their individual records and have the same extent, without loss of information. This process can be repeated for multiple changes in resolution.

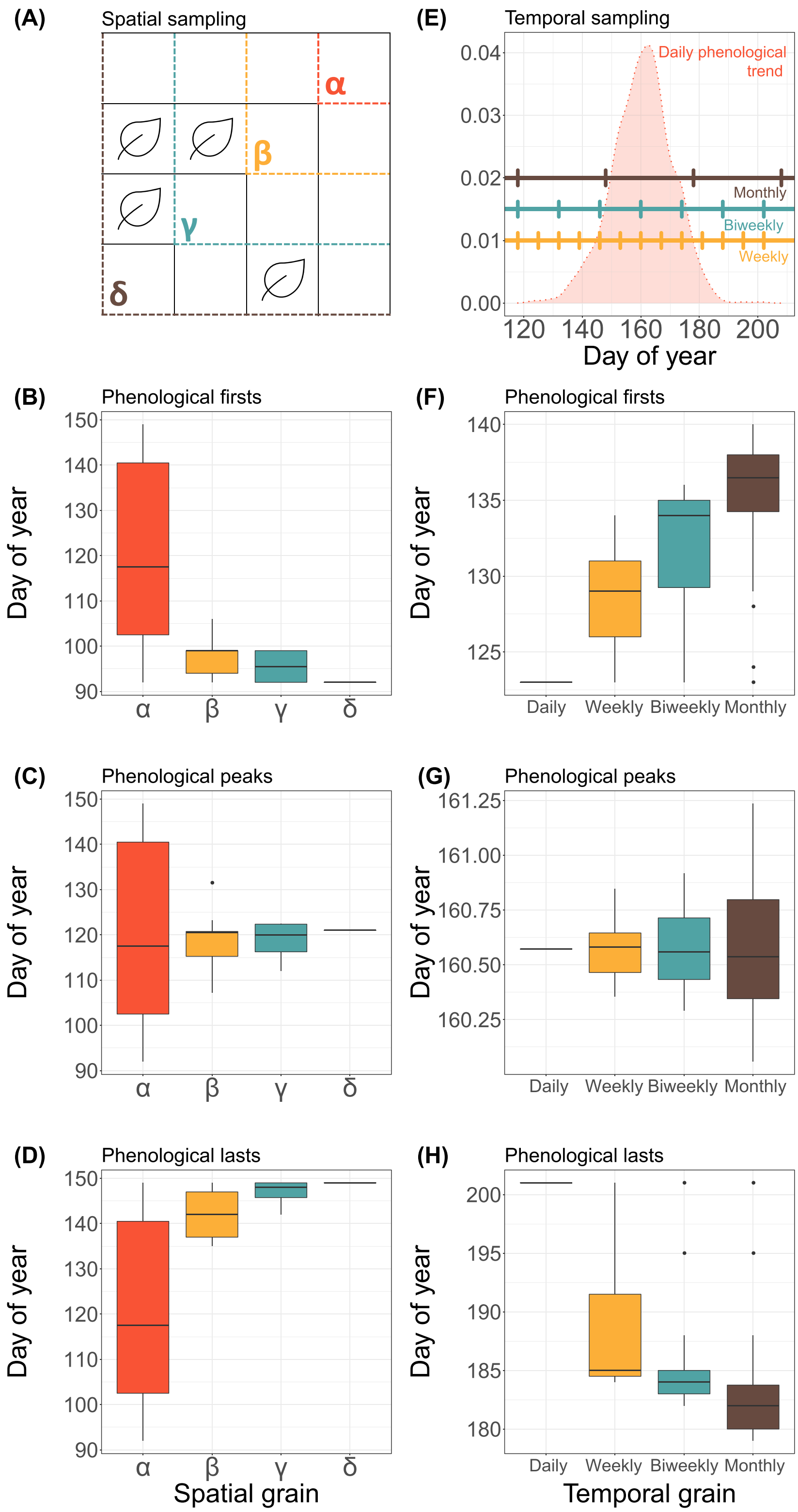
Land surface phenology: variation in vegetation patterns across land surfaces due to seasonality, as observed from remote sensing.

Landscape phenology: the study of the causes and consequences of spatial variation in the timing of biological events.

Nonstationarity: a property of system variables in which modeled relationships or parameter choices are valid in one environment (or at a particular scale) but may not hold when applied to substantially different systems or future environments.

Phenological firsts and lasts: Of the many available phenological metrics that can be measured or estimated, firsts and lasts represent the beginning and ending of processes, respectively (rather than the peak or duration). Firsts include, for example, first flower, first egg laying, first fruiting, start of (growing) season, and onset of fall leaf color and their associated dates. Lasts include (dates of) senescence, end of (growing) season, and last frost. Phenological firsts are not uniquely associated with spring phenology, and phenological lasts are not uniquely associated with fall phenology, which makes these concepts useful even in aseasonal tropical systems.

Phenological metrics: measures of the timing of biological events, including their beginning, peak, end, duration, and



(See figure legend at the bottom of the next page.)

velocity; these can be extracted from phenological time series.

Phenological peaks: phenological metrics that represent the highest amplitude of a process (or occasionally the estimated highest point based on the midpoint of the duration of the process). Phenological peaks include, for example, (dates of) peak greenness, peak flowering, peak Normalized Differential Vegetation Index (NDVI), and peak autumn leaf color. Timing of phenological peaks is not necessarily correlated with timing of phenological firsts and lasts.

Phenological sensitivity: differences in phenological responses among groups of organisms to climate change or other exogenous factors.

Phenological time series: records of biological events or other seasonal variables (e.g., greenness, productivity, presence of snowpack, multiyear observations of bird arrivals during migrations, leaf senescence, phytoplankton blooms) that are tracked through time. Phenological metrics can be extracted from these time series; for example, start of season and end of season dates can be estimated from the Normalized Differential Vegetation Index (NDVI). Phenological time series are often regressed against time series of other variables, such as climate, to understand the drivers of phenology.

Phenology: the timing of life-history events and other seasonal phenomena.

Scale dependence: the property of metrics or patterns (including ecological ones) to change with the scale of observation. Scale dependence is a universal property of complex systems.

Scaling laws: mathematical relationships between the expected value of a metric (such as a phenological metric) and the spatial or temporal scale at which it is measured. Scaling laws are important in landscape ecology and macroecology to bridge scales and predict metrics at scales that have not been measured.

possible to simply extrapolate phenological knowledge across existing scale gaps, and scale gaps present in the literature directly represent gaps in our knowledge of phenology across the globe.

