*J Biosci* (2025)50:63
DOI: 10.1007/s12038-025-00537-z


# Effects of sampling methodology on phenology indices: Insights from sites across India and modelling

Yadugiri V Tiruvaimozhi[1]*, Karthik Teegalapalli[2], Abinand Reddy[3], Akhil Murali[4], Aparajita Datta[4], Aparna Krishnan[4,5], Jayashree Ratnam[5], Mahesh Sankaran[5], Shasank Ongole[6], Srinivasan Kasinathan[4], TR Shankar Raman[4] and Geetha Ramaswami[4]*

[1]*Krea University, Sri City 517 646, India*

[2]*Independent researcher, Candolim 403 515, India*

[3]*Centre for Research into Ecological and Environmental Modelling, The University of St. Andrews, College Gate, St. Andrews KY16 9AJ, Scotland*

[4]*Nature Conservation Foundation, Mysore 570017, India*

[5]*National Centre for Biological Sciences, TIFR, Bellary Road, Bengaluru 560065, India*

[6]*Plant Ecology, University of Bayreuth, Universitaetsstrasse 30, 95447 Bayreuth, Germany*

**Corresponding authors (Emails, vtyadu@gmail.com; geetha@ncf-india.org)*

*MS received 19 January 2024; accepted 5 February 2025*

Plant phenology is the study of the timing and extent of leaf, flower, and fruit production. Phenology data are used to study the drivers of cyclicity and seasonality of plant life-history stages, interactions with organisms such as pollinators, and effects of global change factors. Indices such as the timing of phenological events, and the proportion of individuals in a particular phenophase, seasonality, and synchrony have often been used to summarise plant phenology data. However, these indices have specific utilities and limitations and may be sensitive to sampling methodology, making cross-site comparisons challenging, particularly when data collection methods vary in terms of sample size, observation frequency, and the resolution at which phenophase intensity scores/values are recorded. We used fruiting phenology data from tropical trees across five sites in India to study the effects of sampling methodology on two indices: population-level synchrony (overlap) and seasonality. We supplemented these results with simulations of fast- and slow-changing phenologies to test for the effects of sampling methodology on these indices. We found that the overlap index is sensitive to the resolution of phenophase intensity measurement, with coarser intensity measures leading to overestimation of the overlap index. The seasonality index, on the other hand, was not affected by intensity resolution. Simulations indicated that finer intensity resolution is more important than frequency of observation to accurately estimate population synchrony and seasonality for fast- and slow-changing phenophases. Based on our findings, we provide recommendations for future study designs of tropical tree phenology research, particularly for long-term or cross-site studies.

**Keywords.** Biological rhythms; citizen science; long-term monitoring; seasonality; tropical forest

*Supplementary Information: The online version contains supplementary material available at https://doi.org/10.1007/s12038-025-00537-z.*

**http://www.ias.ac.in/jbiosci**
Published online: 22 July 2025

## 1. Introduction

Phenological research involves monitoring the timing of cyclic biological events which have ecological and evolutionary relevance (Augspurger 1983). It also involves understanding the proximate and ultimate causes of the observed phenological patterns. The timing of events such as leafing, flowering, and fruiting reflects the link between plant growth and reproduction, and ecosystem characteristics including biotic variables such as pollination and dispersal, and abiotic variables such as temperature, rainfall, and other factors (Hemingway and Overdorff 1999). Insights from long-term plant phenological research can also enable the prediction of the effects of climatic changes on plants since climate change-driven phenological shifts can disrupt biotic interactions and thereby affect entire ecosystems (Fisogni *et al*. 2022).

Phenological synchrony (the overlap and similarity of a phenophase among individuals) and seasonality (for instance, the peak of a phenophase) have ecological and evolutionary relevance at various scales ranging from an individual to a community. Several indices of synchrony have been reported in the phenology literature, and these are typically calculated at the scale of an individual plant (e.g., Augspurger 1983; Freitas and Bolmgren 2008; Bogdziewicz *et al*. 2020). A synchrony index may compare only the occurrence of a phenophase in an individual with respect to every other individual in a population (e.g., the Augspurger index; Augspurger 1983). It may also be a composite measure comparing both phenophase occurrence and intensity levels of an individual with those of all other individuals in the population (e.g., the Freitas–Bolmgren index; Freitas and Bolmgren 2008). In the latter approach, maximum synchrony is achieved when an individual's phenophase co-occurs with those of all other individuals in a population, and all individuals show maximum phenophase (e.g., if all individuals are in full bloom). Synchrony declines as individuals in a population vary in their timing (and/or intensity) of a phenophase. Both aseasonal and seasonal populations can display high synchrony. Within a species, asynchronous ripening of fruits may reduce competition for dispersers or consumption by seed predators (McKey 1975; Wheelwright 1985; Terborgh 1990), whereas flowering synchrony may increase pollinator visitation and thereby seed set (Freitas and Bolmgren 2008; Bogdziewicz *et al*. 2020).

The quantification of phenological synchrony and seasonality can therefore provide insights into potential underlying evolutionary mechanisms. These indices can also provide insights into plant responses to climate change, such as spring advancement and autumn postponement, which can be identified based on the annual changes in the seasonality index (Piao *et al*. 2019). However, values of these indices can be expected to vary based on the sampling methods used, particularly sample size, frequency of observation, and the phenophase intensity scores/values (see figure 1 for an illustration), and this can further influence the interpretation of phenological events of a species or populations. Nevertheless, only a few studies have focused on the variation in phenological indices caused by differences in sampling methodology (e.g., Chapman *et al*. 1992, 1994; Hemingway and Overdorff 1999; Freitas and Bolmgren 2008; Morellato *et al*. 2010). In this context, combining long-term data from several sites is the key to understanding the variation in phenology responses at larger spatial scales to habitat, climate, and other factors. Further, data collection methods may vary in different sites according to the specific research questions being addressed, consequently influencing the values of the indices.

We leveraged long-term phenological data on fruiting of trees and shrubs from field studies in four sites in India and from a citizen science effort for the entire State of Kerala (SeasonWatch Citizen Science Network 2022) to study the effects of sampling methodology on selected phenology indices. To supplement this, we also investigated the effects of variations in phenological methods on the indices using simulated data. In the simulations, the variations in the methods could be more robustly represented to arrive at implications for future phenological research. Our specific objectives were to address the following questions: (i) What are the effects of differences in sampling methods on indices of phenological synchrony (overlap index) and seasonality? (ii) What are the effects of differences in sampling methods on the indices of synchrony and seasonality, based on simulations of populations with varying rates of phenophase change (slow- and fast-changing phenophases)? For instance, in a population with a short-lived phenophase (fast-changing phenophase), phenological indices may be more sensitive to the frequency of observation, resolution of phenology intensity measurement, and/or sample size.

Based on our results, we suggest optimal sampling parameters and methods to study tropical tree phenology across populations and communities.

## 2. Materials and methods

### 2.1 *Study sites and datasets*

We used fruiting phenology data (unripe and/or ripe fruits) from four tree phenology monitoring sites across India, and from a pan-India citizen science effort – SeasonWatch (table 1). From SeasonWatch, data for the state of Kerala was chosen since it had the best spatial spread and representation of species compared with other states of India. These sites differ in their climates with Pakke, the Anamalais, and Kerala being wetter sites, and Rishi Valley (hereafter, RV) and the Nagarjunasagar Srisailam Tiger Reserve (hereafter, NSTR) being drier sites. The sites also

