New Phytologist

Research

Methods

On quantifying the apparent temperature sensitivity of plant phenology

Trevor F. Keenan[1,2], **Andrew D. Richardson**[3,4] and **Koen Hufkens**[5,6]

[1]Department of Environmental Science, Policy and Management, UC Berkeley, Berkeley, CA 94720, USA; [2]Earth and Environmental Science Area, Lawrence Berkeley National Lab., Berkeley, CA 94720, USA; [3]School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86004, USA; [4]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86004, USA; [5]Department of Applied Ecology and Environmental Biology, Ghent University, Ghent, Belgium; [6]INRA Aquitaine, UMR ISPA, Villenave d'Ornon, France

Author for correspondence:
Trevor F. Keenan
Tel: +1 510 486 5537
Email: TrevorKeenan@berkeley.edu

Received: *6 May 2019*
Accepted: *31 July 2019*

New Phytologist (2020) **225:** 1033–1040
doi: 10.1111/nph.16114

Key words: autumn, bud-burst, ecosystem, growing degree days, land surface, spring, warming.

Summary

- Many plant phenological events are sensitive to temperature, leading to changes in the seasonal cycle of ecosystem function as the climate warms. To evaluate the current and future implications of temperature changes for plant phenology, researchers commonly use a metric of temperature sensitivity, which quantifies the change in phenology per degree change in temperature.
- Here, we examine the temperature sensitivity of phenology, and highlight conditions under which the widely used days-per-degree sensitivity approach is subject to methodological issues that can generate misleading results. We identify several factors, in particular the length of the period over which temperature is integrated, and changes in the statistical characteristics of the integrated temperature, that can affect the estimated apparent sensitivity to temperature.
- We show how the resulting artifacts can lead to spurious differences in apparent temperature sensitivity and artificial spatial gradients. Such issues are rarely considered in analyses of the temperature sensitivity of phenology.
- Given the issues identified, we advocate for process-oriented modelling approaches, informed by observations and with fully characterised uncertainties, as a more robust alternative to the simple days-per-degree temperature sensitivity metric. We also suggest approaches to minimise and assess spurious influences in the days-per-degree metric.

Introduction

Changes in the timing of plant phenological events have long held the fascination of ecologists (de Reaumur, 1735; Leopold & Jones, 1947). Events such as bud-burst, flowering and leaf senescence play an important role in global ecosystems, affecting multiple aspects of ecosystem function in addition to feedbacks to the atmosphere and climate system (Richardson *et al.*, 2013). Phenological events are inherently sensitive to changes in weather, and recent climate warming has caused an appreciable extension of the growing season (Linderholm, 2006). Accurately characterising the controls of phenology is therefore a necessity in order to predict phenological responses to climate variability and future change.

Temperature is widely recognised as the dominant control of spring phenology (P) in temperate and boreal ecosystems (de Reaumur, 1735; Leopold & Jones, 1947; Leith, 1974). The nature of the temperature response (Parmesan, 2007; Morin *et al.*, 2010; Hänninen *et al.*, 2011; Clark *et al.*, 2014) is known with less certainty, however. There is indeed no known global response, with large differences apparent between species, locations and populations (Parmesan, 2007; Zohner & Renner, 2014). The response to temperature is also modulated by a host of factors such as photoperiod, latitude, humidity, chilling and dormancy requirements, and the timing of warming (Murray *et al.*, 1989; Myking & Heide, 1995; Morin *et al.*, 2009; Basler & Korner, 2012; Friedl *et al.*, 2014; Laube *et al.*, 2014).

In order to study the response of phenology to spatial and temporal changes in temperature, researchers have used a metric referred to as the temperature sensitivity (e.g. Wolkovich *et al.*, 2012; Chapman, 2013; Wang *et al.*, 2014, 2018; Fu *et al.*, 2015; Keenan, 2015; Thackeray *et al.*, 2016; Güsewell *et al.*, 2017), defined simply as the change in the date of a phenological event per change in temperature over a given period:

$$S_T = \frac{\Delta P}{\Delta T} \qquad \text{Eqn 1}$$

where ΔP is the anomaly in the date of a specific phenological transition, and ΔT is the anomaly in temperature over some integrating period (e.g. mean spring temperatures). The S_T metric is convenient and intuitive, and widely used due to its simplicity and apparent

tractability. The simplicity of S_T allows it to be applied to any dataset for which a metric of temperature change exists, for example for large datasets for which only highly temporally aggregated weather data have been historically available (e.g. Miller-Rushing & Primack, 2008; Primack *et al.*, 2009). It is tractable in the sense that more detailed analytical methods such as process-oriented models (e.g. Chuine *et al.*, 1999; Migliavacca *et al.*, 2012; Melaas *et al.*, 2013) require multiple parameters to be estimated, and often involve multiple interacting processes. Despite the appeal of the S_T metric, however, its simplicity could potentially hide important caveats. Here we examine potential statistical and methodological issues, and show how they can lead to biased results.

