

The spatial and temporal domains of modern ecology

Lyndon Estes ^{1,2*}, Paul R. Elsen ^{3,4}, Timothy Treuer⁴, Labeeb Ahmed⁵, Kelly Caylor^{6,7}, Jason Chang⁵, Jonathan J. Choi⁴ and Erle C. Ellis ⁵

To understand ecological phenomena, it is necessary to observe their behaviour across multiple spatial and temporal scales. Since this need was first highlighted in the 1980s, technology has opened previously inaccessible scales to observation. To help to determine whether there have been corresponding changes in the scales observed by modern ecologists, we analysed the resolution, extent, interval and duration of observations (excluding experiments) in 348 studies that have been published between 2004 and 2014. We found that observational scales were generally narrow, because ecologists still primarily use conventional field techniques. In the spatial domain, most observations had resolutions $\leq 1 \text{ m}^2$ and extents $\leq 10,000 \text{ ha}$. In the temporal domain, most observations were either unreplicated or infrequently repeated (> 1 month interval) and ≤ 1 year in duration. Compared with studies conducted before 2004, observational durations and resolutions appear largely unchanged, but intervals have become finer and extents larger. We also found a large gap between the scales at which phenomena are actually observed and the scales those observations ostensibly represent, raising concerns about observational comprehensiveness. Furthermore, most studies did not clearly report scale, suggesting that it remains a minor concern. Ecologists can better understand the scales represented by observations by incorporating autocorrelation measures, while journals can promote attentiveness to scale by implementing scale-reporting standards.

The scales at which ecosystems are observed play a critical role in shaping our understanding of their structure and function^{1–3}. Ecological patterns emerge from temporal and spatial domains that may be coarser or finer than the processes that shape them, which means that investigation across multiple scales is essential for understanding ecological phenomena^{1,4}. This awareness has grown rapidly since the 1980s⁵, accelerated by the need to understand how changes in the global climate, ocean and land systems are affecting everything from individual populations⁶ to entire biomes⁷, while technological advances in areas such as remote sensing and genetics are making it ever-easier to quantify ecological features across a broad and increasing range of scales^{2,5}.

Given the growing awareness of scale, expanding data-gathering capabilities and the fact that the most comprehensive (and arguably best-known) meta-analyses^{8,9} of ecological research scales were published nearly 30 years ago (but see refs^{4,10} for more recent reviews), it is both timely and important to assess the scales of contemporary ecological investigation. To address this need, we quantified the spatial and temporal domains of empirical observations that were reported within recently (2004–2014) published ecological studies. We define domain as the distribution of observations within the spectrum of one or more scale dimensions (note: this definition differs from the ‘domain of scale’³, which is ‘a portion of the scale spectrum within which process–pattern relationships are consistent regardless of scale’), and empirical observations as ecological observations collected under uncontrolled or non-manipulated conditions. Empirical observations are critical for developing and testing the models that explain why ecological patterns vary in time and space^{1,8}; therefore, the spatio-temporal domains of observations

provide an important indicator of the field’s progress towards achieving a holistic, predictive understanding of ecosystems^{1,2}.

Our study focused on two dimensions of spatial scale (that is, resolution (grain) and extent) and two of temporal scale (that is, interval and duration) (Table 1). We analysed the observational domains within each of these four dimensions and between pairs of these dimensions. We also assessed two additional dimensions—actual extent (the summed area of spatial replicates) and actual duration (the summed observational time of temporal replicates)—which we used to evaluate how much the actual scales of observation (that is, how much space and time are covered by the measurement) differ from the scales they ostensibly represent. These differences may impact how effectively observations characterize ecological phenomena. For one, an increasing gap between actual and ostensible observational scales implies greater interpolation or extrapolation of observed measurements, raising the odds of over-leveraging data. Furthermore, since natural systems are frequently complex, non-linear and non-random^{11–13}, a larger gap increases the likelihood of data challenges such as censoring (*sensu*¹⁴) as phenomena may resolve themselves in the space or time between replicates.

Results

We reviewed 348 papers randomly selected from 42,918 published between 2004 and 2014 in the top 30 ecology-themed journals. We extracted scale data from 378 observations of ‘natural’ (that is, non-experimentally manipulated) ecological features reported within 133 of the reviewed papers (plus an additional 62 cited as the source of observations). Most sampled observations were collected using conventional field methods (80%), followed by automated *in situ*

¹Graduate School of Geography, Clark University, Worcester, MA, USA. ²Woodrow Wilson School, Princeton University, Princeton, NJ, USA.

³Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA, USA. ⁴Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. ⁵Department of Geography and Environmental Systems, University of Maryland, Baltimore County, Baltimore, MD, USA. ⁶Department of Geography, University of California, Santa Barbara, Santa Barbara, CA, USA. ⁷Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, USA. *e-mail: LEstes@clarku.edu

Table 1 | Scale dimensions of ecological observations assessed in this meta-analysis

Component		Units	Description
Spatial	Resolution	m ²	Area of an individual spatial replicate (for example, plot)
	Extent	ha	Area encompassed by all spatial replicates
	Actual extent	ha	Summed area of all spatial replicates
Temporal	Interval	days	Time elapsed between successive temporal replicates
	Duration	days	Time elapsed between first and last temporal replicates
	Actual duration	days	Summed observational time of all temporal replicates

sensing techniques (12.4%), remote sensing (6.9%) and palaeo-reconstruction (<0.8%).

Observational domains within individual dimensions. In terms of resolution, spatial replicates for most (67%) observations were ≤ 1 m², 24% were 1 m² up to 1 ha, and 9% were >1 ha (Fig. 1a). These distributions primarily reflect those of field observations, the dominant observational methodology. Automated sensing and palaeo-reconstruction observations had resolutions that were generally finer (85% or more ≤ 0.1 m²) than field observations (47% ≤ 0.1 m²), while most remote observations were much coarser (70% >100 m²; Supplementary Fig. 1).

The extent of 19% of the observations was ≤ 10 ha, 23% covered 10–1,000 ha, 11% covered 1,000–10,000 ha, 19% covered 10,000–100,000 ha, 12% covered 100,000–1,000,000 ha and 15% covered >1,000,000 ha (Fig. 1b). As with resolution, the extent covered by automated sensing methods tended to be smaller (52% ≤ 100 ha) than those of field observations (31% ≤ 100 ha), while 96% of remote and all palaeo-reconstruction observations covered areas >10,000 ha.

