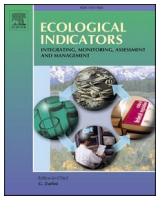
Ecological Indicators 181 (2025) 114484

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



Deriving exploratory temperature-based phenological indicators under data-limited conditions: integrating ERA5 and citizen science☆

Yu-Shiang Huang [a], Chieh-Hsiang Fan [a], Yi-Cian Lu [a,b], Yi-Ta Wu [a], Yu-Cheng Niu [a], Yi-Huan Hsieh [a,c], Yu-Kai Liao [a,*], Syuan-Jyun Sun [a,*]

[a] International Degree Program in Climate Change and Sustainable Development, National Taiwan University, Taipei 106319, Taiwan
[b] International Doctoral Program in Agriculture, National Chung Hsing University, Taichung 40227, Taiwan
[c] Office of Sustainability, National Taiwan University, Taipei 106319, Taiwan

ARTICLE INFO

Keywords:
Phenology
Temperature indicators
ERA5 reanalysis
Citizen science
Climate change monitoring

ABSTRACT

Phenological shifts are among the most visible biological responses to climate change, with profound implications for ecological interactions, food webs, and predictive modeling. Yet in tropical and subtropical regions, where global biodiversity is concentrated, phenological responses to temperature variability remain poorly quantified due to limited long-term monitoring. This study explores the potential of integrating gridded climate reanalysis datasets with opportunistic citizen science observations to develop exploratory, temperature-based phenological indicators. We propose the Cross-Scale Coarse Indicators (CSCIs), a novel family of indicators designed to address spatial and temporal mismatches in coarse-resolution climate data and unstructured biological records. By incorporating coarse spatial grids and extended temporal windows, CSCIs aim to enhance robustness while preserving biological relevance under uncertain data conditions, as demonstrated in two case studies. Case 1 modeled flowering observations of four tree species (*Crateva religiosa*, *Cornus controversa*, *Firmiana colorata*, *Helicteres isora*) using ERA5-derived temperature indicators. Variable selection was performed using Random Forests, followed by Partial Least Squares regression to assess explanatory power. In Case 2, indicators for *Turpinia formosana* were derived from finer-resolution historical datasets (0.05°–0.01°) to evaluate indicator stability across scales. Our results demonstrate that CSCIs can capture meaningful temperature–phenology relationships, even with coarse and biased conditions. Despite species-specific difference, consistent indicator convergence across scales supports the utility of the CSCI. These findings highlight CSCIs as a scalable and transferable tool for phenological research in data-scarce regions and a practical foundation for climate-biodiversity assessment where structured monitoring is limited.

1. Introduction

Climate change has profoundly influenced phenological patterns in plants, particularly flowering and leaf-out, with cascading effects on species interactions, ecosystem functioning, and agricultural productivity (Walther et al., 2002; Cleland et al., 2007; Panchen et al., 2012). Earlier spring events, in particular, are widely observed in response to warming (Badeck et al., 2004; Cleland et al., 2007), but species-specific responses can desynchronize ecological relationships and consequently affect the functioning of entire ecosystems (Thackeray et al., 2016). To understand the patterns and trends of phenological changes, long-term ecological monitoring is an essential scientific endeavor. The data collected not only reveal transformations within ecosystems but also serve as a crucial foundation for environmental legislation and policy formulation (Lindenmayer and Likens, 2010). Effective ecological monitoring program demands clearly defined objectives, minimization of observer bias and data loss, sufficient spatial and temporal replication, and adequate statistical power to detect significant changes (Legg and Nagy, 2006; Field et al., 2007).

However, while phenology has been extensively studied in temperate systems, tropical-subtropical regions remain substantially underrepresented in the global phenological literature, largely due to

☆ This article is part of a special issue entitled: 'Biodiversity & climate change' published in Ecological Indicators.

* Corresponding authors.
E-mail addresses: r13247011@ntu.edu.tw (Y.-S. Huang), r13247001@ntu.edu.tw (C.-H. Fan), d13248003@ntu.edu.tw (Y.-C. Lu), r13247004@ntu.edu.tw (Y.-T. Wu), r13247018@ntu.edu.tw (Y.-C. Niu), coldfishball@gmail.com (Y.-H. Hsieh), liaoyk@ntu.edu.tw (Y.-K. Liao), sjs243@ntu.edu.tw (S.-J. Sun).

https://doi.org/10.1016/j.ecolind.2025.114484
Received 8 September 2025; Received in revised form 24 November 2025; Accepted 24 November 2025
Available online 2 December 2025

the lack of long-term, large-scale monitoring networks (Cleland et al., 2007; Taylor et al., 2019). Although temperate-zone research has established a relatively comprehensive phenological paradigm, extending this framework to tropical contexts often results in unreliable or even incorrect forecasts of phenological responses to climate change (Davis et al., 2022). This geographic bias in phenological research is driven in part by the global concentration of long-term observational datasets in temperate regions, while tropical systems remain comparatively understudied, both due to limited data availability and a long-standing perception that phenological cycles in the tropics are less pronounced or predictable, a view increasingly challenged by recent evidence (Cleland et al., 2007; Panchen et al., 2012; Taylor et al., 2019).

Understanding how ecological processes and patterns vary across scales remains a central challenge in ecology (Levin, 1992). The mismatch between the spatial and temporal scales at which ecological data are collected and those at which environmental drivers are represented complicates cross-scale inference and modeling. To address these limitations, researchers have increasingly turned to global reanalysis datasets such as the Fifth Generation ECMWF Reanalysis (ERA5), which provide spatially continuous climate estimates at daily to sub-daily temporal resolution (Hersbach et al., 2020). However, ERA5′s coarse native resolution (0.1°, ~11 km^2 in tropical areas) and for forest structural complexity limit its capacity to capture the highly heterogeneous microclimates of tropical forest understories, which can differ substantially from above-canopy or open-area conditions (Ismaeel et al., 2024). Such coarse-to-fine scale mismatches are not unique to natural forest systems but are also evident in urban climatology, where surface heat retention and canopy structure create distinct temperature regimes that influence local vegetation dynamics (Biswas et al., 2022; Patil and Surawar, 2023).

In parallel, citizen science has emerged as a valuable complementary tool for tracking large-scale phenological trends, providing broader spatial and temporal coverage than most formal monitoring networks, while lack of methodological consistency and long-term continuity (Taylor et al., 2019; Klinger et al., 2023). Yet, citizen science records are often perceived as noisy or biased due to inconsistent observer effort, temporal clustering, weekend bias, and taxonomic uncertainty (Panchen et al., 2012; Courter et al., 2013; Baker et al., 2021; Yoder and Butterfield, 2024; Canavan et al., 2025). Despite their limitations, both ERA5 and citizen science data still can offer valuable insights as “secondary resources”. The key challenge lies in developing methodological frameworks capable of deriving robust, cross-scale indicators from inherently noisy, coarse-resolution, and heterogeneous datasets. (Zellweger et al., 2019; Cazalis et al., 2020; Antonelli et al., 2023, Trew et al., 2024).

To address this research gap, this study undertakes a two-part exploratory analysis. First in Case 1, we develop and test the “Cross-Scale Coarse Indicator” (CSCI), a family of indicators designed to derive robust phenological signals by integrating coarse-resolution reanalysis data (ERA5) with opportunistic citizen science records. This case applied the proposed workflow to four species using ERA5 reanalysis data at a 0.1° spatial resolution to examine the feasibility of deriving temperature-based phenological indicators at coarse temporal and spatial scales. The CSCI is guided by three premises to manage data limitations: (1) using coarse grids to provide spatial tolerance for positional uncertainty; (2) applying temporal smoothing to transform short-term noise into seasonal signals; and (3) incorporating biological relevance for interpretability. Second, in Case 2, we utilize historical gridded observation datasets at 0.05° and 0.01° resolutions to investigate the phenological indicators of an evergreen tree species endemic to subtropical islands. This complementary analysis assesses whether indicator trends remain consistent across different temporal and spatial scales, thereby identifying potential biases and reinforcing our overall goal of developing robust indicators for underrepresented species in data-limited regions.