The ecological consequences of scale gaps in phenology

The studies reviewed here each make important contributions to our understanding of landscape phenology, but most are suited to a particular scale. As a consequence of scale dependence, landscape ecology theory [44,48,51,52,61] suggests that most of the climate–phenology relationships that are derived from small-scale field studies should not be directly used to inform models of ecological processes at coarser spatial grains without accounting for the effects of scale mismatch. Similarly, downscaling climate–phenology relationships derived from coarse-grain climate models or remote sensing to the scale of individual organisms or study sites is a nontrivial problem. Indeed, the effects of spatiotemporal observation scale on phenological inferences have been documented in a handful of recent empirical studies. For instance, Tian *et al*. [62] found that rural–urban differences in phenology can be amplified with spatial resolution, and Xie and Wilson [58] demonstrated that 8-day and twice-daily time series can differ in inferences of leaf senescence by over 30 days.

Scale gaps represent a lack of knowledge about landscapes and regions, which may have profound consequences for the study of ecological processes. In particular, because organisms interact with landscapes at fundamentally different spatial and temporal scales, the seasonal progression of biological and ecological events that they experience and how those events are changing in response to forcings such as climate change may differ dramatically [63]. For example, pollinators with very different foraging distances, such as hummingbirds [64] and pollinating insects [65,66], may have different capacities to track resources over heterogeneous landscapes. This means that the spatial scale at which pollinators' phenological synchrony with their floral resources is relevant might differ between these two groups, and they may experience climate change very differently.

The critical links between behavior, dispersal scale, and environmental heterogeneity have been explored in a few systems [67], but more such work is needed across different types of ecosystems and communities to understand the broad consequences of scale-dependent phenology. With the possible exception of overstory trees, the phenology of many organisms remains understudied at landscape to regional spatial scales. Perhaps more troubling is the notion that the sensitivity of **phenological metrics** to **exogenous** forcings such as climate change is also scale dependent [68–70]. As researchers scramble to assemble ecological data into predictive frameworks for anticipating the impacts of global change on biodiversity [71,72], the climate sensitivities of phenological transition dates serve as critical links between climate, ecosystems, and the global carbon cycle [73]. Approaches for reconciling observations made at different taxonomic and spatiotemporal resolutions and extents are therefore at the forefront of current landscape phenology research [74].

Figure 1. The effects of spatial and temporal sampling grain on phenological observations due to aggregation. (A) A large number of small grain size (or high-resolution) 'pixels' (α) covering a given spatial extent will capture greater variance in a phenological metric than a larger grain size (α < β < γ < δ). However, as grain size increases, (B) individual phenological firsts within each pixel (or grain) are tallied earlier and (D) phenological lasts are tallied later, while (C) phenological peaks, defined here as the mean date of a phenological event, remain relatively constant. (E) The temporal grain of a study represents the frequency of observation. As observations become less frequent, (F) phenological firsts tend to be recorded later and (H) phenological lasts earlier, while (G) phenological peaks are observed relatively consistently. These phenological trends are based on randomly generated distributions of numbers and thus reflect scale dependence purely due to statistical aggregation. This illustrates that if detection is perfect, observations made at high spatial resolution will report less extreme phenological firsts and lasts over the average pixel, while higher-frequency observations will report more extreme dates for firsts and lasts.

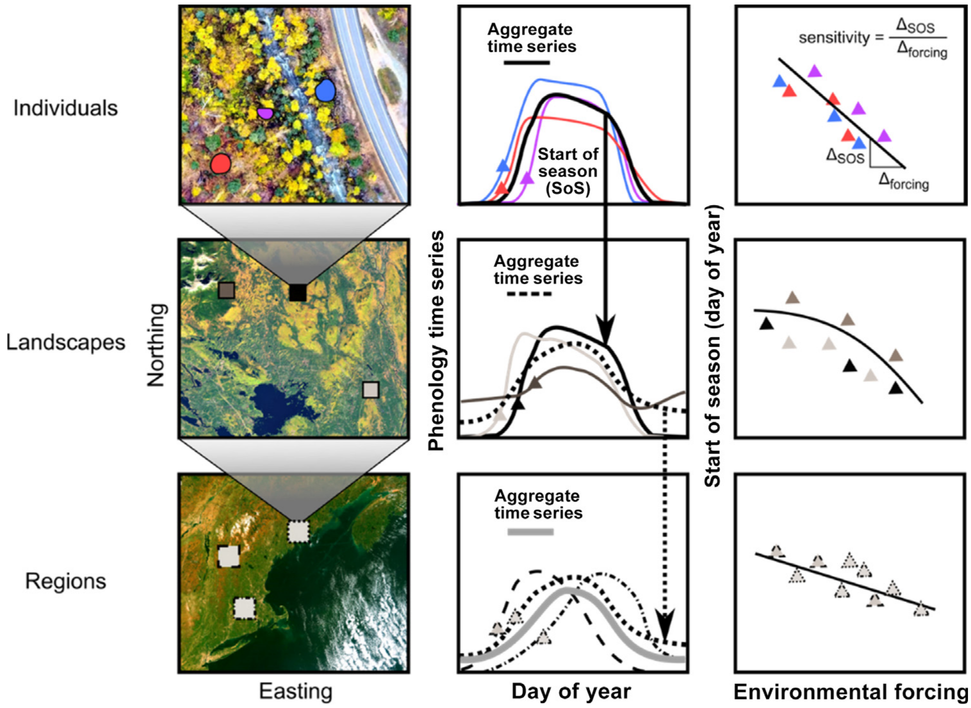


Figure 2. Measures of phenology, and the relationships between those measures and environmental forcing, are inherently scale dependent. At small spatial scales (top row), study units are often individual organisms (left panel), which have unique trajectories of attributes such as greenness across the season (phenology time series; colored curves in the center panel). Measures that are extracted from these time series, such as the timing of greenup onset [start of season (SOS); triangles] are often related to environmental forcing to derive measures of **phenological sensitivity**, expressed as the slope of the best-fit line relating environmental forcing to the phenology measure. As the spatial grain of observation becomes coarser and the extent larger, encompassing landscapes (center row) or regions (bottom row), phenology time series are aggregated over increasingly heterogeneous individuals and environments. This can change the shape of the phenology time series and measurements (such as SOS) that are extracted from them. Changes in observation scale can change both the absolute values of phenology measures and the relationships between those measures and environmental forcing, resulting in changes from linear to nonlinear relationships [68,69,91] (center right panel) and changes in phenological sensitivity (bottom right panel).