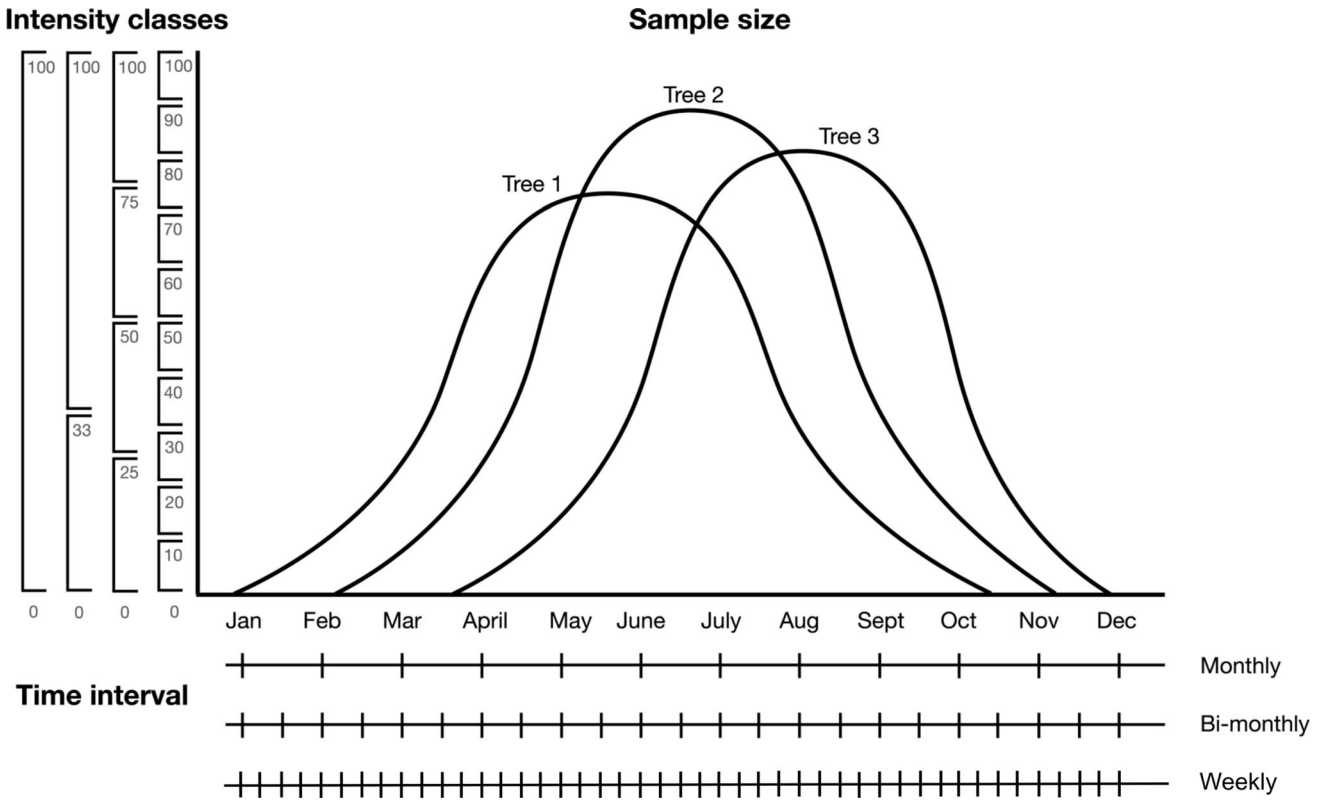



**Figure 1.** Variables related to sampling methods that affect phenology indices: Resolution of intensity of phenology measurements (usually expressed as a percentage), sample sizes (number of individual trees observed per species), and time interval (frequency of phenology measurements).

**Table 1.** Description of the sites and datasets included in this study

| Site | Date range | Number of species monitored | Sample size (no. of individual trees per species) | Frequency of sampling | Resolution of phenology intensity measures | Elevation range (m) | Average annual rainfall (mm) |
|---|---|---|---|---|---|---|---|
| Rishi Valley | 2007–2016 | 18 | 10–40 | Fortnightly | 0, 33%, 100% (none, few, many) | ∼700 | 690 |
| NSTR (Andhra Pradesh) | 2018 | 44 | 10 | Fortnightly | 10% Intervals | 500–850 | 713 |
| Pakke (Arunachal Pradesh) | 2011–2019 | 35 | 10–25 | Monthly | 25% Intervals | 100–300 | 2500 |
| The Anamalais (Tamil Nadu) | 2017–2020 | 172 | 1–41 | Monthly | 25% Intervals | 800–1400 | 2500 |
| Kerala State (SeasonWatch) | 2014–2022 | 150 | Variable | Weekly | 0, 33%, 100% (none, few, many) | 0–2500 | 3348 |

NSTR, Nagarjunsagar Srisailam Tiger Reserve.

differ in seasonality and tree species composition (table 1). Details of the habitat, species, sampling effort, and methods are provided in the appendix.

We used fruiting phenology as the response variable in this study because synchrony and seasonality of fruiting and their responses to climatic factors have consequences for the reproductive fitness of the plant. It also has community-level effects, impacting fruit and seed predators and dispersers. While flowering phenology is similar in its impact on plant–pollinator interactions and reproductive fitness, our datasets from different monitoring sites have the best resolution and the most data for fruiting phenology. We focused on fruiting phenology because these data were best suited to test the effects of differences in phenology data collection methods on phenological indices.

### 2.2 *Methodological aspects of phenology data collection*

The methods used to collect fruiting phenology data vary across the sites included in our study. Here we focused on three key methodological properties: (i) Phenology intensity, which is a measure of the fraction of the tree crown showing a given phenology. For example, fruiting intensity is 100% if the entire tree crown shows the presence of (ripe and/or unripe) fruits. In tropical trees, it is possible, however, that some species or individuals do not reach 100% fruiting intensity in some years. The resolution of phenology intensity measurements varies, and can range from a presence/absence measure to values comprising 10% intervals in our datasets. (ii) Sample size, which ranges from 2 to 41 individuals of a species in the site-based phenology datasets to several thousand in SeasonWatch; and (iii) Time interval or frequency of observations, which ranges from weekly to monthly (table 1).

### 2.3 *Phenology indices*

We focused on two frequently used indices in phenology studies to demonstrate the effects of the methods used in phenology data collection on the final metric and consequently on potential inferences: the overlap index and the seasonality index.

In this study we considered an overlap index that is intermediate between the Augspurger and Freitas–Bolmgren synchrony indices mentioned in the introduction. This is derived from the overlap indices in published literature (Ridout and Linkie 2009; Fisogni *et al*. 2022) which are defined for the overlap of two probability distribution functions. Here, we modified it to be used with discrete phenology measurements. In the overlap index, maximum synchrony (i.e., maximum values of the overlap index) is achieved when an individual's phenophase co-occurs with that of every other individual at the same intensity, even if it is not the maximum intensity of the phenophase (e.g., if all individuals have 50% flowering at a given time). In effect, a population in which all individuals show a particular phenophase simultaneously at the exact same intensity (i.e., zero variation in the timing and intensity of a phenophase) has maximum synchrony. Average overlap synchrony across all the individuals in a sample provides an estimate of population-level overlap synchrony.

Specifically, the overlap index ($O$) of individuals $i, j$ in a population of $S$ individuals is

$$O_S(i,j) = O_S(j,i) = \frac{\int_T min(p_{i,t}, p_{j,t})dt}{\int_T max(p_{i,t}, p_{j,t})dt} \quad (1)$$

where $p$ is phenology measurement (e.g., intensity of fruiting) over time $T$.

Given that phenology measurements are taken at discrete time points in practice, the above formula can be written as follows, for time ($t$) increasing from 1 to $T$:

$$O_s(i,j) = O_s(j,i) = \frac{\sum_{t=1}^{t=T} min(p_{i,t}, p_{j,t})}{\sum_{t=1}^{t=T} max(p_{i,t}, p_{j,t})} \quad (2)$$

In other words, if the phenology intensity measurement (e.g., intensity of fruiting) of individuals $i$ and $j$ are identical at all time points $t$ from $t$= 1 to $t$= $T$, their phenologies are perfectly overlapping. This means that at every time point $t$, $min(p_{i,t}, p_{j,t})$ and $max(p_{i,t}, p_{j,t})$ are identical, resulting in an overlap index $O_s(i,j) = O_s(j,i) = 1$.

The overlap index of individual $i$ in a population of $S$ individuals is a composite value of the overlap of each individual $i$ with every other individual in the population (or sample), and can be written as

$$O_s(i) = \frac{\sum_{j=1, j\neq i}^{S} O_s(i,j)}{S-1} \quad (3)$$

And the population overlap is

$$O_S = \frac{\sum_i^S O_S(i)}{S} \quad (4)$$

The seasonality of a phenophase for a given population is described by the average timing (day of

the year) of the phenophase (which is the 'peak' of the given phenophase), and an index of amplitude, ρ, which describes the seasonality of a given phenophase, with larger ρ values indicating a highly seasonal phenophase. In this study, we refer to the index of amplitude, ρ, as the seasonality index. We use the function *circ.stats* in the R package CircStats to calculate both the timing and amplitude of seasonality for our datasets (Agostinelli and Agostinelli 2001). Briefly, this package considers all the occurrences of a given phenophase in a year as vectors on a unit circle and calculates the mean direction (which is the average day of the year, or 'peak' season of the phenophase) and the mean vector length (which is the 'seasonality index', or ρ). Of these, we focused on the seasonality index, ρ, in this study. The computation of both the seasonality index (ρ) and the overlap index used data on the occurrence of a phenophase, the intensity of the phenophase, and sample size. In this study, we assessed the degree of influence of these factors on the computed indices.