2. Methods

2.1. The three ecological premises of Cross-Scale Coarse Indicator (CSCI)

In designing the indicators, it is essential to consider the intrinsic limitations of both the ERA5 reanalysis data and the citizen science records. The ERA5 reanalysis offers globally accessible and spatially consistent climate information, enabling the development of cross-regional and comparable indicators. However, its spatial resolution (0.1°) remains too coarse relative to individual plants. Meanwhile, citizen science data are not derived from standardized phenological monitoring protocols; uncertainties may arise from the timing of phenophase observation, the precision of geolocation, or even the delay between observation and upload to the platform. These mismatches between the spatial and temporal scales of climatic drivers and biological observations exemplify the long-recognized “problem of pattern and scale” in ecology (Levin, 1992). To comprehensively account for these sources of bias and uncertainty, we propose the Cross-Scale Coarse Indicator (CSCI). Here, we define CSCI as a family of exploratory, coarse-resolution and noisy data derived indicators that operate under three widely recognized ecological premises:

The first premises concerns the “Spatial Tolerance to Positional Uncertainty.” Both ERA5 and finer-scale grids suffer from a mismatch between climatic and organismal spatial scales, where coarse-resolution grids may provide a form of spatial tolerance to positional uncertainty, especially in heterogeneous terrains (Bock and Parracho, 2019). Similarly, the bias of geolocation precision is inherent to heterogeneous observational records. Even on structured citizen science platforms such as iNaturalist, the coordinate accuracy varied among each record. Under these circumstances, the coarse spatial grids may provide a form of spatial tolerance that accommodates such geographic uncertainty. The second premises is “Temporal Smoothing of Spatial Bias”. As noted above, coarse-resolution datasets such as ERA5 cannot capture the fine-scale microclimatic conditions experienced by individual organisms. Consequently, if indicators were designed at short temporal scales, such as daily intervals, the resulting mismatch between modeled and experienced temperature would likely be unacceptably large. To mitigate this issue, the temporal windows of CSCIs in this study were intentionally extended to month periods, aggregating temperature over longer intervals and transforming short-term noise into seasonal-scale signals. Last but not least, “Consideration of Biological Mechanisms” also imperative when designing CSCI. Although machine learning models often achieve high predictive performance, their limited interpretability can hinder ecological understanding and the transferability of results across systems (Pichler and Hartig, 2023). As purely statistical correlative models may have limited explanatory power under climate change, biological invasions, or novel environmental conditions, incorporating mechanistic considerations is essential for improving interpretability and transferability (Kearney and Porter, 2009). The rationale behind this consideration is not necessarily focused on the physiological mechanisms of a specific plant species, rather, it may serve as clues to overall environmental variations.

2.2. Case 1: exploratory phenological indicators based on ERA5 coarse-scale grids

In this case, we explored flowering responses to the seasonal thermal environment by selecting four tree species: *Crateva religiosa* (n = 92), *Cornus controversa* (n = 105), *Firmiana colorata* (n = 112), and *Helicteres isora* (n = 186), located in different zones and possessing distinct characteristics (Fig. 1). These species each possesses medicinal and cultural significance, yet they have rarely been discussed in the context of climate–phenology interactions. All the observation records were obtained from public accessible platforms, basically iNaturalist. Species-specific description and sample selection approach are providing in Supplementary Methods (see Supplementary Material).

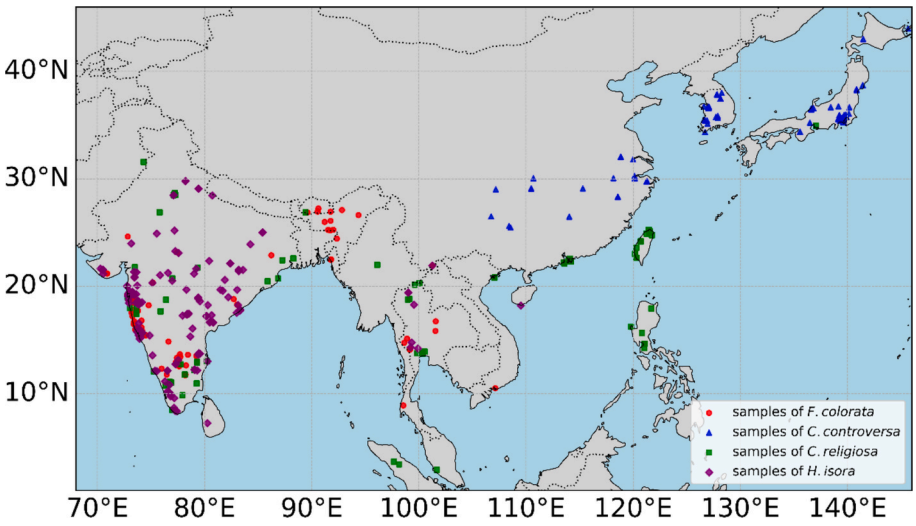


Fig. 1. Geographic distribution of citizen science records used in Case 1 of the four studied species across South, Southeast, and East Asia: *F. colorata* (red circles), *C. controversa* (blue triangles), *C. religiosa* (green squares), and *H. isora* (purple diamonds). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To characterize local thermal conditions preceding flowering, we obtained daily 2 m air temperature data from the ERA5-Land reanalysis product (Muñoz Sabater, 2019), which provides global coverage at a 0.1° spatial resolution (~11 km) in tropical areas. We then extracted the daily minimum and maximum temperatures for each georeferenced flowering record for growing degree days (GDD)-based variables calculation (McMaster and Wilhelm, 1997). In parallel, we calculated the 24-hour average temperature and used it to derive temperature threshold and standard deviations. Daily temperature values from surrounding grid points were used to estimate local thermal conditions via inverse distance weighting (IDW), a commonly used method for spatial interpolation of environmental indicators (Shepard, 1968). To account for missing data, we restricted the IDW procedure to valid neighboring points, thereby ensuring the reliability of interpolated temperature estimates.

We derived 61 biologically motivated CSCIs and grouped them into six categories (Table 1), each representing a distinct ecological aspect of thermal influence: (1) mean daily temperature, summarized across five fixed pre-flowering windows (90, 120, 180, 270, and 365 days); (2) low- and high-temperature thresholds, including the 1st and 10th percentiles of daily minima and maxima; (3) absolute deviations between mean temperatures and their corresponding low and high percentiles; (4) standard deviations, computed over fixed three-month windows to represent thermal variability; (5) cumulative growing degree days (GDD), the standard threshold-based approach that accumulates heat units above a defined base temperature of 15 °C on a daily basis (McMaster and Wilhelm, 1997); and (6) GDD residuals, defined as the difference between cumulative GDD and fixed thermal flowering thresholds ranging from 500 to 3500 °C-days (in 5-unit increments). For CSCIs involving GDD threshold values, we selected the optimal value for each record by identifying the one that maximized the coefficient of determination (R^2) in a simple regression, following established practice for retrospective calibration (McMaster and Wilhelm, 1997). While designing the indicators, we considered approaches previously used to establish indicators in past phenological and interdisciplinary studies. Although the same indicators may not capture identical phenomena in the context of this research, they remain highly valuable for enhancing the interpretability of the indicators and for planning future field validation (see Supplementary Information in Supplementary Material).

To identify the CSCIs most strongly associated with flowering records, we employed a Random Forest algorithm to rank the relative importance of all 61 CSCIs. This ensemble method is well suited to ecological datasets where non-linear relationships and correlated variables are common (Cutler et al., 2007). We then applied a heuristic selection approach, retaining the top quartile (15 indicators) to balance dimensionality reduction with explanatory coverage, which were subsequently used to fit a preliminary PLS regression model. Centering and scaling were applied to ensure that all predictors contributed equally to the covariance structure underpinning the PLS components. Variable importance in projection (VIP) scores was computed, and a threshold of

Table 1
Summary of the 61 temperature-base CSCIs used. Each category represents a distinct group of temperature-based or GDD-based indicators, computed over various time windows prior to flowering events.

Category	Description	Window (days)	Unit	Amount
Mean Temperature	Mean daily temperature over previous days	0–90, 0–120, 0–180, 0–270, 0–365	°C	5
Temperature Extremes	Percentile-based high and low temperature indicators	0–90, 0–120, 0–180, 0–270, 0–365	°C	20
Thermal Deviation	Absolute difference between mean and extreme temperatures	0–90, 0–120, 0–180, 0–270, 0–365	°C	20
Temperature Variability	Temperature standard deviation over 3 months durations	0–90, 90–180, 120–210, 180–270, 210–300, 270–365,	°C	6
GDD Sum	Cumulative growing degree days (base 15 °C) over previous days	0–90, 0–120, 0–180, 0–270, 0–365	°C-days	5
GDD Residual	Deviation from optimal GDD threshold over previous days	0–90, 0–120, 0–180, 0–270, 0–365	°C-days	5

VIP > 0.8 was applied to identify a reduced set. A refined PLS model was then fitted using only selected CSCIs. Model performance was assessed using R^2 and root mean square error (RMSE), both for the full models and under five-fold cross-validation. Besides, we applied correlation metrics to understand the association between the selected CSCIs. All analyses were conducted in Python 3.10 using the scikit-learn library (Pedregosa et al., 2011).

2.3. Case 2: deriving phenological indicators of Turpinia formosana using gridded temperature observation data

In this case, we used gridded temperature observation data to explore the temperature-based CSCIs of *Turpinia formosana* (Staphyleaceae), a subtropical, island-endemic evergreen tree species. *T. formosana* is an important subcanopy species in the evergreen broad-leaved forest communities, showing a stable occurrence across various plant associations in the Nan-Tze-Shian watershed (Chou et al., 2007) and other broad-leaved forest across Taiwan. Isolated components of this species exhibit significant osteogenic potential, suggesting their possible use as anti-osteoporotic agents (Imtiyaz et al., 2019). We collected 195 citizen-science flowering observation records from iNaturalist (mean day-of-year (DOY) = 105.39, median DOY = 106, SD = 14.49), and only those records containing identifiable flowers and valid geographic coordinates were included. These observations, spanning from April 2016 to May 2023, were concentrated between March and May (DOY 60–150) (Fig. 2). It is noteworthy that 128 observations (65.7 %) in this dataset were recorded on weekends or public holidays, reflecting a form of bias observed in citizen science data collection.