In the temporal dimensions, 37% of observations were not repeated (Fig. 1c), 17% were repeated at short intervals (sub-second to daily), 20% were repeated at daily to monthly intervals, 18% were repeated at monthly to yearly intervals, 6% were repeated at yearly to decadal intervals and 2% were repeated at decadal or greater intervals. Among temporally replicated observations (Supplementary Fig. 1), automated sensing had the finest intervals (61% ≤ 1 day and 100% ≤ 1 year), followed by remote observation (37% ≤ 1 day and 78% ≤ 1 year), field methods (17% ≤ 1 day and 86% ≤ 1 year) and palaeo-reconstructions (21% ≤ 1 decade).

The duration was ≤ 1 day for 31% of the sampled observations (due to lack of temporal replication), while 10% covered 1 day to 1 month, 23% covered 1 month to 1 year, 27% covered 1–10 years and 9% covered >1 decade (Fig. 1d). Palaeo-reconstructions naturally had the longest duration (67% >1 decade), while only ~40% of field, automated and remote observations had durations exceeding 1 year.

Observational domains within two dimensions. Contrasting resolution with interval revealed that most temporally replicated observations had resolutions of 10 cm² to 1 m² and were revisited at daily to yearly intervals (Fig. 2a). A less dense, oblong concentration of observations bounded on the upper left by monthly to yearly observations at 100 m² resolution and on the lower right by near-daily to monthly observations with 1–10 ha resolution is also evident. The four observational methods had substantially different domains, as indicated by the locations of their median values

(see Supplementary Fig. 2): the median domain of field observations had 0.1–1 m² resolution and a monthly interval, whereas remote observations had a coarser median resolution (1,000 m²) but finer median interval (~1 day). Palaeo-reconstructions and automated sensing were both finely resolved (median between 10 cm² and 0.01 m²), but automated approaches had an hourly to daily median interval compared with a multi-decadal interval for palaeo-reconstructions.

Comparing the interval and duration of temporally replicated observations showed that most observations had daily to decadal intervals and durations of one month to one decade (Fig. 2b). Interval appears to increase with duration; observations lasting one month to one year tend to have daily to monthly intervals, while those lasting one year to one decade tend to have yearly to decadal intervals. This tendency is reflected in the domain medians of the primary observational methods: automated sensing had the finest median interval (hour–day) and shortest duration (month–year), followed by remote sensing (~1 day and 1 year, respectively), field observations (1 month and ~1 year, respectively) and finally palaeo-reconstructions (1 decade and millennium, respectively).

Contrasting the two spatial dimensions shows a primary concentration of observations of 10 cm² to nearly 100 m² resolution with extents ranging between ~1,000 and 1,000,000 ha (Fig. 2c). Another prominent concentration consists of higher-resolution (1 cm² to 1 m²), smaller-extent (10–1,000 ha) observations, beneath which lies a third, fainter concentration of 1–1,000 cm² resolution and 1,000 m² to <10 ha extent. These three concentrations suggest that observational extent increases with resolution, which is further evident in the median domain values (and kernel densities; Supplementary Fig. 2) of automated (0.01 m² resolution, 100 ha extent), field (0.1–1 m² resolution, 1,000–10,000 ha extent) and remote (1,000 m² resolution, 1–10 million ha extent) observations. Palaeo-reconstructions were the exception, having very fine median resolution (0.01 m²) but large extent (1 million ha)—a possible artefact of small sample size.

There are two primary observational domains within the contrast between duration and extent. The first consists of observations lasting 1 month to 1 decade with extents of 10–1,000 ha, while the second is defined by observations of 1 year to several decades that cover 10,000–1,000,000 ha (Fig. 2d). Three other notable but lesser concentrations are also evident, including small-area observations (0.1–1 ha) covering 1 month to 1 decade, and short-duration, temporally unreplicated observations (≤ 1 day) of either 1–100 ha or 10,000–1,000,000 ha. The median observation from automated sensing (1 year duration, 100 ha extent) lies near the centre of the first major concentration, while the median extents of field (1,000–10,000 ha) and remote (1–10 million ha) observations bound the second major concentration at its upper and lower extents, with the median duration of both observational types falling between 1 month and 1 year.

Differences between actual and ostensible scales. Observational extent was on average 5.6 orders of magnitude larger than actual extent (Fig. 3a). This difference increased with extent, reaching a maximum of 8.3 between 100 million and 1 billion ha of extent, then falling to 3 orders of magnitude between 1 and 10 billion ha (these extents comprised <2% of observations, which were primarily collected with remote sensing). Remote observations had the smallest mean difference magnitude (1.9), compared with ≥ 5.7 for the other three methods (Supplementary Fig. 3).

The difference magnitudes between observational duration and actual duration were somewhat smaller, averaging 3.4 and ranging from ~2 for the shortest durations (hour–day) to >4 for observations lasting 1 decade to 1 century (Fig. 3b). As with extent, the difference fell substantially for the longest durations (century to millennia), as these domains were covered by palaeo-reconstructions