### 2.4 *Simulated populations*

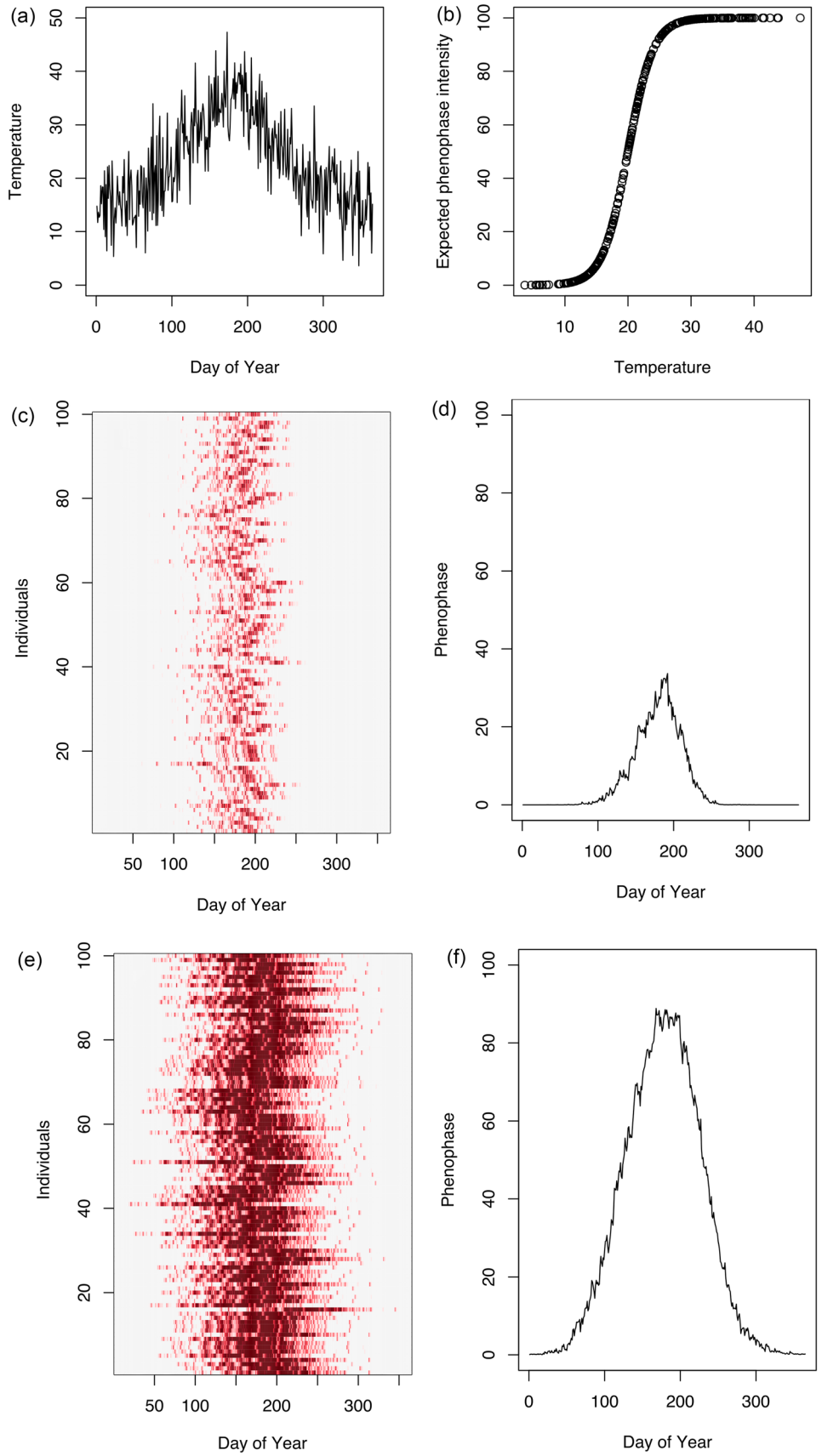
In tropical forests of Southeast Asia and South America, variability in tree phenology has been found to result from cues such as extreme changes in solar radiation, rainfall, or temperature (Sakai *et al*. 1999; Butt *et al*. 2015). Here, we chose temperature as a climatic trigger affecting reproductive phenology. To test the effects of methodological differences in data collection on phenology indices, we simulated two populations of 100 individuals each with a temperature-driven phenology, with temperature, and consequently the phenophase, peaking at about 180 days (figure 2a–b). Temperatures were simulated to fluctuate from day to day. Population phenologies responded to temperature with a lag of ∼20 days. One of these populations was simulated to have a fast-changing phenophase, with individuals showing phenophase for relatively fewer days. This was achieved by setting the sensitivity of the phenophase as 38°C. The second population was modelled to have a slow-changing phenophase, with temperatures above 28°C eliciting the phenophase (figure 2c–f).

While we used temperature as the climatic factor driving phenology in this simulation, similar simulations may be done using solar radiation, precipitation, or other climatic or environmental factors impacting phenology. This may be done by changing the range of values, the day of the year when the climatic factor may peak (e.g., 180 days in this simulation), and the sensitivity of the phenophase (e.g., 38°C and 28°C for the fast- and slow-changing phenophase types in this study). Phenological responses may also be simulated to mirror climatic or environmental drivers more closely by reducing the number of lag days. While not a focus of this study, a modified version of this simulation may also be useful in studying the impacts of multiple interacting climatic and environmental factors on tree phenology.

### 2.5 *Analyses*

For the primary datasets, we tested for the effect of phenology intensity resolution by comparing phenology indices calculated using actual data (fine-scale intensity resolution) with those calculated by setting intensity values greater than 0 to 1, thus reducing it to a presence/absence measurement (see table 1 for details of phenological intensity resolution of the actual data from each site). In the simulated datasets, we compared the effect of recording phenology intensity at 10% or 25% intervals; as none, few, or many (0, 33, and 100%); and as absence/presence (0, 100%). We tested for sample size effects by iteratively sampling fewer individuals from the real and simulated datasets and using these data to compute phenology indices.

The overlap index across sites and species varied considerably. Most of the datasets did not represent the entire plant community, and since phenology is highly species-specific, we chose to examine representative species from these landscapes for further comparisons. To assess the effects of the resolution of phenological intensity on phenology indices, we chose the 10 most observed species in each site (see the appendix for details of representative species). To assess the effects of sample size in phenology data collections on phenology indices, we used phenology data for one representative species from each of the five datasets: *Erythroxylum monogynum* Roxb. from Rishi Valley, *Anogeissus latifolia* (Roxb. ex DC.) Wall. ex Guill. and Perr. from NSTR, *Horsfieldia kingii* (Hook.f.) Warb. from Pakke, *Paracroton pendulus* Miq. from the Anamalais, and *Mangifera indica* L. from SeasonWatch. Each of these selected species are among those with the largest number of individuals sampled in each site. These are also species that are dominant in the landscape, and/or are ecologically interesting in other ways (e.g., *Horsfieldia kingii* is a relatively rare nutmeg species that is dispersed by hornbills in the Pakke site, and hence has high numbers of individuals that are monitored for phenology (Datta and Rane 2013)).


(a)
Temperature
Day of Year
(b)
Expected phenophase intensity
Temperature
(c)
Individuals
Day of Year
(d)
Phenophase
Day of Year
(e)
Individuals
Day of Year
(f)
Phenophase
Day of Year

◀**Figure 2.** Simulation set-up: A temperature-sensitive phenology was modelled, with temperatures peaking at 180 days (**a**, **b**). Phenologies of individuals and populations were simulated with a fast-changing phenophase (**c**, **d**) and a slow changing phenophase (**e**, **f**). The red dots in (**c**) and (**e**) indicate the presence of phenophase in each of the 100 individuals in the simulated population (*y*-axis) in each day of the year (*x*-axis). Note that the phenophase is maintained for longer periods of time in panel (**e**) (slow-changing phenophase) than in panel (**c**) (fast-changing phenophase). Panels (**d**) and (**f**) show how the phenophase peaks around 180 days in both simulated populations with fast- and slow-changing phenophases, although the intensity of the phenophase varies.

## 3. Results

### 3.1 *Phenology across sites – general trends*

Reproductive phenology, both in terms of overlap and seasonality indices, was site- and species-specific. In the two wetter sites (Pakke and the Anamalais) and one seasonally dry site (NSTR), for instance, fruiting in different species occurred almost throughout the year (figure 3). In the two other datasets (Rishi Valley and SeasonWatch), many of the observed species fruited around the same time of the year.

### 3.2 *Intensity resolution effects on phenology indices*

The overlap index of fruiting in 10 representative species of Pakke and Rishi Valley varied considerably, from values close to 0 overlap to high overlap (0.05–0.78 in Pakke and 0.11–0.72 in Rishi Valley). Overlap indices in the SeasonWatch and Anamalais datasets tended to be lower than 0.5 (0.01–0.2 in SeasonWatch, and 0.04–0.36 in the Anamalais), while ranging from 0.02 to 0.53 in the NSTR dataset. In four of the five datasets, the phenological overlap index for fruiting was marginally higher when coarser intensity measurements were used for the 10 most observed species as compared with finer intensity measures (figure 4). The effect of resolution of intensity measurements on overlap was most apparent in the NSTR dataset, which had the finest resolution of intensity estimates for phenology (10% intervals). Here, the overlap index was substantially overestimated when intensity was measured at a coarse (presence/absence) scale.

The seasonality index for 10 representative species varied considerably in four of the five datasets but, on average, tended to be above 0.5, indicating that species were seasonal in their fruiting phenology. The Season Watch species, on average, had a lower seasonality index because of the wider geographic scale of measurement. The seasonality index across datasets was not affected by the resolution of phenology intensity measurements (figure 4).