We used the Taiwan Climate Change Projection Information and Adaptation Knowledge Platform (TCCIP) Gridded Observation dataset (version V2.8) (Lin and Yang, 2025) for daily mean air temperature and its provided daily standard deviation, at 0.01° (~1 km; grid-center elevation) and 0.05° (~5 km; averaged from 0.01° cells) spatial resolutions, referenced to WGS84 (GCS_WGS_1984) and covering 1960–2023. This Gridded Observation dataset is based on historical weather station observational data. After quality control and a daily lapse-rate elevation adjustment, each 0.01° grid cell is estimated as a weighted average of nearby reference stations, with weights jointly reflecting distance, elevation difference via the daily lapse rate, and azimuth; the five highest-weight stations are used (Yang et al., 2024).

We compared four configurations of grid resolution and temporal scale of CSCIs: (i) Resolution: 0.01° (~1 km) vs 0.05° (~5 km) grids. Each observation was snapped to the nearest grid center (tolerance ≈ 0.005° and 0.025°, respectively). (ii) Temporal scale: three-month windows (90-days) vs one-month (30-days) cumulative windows computed over the 365 days prior to the flowering date. We directly

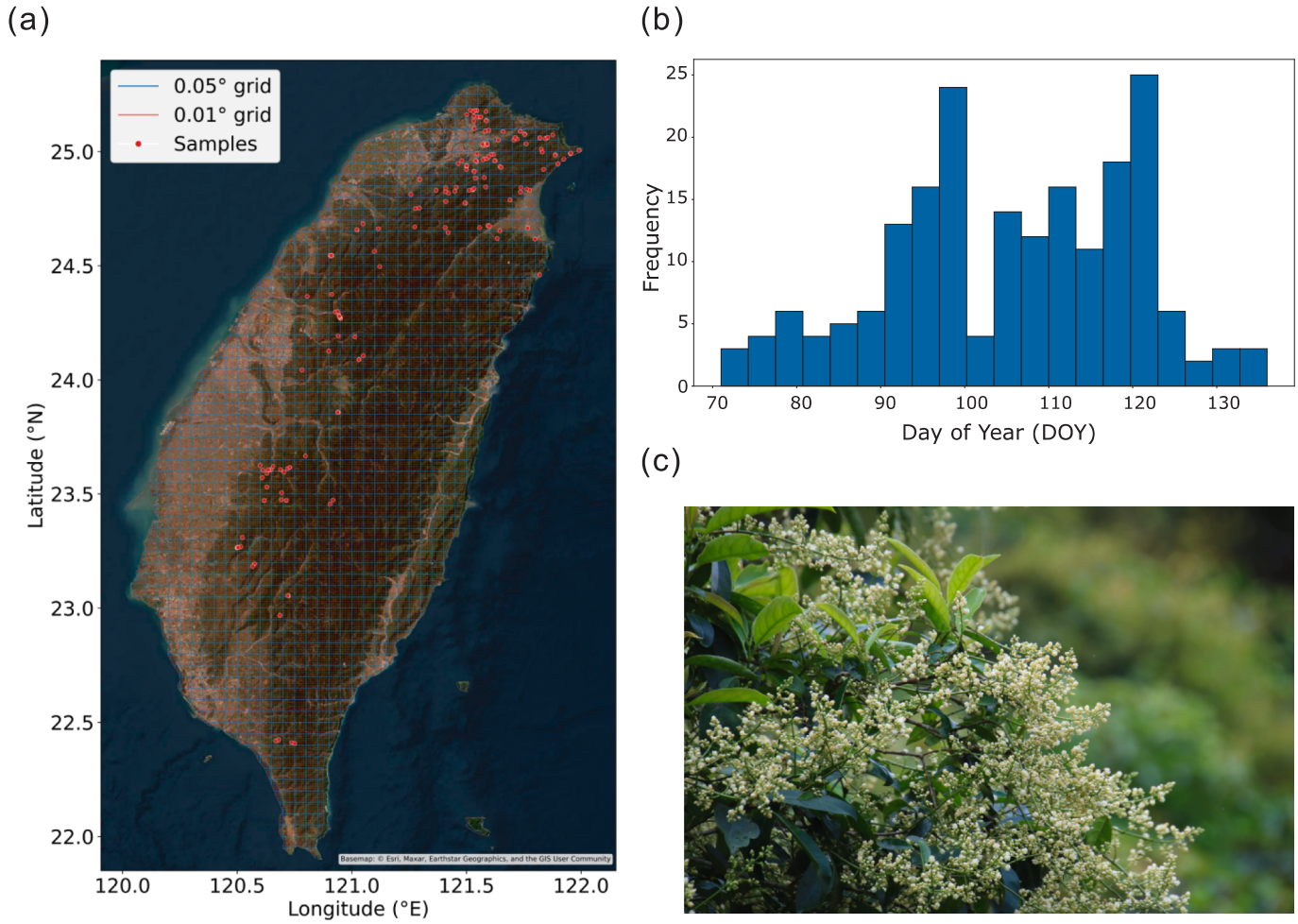


Fig. 2. (a) Distribution of *T. formosana* samples collected from iNaturalist in Case 2 (red circles). Blue lines on the map represent 0.05° grid cells (~5 km), while red lines indicate 0.01° grid cells (~1 km). This illustrates the scale mismatch between gridded climatic data and individual observation records. Basemap: Esri (2025), Maxar, Earthstar Geographics, and the GIS User Community (World Imagery). (b) Distribution of flowering dates for *T. formosana* samples, where the x-axis represents the Day of Year (DOY) and the y-axis indicates the frequency of sample occurrences. The histogram reveals a concentration trend in flowering records. (c) *T. formosana* producing small and densely arranged flowers during spring (April 2023, photo by first author). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

obtain the daily mean temperature and standard deviation from the gridded observation dataset and calculate the average value within the window. For each window we used daily mean temperature to derive low- and high-temperature thresholds. Since we aim to examine potential differences in CSCIs across different time windows, we introduced high and low temperature thresholds within specific temporal periods.

Based on the aforementioned premises, 120 1-month–scale CSCIs and 55 3-month–scale CSCIs were designed (Table 2). Our response variable was day-of-year (DOY) of flowering records. We first ran a Random Forest algorithm and retained the top quartile by importance. We then fit a preliminary PLS and computed VIP; variables with VIP > 0.8 form the refined PLS feature set. The number of PLS components was chosen by five-fold cross-validation, optimizing CV R^2 and reporting RMSE. Finally, each CSCI used in the refined PLS was also tested via univariate linear regression against DOY with R^2 and RMSE reported.

3. Results

3.1. Case 1: exploratory phenological indicators based on ERA5 coarse-scale grids

For all four species, the identical workflow was applied (Fig. 3). In the preliminary PLS model, 15 temperature CSCIs selected through Random Forest were employed (Fig. S1), yielding an R^2 value of 0.6631, 0.893, 0.4345, 0.6184 and an RMSE of 13.1, 4.8, 16.55 and 18.22 for *C. religiosa*, *C. controversa*, *F. colorata* and *H. isora*, respectively. Subsequently, by selecting variables with variable importance in projection (VIP) scores > 0.8, 10, 13, 12, and 12 CSCIs were retained, respectively (Fig. S2), and the refined PLS model achieved an R^2 of 0.6997, 0.8886, 0.4418, 0.6266 and a root mean square error (RMSE) of 12.37, 4.9, 16.44 and 18.02 days (Fig. S3). A five-fold cross-validation was implemented to confirm the robustness of the refined PLS model, yielding a mean R^2 of 0.6088, 0.8655, 0.235, 0.6054 and a mean RMSE of 12.85, 5.22, 17.63 and 18.29 across folds.

Through univariate regression analysis, the explanatory power of each individual CSCI for DOY was assessed. For *C. religiosa*, the five CSCIs with the highest explanatory power were identified. The CSCI with the greatest explanatory strength was std_temp_120_210 (R^2 = 0.565, RMSE = 14.89), followed by low10pct_temp_270d (R^2 = 0.562, RMSE = 14.94), low1pct_temp_270d (R^2 = 0.501, RMSE = 15.95), mean_temp_270d (R^2 = 0.400, RMSE = 17.48), and GDD_sum_0_270 (R^2 = 0.397, RMSE = 17.53). For *C. controversa*, the five CSCIs with the highest explanatory power were diff_abs_mean_vs_low10pct_180d (R^2 = 0.749, RMSE = 7.36), low10pct_temp_180d (R^2 = 0.676, RMSE = 8.36), low1pct_temp_180d (R^2 = 0.608, RMSE = 9.19), diff_abs_mean_vs_low1pct_180d (R^2 = 0.564, RMSE = 9.69), and diff_abs_mean_vs_high10pct_270d (R^2 = 0.482, RMSE = 10.57). Among the two species mentioned above, all CSCIs exhibited statistically significant relationships ($p < 0.0001$).