What are the spatial and temporal scale gaps in landscape phenology?

Spatial and temporal grains surveyed in landscape phenology research have historically been constrained by logistical considerations and the availability of sensing platforms and datasets, resulting in dramatic scale gaps in space and time. We assessed the multitude of scales and methods used through a systematic review of the recent landscape phenology literature (Box 1) and identified gaps and biases present in landscape-level examinations of phenological patterns in natural systems.

Despite the increasingly global coverage of freely available remote sensing data (i.e., satellite imagery), landscape phenology studies were concentrated in the temperate Northern Hemisphere, with the vast majority of studies from northeastern North America, Western Europe, and Northeast Asia (Figure 3). Less than one-third of the studies examined phenology in the Southern Hemisphere (26%), and fewer still took place in the tropics (~15%). In terms of spatial scale, the literature was dominated by field studies of individual organisms, small study plots, and coarse-grain remote sensing studies, with relatively few studies dealing with phenology at intermediate spatial grains (10–250 m; 12%). Likewise, <2% of the studies that reported sampling frequency used

Box 1. Systematic review of the recent landscape phenology literature

Our systematic and quantitative review of landscape phenology literature consisted of a survey of the Web of Science on August 1, 2020, using the Boolean search term '((phenology OR phenological) AND climat* AND (scal* OR spatial OR landscape*)) NOT ALL=(agricult* OR crop* OR agronom* OR horticult* OR cultivar* OR wine OR viticult* OR cultivation* OR farm*) AND LANGUAGE: (English),' excluding reviews and non–peer-reviewed literature. This resulted in 754 articles across the 12 journals that had the most articles meeting these initial criteria. We excluded studies that solely involved human-controlled and/or human-manipulated systems in order to focus our review on natural systems and responses. We removed studies not primarily focused on phenological patterns, those focused only on abiotic systems, and those focused on the changing migration patterns of highly vagile organisms, as it is difficult to assign spatial scale to such systems. As a result, of the 258 studies that met our criteria, only 19 studies examining the phenology of organisms other than terrestrial plants were included (10 animal, 1 fungal, and 8 plankton). Phenological metrics (i.e., dates or durations of specific events) were extracted from phenological time series (e.g., greenness, color, flowering, and measures of abundance). Most of the studies reviewed used remote sensing techniques and used spectral indices to infer the onset of phenological events and processes (Figure I). This trend has led to distinct gaps in spatiotemporal resolution across the literature.

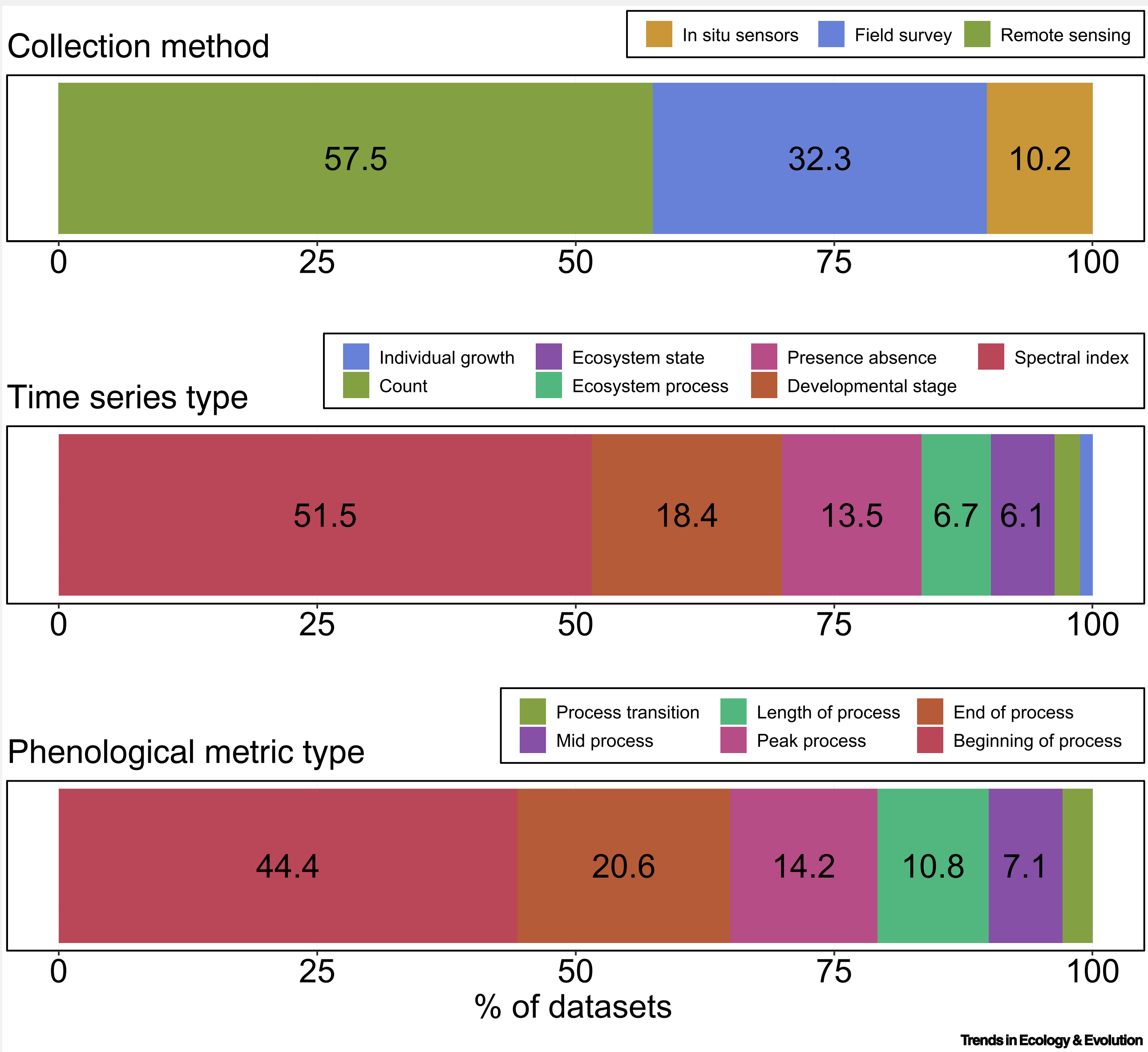


Figure I. Variation in study method, phenological time series examined, and phenological metric of concern across the reviewed literature. Numbers represent percentages among all datasets examined by the reviewed studies for which information was reported.