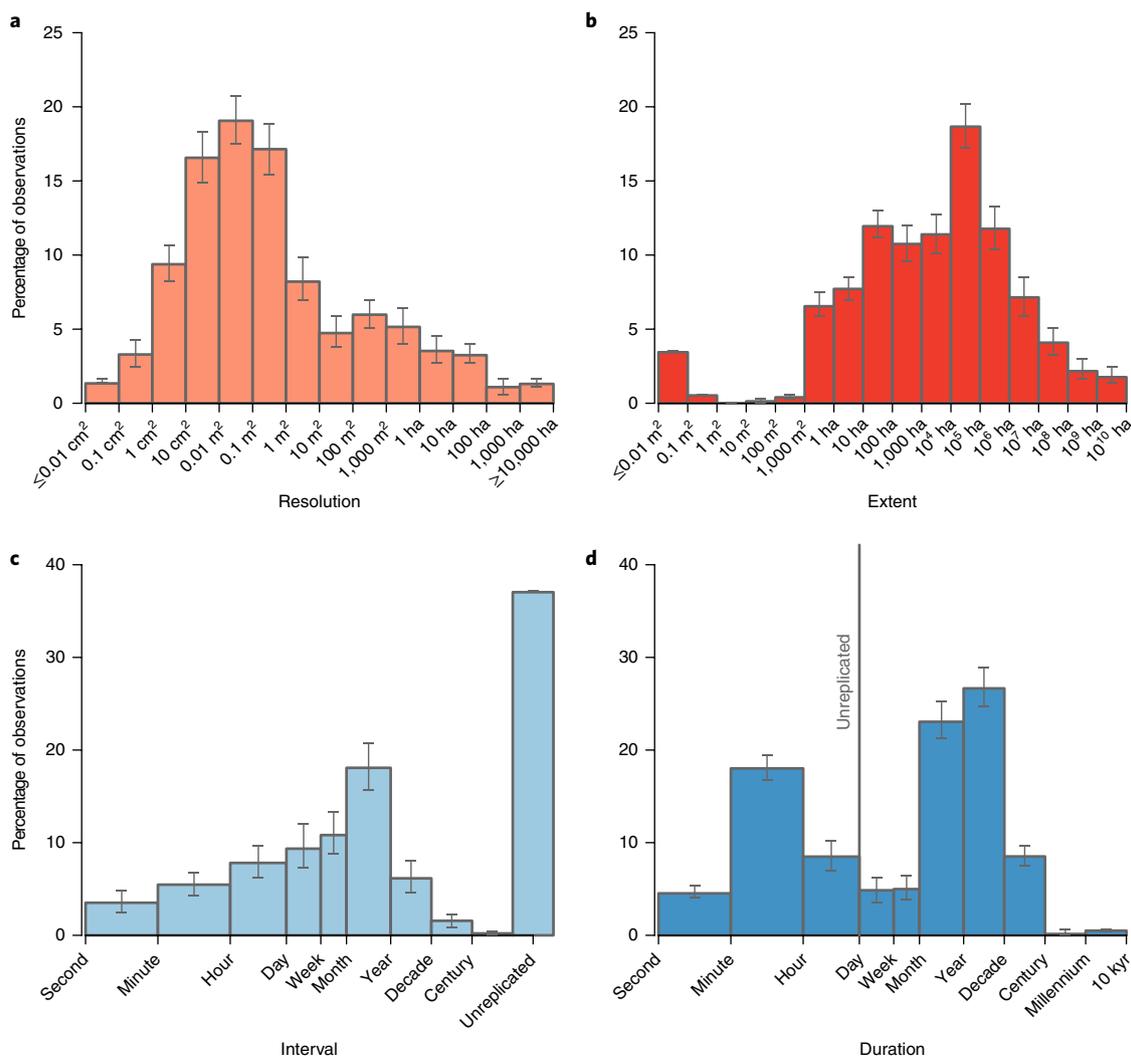


Fig. 1 | Observational domains within individual dimensions. **a–d**, Histograms of the resolution (**a**), extent (**b**), interval (**c**) and duration (**d**) of observations collected from the surveyed ecological studies. Bars represent the average percentages for each bin realized after 1,000 perturbed resamples, while the grey error bars indicate 95% confidence intervals. The bar widths in **c** and **d** indicate differences in scale between the x axis labels. The grey vertical line in **d** indicates that the majority (>95%) of observations of ≤ 1 day duration were temporally unreplicated. kyr, thousand years.

(Supplementary Fig. 3), which show little difference between actual and ostensible duration because coring techniques capture continuous temporal records. The mean difference magnitudes for the other three observational methods ranged from just over 3 (field and automated sensing) to nearly 6 (remote observations).

Potential biases and uncertainties in quantifying scales. Our results were potentially influenced by several methodological issues. First, most studies did not precisely report observational scales, thus we had to estimate, rather than simply record, scale values for most observations (we estimated 63, 60, 69, 36, 64 and 83% of resolution, extent, actual extent, interval, duration and actual duration values, respectively). Estimation errors may therefore have biased our findings. We attempted to quantify and account for this error by assessing between-observer variability and incorporating this uncertainty into our resampling methodology (Supplementary Results). The resulting confidence intervals (Fig. 1) suggest that estimation errors did not unduly influence our findings.

Our scale-estimation protocols may also have introduced bias—particularly our protocol for estimating resolution (the smallest areal unit of complete measurement). We selected this definition

for the sake of consistency, but some papers reported resolution as a larger area in which sub-samples were taken. For these, our estimates were finer than what the studies' authors considered to be the resolution. Our results would also be somewhat different if we had included observations from experiments. For example, average resolution and duration would probably be finer^{8,9}. Additionally, the token one second (Supplemental Methods) we used to represent the duration of remotely sensed temporal replicates (which are effectively instantaneous) caused us to underestimate the differences between their durations and actual durations (Supplementary Fig. 3). However, the relatively small number of remote observations suggests that the impact of this bias on our overall findings was negligible.

It is also possible that our findings misrepresent observational domains because of sampling error. Although we randomized our sample to ensure representativeness, we reviewed just 0.8% of the papers published during our study period. Our sample may therefore under- or over-represent observational coverage in certain domains, particularly for specific methods. This possibility is greatest for palaeo-reconstructions, where the small sample size probably resulted in an overestimate of typical observational extent