In contrast, for *F. colorata*, only 6 out of the 12 CSCIs reached statistical significance ($P < 0.05$). These CSCIs, ordered by their explanatory power (R^2), were as follows: diff_abs_mean_vs_low1pct_120d (R^2 = 0.133, RMSE = 20.49, p = 0.0001), high10pct_temp_365d (R^2 = 0.113, RMSE = 20.73, p = 0.0004), std_temp_90_180 (R^2 = 0.097, RMSE = 20.91, p = 0.0010), low1pct_temp_120d (R^2 = 0.091, RMSE = 20.98, p = 0.0014), diff_abs_mean_vs_high10pct_90d (R^2 = 0.073, RMSE = 21.18, p = 0.0044), and diff_abs_mean_vs_high1pct_90d (R^2 = 0.048, RMSE = 21.47, p = 0.021). Similarly, For *H. isora*, 11 out of the 12 CSCIs exhibited statistically significant relationships ($P < 0.05$). The five CSCIs with the highest explanatory power were as follows: std_temp_270_365 (R^2 = 0.361, RMSE = 23.59), low1pct_temp_90d (R^2 = 0.208, RMSE = 26.25), low10pct_temp_90d (R^2 = 0.207, RMSE = 26.26), mean_temp_120d (R^2 = 0.170, RMSE = 26.88), and diff_abs_mean_vs_high1pct_180d (R^2 = 0.127, RMSE = 27.55).

Based on the correlation matrix of the CSCIs selected above (Fig. S4, Fig. S5, Fig. S6, Fig. S7), the temperature-based CSCIs intercorrelations were identified through VIP screening. For *C. religiosa*, the two CSCIs with the highest explanatory power, std_temp_120_210 and low10pct_temp_270d, showed a strong negative correlation (r = –0.85). Moreover, these two CSCIs were moderately negatively correlated (r = –0.51) and strongly positively correlated (r = 0.78), respectively, with the third most explanatory CSCI, mean_temp_270d. Notably, although GDD_sum_0_270 represents an CSCI of cumulative thermal energy, it exhibited strong positive correlations with the low temperature threshold CSCIs low10pct_temp_270d and low1pct_temp_270d (r = 0.77 and 0.71, respectively), while showing a moderate negative correlation with the temperature variability indicator std_temp_120_210 (r = –0.50).

For *C. controversa*, the CSCI with the highest explanatory power, diff_abs_mean_vs_low10pct_180d, showed strong negative correlations with the second and third most explanatory CSCIs, low10pct_temp_180d

Table 2
Summary of the two temporal scale temperature-based CSCIs used. Each category represents a distinct group of temperature-based indicators, computed over different temporal windows prior to flowering events.

Category	Description	Unit	Scale	Window (days)	Amount
Mean Temperature	Mean daily temperature over previous days	°C	3-Month	0–90, 0–120, 0–180, 0–270, 0–365	5
			1-Month	0–30, 0–60, 0–90,0–120, 0–150, 0–180,0–210, 0–240, 0–270, 0–300, 0–330, 0–365	12
Temperature Extremes	Percentile-based high and low temperature indicators over specific durations	°C	3-Month	0–90, 90–180, 120–210, 180–270, 210–300, 270–365	24
			1-Month	0–30, 30–60, 60–90, 90–120, 120–150, 150–180, 180–210, 210–240, 240–270, 270–300, 300–330, 330–365	48
	Percentile-based high and low temperature indicators over previous days	°C	3-Month	0–90, 0–120, 0–180, 0–270, 0–365	20
			1-Month	0–30, 0–60, 0–90, 0–120, 0–150, 0–180, 0–210, 0–240, 0–270, 0–300, 0–330, 0–365	48
Temperature Variability	Temperature standard deviation	°C	3-Month	0–90, 90–180, 120–210, 180–270, 210–300, 270–365	6
			1-Month	0–30, 30–60, 60–90, 90–120, 120–150, 150–180. 180–210, 210–240, 240–270, 270–300, 300–330, 330–365	12

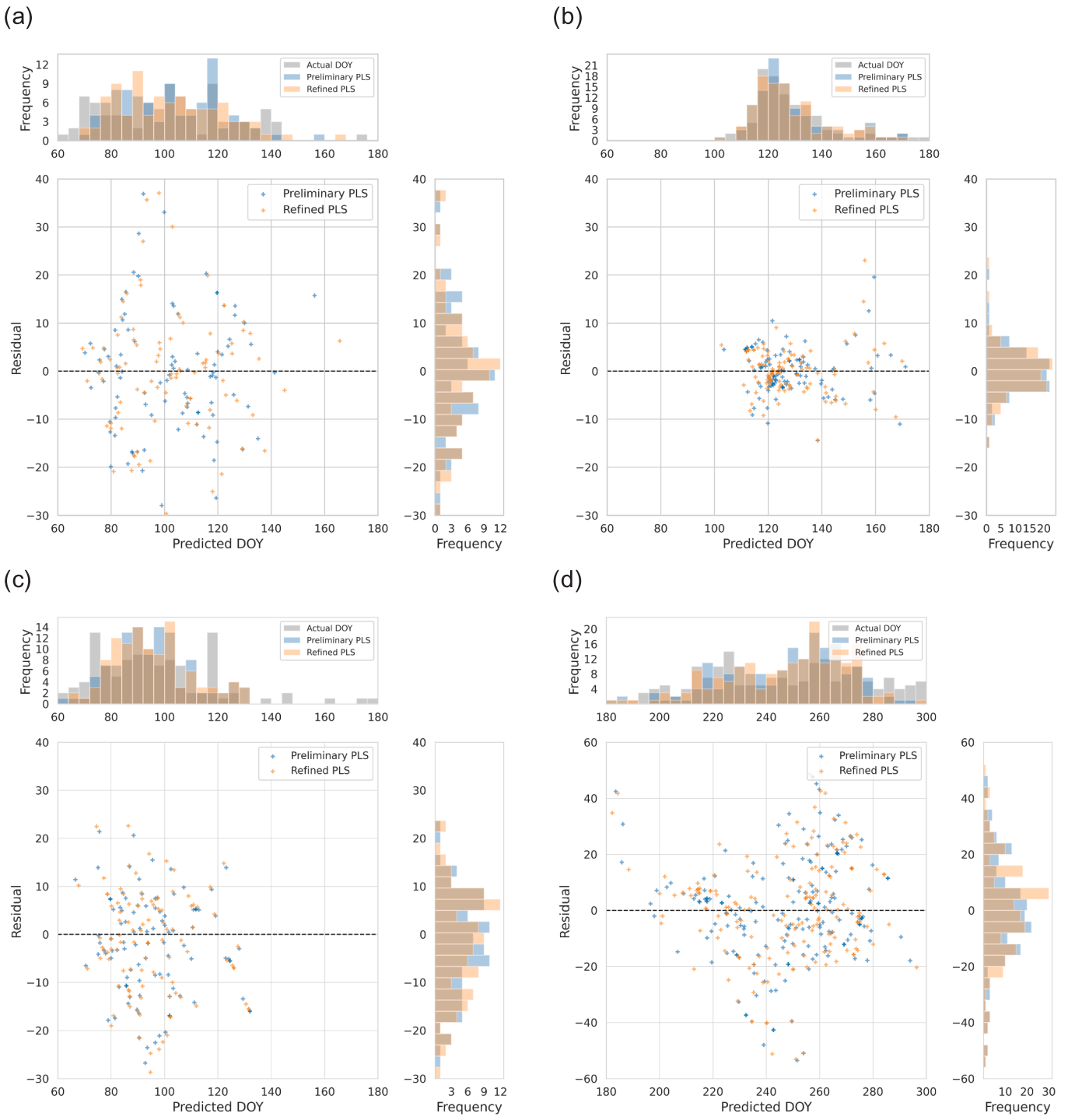


Fig. 3. Comparison of predictive performance between the preliminary and refined PLS models for (a) *C. religiosa* (mean actual DOY = 101.9; median actual DOY = 102; mean DOY predicted by preliminary PLS = 101.9; median DOY predicted by preliminary PLS = 102; mean DOY predicted by refined PLS = 101.9; median DOY predicted by refined PLS = 99), (b) *C. controversa* (mean actual DOY = 128.4; median actual DOY = 124; mean DOY predicted by preliminary PLS = 128.4; median DOY predicted by preliminary PLS = 125; mean DOY predicted by refined PLS = 128.4; median DOY predicted by refined PLS = 124), (c) *F. colorata* (mean actual DOY = 95.7; median actual DOY = 92; mean DOY predicted by preliminary PLS = 95.7; median DOY predicted by preliminary PLS = 94; mean DOY predicted by refined PLS = 95.7; median DOY predicted by refined PLS = 94) and (d) *H. isora* (mean actual DOY = 248.6; median actual DOY = 250; mean DOY predicted by preliminary PLS = 248.6; median DOY predicted by preliminary PLS = 252; mean DOY predicted by refined PLS = 248.6; median DOY predicted by refined PLS = 254). The central panel shows the residual distribution (observed – predicted DOY) for both models, illustrating the spread and bias of predictions. The right panel presents the frequency distribution of residuals, enabling a visual comparison of prediction errors across models. The upper panel displays the distribution of predicted DOY values from the two PLS models alongside the observed DOY distribution in the dataset.

and low1pct_temp_180d (r = –0.90 and –0.81, respectively). Notably, diff_abs_mean_vs_low10pct_180d exhibited strong positive correlations with other deviation indicators, diff_abs_mean_vs_high10pct_270d and diff_abs_mean_vs_high10pct_180d (r = 0.80 and 0.83, respectively). For *F. colorata*, most explanatory indicators exhibit moderate to high positive correlations (r ≈ 0.5–0.9), indicating strong linear associations among these temperature variables. For *H. isora*, the CSCI with the strongest explanatory power, std_temp_270_365, exhibited only weak correlations with other statistically significant CSCIs, low1pct_temp_90d, low10pct_temp_90d, mean_temp_120d, and high1pct_temp_180d (r = 0.10, 0.12, 0.10, and –0.14, respectively).