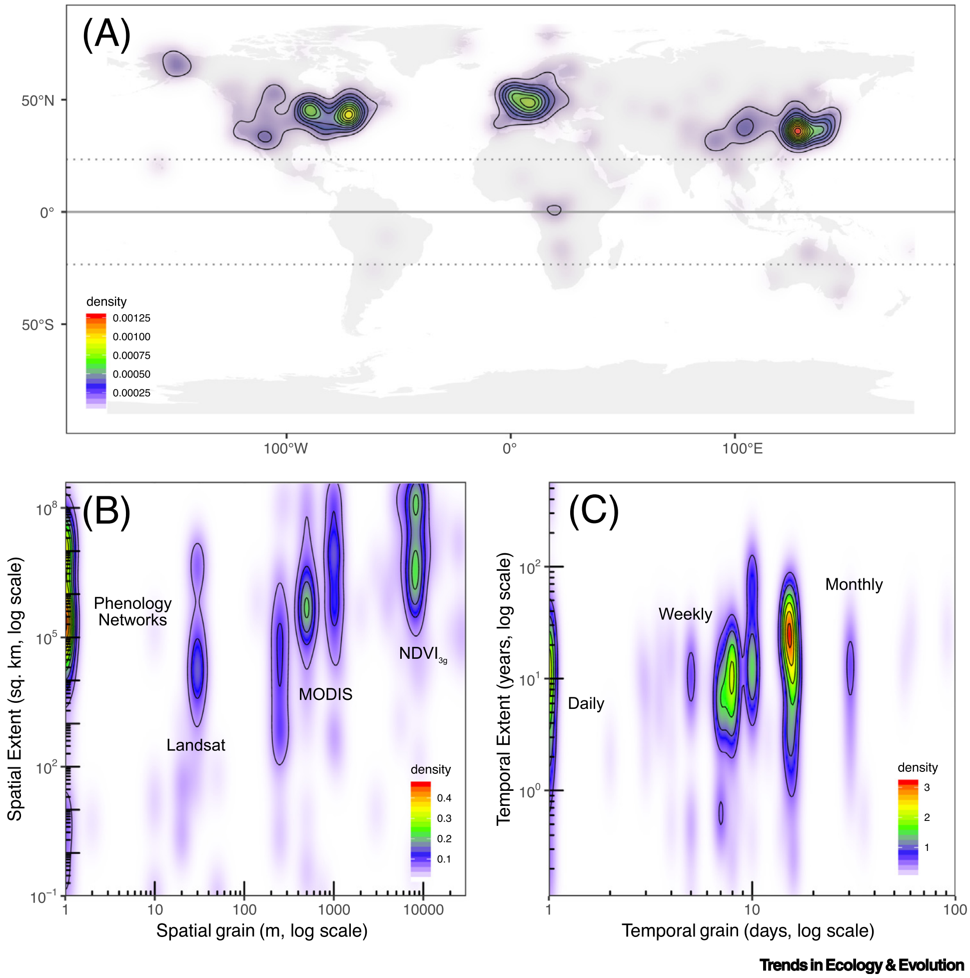


Figure 3. Heatmaps of the geographic distributions and spatial and temporal scales of analysis in the papers included in our review of the landscape phenology literature. Top panel shows the density of study area centroids, highlighting hotspots in the temperate Northern Hemisphere (northeastern North America, Northern Europe, and Northeast Asia) and a dearth of studies in the tropics and southern latitudes. The map does not include studies with a global or hemispheric extent. The bottom left panel shows the spatial grain and extent of the studies, highlighting clusters of studies using point measurements from phenological field observations ('Phenology Networks') and remote sensing observations at Landsat (~30 m), MODIS (Moderate Resolution Imaging Spectroradiometer) (250, 500, 1000 m), and AVHRR (Advanced Very High Resolution Radiometer) imagery aggregated at ~8 km as part of the Normalized Differential Vegetation Index (NDVI3g) dataset. The bottom right panel shows the temporal grain (sampling interval) and temporal extent (study duration) of the studies, highlighting clusters of studies at daily, weekly, and monthly sampling intervals.

intermediate temporal grains (2–7-day sampling intervals), despite the temporal durations of studies ranging from weeks to >100 years.

These sampling patterns are driven by the wide availability of coarse-grain satellite imagery and data products [75], and recent methodological and technical advances that have facilitated their analysis [76,77]. Similarly, although the rapid digitization and online mobilization of natural history collections and historical surveys have greatly increased the temporal span of phenological studies, the temporal distribution of data points assembled from such sources is often irregular and biased [78,79]. Furthermore, only a handful of studies incorporate datasets at multiple spatial and temporal scales or explicitly address issues of scale in their analysis, and a large proportion of field studies did not clearly report the spatial aspects (extent and resolution) of the data used. Compounded by the general caveats and limitations of integrating information collected with different methods and standards, this further hinders our ability to link phenological knowledge across scales, as different scales of observation may result in fundamentally different observations of phenological timing.

Addressing the gaps: beyond spectral properties of landscapes and phenological firsts

A majority (59%) of landscape phenology studies in our review use time series of the spectral properties of ecosystems (e.g., Normalized Differential Vegetation Index) as a proxy for ecological processes (see Figure I in Box 1,). Although these spectral properties are simple to calculate, they often relate only indirectly to the ecological processes of interest. Platforms such as tower-mounted sensors (e.g., phenocams) and unmanned aerial vehicles (e.g., drones) can provide alternatives to gather near-surface, intermediate-scale data at fine spatiotemporal resolutions that can both address the scale gaps between plot-scale observations and regional or global scale remote sensing, and they can be used to more efficiently link **environmental forcings** to changes in ecological processes [60,80,81]. It has also become increasingly feasible to gather on-the-ground phenological data across larger spatial extents at relatively fine resolutions through community science efforts such as the USA National Phenology Network and iNaturalist. Furthermore, the rapid digitization of museum/herbarium collections has facilitated the extraction of phenological information across great spatiotemporal extents [64], aided by advances in machine learning and crowdsourcing platforms [24,82,83].