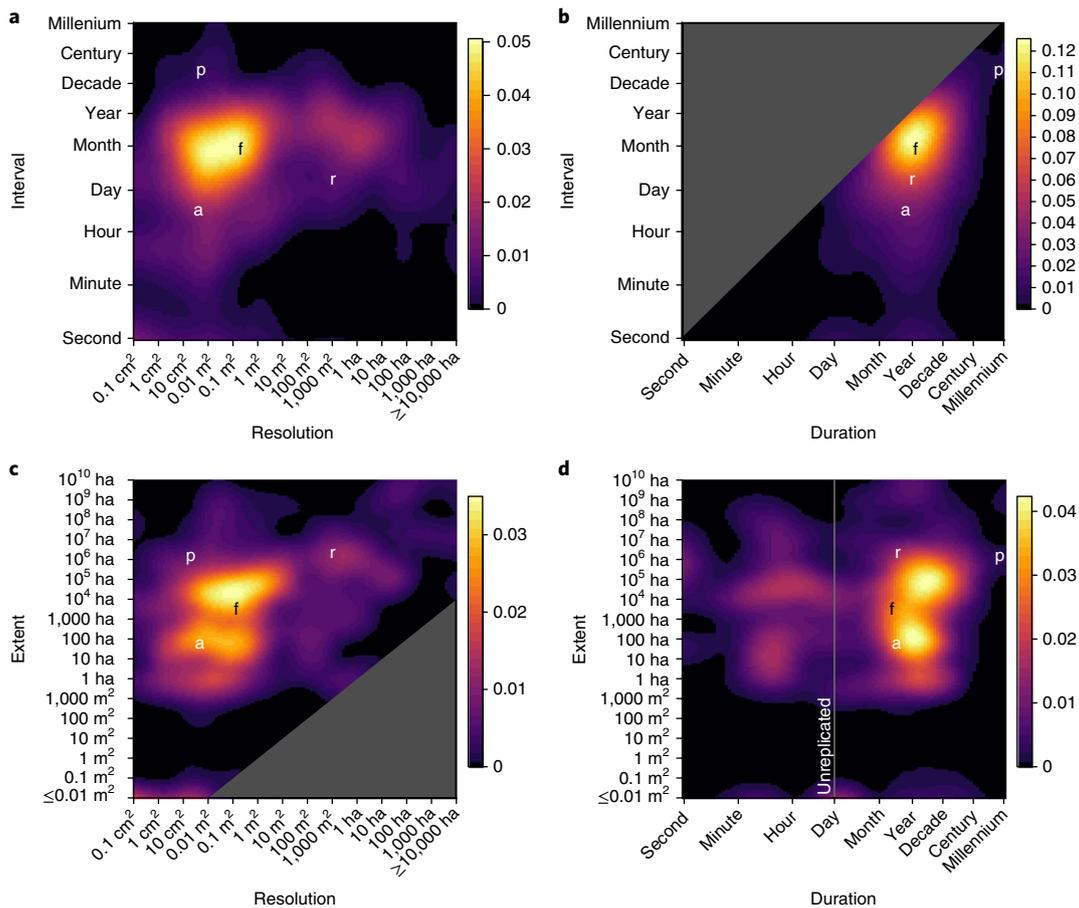


Fig. 2 | Observational domains within two dimensions. a–d, Kernel density estimates of observational densities within the domains defined by: resolution and interval (of temporally replicated observations) (**a**), duration and interval (of temporally replicated observations) (**b**), resolution and extent (**c**), and duration and extent (**d**). Density estimates were applied to the log-transformed values of each observational dimension, and density estimates were rescaled to represent percentages. The letters in the plots denote the median values of different observational methods (a, automated sensing; f, field observations; p, palaeo-observations; r, remote sensing). The grey shaded areas represent physically impossible domains (intervals greater than duration and resolutions greater than extent). Density values below the lower third percentile fall within the darkest portion of the colour gradient. The grey vertical line in **d** indicates that the majority (>95%) of observations of ≤ 1 day duration were temporally unreplicated.

(for example, Fig. 2b,c; however, the interval and duration values are probably more representative).

Finally, our omission of papers published after 2014 could also have biased our findings. Although our sample size was too small to assign statistical significance, we found a possible positive trend in the use of remote observations and a corresponding decline in field observations over the course of our study period. If these trends were not spurious, they suggest that including studies from 2015–2017 would result in a somewhat larger relative sample of remote observations, which could slightly increase the mean observational extent (see Supplementary Results).

Discussion

Our results suggest that modern ecology's observational domains are fairly narrow and that ecologists still primarily rely on conventional field-based observational techniques. In the spatial dimensions, most observations have resolutions $\leq 1 \text{ m}^2$ and extents $\leq 10,000 \text{ ha}$ (Fig. 1a,b). In the temporal dimensions, most observations are either unreplicated or relatively infrequent (>1 month interval; Fig. 1c), and have relatively short durations (≤ 1 year; Fig. 1d).

Contrasting observational dimensions reveals that larger extents are associated with larger spatial replicates (Fig. 2c), while longer durations are associated with longer intervals (Fig. 2b). The latter

association reflects a cost-imposed tradeoff between sampling frequency and temporal duration that is characteristic of field observations, but also appears to affect the other three methods, as evidenced by their relative domain locations. A similar tradeoff is illustrated by the inverse relationship between resolution and interval (Fig. 2a), which primarily relates to field observations, where larger spatial replicates demand greater effort, reducing sampling frequency⁹. Less obvious is the opposite tradeoff that affects remote observation (Supplementary Fig. 2), where finer resolution (necessary for detail) typically necessitates longer intervals¹⁵.

As a result of these tradeoffs, there are several notable observational gaps, specifically within the domains defined by high-frequency (daily to sub-daily intervals) observations with high to moderate resolutions ($>1 \text{ m}^2$ to 100 ha ; Fig. 2a) and decadal or longer durations (Fig. 2b). Another gap is evident in the high-to-moderate-resolution, large-extent (1 million to 10 billion ha) domain (Fig. 2c).

Have these domains changed since the seminal papers on scale first appeared in the late 1980s?^{1,3,8} A comprehensive answer would require a similar analysis focused on earlier literature, but the data provided by three previous studies provide partial insight. The first dataset consists of duration values that ref.⁸ extracted from 623 studies published in *Ecology* between 1977 and 1987. The mean duration of the most comparable subset of those values ($n = 419$) was