### 3.3 *Sample size effects on phenology indices*

In the SeasonWatch dataset, in which the sample size of *M. indica* was 1824, sample size effects were strongly apparent; the overlap index was higher at low sample sizes as compared with mid- and high-sample sizes (figure 5). In all other datasets, the overall sample size for observation was between 5 and 41 trees, and therefore, sample-size effects were not as evident (figure 5). Effects of sample size on seasonality index were also comparable across datasets with marginal variation between lower, mid, and maximum sample sizes in Pakke, Rishi Valley, and the Anamalais. Lower sample sizes yielded considerably higher seasonality index values only in the SeasonWatch data (figure 5; supplementary figure 1).

### 3.4 *Effects of intensity resolution, sample size, and observation frequency in simulated populations*

Simulated tree populations are of two types depending on the duration and amplitude of their phenophases. The simulated population with a fast-changing phenophase had short-duration, small-amplitude peaks (figure 2c and d), and the population with a slow-changing phenophase had longer-duration, higher-amplitude peaks (figure 2e and f). Differences in phenology measurement methods affected the overlap index estimates differently in these two populations. The overlap index for populations with a fast-changing phenophase was higher than the true value at low-resolution measurements of phenophase intensity, but was relatively uninfluenced by frequency of observation or sample size (figure 6a). The overlap index estimates were closest to the true value at the highest resolution of intensity measurements, irrespective of sample size or frequency of observation. The overlap index for populations with a slow-changing phenophase was also higher than the true value at low resolutions, but progressively approached the true value with increasing resolution of phenophase intensity measurements. Similarly, differences in sample size and frequency of observation showed no effect on estimates of overlap

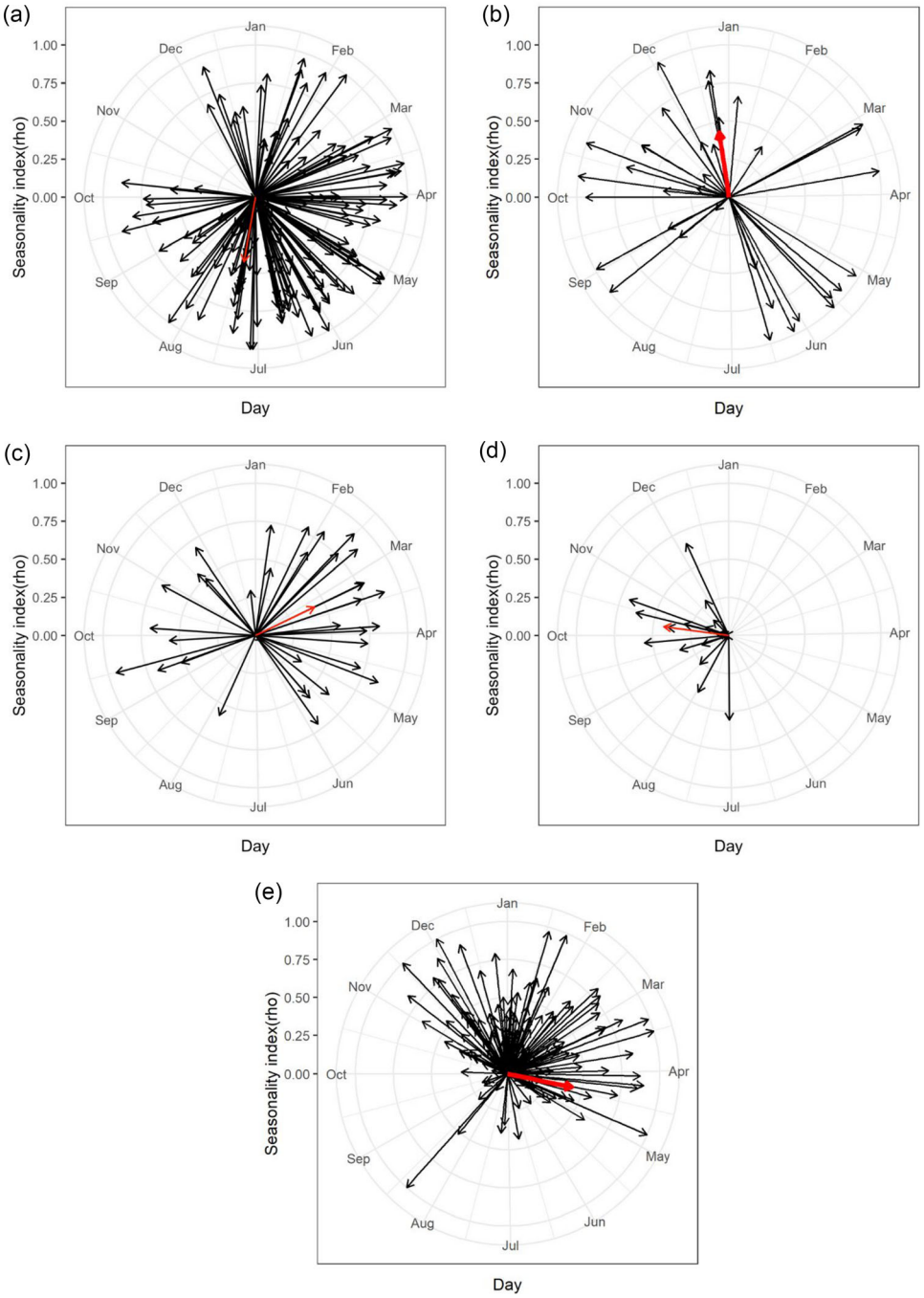



**Figure 3.** Seasonality (ρ) and timing of fruiting across five datasets from the tropics: (**a**) the Anamalais, (**b**) NSTR, (**c**) Pakke, (**d**) Rishi Valley, and (**e**) SeasonWatch. Each arrow represents a species observed in the dataset, with the direction of the arrow indicating the mean season of the phenophase, while the length of the arrow indicates the mean amplitude of the phenophase (longer arrow indicates more seasonal response). The red arrow represents seasonality values for a selected species from each dataset – *Paracroton pendulus* (the Anamalais), *Anogeissus latifolia* (NSTR), *Horsfieldia kingii* (Pakke), *Erythroxylum monogynum* (Rishi Valley), and *Mangifera indica* (SeasonWatch).

index. In both cases, greater resolution of phenophase intensity measurement captured the true overlap most accurately.

As in the case of the overlap index, seasonality estimates were mainly affected by the resolution of phenophase intensity measurements (figure 7a and b); they were underestimated at coarser intensity measures, and converged on the true value at finer-scale intensity measurements. This is in line with the results based on seasonality index calculations using site-based data as