Despite the wide variety of **phenological time series** examined, start-of-season measures (e.g., start of spring or start of growing season extrapolated via greenness indices) were by far the most common measures in our review (Box 1). Comparatively fewer studies examined the durations or peaks of phenological processes (11%). While it can be easier and more efficient to detect and record the beginning of a process, phenological 'firsts' represent one extreme of the process and may be affected by population size and sampling effort [84]. Although biologically and ecologically meaningful, the timing of the start of a phenological event such as flowering does not always correlate with the timing of the peak or duration of the process [85], which may better determine ecologically critical factors such as reproductive success in plants [86]. Finally, measures of phenological extremes (firsts and lasts) and their sensitivity to environmental forcings can be especially sensitive to the spatiotemporal scale of observation – when we vary the grain of observation by aggregating phenological metrics, we can see that there is a tendency for later **phenological firsts** and earlier **phenological lasts** to be estimated at smaller spatial and larger temporal grains due to statistical aggregation (Figure 1). Measures of **phenological peaks** tend to be significantly less sensitive to spatiotemporal grain and are as efficient to estimate as firsts and lasts in some situations, such as remote sensing studies and those incorporating natural history collections [20,87].

Addressing the gaps: putting scaling rules to work

Mathematical scaling relationships (or rules) allow us to form null expectations about how spatial and temporal resolution influence conclusions drawn from particular scales of measurement. If scaling rules can be constructed to explain the relationship between a particular phenological metric and the landscape in which it is being evaluated, we can expect and explain mismatches between data measured at different scales (e.g., field surveys and remote sensing). Scaling rules can allow researchers to use small-scale data with **fully nested data** structure to go beyond scale-specific measures to predict relationships between phenological time series and scale at scales that are not measured empirically, as is often done in macroecological studies [88]. Data collected from larger scales can then be used to validate scaling relationships. Simulation studies, especially those that are directly informed by empirical data, are a powerful way to construct and investigate scaling relationships and to investigate phenological metrics' sensitivities to exogenous factors [49,70,89,90].

Ecosystems are expected to contain different scaling relationships with respect to every landscape phenology metric being estimated or measured. Beyond the inherent scale dependence of the metrics themselves, exogenous forcings such as climate change are expected to change the timing of phenological events. Climate change will also alter the phenological scaling relationships themselves (through a process called '**nonstationarity**') in ways that have not yet been investigated. This constitutes a fundamental challenge in applying model outcomes to unstudied or future systems [69,91]. Multiscale investigations of landscape metrics, the construction of scaling relationships for individual phenological metrics, and the application of those relationships (or **scaling laws**) for prediction in unmeasured scales are therefore critical in synthesizing expected changes to phenological metrics across scales and levels of statistical aggregation [61,70,92].

In the simplest case, scaling rules provide rigorous tests of the robustness of measurements to scale effects. If scale effects are not negligible, then several strategies can be effective. In the case of upscaling (where dependent variables are collected at a finer grain than the scale of desired inference), empirical scaling relationships provide a direct basis for adjusting estimates of the dependent variable to those expected for the appropriate spatial grain. In this case, scaling rules can alleviate bias due to scale effects, but biases and loss of information may persist unless small-scale samples are representative and appropriate aggregation methods are used [57,93]. Adjustments based on estimated scaling relationships are also useful in the case of downscaling (where the dependent variable is collected at coarser spatial or taxonomic grain than desired), and several techniques can also be used to come up with estimates at finer spatial or taxonomic scales, including mixture models [94], constructed analogs [95], and machine learning [96]. Often, such adjustments can take the form of simple mathematical corrections or altered priors. These techniques have seen less frequent use in landscape phenology research, but they are promising avenues for enhancing coarse-grained observations in heterogeneous environments.

Establishing explicit scaling rules can require scale-explicit multivariate data collected across multiple scales [68,80]. Such data is not always available and can be costly to collect. However, studies are increasingly taking advantage of new data collection methods such as ground-based cameras ('phenocams') [81], citizen science efforts [30], and drones [60], enabling data collection at fine spatial and temporal grains across large spatial extents. In addition to facilitating the establishment of scaling rules, such methods, when applicable, enable researchers to tailor the spatial and temporal properties of sampling to the ecological question rather fitting the question to the available data.

Even when multivariate data are not fully available across the scales of intended inference, it is still possible to account for certain aspects of scale dependence and increase the accuracy and validity of phenological extrapolation. For instance, as we demonstrate in Figure 1, the effects of statistical aggregation on phenological metrics can be relatively easily defined within a spatiotemporal extent by varying the grain of observation, using either real or simulated data. Likewise, as high-resolution geoclimatic data have become increasingly available, it is possible to estimate how environmental heterogeneity across a spatiotemporal extent may affect phenological aggregation or prediction across grains by varying the degree of heterogeneity across space and time, especially when species' or ecosystems' phenological responses to environmental cues are known. Furthermore, the uncertainty caused by spatiotemporal gaps can directly inform predictions. Along these lines, Pearse *et al*. [87] demonstrated that the timing of phenological firsts can be accurately estimated from sparsely and irregularly collected data by drawing information from the sampled distribution of observations as opposed to just the first observation. Thus, accounting for even a subset of scaling effects can increase the accuracy of phenological inference across scales and allow us to formulate null expectations about what is driving variation in phenology.

Concluding remarks and future perspectives

Landscape phenology is poised to revolutionize our understanding of the interplay between environmental heterogeneity and seasonal ecological processes. Historically, this type of research has been constrained by a limited set of observational tools that collect data at a small number of discrete spatiotemporal scales, leading to large gaps in our understanding of this complex set of phenomena. New observational and analytical methods make it increasingly possible to tailor the grain and extent of observations to the ecological question at hand and examine how phenological landscapes vary across a wide continuum of spatial and temporal scales. We draw attention to critical scale gaps in the literature that can be bridged through the construction of individual metric–level scaling laws and met by developing technologies that are increasingly capable of generating fully nested, high-resolution visual data and artificial intelligence that can analyze such imagery. The application of such methods and technologies to phenological studies is still limited but nonetheless demonstrates great promise. Along these lines, it is important to continue to support and expand large-scale field surveys and community-driven efforts to monitor on-the-ground phenology, preferably targeting existing scale gaps. As the number of observations and observation types increases and as researchers attempt to synthesize phenology studies into generalizable understanding across scales and geographies, it is critical to recognize and report the scale dependence of phenology time series and the measures we extract from them. Such efforts also facilitate the integration of existing methods and datasets,

Outstanding questions

What are the relative effects of climate change on landscape phenology in temperate regions, where climatic variability is low across space but high across (seasonal) time, versus in tropical regions, where climatic variability can be high across space but low across time?

How does phenological sensitivity to climate vary across a species's range?

Do phenological shifts at the local community scale translate to consistent directional shifts across the entire range of a species and vice versa?

How do density of species and diversity of taxa present affect phenological inference across scales?

How has the duration of phenological events such as flowering responded to climate change across species' ranges?

At what spatiotemporal scales can the constituents of mutualisms and trophic interactions track each other's phenological changes?