3.2. Case 2: deriving phenological indicators of Turpinia formosana using gridded temperature observation data

Across all configurations, both Random Forest selection and PLS modeling demonstrated predictive performance. For the 0.01°, 3-month configuration, 14 of 55 CSCIs were retained after Random Forest. The preliminary and refined PLS models (A = 6) both achieved high performance (in-sample R^2 = 0.954, RMSE = 3.11; CV R^2 = 0.943, RMSE = 3.44). The top five single-variable regressions showed R^2 values ranging from 0.61 (low10pct_temp_120_210) to 0.48 (high10pct_temp_120_210), all showing highly significant relationships (p < 0.0001) (Table 3).

For the 0.01°, 1-month configuration, 30 of 120 CSCIs were retained after Random Forest filtering, and 29 remained after applying the VIP > 0.8 criterion. The refined PLS model (A = 12) achieved strong performance (in-sample R^2 = 0.977, RMSE = 2.17; CV R^2 = 0.967, RMSE = 2.61). The top five single-variable regressions yielded R^2 values of 0.59–0.50, with low10pct_temp_120_150 (R^2 = 0.592, RMSE = 9.23) and low1pct_temp_180_210 (R^2 = 0.50, RMSE = 10.26), all showing significant relationships (p < 0.0001) (Table 3).

Similarly, for the 0.05° configurations, Random Forest retained 14 of 55 CSCIs for the 3-month setting and 30 of 120 CSCIs for the 1-month setting, with 25 features remaining after applying the VIP > 0.8 criterion. The refined PLS models achieved comparably high performance (3-month: in-sample R^2 = 0.956, RMSE = 3.02; CV R^2 = 0.946, RMSE = 3.35; 1-month: in-sample R^2 = 0.975, RMSE = 2.30; CV R^2 = 0.964, RMSE = 2.74). Among the top five single-variable regressions, low1pct_temp_120_210 showed the highest explanatory power for the 3-month configuration (R^2 = 0.544, RMSE = 9.76), while low10pct_temp_120_150 ranked highest for the 1-month configuration (R^2 = 0.559, RMSE = 9.59). All these relationships were highly significant (p < 0.0001) (Table 3).

Table 3
The top 5 CSCIs selected in Case 2 were ranked based on their R^2 from univariate regression analyses. In both the 0.01° and 0.05° grid groups, identical indicator combinations with the top 5 highest explanatory power (R^2) were identified for the two temporal scales. Although the RMSE values between the two groups showed slight differences, the regression slopes consistently reflected the same trend.

Grid	Scale	CSCI	R^2	RMSE	Slope
0.01° (~1km)	1-Month	low10pct_temp_120_150	0.592	9.236	−2.926
		low1pct_temp_120_150	0.5897	9.259	−2.693
		high10pct_temp_120_150	0.536	9.845	−3.489
		low10pct_temp_180_210	0.504	10.176	−3.161
		low1pct_temp_180_210	0.496	10.261	−3.102
	3-Months	low10pct_temp_120_210	0.613	8.993	−3.733
		low1pct_temp_120_210	0.583	9.333	−2.879
		low10pct_temp_180_270	0.49	10.32	−3.463
		low1pct_temp_180_270	0.489	10.328	−3.133
		high10pct_temp_120_210	0.447	10.79	−3.32
0.05° (~5km)	1-Month	low10pct_temp_120_150	0.559	9.594	−2.969
		low1pct_temp_120_150	0.558	9.614	−2.738
		high10pct_temp_120_150	0.486	10.364	−3.464
		low10pct_temp_180_210	0.451	10.714	−3.175
		low1pct_temp_180_210	0.424	10.97	−3.008
	3-Months	low1pct_temp_120_210	0.544	9.76	−2.905
		low10pct_temp_120_210	0.544	9.763	−3.657
		low10pct_temp_180_270	0.439	10.831	−3.522
		low1pct_temp_180_270	0.423	10.981	−3.092
		high10pct_temp_120_210	0.397	11.222	−3.337

4. Discussion

This study demonstrates that temperature-driven phenological indicators can be derived from citizen science records and coarse-resolution climate reanalysis data, even in the absence of formal long-term monitoring networks. Case 1 showed that exploratory Cross-Scale Coarse Indicators (CSCIs) achieved moderate to strong explanatory power for flowering observations across four deciduous species, particularly when phenological cues were closely aligned with seasonal temperature exposure. Case 2 further showed that even under coarse spatial resolution and weekend-biased data, a consistent set of CSCIs retained high predictive performance across multiple configurations. These patterns suggest the feasibility of detecting biologically relevant climatic signals from opportunistic and coarse data sources.

The refined PLS models in the Case 1 exhibited varying degrees of performance, and a similar pattern was observed for the individual CSCI. Although the CSCI cannot be directly linked to phenological mechanisms, these exploratory findings may offer valuable insights for future research. From the exploratory findings of Case 1, we can formulate hypotheses from several perspectives. First is the temperature sensitivity of the phenology of individual species. In this study, all CSCI were designed on a temperature basis, which suggest that the performance of the refined PLS model can serve as a preliminary reference. For temperate-subtropical tree species *C. controversa*, the refined PLS model demonstrates high explanatory power. For the subtropical–tropical deciduous tree *C. religiosa* and the sub-deciduous tree *H. isora*, the refined PLS models showed moderate performance, suggesting that, in addition to temperature, other factors also influence their phenological patterns. The phenology of tropical deciduous trees is closely associated with factors such as rainfall, temperature, and solar radiation, and their deciduous habit represents a key adaptive mechanism for coping with drought stress (Elliott et al., 2006; Singh and Kushwaha, 2016). Therefore, relying solely on temperature-based considerations can be inherently biased. This may also explain the relatively low explanatory power of the refined PLS model for the tropical deciduous species *F. colorata*. Second, these exploratory results also offer critical insights into the key temporal windows that shape the phenological responses of individual species. For instance, temperature exposure within 270-day and localized warming occurring approximately 4–7 months prior to flowering may represent a critical window in the phenological cycle of *C. religiosa*. In addition, these CSCIs should be interpreted with caution, as they may not purely reflect temperature signals but could also incorporate influences from other environmental or ecological factors. In semi-arid systems, temperature seasonality and the timing of evaporative demand may interact with rainfall patterns to influence vegetation dynamics, rather than acting independently (Deblauwe et al., 2008). This suggests that some temperature-based CSCIs may also respond to broader climatic variability, such as changes in aridity and diurnal temperature range (Vinnarasi et al., 2017), rather than to seasonal temperature alone.

In a correlation matrix, two variables with lower correlation provide information from distinct dimensions. The correlations among CSCIs can provide researchers with an initial understanding of the temperature-driven factors involved. For instance, a strong positive or negative correlation was observed among the major CSCIs of *C. religiosa* and *C. controversa*, indicating that their thermal response is primarily governed by the same set of temperature conditions. For *C. controversa*, the most influential thermal deviation CSCI exhibits a strong negative correlation with low-temperature threshold and a positive correlation with high-temperature deviation. This supports the idea that the climatic

suitability of this species is shaped by the magnitude of low-temperature deviations. For *F. colorata*, the thermal deviation CSCIs were highly intercorrelated, supporting that they represent similar aspects of temperature departure from the mean. In contrast, for *H. isora*, the primary standard deviation-based CSCI shows weak correlation coefficients with other indicators, suggesting a more diffuse or multi-dimensional climatic response rather than sensitivity to a single temperature characteristic.

Although we could not directly link CSCI to specific physiological mechanisms, these exploratory results are sufficient to support hypothesis generation, which can then be tested through field studies. This is particularly valuable in tropical-subtropical regions where long-term biodiversity monitoring networks are lacking, as it can substantially reduce the high costs associated with preliminary trend observations and data collection. The method demonstrated in our workflow employs a linear approach to capture the relationship between CSCI and phenological observations. However, plant phenological responses may also exhibit nonlinear relationships with temperature. As such, the use of linear indicators may underestimate certain phenological responses (Jochner et al., 2016; Meng et al., 2016). In addition, species exhibiting multiple phenological events within a year or extended phenological cycles cannot be directly applied to this workflow. The CSCI and response variables should be tailored to the specific traits of each species, and appropriate statistical methods should be selected accordingly.

In Case 2, the refined PLS model exhibited unusually high performance, with cross-validated R^2 values exceeding 0.94 across all four configurations. However, the five-fold cross-validation results also ruled out the possibility of overfitting in the refined PLS model. Among the four configurations, the top five CSCIs with the highest explanatory power were identical, while the 3-month CSCI encompasses the temporal window of the 1-month CSCI, indicating that the workflow indeed captured the trends of phenological signals rather than random statistical patterns. Therefore, it can be inferred that the flowering phenology of *T. formosana* is primarily driven by seasonal temperature, and robust signals can still be extracted even under conditions of coarse resolution and pronounced weekend bias. Among the CSCIs identified for *T. formosana*, the top five in explanatory power are all associated with temperature thresholds, particularly temperature exposure during the periods 4–5 months before flowering (DOY 120–150) and 6–7 months before flowering (DOY 180–210). This exploratory finding provides valuable insights for formulating hypotheses in future field studies.