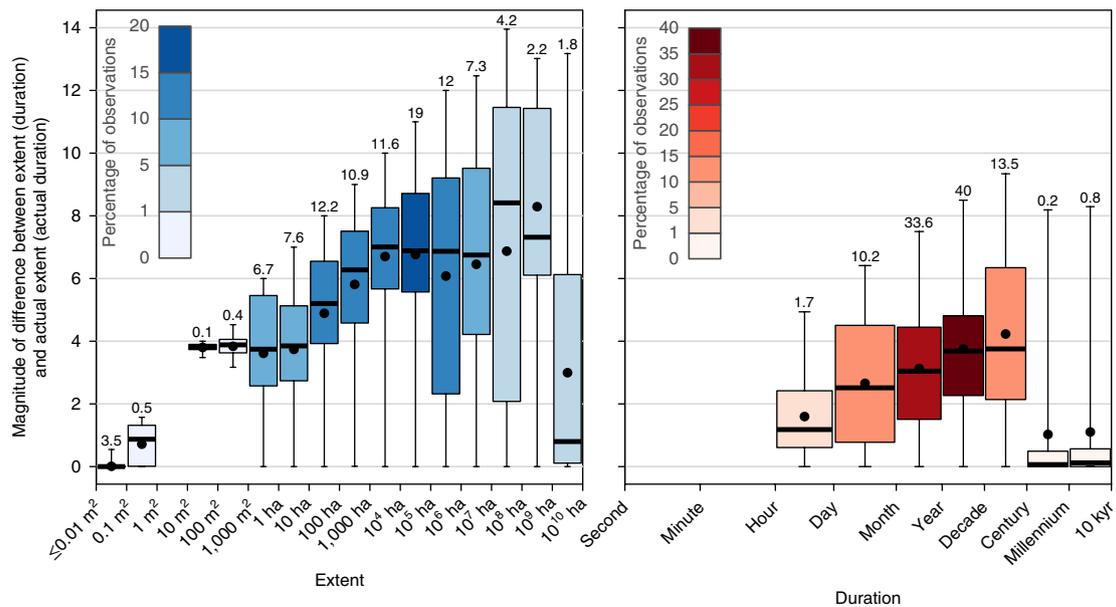


Fig. 3 | The magnitudes of the difference between actual and ostensible scales. **a, b**, Difference between extent and actual extent (the summed area of spatial replicates) (**a**), and duration and actual duration (the summed sampling duration across temporal replicates) (**b**). Difference values are expressed in terms of how many orders of magnitude larger (longer) extent (duration) is than actual extent (actual duration), and are summarized (as box plots, with the circle in the box representing the mean and the line the median) in bins representing the increasing scales of extent (duration). The percentages of observations falling within each bin are indicated by the colour of the interquartile and the numerical value above the upper whisker.

3.6 years, versus 3.3 years in our sample (or 5.1 year when excluding temporally unreplicated observations). The second dataset is found in ref. ⁹, which assessed the resolutions of 97 community ecology experiments published in *Ecology* between 1980 and 1986. The average of those (12,657 m²) was substantially smaller than the mean of our sample (1,479,465 m²), but comparing the eightieth percentile value (197 m²) in ref. ⁹ with ours (115 m²) shows that most contemporary observations are finer-grained than most 1980s-era experiments. Ref. ¹⁰ provides the third dataset, which compares the extent and interval of 25 studies published in 2003–2004 in *Ecology*. The mean interval was 178 days, compared with 684 days in our sample, but the eightieth percentile value in our study was 169 days compared with 329 days in theirs. Extent in our sample was substantially larger according to multiple summary statistics, including the mean (368,403 ha versus 114,965,072 ha in our study), median (9 ha versus 5,051 ha) and ninetieth percentile (136,000 ha versus 46,424,808 ha; this value is smaller than the mean, which is skewed by a small number of very-large-extent observations).

Although limited due to methodological differences (for example, a focus on experiments versus unmanipulated systems), these comparisons suggest that the duration and resolution of ecological observations have changed little in the past 30 years, but observational frequency and extent have both increased. A weak positive trend in our data also suggests that the mean extent of ecological observations is steadily increasing (Supplementary Fig. 5), which probably corresponds to increasing use of remote sensing (Supplementary Fig. 4).

Despite this apparent increase in observational extent, there remains a large gulf between the areas that ecologists actually observe and the areas their observations are intended to represent (Fig. 3a). A substantial discrepancy also exists between the amount of time spent observing phenomena and the time spans those observations theoretically represent (Fig. 3b). These differences between the actual and ostensible scales of observation have implications for ecological understanding, as the unobserved portions of space and time may contain important patterns and processes that are not captured by replicates, due to phenomenon-dependent factors

such as autocorrelation and representativeness of the sampling scheme^{16–20}. Brief, infrequent snapshots, or fine-grained, spatially sparse replicates, may be sufficient to characterize many phenomena (for example, annual changes in tree cover are well-represented by low-frequency satellite imaging²¹), but may be inadequate for more dynamic phenomena. For example, wildfire extent and duration can be mapped by daily return satellites^{22,23}, but the instantaneous nature of the imaging means that they cannot be used to observe fire behaviour²⁴. To capture such behaviour, long periods of continuous observation may be more important than frequent repeats for understanding the dynamics.

It is therefore important to examine whether the scales of the phenomena being observed are adequately captured by the design of replicates. Our methods suggest one possible procedure for assessing the scale representativeness of observations, which is to (1) calculate the autocorrelation (spatial or temporal) within the observations (for example, using a semi-variogram), (2) find the threshold distance (or time) below which a suitably strong correlation (for example, $r = 0.7$) will exist between neighbouring sampled values, (3) add that distance (or time) to the sample resolution (or duration) and (4) recalculate actual extent (or duration) using the adjusted resolution (or sampling duration). The difference between this autocorrelation-adjusted actual extent (or actual duration) and extent (or duration) may provide a useful additional measure of how well the replicates represent the intended scale of observation. Although increasing spatial or temporal coverage may not always be the goal of a study, if the gap between actual and ostensible values remains large, alternative sampling methods may be used to close it. For example, remote sensing provides wall-to-wall spatial coverage of a study area, erasing the difference between actual extent and extent. Furthermore, the interval of high-resolution imaging (higher resolution is preferred in images as it allows individual features to be better discerned^{25,26}) is now approaching daily to sub-daily scales^{27,28}, allowing improved representation of spatial and temporal dynamics. For phenomena that cannot be measured from space—either because they are not visible or because they

require continuous observation—new approaches for collecting in situ or near-surface observations (for example, low-cost wireless sensors^{10,29,30}, citizen observers³¹ and autonomous vehicles³²) can be used to increase the spatial and temporal coverage of observations.