How can we better apply artificial intelligence to integrate phenological observations made with different technologies (at different scales)?

Box 2. Recommendations for future work

(i) Standardize terminology: Researchers should be explicit about the phenological time series that was tracked and report all the phenological metrics extracted. For example, instead of reporting that a phenocam project tracked 'greenness,' the report could specifically mention the Normalized Differential Vegetation Index (NDVI) as the time series tracked, with start of season, peak greenness, and end of season as the metrics extracted. A standardized ontology of phenological terms and concepts is available [97].

(ii) Make sampling units spatially explicit: Researchers should report the spatial units, sampling interval, and spatiotemporal grain and extent of the study.

(iii) Increase geographic specificity: Field studies should record and report the Global Positioning System (GPS) coordinates of all individuals that were tracked when possible and report the boundaries of study areas over which measurements were aggregated.

(iv) Where appropriate to the hypothesis, select metrics that are more independent of spatial and temporal scaling effects than others, such as measures of the middle or the peak of a process: Analyses that use these are more likely to generalize across scales.

(v) Incorporate multiscale data when possible: The rise of new platforms for collecting phenology data over the past decade, including digitized museum collections [64], citizen-scientist networks [86], ground-based cameras [98], drones [71], and higher-resolution satellite sensors [72], present unique opportunities to bridge gaps in spatiotemporal scales. These platforms allow researchers to measure the spatial variation and scale dependence of phenology at scales relevant to individual organisms, a key advance for understanding how ecological interactions scale from individuals to populations and communities.

(vi) Summarize scale dependence in phenology as empirical scaling rules and use their mathematical forms to address unmeasured gaps and extrapolate to unmeasured scales: This could be achieved by making use of a fully nested data structure, an approach common in macroecology [99,100]. The resources and technologies required for this are not yet widely accessible, with some exceptions such as multiscale observations of greenness in dominant canopy tree species or systems where the maximum spatiotemporal extent is relatively narrow.

(vii) Target unmeasured spatial and temporal scales for studies of phenological metrics: These intermediate-scale results can be used to validate predictions from scaling laws developed from fully nested data at smaller scales.

which remain largely heterogeneous in their application and presentation. Although the majority of the studies we reviewed were focused on terrestrial plants, the issues and concepts we illustrate have generality for other kinds of organisms, even those with more complex and dynamic ecologies (e.g., highly vagile organisms that may lack data at fine temporal grains and large spatial extents). Neglecting the important issues associated with scale dependence can cause extrapolation errors with large consequences for our understanding of how organisms respond to environmental change. It is therefore critical that we understand how to best capture biologically and ecologically meaningful data for ecosystems and species, including which resolutions and extents of data are relevant and how to integrate multiple data sources without introducing statistical bias associated with scale. Future studies may be able to use scaling rules to derive new insights into the causes and consequences of shifting phenological landscapes as well as broader questions about the spatial and temporal organization of ecological communities (see Outstanding questions). To this end, we present recommendations for maximizing the potential and applicability of future work in landscape phenology (Box 2).

Acknowledgments

We express gratitude to D.H. Hembry and A.L. Rhodes for aiding literature review. We also thank B.J. Enquist and members of the Bridging Biodiversity and Conservation Science (BBCS) initiative at the University of Arizona Institutes for Resilience for their insights and comments. D.S.P. was supported by NSF-DEB 1754584, and I.K.B. was supported by a National Science Foundation Postdoctoral Fellowship in Biology (PRFB-1711936).

Declaration of interests

The authors have no interests to declare.

References

1. Morisette, J.T. *et al.* (2009) Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. *Front. Ecol. Environ.* 7, 253–260
2. Yang, L.H. and Rudolf, V.H.W. (2010) Phenology, ontogeny and the effects of climate change on the timing of species interactions. *Ecol. Lett.* 13, 1–10
3. Richardson, A.D. *et al.* (2013) Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agric. For. Meteorol.* 169, 156–173
4. Keenan, T.F. *et al.* (2014) Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nat. Clim. Chang.* 4, 598–604
5. Root, T.L. *et al.* (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421, 57–60
6. Willis, C.G. *et al.* (2008) Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17029–17033
7. Van Buskirk, J. *et al.* (2009) Variable shifts in spring and autumn migration phenology in North American songbirds associated with climate change. *Glob. Chang. Biol.* 15, 760–771
8. Inouye, D.W. (2008) Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology* 89, 353–362
9. Miller-Rushing, A.J. *et al.* (2008) How well do first flowering dates measure plant responses to climate change? The effects of population size and sampling frequency. *J. Ecol.* 96, 1289–1296
10. Primack, D. *et al.* (2004) Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *Am. J. Bot.* 91, 1260–1264
11. Parmesan, C. and Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421, 37–42
12. Corlett, R.T. and Lafrankie, J.V. (1998) Potential impacts of climate change on tropical Asian forests through an influence on phenology. *Clim. Chang.* 39, 439–453
13. Edwards, M. and Richardson, A.J. (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430, 881–884
14. Sackett, T.E. *et al.* (2011) Response of macroarthropod assemblages to the loss of hemlock (*Tsuga canadensis*), a foundation species. *Ecosphere* 2, art74
15. Kudo, G. and Ida, T.Y. (2013) Early onset of spring increases the phenological mismatch between plants and pollinators. *Ecology* 94, 2311–2320
16. Kendrick, J.A. *et al.* (2015) Changes in canopy structure and ant assemblages affect soil ecosystem variables as a foundation species declines. *Ecosphere* 6, 1–20
17. Delmas, E. *et al.* (2019) Analysing ecological networks of species interactions. *Biol. Rev. Camb. Philos. Soc.* 94, 16–36
18. Price, M.V. and Waser, N.M. (1998) Effects of experimental warming on plant reproductive phenology in a subalpine meadow. *Ecology* 79, 1261–1271
19. Carter, S.K. and Rudolf, V.H.W. (2019) Shifts in phenological mean and synchrony interact to shape competitive outcomes. *Ecology* 100, e02826
20. Davis, C.C. *et al.* (2015) Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms. *Am. J. Bot.* 102, 1599–1609
21. Beaubien, E.G. and Hamann, A. (2011) Plant phenology networks of citizen scientists: recommendations from two decades of experience in Canada. *Int. J. Biometeorol.* 55, 833–841
22. Zhang, X. *et al.* (2003) Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* 84, 471–475
23. Park, I.W. (2012) Digital herbarium archives as a spatially extensive, taxonomically discriminate phenological record; a comparison to MODIS satellite imagery. *Int. J. Biometeorol.* 56, 1179–1182
24. Willis, C.G. *et al.* (2017) CrowdCurio: an online crowdsourcing platform to facilitate climate change studies using herbarium specimens. *New Phytol.* 215, 479–488
25. Brenskelle, L. *et al.* (2019) Integrating herbarium specimen observations into global phenology data systems. *Appl. Plant Sci.* 7, e01231
26. Zohner, C.M. and Renner, S.S. (2014) Common garden comparison of the leaf-out phenology of woody species from