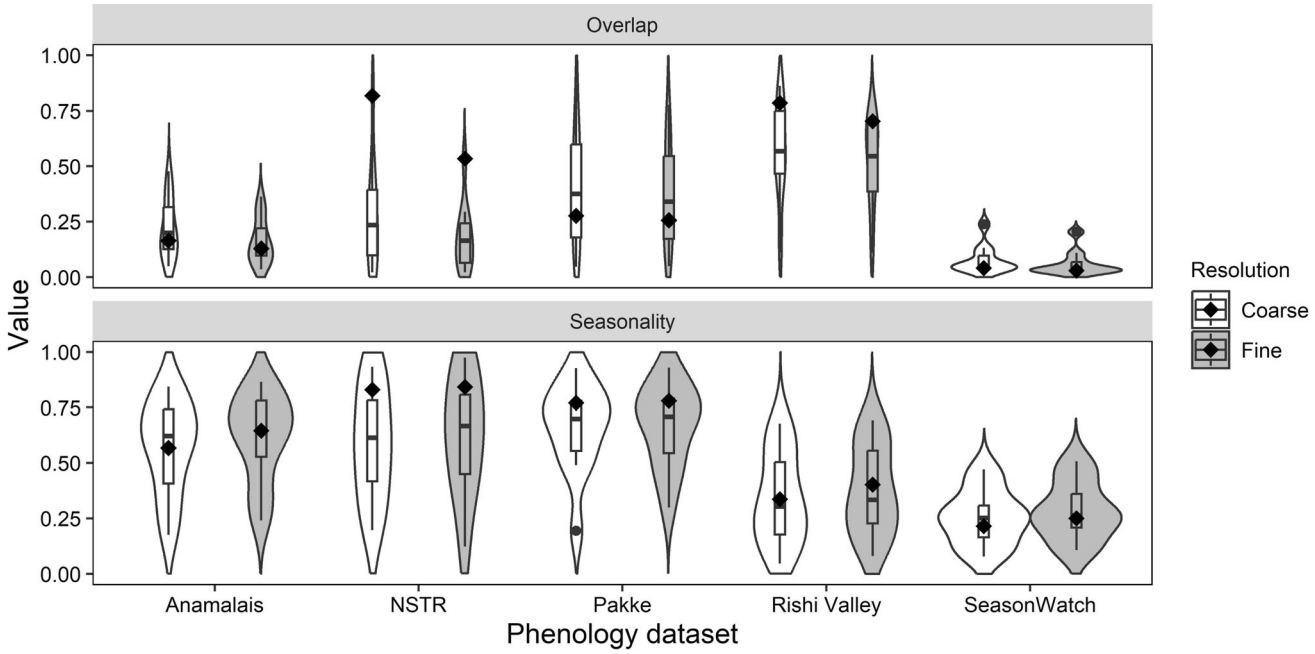



**Figure 4.** Overlap and seasonality indices for fruiting in the 10 most observed or dominant species from five phenology datasets with two levels of phenology intensity resolution. The diamond within each violin plot indicates the mean overlap and seasonality index of one representative species from each site: *Paracroton pendulus* (the Anamalais), *Anogeissus latifolia* (NSTR), *Horsfieldia kingii* (Pakke), *Erthroxylum monogynum* (Rishi Valley), and *Mangifera indica* (SeasonWatch). 'Fine' corresponds to the finest level of phenology intensity measurement, i.e., 0, 10%, 20% to 100% in NSTR; 0, 25%, 75%, 100% in Pakke and the Anamalais; and 0, 33%, 100% in the Rishi Valley and SeasonWatch datasets, respectively. For all datasets, values above 0 were set to 1 in order to derive the 'Coarse' intensity measurement resolution, corresponding to a phenophase being present or absent. Error bars on the overlap index indicate standard errors across the population of the species observed in the dataset.

well (figure 4). In the simulated data, frequency of observation and sample size do not appear to influence the estimated seasonality index for fast- and slow-changing phenologies. For both overlap and seasonality, larger sample sizes reduce the error around the average index estimates. Populations with fast-changing phenologies have much larger errors around seasonality estimates than populations with slow-changing phenologies. This is mitigated to some extent at higher sample sizes. Overall, irrespective of whether observations were made at the weekly, fortnightly, or monthly scales, measuring intensity at a finer scale yielded better estimates of overlap and seasonality, whereas measuring more individuals yielded better precision of the estimates.

## 4. Discussion

In this study, we sought to understand the effects of differing methodologies on the estimation of two measures frequently used in plant phenological studies: overlap and seasonality indices. The primary motivation to conduct this study was to infer generalizable phenology trends across five datasets differing in methods of quantification, geographic location, and scale of phenological data collection. Overall, based on an analysis focused on a representative set of species per site, we found that only the finest level of intensity estimation (10% intervals) offered any advantage in the accurate estimation of the overlap index. The seasonality index across all datasets was not affected by the intensity resolution. As expected, depending on the overall variability in the population of interest, larger sample sizes of a single representative species allowed for higher precision in the estimation of overlap and seasonality across datasets. In simulated fast- and slow-changing populations as well, finer intensity resolution (10% intervals) and higher sample sizes (at least 40 individuals of each species) were required to accurately estimate these indices.

The NSTR dataset had the highest average overlap index in the fruiting phenology for representative species. This site is highly seasonal with a distinct wet period that lasts $\sim$4 months with an average rainfall of

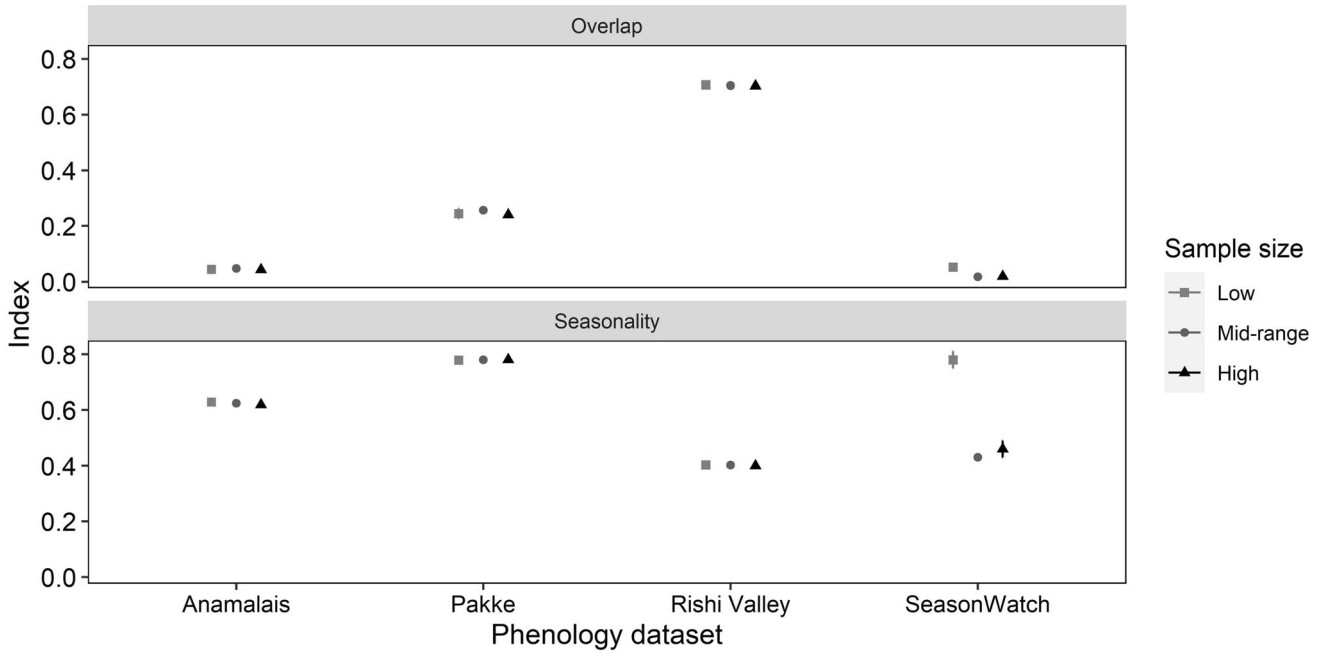



**Figure 5.** Effect of sample size (number of unique individuals monitored for phenology) on the overlap and seasonality indices of fruiting in one representative species each across different phenology datasets. 'Low', 'mid-range', and 'maximum' in different datasets are as follows – the Anamalais: 10, 20, 41; Pakke: 5, 10, 24; Rishi Valley: 10, 20, 40; and SeasonWatch: 30, 550, 1824. Error bars are based on 1000 random samples of low- and mid-range sample sizes. The effect of sample size was most discernible in the dataset with the highest maximum sample size (SeasonWatch) for overlap and seasonality, with both indices being overestimated at lower sample sizes as compared with mid-range sample sizes. NSTR had overall species sample sizes of only 10 trees and was therefore excluded from this assessment.

713 mm (1970–2015). Lobo *et al.* (2003) suggested that in forests with a distinct dry season, abiotic factors are possibly the proximate causes affecting phenological patterns, leading to a higher value of overlap. In Pakke and the Anamalais, the overall higher moisture availability over a longer period possibly resulted in a wider seasonal window for fruiting, resulting in lower average overlap indices. Similar seasonal patterns of low overlap and wide distribution of fruiting across the year have been recorded in other tropical moist forest communities (e.g., Wheelwright 1985; Chapman *et al.* 2005).

Seasonality, on the other hand, was estimated as >0.4 in four of the five datasets, indicating that reproductive phenology in these species 'peaked' at certain times in the year. In the wetter sites, reproductive phenology was highly seasonal on average, while among the drier sites, NSTR was more seasonal than Rishi Valley. In both dry and wet sites, fruiting seasonality appeared related to the season of maximum rainfall, with the peak (as indicated by mean day of the year) occurring either just before (Pakke) or during (the Annamalais) the main summer or southwest monsoon in the two wet sites.

We regarded the overlap and seasonality index values of the SeasonWatch dataset with more caution than other datasets. Owing to the large spatial spread, and inconsistent measurement of trees in the citizen science project, undocumented sources of variation in phenology overlap may make these indices inappropriate at this spatial scale. In a highly heterogeneous dataset such as SeasonWatch, it may be more reliable to use simple measures of phenology such as the percentage of trees showing a phenophase across the spatial extent of the dataset. Other studies that have compared methodological effects on phenology concur that sampling strategies for site-specific studies need to be different from that of citizen science data (Morellato *et al.* 2018).