While CSCIs offer a practical screening tool, their interpretability is constrained by several key limitations. First, low RMSE values may not always reflect high model quality, particularly given the limited information on full flowering durations. In our case, the RMSE values of both the PLS model and the single-variable regression model were smaller than the standard deviation of the original collected samples, with the former exhibiting an even lower RMSE than the latter. Considering that citizen science observations may concentrate around the peak flowering period rather than evenly distributed throughout the entire flowering phase (Canavan et al., 2025), having an RMSE smaller than the actual range can be regarded as acceptable. However, particular caution is required when interpreting such unusually small RMSE values. The most reliable means of validation remains field monitoring networks, which can help determine reasonable flowering durations and assess the prediction error of the CSCI. Second, coarse-resolution reanalysis data might not capture microclimatic variability experienced by local species. Empirical evidence from temperate forests shows that understory microclimates differ substantially from macroclimatic conditions, and such discrepancies can influence plant responses to climate change, leading to apparent climatic lags (De Frenne et al., 2013). Recent studies have also confirmed that microclimatic changes induced by urbanization, such as the urban heat island effect and anthropogenic disturbances, may influence plant flowering phenology and interpopulation synchrony (Fujiwara et al., 2025). These findings are consistent with regional-scale modeling studies showing that built-environment heat accumulation and surface energy imbalance substantially modify near-surface thermal regimes (Biswas et al., 2022; Patil and Surawar, 2023; Sadat et al., 2024). Furthermore, habitat and topography can generate microclimatic temperature variations of a magnitude comparable to projected global warming, meaning that species distributions are often determined by fine-scale microclimatic conditions (Suggitt et al., 2011). Therefore, when interpreting the CSCI, it should be regarded as reflecting seasonal climatic trends rather than being directly linked to the microclimate actually experienced by organisms. Another major limitation of our workflow is that it reconstructs phenological trends based on data from a specific period of the past. However, this approach cannot capture the profound consequences that may arise under scenarios of rapid and intensified climate change. Once an ecosystem crosses a tipping point, it may undergo abrupt and potentially irreversible changes in its state and functioning (Scheffer et al., 2001; Lenton et al., 2008). Therefore, relying solely on linear predictions based on CSCI may lead to an overly optimistic outlook on climate change impacts. It should be regarded as a short- to medium-term reference indicator, rather than a projection tool for long-term climatic scenarios.

Recent work using Sentinel-2 imagery has shown that satellite-derived phenology metrics can complement ground-based networks and provide species-specific insights in temperature area at high spatial resolution (Koch et al., 2025). Studies in tropical America have also shown that higher-resolution remote sensing data can better capture phenological differences among individual tree crowns (Song et al., 2024). Furthermore, in temperate ecosystems, remote sensing data have been integrated with plant physiological processes to simulate seasonal NDVI variations and capture regional phenological patterns (Bajocco et al., 2025). In contrast, the present study extends this integrative perspective to data-limited tropical-subtropical systems, leveraging coarse-scale reanalysis datasets and opportunistic citizen science records. While not requiring high-resolution data, the exploratory indicators still balance interpretability and comparability. In parallel, ERA5 reanalysis data has been applied to research in plant phenology, with a particular focus on agricultural science. For example, a study conducted in Brazil demonstrated that the ERA5-Land reanalysis data can serve as a viable substitute for meteorological station records and be utilized for climatic risk zoning analyses of maize production (Matsunaga et al., 2024). Researchers in Latvia combined long-term (1959–2019) phenological records from the Pūre apple orchard with E-OBS and ERA5 gridded data to investigate the effects of climate change on the flowering period of apple trees (Kalvāne et al., 2021). In addition, the ERA5 dataset can be used as the temperature background information in remote sensing algorithms to predict rice-crop intensity in China (He et al., 2024).

In this study, we propose a novel approach to utilizing ERA5 reanalysis data for phenological assessment. The focus of phenological research is expanded beyond specific agricultural crops to include wildlife species. Unlike traditional long-term monitoring networks, our approach relies on opportunistic observations that can be repurposed for phenological inference. Although the discrepancy between the proposed CSCI and field-based phenological observations remains to be validated, the comparability and scalability of indicators constructed from globally consistent ERA5 reanalysis data offer substantial potential for practical applications. Future research should focus on the field validation of exploratory CSCI, for instance by integrating long-term biodiversity monitoring networks to evaluate and correct potential biases in the CSCI. In addition, other phenological drivers beyond temperature should be taken into consideration, particularly in tropical-subtropical systems. For example, reduced flower production under warmer and more humid atmospheric conditions in Amazonian forests (Vleminckx et al., 2024), and the dependence of Southeast Asian mass flowering events on the combined occurrence of drought and low temperature cues (Numata et al., 2022). In addition, regional climatic interactions between temperature and precipitation have been well documented in South Asian tropical–subtropical cities, where heat island intensification

alters rainfall formation and local humidity regimes (Patil and Surawar, 2023), suggesting that similar processes may influence flowering cues in semi-urban or fragmented habitats. These findings suggest that rainfall, irradiance, and other interacting factors can introduce gaps between statistical prediction and actual flowering onset. Even with these caveats, our results demonstrate that exploratory signal detection can provide a feasible alternative where fundamental monitoring data are scarce, offering an entry point for biodiversity monitoring and climate-change impact assessment at a time when most tropical phenological responses remain poorly understood.

5. Conclusion

In summary, this study highlights how exploratory phenological indicators can be developed from open-access ERA5 climate reanalysis data and citizen science records. By tolerating spatial uncertainty and temporally aggregating temperature signals, the Cross-Scale Coarse Indicator (CSCI) provides a novel approach to screening phenological patterns in species lacking long-term observational data. While future validation is essential, especially in tropical-subtropical systems, this approach may help address a persistent data gap in global change biology. By extending phenological research beyond traditional monitoring networks, it offers a flexible, scalable, and cost-efficient strategy for biodiversity tracking in a changing climate.

Ethics statement

The authors declare no ethical issues rising from this work.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

CRediT authorship contribution statement

Yu-Shiang Huang: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Chieh-Hsiang Fan:** Software, Investigation, Data curation. **Yi-Cian Lu:** Supervision, Project administration. **Yi-Ta Wu:** Project administration. **Yu-Cheng Niu:** Project administration. **Yi-Huan Hsieh:** Supervision, Conceptualization. **Yu-Kai Liao:** Supervision, Funding acquisition, Conceptualization. **Syuan-Jyun Sun:** Writing – original draft, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Y.-K. L. and S.-J.S. were supported by NTU New Faculty Founding Research Grant. S.-J.S. was supported by National Science and Technology Council 2030 Cross-Generation Young Scholars Program (112-2628-B-002-013-; 113-2628-B-002-028-; 114-2628-B-002-027-), and Yushan Fellow Program provided by the Ministry of Education (MOE-111-YSFAG-0003-002-P1). Y.-C. L. was supported by National Science and Technology Council Graduate Research Fellowship (114WFA0112693).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2025.114484.

Data availability

The data and codes are openly available at https://github.com/brainbrian2000/Temperature_based_phenological_indicator_under_data_limited_condiction_ERA5_and_citizen_science.

References

Antonelli, A., Dhanjal-Adams, K.L., Silvestro, D., 2023. Integrating machine learning, remote sensing and citizen science to create an early warning system for biodiversity. Plants People Planet 5 (3), 307–316. https://doi.org/10.1002/ppp3.10337.

Askeyev, O., Askeyev, A., Askeyev, I., Sparks, T., 2022. Extreme temperatures help in identifying thresholds in phenological responses. Glob. Ecol. Biogeogr. 31 (2), 321–331. https://doi.org/10.1111/geb.13430.

Badeck, F.W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., Sitch, S., 2004. Responses of spring phenology to climate change. New Phytol. 162 (2), 295–309. https://doi.org/10.1111/j.1469-8137.2004.01059.x.

Bajocco, S., Ricotta, C., Bregaglio, S., 2025. Bridging the gap between remotely sensed phenology and the underlying ecophysiological processes: the SWELL model. Methods Ecol. Evol. https://doi.org/10.1111/2041-210X.70067.

Baker, E., Drury, J.P., Judge, J., Roy, D.B., Smith, G.C., Stephens, P.A., 2021. The verification of ecological citizen science data: current approaches and future possibilities. Citiz. Sci.: Theory Pract. 6 (1). https://doi.org/10.1002/ppp3.10337.

Beil, I., Kreyling, J., Meyer, C., Lemcke, N., Malyshev, A.V., 2021. Late to bed, late to rise—warmer autumn temperatures delay spring phenology by delaying dormancy. Glob. Chang. Biol. 27 (22), 5806–5817. https://doi.org/10.1111/gcb.15858.

Bhattacharjee, S., Ghosh, C., Sen, A., Lala, M., 2023. Characterization of *Firmiana colorata* (Roxb.) R. Br. leaf extract and its silver nanoparticles reveal their antioxidative, anti-microbial, and anti-inflammatory properties. Int. Nano Lett. 13 (3), 235–247. https://doi.org/10.1007/s40089-023-00392-6.

Biswas, M.H.A., Dey, P.R., Islam, M.S., Mandal, S., 2022. Mathematical model applied to green building concept for sustainable cities under climate change. J. Contemp. Urban Aff. 6 (1), 36–50. https://doi.org/10.25034/ijcua.2022.v6n1-4.