different native climates, combined with herbarium records, forecasts long-term change. *Ecol. Lett.* 17, 1016–1025
27. Zhang, H. *et al.* (2015) Sensitivity of flowering phenology to changing temperature in China. *J. Geophys. Res. Biogeosci.* 120, 1658–1665
28. Cole, E.F. and Sheldon, B.C. (2017) The shifting phenological landscape: within-and between-species variation in leaf emergence in a mixed-deciduous woodland. *Ecol. Evol.* 7, 1135–1147
29. Asam, S. *et al.* (2018) Relationship between spatiotemporal variations of climate, snow cover and plant phenology over the Alps – an earth observation-based analysis. *Remote Sens.* 10, 1757
30. Park, D.S. *et al.* (2018) Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity across the eastern United States. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 374, 20170394
31. Park, D.S. *et al.* (2020) Phenological displacement is uncommon among sympatric angiosperms. *bioRxiv* Published online August 5, 2020. https://doi.org/10.1101/2020.08.04.236935
32. Liang, L. and Schwartz, M.D. (2009) Landscape phenology: an integrative approach to seasonal vegetation dynamics. *Landsc. Ecol.* 24, 465–472
33. Wang, J. *et al.* (2020) Multi-scale integration of satellite remote sensing improves characterization of dry-season green-up in an Amazon tropical evergreen forest. *Remote Sens. Environ.* 246, 111865
34. Olliff-Yang, R.L. and Ackerly, D.D. (2020) Topographic heterogeneity lengthens the duration of pollinator resources. *Ecol. Evol.* 10, 9301–9312
35. Koebsch, F. *et al.* (2020) Refining the role of phenology in regulating gross ecosystem productivity across European peatlands. *Glob. Chang. Biol.* 26, 876–887
36. Zohner, C.M. *et al.* (2020) Late-spring frost risk between 1959 and 2017 decreased in North America but increased in Europe and Asia. *Proc. Natl. Acad. Sci. U. S. A.* 117, 12192–12200
37. Crawley, M.J. and Harral, J.E. (2001) Scale dependence in plant biodiversity. *Science* 291, 864–868
38. Rahbek, C. and Graves, G.R. (2001) Multiscale assessment of patterns of avian species richness. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4534–4539
39. Davies, K.F. *et al.* (2005) Spatial heterogeneity explains the scale dependence of the native–exotic diversity relationship. *Ecology* 86, 1602–1610
40. Park, D.S. *et al.* (2020) Darwin's naturalization conundrum can be explained by spatial scale. *Proc. Natl. Acad. Sci. U. S. A.* 117, 10904–10910
41. Primack, R.B. and Miller-Rushing, A.J. (2012) Uncovering, collecting, and analyzing records to investigate the ecological impacts of climate change: a template from Thoreau's Concord. *Bioscience* 62, 170–181
42. Bigler, C. and Vitasse, Y. (2019) Daily maximum temperatures induce lagged effects on leaf unfolding in temperate woody species across large elevational gradients. *Front. Plant Sci.* 10, 398
43. McCulloch, G.A. and Waters, J.M. (2018) Testing for seasonality in alpine streams: how does altitude affect freshwater insect life cycles? *Freshw. Biol.* 63, 483–491
44. Levin, S.A. (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur Award Lecture. *Ecology* 73, 1943–1967
45. Schneider, E.E. *et al.* (2019) Tree spatial patterns modulate peak snow accumulation and snow disappearance. *For. Ecol. Manag.* 441, 9–19
46. Inouye, B.D. *et al.* (2019) Phenology as a process rather than an event: from individual reaction norms to community metrics. *Ecol. Monogr.* 89, e01352
47. Heim, K.C. *et al.* (2019) A general model of temporary aquatic habitat use: water phenology as a life history filter. *Fish Fish.* 20, 3
48. Turner, M.G. *et al.* (1989) Effects of changing spatial scale on the analysis of landscape pattern. *Landsc. Ecol.* 3, 153–162
49. Wagner, H.H. and Fortin, M.-J. (2005) Spatial analysis of landscapes: concepts and statistics. *Ecology* 86, 1975–1987
50. Chen, X. *et al.* (2018) The mixed pixel effect in land surface phenology: a simulation study. *Remote Sens. Environ.* 211, 338–344
51. Woodcock, C.E. and Strahler, A.H. (1987) The factor of scale in remote sensing. *Remote Sens. Environ.* 21, 311–332
52. Wiens, J.A. (1989) Spatial scaling in ecology. *Funct. Ecol.* 3, 385–397
53. Sandel, B. and Smith, A.B. (2009) Scale as a lurking factor: incorporating scale-dependence in experimental ecology. *Oikos* 118, 1284–1291
54. Chuine, I. *et al.* (2000) Scaling phenology from the local to the regional level: advances from species-specific phenological models. *Glob. Chang. Biol.* 6, 943–952
55. Badeck, F.-W. *et al.* (2004) Responses of spring phenology to climate change. *New Phytol.* 162, 295–309
56. Hamil, K.-A.D. *et al.* (2016) Cross-scale contradictions in ecological relationships. *Landsc. Ecol.* 31, 7–18
57. Zhang, X. *et al.* (2017) Exploration of scaling effects on coarse resolution land surface phenology. *Remote Sens. Environ.* 190, 318–330
58. Xie, Y. and Wilson, A.M. (2020) Change point estimation of deciduous forest land surface phenology. *Remote Sens. Environ.* 240, 111698
59. Peng, D. *et al.* (2017) Scaling effects on spring phenology detections from MODIS data at multiple spatial resolutions over the contiguous United States. *ISPRS J. Photogramm. Remote Sens.* 132, 185–198
60. Klosterman, S. *et al.* (2018) Fine-scale perspectives on landscape phenology from unmanned aerial vehicle (UAV) photography. *Agric. For. Meteorol.* 248, 397–407
61. Wu, J. (2004) Effects of changing scale on landscape pattern analysis: scaling relations. *Landsc. Ecol.* 19, 125–138
62. Tian, J. *et al.* (2020) Coarse-resolution satellite images overestimate urbanization effects on vegetation spring phenology. *Remote Sens.* 12, 117
63. Beard, K.H. *et al.* (2019) The missing angle: ecosystem consequences of phenological mismatch. *Trends Ecol. Evol.* 34, 885–888
64. Araujo, A.C. *et al.* (2018) Spatial distance and climate determine modularity in a cross-biomes plant–hummingbird interaction network in Brazil. *J. Biogeogr.* 45, 1846–1858
65. López-Uribe, M.M. *et al.* (2015) Nest suitability, fine-scale population structure and male-mediated dispersal of a solitary ground nesting bee in an urban landscape. *PLoS One* 10, e0125719
66. Memmott, J. *et al.* (2007) Global warming and the disruption of plant-pollinator interactions. *Ecol. Lett.* 10, 710–717
67. Hindle, B.J. *et al.* (2015) Topographical variation reduces phenological mismatch between a butterfly and its nectar source. *J. Insect Conserv.* 19, 227–236
68. Gonzalez, A. *et al.* (2020) Scaling-up biodiversity-ecosystem functioning research. *Ecol. Lett.* 23, 757–776
69. Newman, E.A. *et al.* (2019) Scaling and complexity in landscape ecology. *Front. Ecol. Evol.* 7, 293
70. Newman, E.A. *et al.* (2021) Disentangling the effects of climate change, landscape heterogeneity, and scale on phenological metrics. *bioRxiv* Published online February 8, 2021. https://doi.org/10.1101/2021.02.05.429398
71. Bradford, J.B. *et al.* (2018) Anticipatory natural resource science and management for a changing future. *Front. Ecol. Environ.* 16, 295–303
72. Dietze, M.C. *et al.* (2018) Iterative near-term ecological forecasting: needs, opportunities, and challenges. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1424–1432
73. Girardin, C.A.J. *et al.* (2016) Seasonal trends of Amazonian rainforest phenology, net primary productivity, and carbon allocation. *Glob. Biogeochem. Cycles* 30, 700–715
74. Cleland, E.E. *et al.* (2007) Shifting plant phenology in response to global change. *Trends Ecol. Evol.* 22, 357–365
75. Anderson, C.B. (2018) Biodiversity monitoring, earth observations and the ecology of scale. *Ecol. Lett.* 21, 1572–1585
76. Hufkens, K. *et al.* (2018) An integrated phenology modelling framework in R. *Methods Ecol. Evol.* 9, 1276–1285
77. Duarte, L. *et al.* (2018) QPhenoMetrics: an open source software application to assess vegetation phenology metrics. *Comput. Electron. Agric.* 148, 82–94
78. Hedrick, B.P. *et al.* (2020) Digitization and the future of natural history collections. *Bioscience* 70, 243–251