The simulation results indicate that standard indices of phenology can be affected by underlying species-specific phenology patterns. High variability in the timing and persistence of phenophases may be quantifiable using finest scales of intensity estimates, irrespective of frequency of observation and low sample sizes. Overlap and seasonality indices are both not ideal descriptors of phenology in rapidly changing and short-duration phenology, except at the highest resolutions of phenology intensity measurements. An alternative to

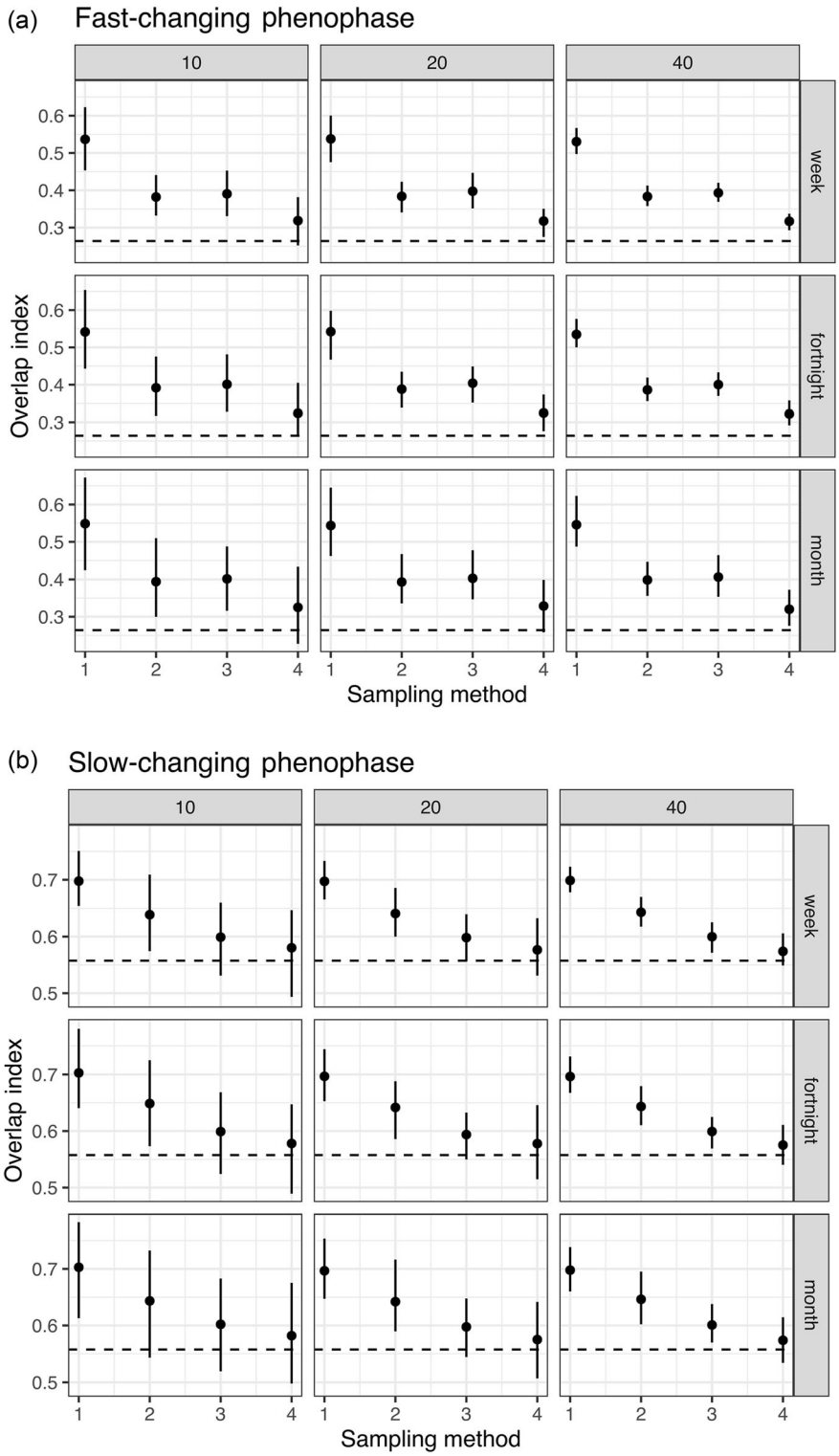



**Figure 6.** Sampling methodology effects on overlap synchrony index in simulated populations showing (**a**) fast-changing and (**b**) slow-changing phenology. Sampling methods (1, 2, 3, 4) correspond to phenology intensity measurements as follows: 1, presence/absence; 2, None, Few, Many; 3, 25% intervals; and 4, 10% intervals. Each column of panels corresponds to different sample sizes (10, 20, and 40 individuals), and rows represent frequency of sampling (i.e., whether the phenology data are recorded at weekly, fortnightly, or monthly intervals). The dashed horizontal lines represent the 'true' overlap synchrony of the two populations (0.264 for the population with the fast-changing phenology, and 0.557 for the slow-changing phenology population). The dots indicate mean population-level overlap and error bars are confidence intervals calculated over 100 iterations.

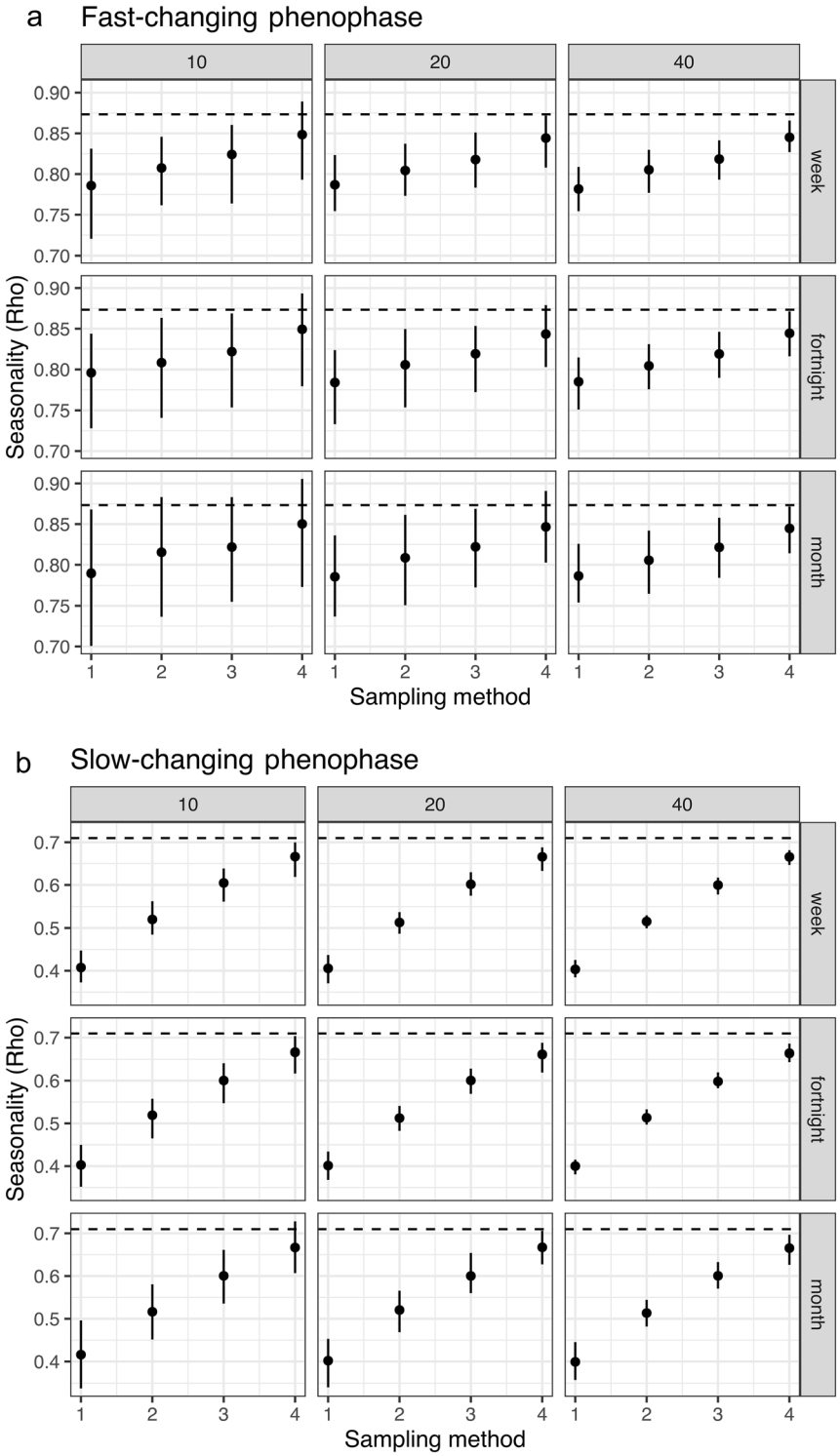



**Figure 7.** Sampling methodology effects on seasonality (ρ) index in simulated populations showing fast-changing (**a**) and slow-changing phenology (**b**). Sampling methods (1, 2, 3, 4) correspond to phenology intensity measurements as follows: 1, presence/absence; 2, None, Few, Many; 3, 25% intervals; and 4, 10% intervals. Each column of panels corresponds to different sample sizes (10, 20, and 40 individuals), and rows represent frequency of sampling (i.e., whether the phenology data are recorded at weekly, fortnightly, or monthly intervals). The dashed horizontal lines represent the 'true' ρ values of the two populations (0.873 for the population with the fast-changing phenology, and 0.710 for the slow-changing phenology population). The dots indicate mean population-level seasonality and error bars are confidence intervals calculated over 100 iterations.

using absolute scales of intensity resolution is to consider median values of intensity categories (e.g., 0, 0.16, 0.66 instead of 0, 0.33, 1). At median-intensity resolution scales, all scales of intensity measurement other than presence/absence yield more accurate estimates of the true overlap (supplementary figure 2).