Bock, O., Parracho, A.C., 2019. Consistency and representativeness of integrated water vapour from ground-based GPS observations and ERA-Interim reanalysis. Atmos. Chem. Phys. 19 (14), 9453–9468. https://doi.org/10.5194/acp-19-9453-2019.

Carrascal, L.M., Galván, I., Gordo, O., 2009. Partial least squares regression as an alternative to current regression methods used in ecology. Oikos 118 (5), 681–690. https://doi.org/10.1111/j.1600-0706.2008.16881.x.

Canavan, S., Rodríguez, J., Gervazoni, P., Pipek, P., Le Roux, J.J., Castillo, M.L., Novoa, A., 2025. iEcology reveals the importance of geography and genetic makeup in the flowering phenology of invasive Carpobrotus taxa. Ecol. Solutions Evidence 6 (4), e70122. https://doi.org/10.1002/2688-8319.70122.

Cazalis, V., Princé, K., Mihoub, J.B., Kelly, J., Butchart, S.H., Rodrigues, A.S., 2020. Effectiveness of protected areas in conserving tropical forest birds. Nat. Commun. 11 (1), 4461. https://doi.org/10.1038/s41467-020-18230-0.

Chang-Yang, C.H., Chiang, P.H., Wright, S.J., Hsieh, C.F., Sun, I.F., 2024. Proximate cues of flowering in a subtropical rain forest. Biotropica 56 (1), 78–89. https://doi.org/10.1111/btp.13282.

Chavre, B., 2024. Helicteres isora L.: a traditional medicinal plant. Int. J. Plant Environ. 10 (03), 179–183. https://doi.org/10.18811/ijpen.v10i03.23.

Yang, C.-T., Wang, C.-Y., Lin, S.-Y., 2024. Gridded observation temperature V2 data production history. Taiwan Climate Change Projection Information and Adaptation Knowledge Platform (TCCIP) (Accessed on 12-10-2025) Retrieved from the TCCIP Website https://tccip.ncdr.nat.gov.tw/upload/data_profile/20220706114121.pdf.

Chou, F.S., Liao, C.K., Yang, Y.P., 2007. Classification and ordination of evergreen broad-leaved forest in the middle and upper watershed of the Nan-Tze-Shian Stream in southwestern Taiwan. Taiwania 52 (2), 127–144. https://taiwania.ntu.edu.tw/pdf/tai.2007.52.127.pdf.

Chuine, I., Régnière, J., 2017. Process-based models of phenology for plants and animals. Annu. Rev. Ecol. Evol. Syst. 48 (1), 159–182. https://doi.org/10.1146/annurev-ecolsys-110316-022706.

Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A., Schwartz, M.D., 2007. Shifting plant phenology in response to global change. Trends Ecol. Evol. 22 (7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003.

Courter, J.R., Johnson, R.J., Stuyck, C.M., Lang, B.A., Kaiser, E.W., 2013. Weekend bias in citizen science data reporting: Implications for phenology studies. Int. J. Biometeorol. 57, 715–720. https://doi.org/10.1007/s00484-012-0598-7.

Cutler, D.R., Edwards Jr, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88 (11), 2783–2792. https://doi.org/10.1890/07-0539.1.

Davis, C.C., Lyra, G.M., Park, D.S., Asprino, R., Maruyama, R., Torquato, D., Ellison, A. M., 2022. New directions in tropical phenology. Trends Ecol. Evol. 37 (8), 683–693. https://www.cell.com/trends/ecology-evolution/abstract/S0169-5347(22)00112-4.

De Frenne, P., Rodríguez-Sánchez, F., Coomes, D.A., Baeten, L., Verstraeten, G., Vellend, M., Verheyen, K., 2013. Microclimate moderates plant responses to