The aforementioned insights regarding modern observational domains must be tempered by the uncertainty within our own scale estimates, as detailed above. However, most of this uncertainty is attributable to unclear reporting of scale values in the majority of papers we reviewed (a problem also noted in geography studies³³). This tendency towards vague documentation offers one final insight: despite decades of accumulated knowledge regarding its importance^{1–3,34}, scale appears to remain a low priority throughout much of the ecological discipline. Beyond contributing to the broader problem of scientific reproducibility³⁵, inattentiveness to scale increases the risk that observations inadequately represent the phenomenon of interest, thereby limiting the generalizability of any derived ecological knowledge^{3,33,34}. To mitigate this problem, we recommend that ecological journals require authors to quantify and clearly report the values of resolution, extent, interval and duration. Fortunately, some journals already appear to be implementing such policies. For example, *Global Ecology and Biogeography* now requires information on the spatial, temporal and taxonomic scale of studies to be in the abstract (a policy adopted in early 2016).

Looking forwards. Our study suggests that the concept of scale has yet to fully permeate the discipline of ecology. Evidence for this assertion lies in the continued narrowness of ecology's observational scale domains and the poor documentation of scale dimensions in the literature. However, the increasing extent of ecological observations, enabled by remote sensing and presumably motivated by many ecologists' appreciation of scale-related issues, suggests that ecology's scale domains are gradually changing. In the coming years, the accelerating gains in technology and analytical methods will allow researchers new and unprecedented capabilities to peer into, and thus close, the prominent holes in observational domains. A renewed, discipline-wide focus on scale's importance, including the adoption of stricter scale-reporting standards by journals, will help to spur ecologists to address these gaps, while fostering the improved transferability of knowledge within the discipline.

Methods

Paper selection and review. We used the 2012 Web of Science impact factors to select the 30 highest-ranked ecology-themed journals that published studies with an observational component, excluding journals devoted to reviews, meta-analyses, or laboratory, cellular or experimental studies. To select a representative sample of recent ecology studies, we downloaded the metadata for all papers published in the selected journals (Supplementary Table 1) between 2004 and 2014.

Our study involved 6 different observers (those reviewing the papers to extract the observational scales), each of whom was given a randomly selected batch of 500 titles. A separate set of 20 papers was also randomly selected and given to all observers to review independently. This was to (1) calibrate the interpretations and extraction of scale-related information between observers and (2) estimate between-observer variance.

Each observer first reviewed the papers in the calibration set and then commenced reviewing papers in their individual random draws, beginning at the top of the list and then proceeding until at least 20 eligible papers describing ecological observations were reviewed. In cases where the reviewed papers used observations that were described in another publication, we reviewed those source papers to extract the observational dimensions. We excluded papers that were opinion or perspectives pieces (unless they presented or used existing observational data), or theoretical studies based on generated data. We also did not collect scale information from papers (or the relevant parts of papers) describing experimental manipulations because experiments tend to be of limited extent, duration and resolution due to their higher logistical costs^{8,9}. Including data from experiments would therefore probably have biased our findings towards finer scales, while minimizing the impact that new observing methods (for example, satellite imaging and wireless sensing) may have had in expanding the scales of ecological investigation^{10,36,37}. A bibliography of the reviewed papers appears in the Supplementary Information.

Estimating observational scales. We recorded six primary dimensions of ecological observations—three related to space and three related to time. The space-related dimensions were resolution, extent and actual extent. Here, extent was primarily defined as the area falling within a perimeter defined by the outermost spatial replicates, while actual extent was the summed area of all spatial replicates (that is, $N \times$ resolution, where N is the number of spatial replicates, which we also recorded), or the area that ecologists observe in practice. In assessing spatial scales, our analysis only considered the Cartesian plane; we did not calculate the z (or depth) dimension, although this dimension is of greater importance for certain sub-disciplines of ecology (for example, depth profiles in marine ecology). In some cases (primarily palaeoecological studies), values extracted from the z dimension provided temporal information that was used to calculate both the interval and the duration of the observation.

For time dimensions, we extracted information related to the observational interval, duration and actual duration. Duration was defined as the time between the first and last temporal replicate, whereas actual duration quantifies the amount of time spent observing a particular location, which we calculated by multiplying the sampling duration (the time spent collecting a single temporal replicate) by the number of temporal replicates.

A full definition of all dimensions and how they were recorded is contained within a list of frequently asked questions (see Supplementary Methods), which was provided to each observer for initial study and reference, and adapted as necessary during the course of the study to ensure methodological consistency.

To account for potential differences in scales related to methodology, we classified each observation according to the following broad categories: field methods (manual in situ data collection), automated (in situ) sensing, remote sensing/other geographic data (hereafter remote observations) and palaeo-reconstruction approaches. We also recorded when observations were reported in any study with an unclear or missing scale value.

Calibration and consistency. Most studies did not explicitly report values for all the assessed scales, and thus interpretation and judgement had to be applied to develop reasonable estimates for their values. The frequently asked questions (Supplementary Methods) provided the protocol we followed, and were initially developed following consultation between observers before reviewing commenced. We conducted an iterative process of calibration to ensure consistency and reliability of the estimates. First, we used the calibration set to calculate between-observer variability with respect to paper selection/rejection and the estimation of scales. Based on this, the lead author reviewed individual records in each observer's calibration set, flagged values where the estimation procedure departed from the protocol and returned these to observers for re-estimation without providing an estimate of the actual value. Instead, the relevant section of the protocol was highlighted, and further explanation and clarifying discussion were undertaken as needed. The protocol language was adjusted for clarity during this process, and new items were added to cover circumstances that had not been addressed by the initial version. The variability measures were recalculated after each iteration.

To ensure consistency within the main analysis, the lead author also reviewed each observer's results from their individual draw of papers and flagged values that appeared to deviate from the protocol for re-review by the observer. Revised values were re-inspected, and in some cases a secondary review of particular papers was undertaken to cross-check the estimated scales.