We infer that an understanding of the underlying rate of phenophase change is essential to accurately estimate population-level synchrony in phenology using the overlap index. While studies elsewhere have highlighted the trade-off between sample size and frequency of observation (Morellato *et al*. 2010 and references therein), we highlight the role of underlying rate of change of a phenophase to be a key factor determining sampling methodology, irrespective of sample size and sampling duration (figures 6 and 7). This is particularly useful in the assessment of tropical tree species characterized by low sample size in forest communities (Morellato *et al*. 2010 and references therein). We found no effect of frequency of observation on the estimation of these two indices.

Both phenological seasonality and synchrony may be sensitive to temperature and precipitation changes and are therefore key in understanding plant responses to climate changes. Some plants may be more sensitive to climatic factors, and thereby have fast-changing phenophases. Other species may have a broader range of climatic conditions during which a phenophase is sustained, and thereby have slow-changing phenophases. We found that larger sample sizes are more important in populations with fast-changing phenologies to accurately estimate seasonality. We expected frequency of observation to also become important in accurate overlap and seasonality estimation in populations with fast-changing phenophases, but we did not find evidence for this in our simulations. Frequency of observation, however, may be an important factor in other phenology descriptors such as start or duration of phenophase, which are often used in cross-species or longitudinal phenology comparisons (Morellato *et al*. 2010).

We can expect these methodological considerations to also affect inferences on plant phenological response to climate or other ecological and evolutionary strategies at the species or community level. For instance, other metrics of detecting within-population concurrences of phenology (e.g., SD in first dates of flowering) have allowed for inferences such as increased phenological synchrony under rapid warming conditions within a season (Wang *et al*. 2016). In life-history strategies such as highly concurrent production of reproductive phenophases (masting), high overlap in flowering has been experimentally shown to be associated with better fruit set (Bogdzeiwicz *et al*. 2020). Although we did not examine community-level patterns in this study, these indices can be extended to encompass multiple species. For instance, a 40-year study on 51 perennial plant species in the Iberian Peninsula revealed substantial changes in the co-flowering patterns of these species, driven primarily by changing climatic conditions (Pareja-Bonilla *et al*. 2023). A substantial body of literature focuses on yet another level of phenological synchrony in relation to climate change — the phenological synchrony between trophic levels (Renner and Zohner 2018 and references therein). For instance, overlap and start dates of plant leaf-out and invertebrate herbivore phenophases have been reported to have changed under climate warming, resulting in consequences for invertebrate populations (e.g., van Dis *et al*. 2023). To make comparable inferences on climate impacts using the synchrony and overlap indices examined in this study, it is important to have larger sample sizes, especially for faster-changing plant populations, and to keep these considerations in mind while examining community-level patterns as well.

We conclude that varying methodology and underlying phenology dynamics affect estimation of standard indices. Based on results from five diverse datasets and two simulated populations, we recommend the following best practices for analyzing phenological patterns using the overlap and seasonality indices:

(i) For accurate estimation of the overlap index, finest intensity resolutions, at least at 10% intervals, are needed; resolutions coarser than 10% intervals perform similarly to presence/absence data in the estimation of the overlap index.
(ii) Depending on the larger ecological question being addressed, seasonality may be a more appropriate index to characterise species phenology if fine-level intensity measurements are not logistically possible.
(iii) Pilot studies and examination of regional floras for approximate timing of phenophases can be used to determine whether the population of interest has fast- or slow-changing phenophases. In species with fast-changing phenophases, prioritizing higher sampling effort can yield more accurate estimates of the seasonality index.

## Appendix

### *Summary of sampling effort and methods of phenology observation across the five datasets*

*Pakke:* A long-term tree phenology monitoring study of 722 reproductive trees of 53 selected species is

underway in the tropical semi-evergreen forests of the Pakke Wildlife Sanctuary and Tiger Reserve (see Datta and Rane 2013 for more details regarding the study site). The overall phenology tree sample represents 55% of the adult tree density and species composition at the study site based on prior studies (Datta 2001; Datta and Rawat 2008). Fifteen of the 20 top-ranked species are represented in the phenology sample. Other vegetation and climatic parameters are also recorded. Two to four observers carry out the monitoring every fortnight using binoculars to assess the presence of each phenophase in the tree canopies, along marked trails for each species. The phenophases recorded are the presence/absence of young leaf, mature leaf, senescent or old leaf, flower buds, open flowers, and unripe and ripe fruits. Unripe and ripe fruits are scored on a scale of 0–4, where 0 indicates no fruits and 1 denotes 25% of canopy in fruit, 2 refers to 25–50% of the canopy in fruit, 3 indicates 50–75%, and 4 denotes 100% of canopy in fruit. For this study, we used the monthly phenology data for 670 individual trees of 35 species (10–25 individuals per species) from April 2011 to September 2019. From this, data from 255 individuals of 12 species with 20–25 individuals per species were selected for representative species analyses. These representative species include *Aglaia* sp., *Ailanthus grandis, Dalrympelea pomifera, Dysoxylum cauliflorum, Dysoxylum gotadhora, Gynocardia odorata, Horsfieldia kingii, Livistona jenkinsiana, Pterospermum acerifolium, Sterculia villosa, Stereospermum chelonoides,* and *Tetrameles nudiflora*.

*The Anamalais:* In total, 1376 individual trees of 172 species were observed every month from March 2017 to December 2020. Phenological observations were made on 1–41 individuals per species along seven trails located in the Anamalais Plateau and Anamalai Tiger Reserve (10.3021°N, 76.8298°E – 10.4011°N, 76.9929°E) in the Anamalai Hills of the Western Ghats. All trails were located in mid-elevation tropical wet evergreen forests of the *Cullenia exarillata–Mesua ferrea–Palaquium ellipticum* type (Pascal 1988). From this, data from 507 individuals of 13 species, with 35–41 individuals per species were selected for representative species analyses. These representative species are *Acronychia pedunculata*, *Antidesma menasu*, *Cullenia exarillata*, *Euodia lunu-ankenda*, *Gomphandra coriacea*, *Litsea stocksii, Mesua ferrea*, *Myristica dactyloides*, *Palaquium ellipticum*, *Paracroton pendulus*, *Persea macrantha*, *Vateria indica*, and *Villebrunea integrifolia*. The trails were surveyed at the start of every month and trees were visually scored for the following phenophases: leaves (young/mature), flowers (buds, open), and fruits (unripe, ripe). Phenophase was scored on a scale of 0 to 4, 0 indicating the absence of the phenophase, and values from 1 to 4 indicating the intensity (from low to high in 25% of classes based on the extent of canopy manifesting the phenophase).

*Rishi Valley:* The study was carried out in a semi-arid tropical scrubland along the Eastern Ghats in the Deccan Plateau. The study area has distinct cool and warm, and dry and wet seasons (Ramaswami *et al.* 2019). Overall, 647 individual trees of 18 species (12–40 individuals per species) were observed from December 2007 to December 2016. From this, data from 600 individuals of 15 species with 40 individuals per species were selected for representative species analyses. These representative species are *Acacia leucophloea*, *Albizia amara*, *Azadirachta indica*, *Chomelia asiatica*, *Delonix regia*, *Erythroxylum monogynum*, *Flacourtia sepiaria*, *Lantana camara*, *Peltophorum pterocarpum*, *Pongamia pinnata*, *Randia dumetorum*, *Santalum album*, *Strychnos nux-vomica*, *Tamarindus indica*, and *Wrightia tinctoria.* The trees in Rishi Valley were monitored every fortnight. Two stages of leaf (young and mature), two stages of flower (bud and open), and two stages of fruit (unripe and ripe) were monitored. The same observer has made these observations since 2007, and phenophase quantities are recorded as 'none', 'few', and 'many', corresponding to the phenophase being absent, detected in 30% of the canopy, or detected in >30% of the canopy.