79. Daru, B.H. *et al.* (2018) Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytol.* 217, 939–955
80. Liu, Y. *et al.* (2017) Using data from Landsat, MODIS, VIIRS and PhenoCams to monitor the phenology of California oak/grass savanna and open grassland across spatial scales. *Agric. For. Meteorol.* 237–238, 311–325
81. Richardson, A.D. (2019) Tracking seasonal rhythms of plants in diverse ecosystems with digital camera imagery. *New Phytol.* 222, 1742–1750
82. Lorieul, T. *et al.* (2019) Toward a large-scale and deep phenological stage annotation of herbarium specimens: case studies from temperate, tropical, and equatorial floras. *Appl. Plant Sci.* 7, e01233
83. Davis, C.C. *et al.* (2020) A new method for counting reproductive structures in digitized herbarium specimens using mask R-CNN. *Front. Plant Sci.* 11, 1129
84. Miller-Rushing, A.J. and Primack, R.B. (2008) Global warming and flowering times in Thoreau's Concord: a community perspective. *Ecology* 89, 332–341
85. Ollerton, J. and Lack, A. (1998) Relationships between flowering phenology, plant size and reproductive success in shape *Lotus corniculatus* (Fabaceae). *Plant Ecol.* 139, 35–47
86. Schmitt, J. (1983) Individual flowering phenology, plant size, and reproductive success in *Linanthus androsaceus*, a California annual. *Oecologia* 59, 135–140
87. Pearse, W.D. *et al.* (2017) A statistical estimator for determining the limits of contemporary and historic phenology. *Nat. Ecol. Evol.* 1, 1876–1882
88. McGill, B. (2003) Strong and weak tests of macroecological theory. *Oikos* 102, 679–685
89. Turner, M.G. *et al.* (1989) Predicting across scales: theory development and testing. *Landsc. Ecol.* 3, 245–252
90. Gonzalez, A. and De Feo, O. (2007) Environmental variability modulates the insurance effects of diversity in non-equilibrium communities. In *The Impact of Environmental Variability on Ecological Systems* (Vasseur, D.A. and McCann, K.S., eds), pp. 159–177, Springer, Netherlands
91. Yates, K.L. *et al.* (2018) Outstanding challenges in the transferability of ecological models. *Trends Ecol. Evol.* 33, 790–802
92. Edwards, C. and Crone, E.E. (2020) Fitting phenological curves with Generalized Linear Mixed Models (GLMMs). *bioRxiv* Published online August 19, 2020. https://doi.org/10.1101/2020.06.01.127910
93. Zhao, B. *et al.* (2020) Evaluating autumn phenology derived from field observations, satellite data, and carbon flux measurements in a northern mixed forest, USA. *Int. J. Biometeorol.* 64, 713–727
94. Pan, Y. *et al.* (2012) Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sens. Environ.* 119, 232–242
95. Allstadt, A.J. *et al.* (2015) Spring plant phenology and false springs in the conterminous US during the 21st century. *Environ. Res. Lett.* 10, 104008
96. Gobbi, A. *et al.* (2019) High resolution forecasting of heat waves impacts on leaf area index by multiscale multitemporal deep learning. *arXiv* Published online September 13, 2019. arXiv:1909.07786
97. Stucky, B.J. *et al.* (2018) The plant phenology ontology: a new informatics resource for large-scale integration of plant phenology data. *Front. Plant Sci.* 9, 517
98. Jackson, H.B. and Fahrig, L. (2015) Are ecologists conducting research at the optimal scale? *Glob. Ecol. Biogeogr.* 24, 52–63
99. Harte, J. and Newman, E.A. (2014) Maximum information entropy: a foundation for ecological theory. *Trends Ecol. Evol.* 29, 384–389
100. Harte, J. (2011) *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*, Oxford University Press