Scale-estimation uncertainty. Two major and related sources of uncertainty affected our estimation of observational scales: (1) unclear documentation of observational scales in the reviewed studies; and (2) variation between observers in estimating observational scales (largely in cases where scales were not explicitly reported). To account for these uncertainties, we first quantified the between-observer variability in scale estimates (expressed as the coefficient of variation), which was constructed from each observer's final reported calibration set results. We then used the coefficients of variation for each dimension as the basis for randomly perturbing—over the course of 1,000 iterations—the scale values for each of the sampled observations. For each observational dimension at each iteration, we perturbed its observer-estimated scale value by: (1) randomly selecting (from a uniform distribution) a percentage value p that fell between $100 + y$ and $100 - y$ (where y was the dimension-specific coefficient of variation, expressed as a percentage) and (2) multiplying the scale value by the corresponding proportion ($p/100$). The perturbation occasionally resulted in physically impossible values (for example, interval or actual duration longer than duration, or actual extent larger than extent). In these cases, we capped the perturbed value in the smaller of the two dimensions (that is, resolution or interval) so that it equalled the corresponding value in the largest (that is, extent or duration). We used the resulting set of perturbed observations to quantify uncertainty within our scale estimates.

In addition to the scale-estimation coefficient of variation, we also examined how well observers agreed regarding paper inclusion/exclusion, and how many extractable observations there were per included paper (see Supplementary Results).

Analyses. To characterize the scale domains of observations, we first log-transformed (base-10) the scale values within the 1,000 member perturbed

ensemble to account for the large range in values. To examine the distributions of observational scales within individual dimensions (Fig. 1), we first constructed relative frequency histograms for each of the 1,000 transformed ensemble members for each dimension and then plotted the bin means across all members, as well as the upper and lower 2.5th percentile values for each bin. This produced a histogram of observational scales within each dimension that accounted for scale-estimation uncertainty.

To evaluate the distributions of observations within two scale dimensions (Fig. 2), we used the *splancs* package³⁸ of R³⁹ to calculate a kernel density estimate of the log-transformed values across all ensemble members, using a bandwidth of 1 on a 0.1 resolution image to provide a smoothed result that served to more effectively highlight domains in which ecological observations were concentrated. Bandwidths of varying resolutions were tested on kernel density estimates of sampling interval versus plot resolution to test how sensitive our results were to the bandwidth value (see Supplementary Results). For comparisons involving interval, we removed temporally unreplicated observations because these lacked interval values.

To compare the differences between actual extent and extent and actual duration and duration (Fig. 3), we calculated the magnitude of difference (decade) between each pair as:

$$\text{decade} = \log_{10} x - \log_{10} y = \log_{10} \frac{x}{y}$$

Where x is either extent or duration and y is actual extent or actual duration, respectively. We then evaluated how the magnitudes of difference varied with increasing values of extent/duration, using box plots to summarize decades within the same bins used to summarize the frequency distributions of the extent and duration of observations (Fig. 1b,d). Decades were calculated for each pair for all bootstrap replicates. We plotted the box plots against their corresponding bin means to evaluate how these differences varied with scale (Fig. 3).

Trends in methods and scale. To evaluate the potential impact that excluding studies from 2015–2017 would have on our findings, we analysed the trends in (1) ecological observing methods and (2) typical scales of ecological observations over the 10 year period. To undertake the former assessment, we calculated the percentage of observations made using remote sensing, general field methods and automated in situ methods, and fit a linear regression between these percentages and the publication year, weighting the regression by the total number of observations in each year. For the second analysis, we applied the same regression approach to the four primary dimensions (resolution, extent, interval and duration) to assess whether there were any trends in observational scales.

The regressions and resulting code for trend extrapolations can be found in the 'additional analyses' vignette in the accompanying R package/code repository (available at <https://github.com/agroimpacts/ecoscales>).

Extracting and analysing data from earlier meta-analyses. To compare the results of our analysis with the observational scales of earlier ecological studies, we used graph capture software (<https://automeris.io/WebPlotDigitizer/>) to extract the data values from figure 6.1 of ref. ⁸, figure 1 of ref. ⁹ and figure 2 of ref. ¹⁰. To maintain as much comparability as possible with our inclusion criteria, we excluded experimental studies in the data from ref. ⁸, as well as the values of any studies exceeding 100 years' duration (no upper time bound was provided for these), leaving duration values for 419 (out of 623) studies. Since ref. ⁸ presented duration values as a histogram, we calculated the mean duration across all studies as the weighted (by number of observations per bin) mean of bin centre-point values (that is, the weighted mean of the bin means). We also excluded 4 (of 29) observation values from the data in ref. ¹⁰ on observational extent and frequency, which, in contrast with the other 25, were not randomly selected. Ref. ¹⁰ also used irregular scales for both x (frequency) and y (extent) axes; therefore, we had to visually estimate the scale values for each data point after graphical extraction, and converted their extent values (in km) to hectares and their frequency values to intervals. Ref. ⁹ presented resolution as plot diameters (m), which we squared to make comparable to our resolution metric.

Calculations of scale values from these studies can be found in the 'additional analyses' vignette in the accompanying R package/code repository (available at <https://github.com/agroimpacts/ecoscales>).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Code availability. The code supporting this manuscript is available online at <https://github.com/agroimpacts/ecoscales>.

Data availability. The data supporting this manuscript are available online at <https://github.com/agroimpacts/ecoscales>.

Received: 22 August 2016; Accepted: 27 February 2018;
Published online: 2 April 2018