macroclimate warming. Proc. Natl. Acad. Sci. 110 (46), 18561–18565. https://doi.org/10.1073/pnas.1311190110.
Deblauwe, V., Barbier, N., Couteron, P., Lejeune, O., Bogaert, J., 2008. The global biogeography of semi-arid periodic vegetation patterns. Glob. Ecol. Biogeogr. 17 (6), 715–723. https://doi.org/10.1111/j.1466-8238.2008.00413.x.
Elliott, S., Baker, P.J., Borchert, R., 2006. Leaf flushing during the dry season: the paradox of Asian monsoon forests. Glob. Ecol. Biogeogr. 15 (3), 248–257. https://doi.org/10.1111/j.1466-8238.2006.00213.x.
Esri, 2025. World Imagery [Basemap]. ArcGIS Online. (Accessed 2025-11-24).
Field, S.A., O'connor, P.J., Tyre, A.J., Possingham, H.P., 2007. Making monitoring meaningful. Austral. Ecol. 32 (5), 485–491. https://doi.org/10.1111/j.1442-9993.2007.01715.x.
Fu, Y., Zhang, H., Dong, W., Yuan, W., 2014. Comparison of phenology models for predicting the onset of growing season over the Northern Hemisphere. PLoS One 9 (10), e109544. https://doi.org/10.1371/journal.pone.0109544.
Fujiwara, H., Yamaguchi, H., Nakata, K., Katsuhara, K.R., 2025. Urbanised landscape and microhabitat differences can influence flowering phenology and synchrony in an annual herb. J. Appl. Ecol. https://doi.org/10.1111/1365-2664.70159.
Gilman, E.F., Watson, D.G., 1993. Cornus controversa: Giant Dogwood (Fact Sheet ST-183). Environmental Horticulture Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. https://hort.ifas.ufl.edu/trees/CORCONA.pdf.
He, Z., Li, S., Chang, M., Liu, Y., Liu, K., Wan, L., Wang, Y., 2024. Novel harmonic-based scheme for mapping rice-crop intensity at a large scale using time-series Sentinel-1 and ERA5-land datasets. IEEE Trans. Geosci. Remote Sens. 62, 1–23. https://doi.org/10.1109/TGRS.2024.3387559.
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Thépaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorolog. Soc. 146 (730), 1999–2049. https://doi.org/10.1002/qj.3803.
Imtiyaz, Z., Wang, Y.F., Lin, Y.T., Liu, H.K., Lee, M.H., 2019. Isolated compounds from turpinia formosana nakai induce ossification. Int. J. Mol. Sci. 20 (13), 3119. https://doi.org/10.3390/ijms20133119.
Ismaeel, A., Tai, A.P., Santos, E.G., Maraia, H., Aalto, I., Altman, J., Maeda, E.E., 2024. Patterns of tropical forest understory temperatures. Nat. Commun. 15 (1), 549. https://doi.org/10.1038/s41467-024-44734-0.
Jacobs, M., 1964. The genus *Crateva* (Capparaceae). Blumea: Biodivers. Evol. Biogeogr. Plants 12 (2), 177–208.
Jochner, S., Sparks, T.H., Laube, J., Menzel, A., 2016. Can we detect a nonlinear response to temperature in European plant phenology? Int. J. Biometeorol. 60 (10), 1551–1561. https://doi.org/10.1007/s00484-016-1146-7.
Kalvāne, G., Gribuste, Z., Kalvāns, A., 2021. Full flowering phenology of apple tree (Malus domestica) in Pūre orchard, Latvia from 1959 to 2019. Adv. Sci. Res. 18, 93–97. https://doi.org/10.5194/asr-18-93-2021.
Kearney, M., Porter, W., 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecol. Lett. 12 (4), 334–350. https://doi.org/10.1111/j.1461-0248.2008.01277.x.
Klinger, Y.P., Eckstein, R.L., Kleinebecker, T., 2023. iPhenology: using open-access citizen science photos to track phenology at continental scale. Methods Ecol. Evol. 14 (6), 1424–1431. https://doi.org/10.1111/2041-210X.14114.
Koch, T.L., Grubinger, S., Coops, N.C., Damm, A., Morsdorf, F., Waser, L.T., Hobi, M.L., 2025. Assessment of tree species specific phenology metrics from Sentinel-2 data to complement in situ monitoring. Ecol. Ind. 180, 114299. https://doi.org/10.1016/j.ecolind.2025.114299.
Kumar, N., Singh, A.K., 2014. Plant profile, phytochemistry and pharmacology of Avartani (Helicteres isora Linn.): a review. Asian Pac. J. Trop. Biomed. 4, S22–S26. https://doi.org/10.12980/APJTB.4.2014C872.
Lee, D., Kang, S.J., Lee, S.H., Ro, J., Lee, K., Kinghorn, A.D., 2000. Phenolic compounds from the leaves of Cornus controversa. Phytochemistry 53 (3), 405–407. https://doi.org/10.1016/S0031-9422(99)00502-6.
Legg, C.J., Nagy, L., 2006. Why most conservation monitoring is, but need not be, a waste of time. J. Environ. Manage. 78 (2), 194–199. https://doi.org/10.1016/j.jenvman.2005.04.016.
Lenton, T.M., Held, H., Kriegler, E., Hall, J.W., Lucht, W., Rahmstorf, S., Schellnhuber, H. J., 2008. Tipping elements in the Earth's climate system. Proc. Natl. Acad. Sci. 105 (6), 1786–1793. https://doi.org/10.1073/pnas.0705414105.
Levin, S.A., 1992. The problem of pattern and scale in ecology: the Robert H MacArthur award lecture. Ecology 73 (6), 1943–1967. https://doi.org/10.2307/1941447.
Lindenmayer, D.B., Likens, G.E., 2010. The science and application of ecological monitoring. Biol. Conserv. 143 (6), 1317–1328. https://doi.org/10.1016/j.biocon.2010.02.013.
Marcis, P., Vido, J., Kurjak, D., Lestianska, A., Poltak, D., Rybar, J., Bosela, M., 2025. A combined effect of heat and drought limits the growth of central European silver fir. Agric. For. Meteorol. 371, 110610. https://doi.org/10.1016/j.agrformet.2025.110610.
Matsunaga, W.K., Sales, E.S.G., Júnior, G.C.A., Silva, M.T., Lacerda, F.F., de Paiva Lima, E., de Brito, J.I.B., 2024. Application of ERA5-land reanalysis data in zoning of climate risk for corn in the state of Bahia—Brazil. Theor. Appl. Climatol. 155 (2), 945–963. https://doi.org/10.1007/s00704-023-04670-3.
McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87 (4), 291–300. https://doi.org/10.1016/S0168-1923(97)00027-0.
Meenaloshini, P., Thirumurugan, A., Mohan, P.K., Kumar, T.S., 2024. In vitro anticancer efficacy of hydroalcoholic extract from *Crateva religiosa* G. Forst. bark on human ovarian cancer cells. J. Herbal Med. 45, 100870. https://doi.org/10.1016/j.hermed.2024.100870.
Meng, F., Zhou, Y., Wang, S., Duan, J., Zhang, Z., Niu, H., Inouye, D.W., 2016. Temperature sensitivity thresholds to warming and cooling in phenophases of alpine plants. Clim. Change 139 (3), 579–590. https://doi.org/10.1007/s10584-016-1802-2.
Muñoz Sabater, J., 2019: ERA5-land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.e2161bac (Accessed on 18-10-2025).
Muthukumar, M., Kumar, T.S., Rao, M.V., 2017. Phenology and seed germination of the Indian screw tree Helicteres isora L (malvales: Malvaceae). J. Threat. Taxa 9 (12), 11040–11044. https://doi.org/10.11609/jott.3058.9.12.11040-11044.
Nairn, J.R., Fawcett, R.J., 2015. The excess heat factor: a metric for heatwave intensity and its use in classifying heatwave severity. Int. J. Environ. Res. Public Health 12 (1), 227–253. https://doi.org/10.3390/ijerph120100227.
Numata, S., Yamaguchi, K., Shimizu, M., Sakurai, G., Morimoto, A., Alias, N., Satake, A., 2022. Impacts of climate change on reproductive phenology in tropical rainforests of Southeast Asia. Commun. Biol. 5 (1), 311. https://doi.org/10.1038/s42003-022-03245-8.
Oliveira, A., Lopes, A., Soares, A., 2022. Excess heat factor climatology, trends, and exposure across European functional urban areas. Weather Clim. Extremes 36, 100455. https://doi.org/10.1016/j.wace.2022.100455.
Panchen, Z.A., Primack, R.B., Aniśko, T., Lyons, R.E., 2012. Herbarium specimens, photographs, and field observations show Philadelphia area plants are responding to climate change. Am. J. Bot. 99 (4), 751–756. https://doi.org/10.3732/ajb.1100198.
Patil, R., Surawar, M., 2023. Urban heat island impact and precipitation patterns in Indian Western Coastal Cities. J. Contemp. Urban Aff. 7 (2). https://doi.org/10.25034/ijcua.2023.v7n2-3.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830. https://doi.org/10.48550/arXiv.1201.0490.
Pichler, M., Hartig, F., 2023. Machine learning and deep learning—a review for ecologists. Methods Ecol. Evol. 14 (4), 994–1016. https://doi.org/10.1111/2041-210X.14061.
Sadat, N., Sheikh, H., Asaduzzaman, M., 2024. Evaluating urban heat island mitigation strategies in Rajshahi, using ENVI-Met: a remote sensing approach. J. Contemp. Urban Aff. 8 (2), 546–556. https://doi.org/10.25034/ijcua.2024.v8n2-15.
Santisuk, T., Larsen, K., (Eds.), 2001. *Firmiana* (Roxb.) R. Br. In *Flora of Thailand* (Vol. 7, Part 3, p. 557). Bangkok: The Forest Herbarium, Royal Forest Department.
Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. Nature 413 (6856), 591–596. https://doi.org/10.1038/35098000.
Sharma, S.B., Rana, A., Chauhan, S.V.S., 2006. Reproductive biology of *Crataeva religiosa* forst. Curr. Sci. 716–720. https://www.jstor.org/stable/24089121.
Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM National Conference, pp. 517–524. https://doi.org/10.1145/800186.810616.
Lin, S.-Y., Yang, C.-T., 2025. Gridded observations — data description (version V2.8). Taiwan Climate Change Projection Information and Adaptation Knowledge Platform (TCCIP). Retrieved from the TCCIP Website https://tccip.ncdr.nat.gov.tw/upload/data_document/20220708162241.pdf, [accessed on 12-10-2025].
Singh, K.P., Kushwaha, C.P., 2016. Deciduousness in tropical trees and its potential as indicator of climate change: a review. Ecol. Ind. 69, 699–706. https://doi.org/10.1016/j.ecolind.2016.04.011.
Song, G., Wang, J., Zhao, Y., Yang, D., Lee, C.K., Guo, Z., Wu, J., 2024. Scale matters: spatial resolution impacts tropical leaf phenology characterized by multi-source satellite remote sensing with an ecological-constrained deep learning model. Remote Sens. Environ. 304, 114027. https://doi.org/10.1016/j.rse.2024.114027.
Srinivas, V., Surendra, G., Anjana, M., Kiran, A.S., 2018. A scientific review on *Crateva religiosa*. Int. J. Adv. Pharmaceu. Biol. Chem. 7 (1), 11–16. https://www.ijapbc.com/files/23-05-18/02.pdf.
Suggitt, A.J., Gillingham, P.K., Hill, J.K., Huntley, B., Kunin, W.E., Roy, D.B., Thomas, C. D., 2011. Habitat microclimates drive fine-scale variation in extreme temperatures. Oikos 120 (1), 1–8. https://doi.org/10.1111/j.1600-0706.2010.18270.x.
Taylor, S.D., Meiners, J.M., Riemer, K., Orr, M.C., White, E.P., 2019. Comparison of large-scale citizen science data and long-term study data for phenology modeling. Ecology 100 (2), e02568. https://doi.org/10.1002/ecy.2568.
Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S., Wanless, S., 2016. Phenological sensitivity to climate across taxa and trophic levels. Nature 535 (7611), 241–245. https://doi.org/10.1038/nature18608.
Trew, B.T., Edwards, D.P., Lees, A.C., Klinges, D.H., Early, R., Svátek, M., Maclean, I.M., 2024. Novel temperatures are already widespread beneath the world's tropical forest canopies. Nat. Clim. Chang. 14 (7), 753–759. https://doi.org/10.1038/s41558-024-02031-0.
Vareed, S.K., Reddy, M.K., Schutzki, R.E., Nair, M.G., 2006. Anthocyanins in Cornus alternifolia, Cornus controversa, Cornus kousa and Cornus florida fruits with health benefits. Life Sci. 78 (7), 777–784. https://doi.org/10.1016/j.lfs.2005.05.094.
Vinnarasi, R., Dhanya, C.T., Chakravorty, A., AghaKouchak, A., 2017. Unravelling diurnal asymmetry of surface temperature in different climate zones. Sci. Rep. 7 (1), 7350. https://doi.org/10.1038/s41598-017-07627-5.
Vleminckx, J., Hogan, J.A., Metz, M.R., Comita, L.S., Queenborough, S.A., Wright, S.J., Garwood, N.C., 2024. Flower production decreases with warmer and more humid atmospheric conditions in a western Amazonian forest. New Phytol. 241 (3), 1035–1046. https://doi.org/10.1111/nph.19388.
Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J., Bairlein, F., 2002. Ecological responses to recent climate change. Nature 416 (6879), 389–395. https://doi.org/10.1038/416389a.
Wang, T., Ottlé, C., Peng, S., Janssens, I.A., Lin, X., Poulter, B., Ciais, P., 2014. The influence of local spring temperature variance on temperature sensitivity of spring

phenology. Glob. Chang. Biol. 20 (5), 1473–1480. https://doi.org/10.1111/gcb.12509.
Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intel. Lab. Syst. 58 (2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.
Yan, Z., Xia, J., Qian, C., Zhou, W., 2011. Changes in seasonal cycle and extremes in China during the period 1960–2008. Adv. Atmos. Sci. 28 (2), 269–283. https://doi.org/10.1007/s00376-010-0006-3.
Yoder, J.B., Butterfield, H.S., 2024. Reconstructing 120 years of climate change impacts on Joshua tree flowering using citizen science and herbarium data. Ecol. Lett. 27 (2), 205–216. https://doi.org/10.1111/ele.14478.
Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D., Coomes, D., 2019. Advances in microclimate ecology arising from remote sensing. Trends Ecol. Evol. 34 (4), 327–341. https://www.cell.com/trends/ecology-evolution/fulltext/S0169-5347(18)30304-5?sf210261338=1.