*Nagarjunasagar Srisailam Tiger Reserve (NSTR):* The savanna woodland plot in NSTR is part of the long-term ecological monitoring network LEMoN India initiative (*https://lemonindia.weebly.com*). The site is located at an elevation of 700 a.s.l. and receives a long-term mean annual rainfall (1970–2015) of 713 mm, with eight dry months (figure 1). Most of the rainfall occurs between July and October (summer monsoon). Overall, 113 individuals (girth at breast height > 10 cm) belonging to 11 species that comprise 80% of the stem basal area in a 1-ha plot were monitored fortnightly in the year 2018. The following variables were recorded: canopy fullness (L), and the percentage of flushing (FL), mature (M), and senescing leaves (S). During each session, we assigned a visual percentage score of 0–100 for each phenophase (flushing, mature, and senescence) with a resolution of 10%. The species are *Anogeissus latifolia*, *Bridelia retusa*, *Buchanania cochinchinensis*, *Chloroxylon swietenia*, *Dalbergia paniculata*, *Emblica officinalis*,

*Eriolaena quinquelocularis*, *Grewia orbiculata*, *Pterocarpus marsupium*, *Terminalia elliptica*, and *Ziziphus xylopyrus.*

*SeasonWatch:* This is a citizen science project collating information on tree phenology across India. In total, 120,000 individual trees of 170 species were observed either once or repeatedly every week from November 2011 to September 2023. From this, data from 47,224 individuals of 10 species with 1997–10,238 individuals per species (1 Jan 2014–30 May 2022) were selected for representative species analyses. These species are *Artocarpus heterophyllus*, *Bauhinia purpurea*, *Cassia fistula*, *Mangifera indica*, *Mimusops elengi*, *Phyllanthus emblica*, *Syzygium cumini*, *Samanea saman*, *Tamarindus indica* and *Tectona grandis*. More than 85% of these data come from the south Indian state of Kerala (data used in this study). Trees are registered with the programe by citizen scientists and are observed on a weekly basis for three stages of leaf (young, mature, and dying), two stages of flower (bud and open), and two stages of fruit (unripe, ripe, and 'open' in dehiscent fruit species). Observers were trained to report phenophase quantities as 'none', 'few', and 'many', corresponding to the phenophase being absent, detected in 30% of the canopy, or detected in >30% of the canopy.

## Acknowledgements

The authors would like to acknowledge the following people for collecting and collating data from different field sites: G Moorthi, T Sundarraj, T Vanidas, A. Sathish Kumar, Kshama Bhat, P Pavithra (the Anamalais), Venkateshwarlu Byrapoghu, Niharika, and Ramesh (NSTR), Khem Thapa, late Kumar Thapa, late Tali Nabam, Narayan Mogar, Sagar Kino, Arjun Rai, Turuk Brah, Sital Dako, Amruta Rane, Ushma Shukla, Akanksha Rathore, Swati Sidhu, Rohit Naniwadekar, Veena Rai, Chaithra Gowda, Harman Kour, Noopur Borawake, Saniya Chaplod, Devathi Parashuram, Karishma Pradhan, Bibidishananda Basu (Pakke), P Somnath (Rishi Valley), and the SeasonWatch Citizen Scientist Network comprising thousands of volunteers (SeasonWatch). We thank Divya Mudappa for her valuable inputs for this project, and HS Suresh for making available their review paper on phenology. We also wish to thank the Department of Environment, Forests and Climate Change in the states of Tamil Nadu, Arunachal Pradesh, and Andhra Pradesh, and State Forest Departments for providing permissions and local support to conduct phenology studies. The studies used in the paper were funded by Rohini Nilekani Philanthropies, AMM Murugappa Chettiar Research Centre (Anamalais), Department of Science and Technology GoI, National Centre for Biological Sciences, Bengaluru (NSTR), The Serenity Trust, Whitley Fund for Nature, National Geographic Society (Pakke), and Wipro Foundation (SeasonWatch), as well as grants from the Rainmatter Foundation, and philanthropic donations from Mr. Arvind Datar.

## Author contributions

All authors contributed to conceptualization; KT, AR, AM, AD, AK, JR, MS, SO, SK, TRSR and GR collected and curated the data; YT, KT, AR, AK and GR developed the methodology, curated the data, and performed analyses; YT, KT and GR wrote the initial draft of the manuscript; all authors contributed to editing and finalising the manuscript.

### Declarations

**Conflict of interest** The authors declare no conflict of interest.

## References

Agostinelli C and Agostinelli MC 2001 Package 'CircStats' (*https://cran.r-project.org/web/packages/CircStats/CircStats.pdf*).

Augspurger CK 1983 Phenology, flowering synchrony, and fruit set of six neotropical shrubs. *Biotropica* **15** 257–267

Bogdziewicz M, Pesendorfer M, Crone EE, *et al.* 2020 Flowering synchrony drives reproductive success in a wind-pollinated tree. *Ecol. Lett.* **23** 1820–1826

Butt N, Seabrook L, Maron M, *et al.* 2015 Cascading effects of climate extremes on vertebrate fauna through changes to low-latitude tree flowering and fruiting phenology. *Glob. Chang. Biol.* **21** 3267–3277

Chapman CA, Chapman LJ, Wangham R, *et al.* 1992 Estimators of fruit abundance of tropical trees. *Biotropica* **24** 527–531

Chapman CA, Wrangham R and Chapman LJ 1994 Indices of habitat-wide fruit abundance in tropical forest. *Biotropica* **26** 160–171

Chapman CA, Chapman LJ, Struhsaker TT, *et al.* 2005 A long-term evaluation of fruiting phenology: Importance of climate change. *J. Trop. Ecol.* **21** 31–45

Datta A 2001 An ecological study of sympatric hornbills and fruiting patterns in a tropical forest in Arunachal Pradesh, PhD thesis, Saurashtra University, Rajkot, Gujarat, India

Datta A and Rane A 2013 Phenology, seed dispersal and regeneration patterns of *Horsfieldia kingii*, a rare wild nutmeg. *Trop. Conserev. Sci.* **6** 674–689

Datta A and Rawat GS 2008 Dispersal modes and spatial patterns of tree species in a tropical forest in Arunachal Pradesh, northeast India. *Trop. Conserv. Sci.* **1** 163–185

Fisogni A, De Manincor N, Bertelsen CD, *et al.* 2022 Long-term changes in flowering synchrony reflect climatic changes across an elevational gradient. *Ecography* **2022** e06050

Freitas L and Bolmgren K 2008 Synchrony is more than overlap: Measuring phenological synchronization considering time length and intensity. *Rev. Bras. Bot.* **31** 721–724

Hemingway CA and Overdorff DJ 1999 Sampling effects on food availability estimates: phenological method, sample size, and species composition. *Biotropica* **31** 354–364

Lobo JA, Quesada M, Stoner KE, *et al.* 2003 Factors affecting phenological patterns of bombacaceous trees in seasonal forests in Costa Rica and Mexico. *Am. J. Bot.* **90** 1054–1063

McKey D 1975 The ecology of coevolved seed dispersal systems; in *Coevolution of animals and plants: V Symposium, First International Congress of Systematic and Evolutionary Biology* (Eds.) Gilbert LE and Raven PH (New York: University of Texas Press) pp 159–191

Morellato LPC, Camargo MGG, D'Eça Neves FF, *et al.* 2010 The influence of sampling method, sample size, and frequency of observations on plant phenological patterns and interpretation in tropical forest trees; in *Phenological research* (Eds.) Hudson IL and Keatley MR (Dordrecht: Springer Netherlands) pp 99–121

Morellato LPC, Abernethy K and Mendoza I 2018 Rethinking tropical phenology: Insights from long-term monitoring and novel analytical methods. *Biotropica* **50** 371–373

Pareja-Bonilla D, Arista M, Morellato LPC, *et al.* 2023 Better soon than never: climate change induces strong phenological reassembly in the flowering a Mediterranean shrub community. *Ann. Bot.* **135** 239–254

Piao S, Liu Q, Chen A, *et al.* 2019 Plant phenology and global climate change: Current progresses and challenges. *Glob. Chang. Biol.* **25** 1922–1940

Ramaswami G, Datta A, Reddy A, *et al.* 2019 Tracking phenology in the tropics and in India: The impacts of climate change; in *The implications of climate change for Indian biodiversity: An overview* (Eds.) Bhatt JR, Das A, and Shanker K (New Delhi: Ministry of Environment, Forest and Climate Change, Government of India) pp 45–59

Renner SS and Zohner CM 2018 Climate change and phenological mismatch in trophic interactions among plants, insects, and vertebrates. *Annu. Rev. Ecol. Evol. Syst.* **49** 165–182

Ridout MS and Linkie M 2009 Estimating overlap of daily activity patterns from camera trap data. *J. Agric. Biol. Environ. Stat.* **14** 322–337

Sakai S, Momose K, Yumoto T, *et al.* 1999 Plant reproductive phenology over four years including an episode of general flowering in a lowland dipterocarp forest, Sarawak, Malaysia. *Am. J. Bot.* **86** 1414–1436

SeasonWatch Citizen Scientist Network 2022 SeasonWatch - Data on Tree Phenology Dataset (India). v1.4. SeasonWatch India

Terborgh J 1990 Seed and fruit dispersal–commentary. *Reprod. Ecol. Trop. Forest Plants* **7** 181–190

van Dis NE, Sieperda GJ, Bansal V, *et al.* 2023 Phenological mismatch affects individual fitness and population growth in the winter moth. *Proc. R. Soc. B* **290** 20230414

Wang C, Tang Y and Chen J 2016 Plant phenological synchrony increases under rapid within-spring warming. *Sci. Rep.* **6** 25460

Wheelwright NT 1985 Competition for dispersers, and the timing of flowering and fruiting in a guild of tropical trees. *Oikos* **44** 465–477



Corresponding editor: MEGHNA KRISHNADAS