References

- Levin, S. A. The problem of pattern and scale in ecology. *Ecology* **73**, 1943–1967 (1992).
- Chave, J. The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecol. Lett.* **16**, 4–16 (2013).
- Wiens, J. A. Spatial scaling in ecology. *Funct. Ecol.* **3**, 385–397 (1989).
- Sandel, B. & Smith, A. B. Scale as a lurking factor: incorporating scale-dependence in experimental ecology. *Oikos* **118**, 1284–1291 (2009).
- Schneider, D. C. The rise of the concept of scale in ecology. *BioScience* **51**, 545–553 (2001).
- Tingley, M. W., Koo, M. S., Moritz, C., Rush, A. C. & Beissinger, S. R. The push and pull of climate change causes heterogeneous shifts in avian elevational ranges. *Glob. Change Biol.* **18**, 3279–3290 (2012).
- Xiao, J. F. & Moody, A. Photosynthetic activity of US biomes: responses to the spatial variability and seasonality of precipitation and temperature. *Glob. Change Biol.* **10**, 437–451 (2004).
- Tilman, D. in *Long-Term Studies in Ecology* 136–157 (Springer, New York, 1989).
- Kareiva, P. & Andersen, M. *Community Ecology* 35–50 (Springer, Berlin, 1988).
- Porter, J. et al. Wireless sensor networks for ecology. *BioScience* **55**, 561–572 (2005).
- Levin, S. A. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems* **1**, 431–436 (1998).
- Pringle, R. M. & Tarnita, C. E. Spatial self-organization of ecosystems: integrating multiple mechanisms of regular-pattern formation. *Annu. Rev. Entomol.* **62**, 359–377 (2017).
- Rietkerk, M. & van de Koppel, J. Regular pattern formation in real ecosystems. *Trends Ecol. Evol.* **23**, 169–175 (2008).
- Efron, B. The efficiency of Cox's likelihood function for censored data. *J. Am. Stat. Assoc.* **72**, 557–565 (1977).
- Estes, L. D. et al. A platform for crowdsourcing the creation of representative, accurate landcover maps. *Environ. Model. Softw.* **80**, 41–53 (2016).
- Underwood, A. J. *Experiments in Ecology: Their Logical Design and Interpretation Using Analysis of Variance* (Cambridge Univ. Press, Cambridge, 1997).
- Palmer, M. W. & White, P. S. Scale dependence and the species-area relationship. *Am. Nat.* **144**, 717–740 (1994).
- Cao, Y., Williams, D. D. & Larsen, D. P. Comparison of ecological communities: the problem of sample representativeness. *Ecol. Monogr.* **72**, 41–56 (2002).
- Legendre, P. Spatial autocorrelation—trouble or new paradigm? *Ecology* **74**, 1659–1673 (1993).
- Collins, S. L., Micheli, F. & Hartt, L. A method to determine rates and patterns of variability in ecological communities. *Oikos* **91**, 285–293 (2000).
- Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
- Roy, D. P., Jin, Y., Lewis, P. E. & Justice, C. O. Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sens. Environ.* **97**, 137–162 (2005).
- Jones, B. M. et al. Fire behavior, weather, and burn severity of the 2007 Anaktuvuk River tundra fire, North Slope, Alaska. *Arct. Antarct. Alp. Res.* **41**, 309–316 (2009).
- Clements, C. B. et al. Observing the dynamics of wildland grass fires: FireFlux—a field validation experiment. *Bull. Am. Meteorol. Soc.* **39**, 1369–1382 (2007).
- Dark, S. J. & Bram, D. The modifiable areal unit problem (MAUP) in physical geography. *Prog. Phys. Geogr.* **31**, 471–479 (2007).
- Hay, G. J., Blaschke, T., Marceau, D. J. & Bouchard, A. A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS J. Photogramm. Remote Sens.* **57**, 327–345 (2003).
- Drusch, M. et al. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **120**, 25–36 (2012).
- Hand, E. Startup liftoff. *Science* **348**, 172–177 (2015).
- Collins, S. L. et al. New opportunities in ecological sensing using wireless sensor networks. *Front. Ecol. Environ.* **4**, 402–407 (2006).
- Wolf, A., Falusi, J., Caylor, K., Sheffield, J. & Wood, E. F. A GSM-based surface meteorology network in service of improved African hydrological data assimilation and drought forecasting. In *Proc. American Geophysical Union Abstract H13J-01* (AGU, San Francisco, 2012).
- Dickinson, J. L. et al. The current state of citizen science as a tool for ecological research and public engagement. *Front. Ecol. Environ.* **10**, 291–297 (2012).
- Anderson, K. & Gaston, K. J. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front. Ecol. Environ.* **11**, 138–146 (2013).
- Margulies, J. D., Magliocca, N. R., Schmill, M. D. & Ellis, E. C. Ambiguous geographies: connecting case study knowledge with global change science. *Ann. Am. Assoc. Geogr.* **106**, 572–596 (2016).

34. Wheatley, M. & Johnson, C. Factors limiting our understanding of ecological scale. *Ecol. Complex.* **6**, 150–159 (2009).
35. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nat. News* **533**, 452–454 (2016).
36. Turner, D. P., Ollinger, S. V. & Kimball, J. S. Integrating remote sensing and ecosystem process models for landscape- to regional-scale analysis of the carbon cycle. *Bioscience* **54**, 573–584 (2004).
37. Pettorelli, N. et al. Satellite remote sensing for applied ecologists: opportunities and challenges. *J. Appl. Ecol.* **51**, 839–848 (2014).
38. Rowlingson, B. S. & Diggle, P. J. SplanCs: spatial point pattern analysis code in S-plus. *Comput. Geosci.* **19**, 627–655 (1993).
39. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2011).

Acknowledgements

This work was supported by funds from the Princeton Environmental Institute Grand Challenges programme and the NASA New Investigator Program (NNX15AC64G). E.C.E., J.C. and L.A. of the GLOBE Project (<http://globe.umbc.edu>) were supported by the US National Science Foundation (1125210). J. Daskin, J. Socolar, C. Chang, and R. Grossman provided crucial help in developing and testing the sampling methodology.

Author contributions

L.E. and K.C. conceived the study. L.E., T.T., P.R.E., J.J.C. and E.C.E. helped to design the sampling protocol. L.E., T.T., P.R.E., J.J.C., L.A., and J.C. extracted and analysed the scale data. L.E. performed the analyses and wrote the manuscript. All authors assisted with writing the manuscript and provided comments and edits.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-018-0524-4>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to L.E.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

▶ Experimental design

1. Sample size

Describe how sample size was determined.

By random selection of journal titles

2. Data exclusions

Describe any data exclusions.

Studies or observations not meeting our inclusion criteria (non-experimentally manipulated observations) were excluded.

3. Replication

Describe whether the experimental findings were reliably reproduced.

None.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Journal titles were randomly selected from full list of titles downloaded through Endnote.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

N/A

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The <u>exact sample size</u> (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement indicating how many times each experiment was replicated |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The test results (e.g. <i>P</i> values) given as exact values whenever possible and with confidence intervals noted |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A clear description of statistics including <u>central tendency</u> (e.g. median, mean) and <u>variation</u> (e.g. standard deviation, interquartile range) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Clearly defined error bars |

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

R and associated libraries. Full code and data will be released upon acceptance of paper for publication.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

N/A

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

N/A

b. Describe the method of cell line authentication used.

N/A

c. Report whether the cell lines were tested for mycoplasma contamination.

N/A

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

N/A

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

N/A

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

N/A