S_T is estimated as the slope of the linear regression between observed phenology dates (P) and some metric of temperature (T), most commonly mean temperature calculated across different years, along spatial gradients or between experimental treatments. Mathematically, S_T can therefore be expressed as:

$$S_T = \frac{\mathrm{cov}(P, T)}{\mathrm{var}(T)} \quad \text{Eqn 2}$$

$$= \frac{\mathrm{corr}(P, T) \cdot \mathrm{sd}(P)}{\mathrm{sd}(T)} \quad \text{Eqn 3}$$

S_T is therefore determined by the covariance between P and T, and by the variance of T (Eqn 2), or alternatively expressed, by the correlation between P and T, the standard deviation of P, and the standard deviation of T (Eqn 3). In this paper we examine issues affecting the estimation of a temporal and spatial S_T metric using both ground and remote sensing observations, and simulated data. We show that var(T) can and does change independently of P, an observation that has implications for the derived S_T values. Temporally, we show that var(T), and therefore S_T, are highly dependent on the length of the integration period, due to the nature of year-to-year variability in weather patterns, which tends to be lower over longer time periods. Spatially, we show that var(T) can vary over latitudinal gradients due to a relationship with mean temperature. This means that any investigation into the nature of variation in S_T in relation to both temperature variance and mean temperatures could be subject to the common statistical fallacy of spurious ratio, or 'part of whole', correlations (Pearson, 1897; Chayes, 1971), so termed because var(T) is a function of mean temperature, leading to S_T being examined as a function of the part (viz. T). This implies that all terms in Eqn 3 must be examined in order to properly interpret spatial changes. In addition, we examine how another common statistical issue, termed 'error in variables', due to unaccounted for uncertainty in T, can affect estimates of S_T. Each of these issues can lead to biases in the estimated temporal and spatial changes in apparent S_T. Finally, we suggest methods to more accurately quantify the sensitivity of phenological events to environmental drivers.

Materials and Methods

We use ground observations of phenology, obtained from the Hubbard Brook Experimental Forest long-term measurement site in the northeastern USA, where ongoing meteorological and phenological observations have been made for the past two decades (1989–2012, http://hubbardbrook.org/data/dataset.php?xml:id=51)). We used data from the three dominant tree species (Sugar Maple, American Beech and Yellow Birch). Each year, individual trees were visited every 3−5 d throughout spring, and their phenological status recorded. We estimated the mean date of spring phenology as the date at which leaves reach ½ of final length (database flag ≥ 3).

The temporal sensitivity of spring phenology to changes in mean integrated temperature (S_T) was quantified (Eqn 1), by comparing the anomalies in observed spring phenology dates (relative to the mean phenology date over the examined period), with the corresponding anomalies in mean temperature (T) for different temperature integration periods.

To illustrate the potential influence of latitudinal gradients in temperature variance, we used global gridded temperature data obtained from the European Center for Medium Range Weather Forecasting reanalysis product ERA- Interim, a global four-dimensional reanalysis product that provides daily 2 m air temperature to the present day at a resolution of 79 km ((Dee *et al.*, 2011) http://www.ecmwf.int/). We selected pixels identified as deciduous broadleaved forest, according to satellite observations from the MODIS (Moderate Resolution Imaging Spectroradiometer) Land Cover Dynamics phenology product (MCD12Q2 Collection 5 (Ganguly *et al.*, 2010)). The product, which is based on nadir bidirectional reflectance distribution function-corrected MODIS surface reflectance data (MCD43A4 (Schaaf *et al.*, 2002)) with an 8-d temporal resolution and a 500-m spatial resolution, estimates phenological transitions based on temporal changes in surface vegetation as characterised by the enhanced vegetation index and a logistic model approach (Ganguly *et al.*, 2010). We selected all MODIS Land Cover (MCD12Q1) pixels that were classified consistently as International Geosphere-Biosphere Program type deciduous broadleaved forest (MCD12Q1 class 4) between 2001 and 2012 (Supporting Information Fig. S1). To merge the two datasets (ERA, MODIS), we scaled the MODIS phenology pixels to the coarser resolution of the gridded climate cells by taking the median spring onset date over all pixels within a cell, for each year. Cells that contained fewer than 100 MODIS pixels (0.5% of potential) were discarded. We use the ERA and MODIS data in an illustrative analysis focused on the winter deciduous forests of Europe (42°:55°N, 10°W:60°E). We used the ERA and MODIS data to calculate the relationship between temperature variance and latitude, and between latitude and the covariance of temperature and spring phenology.

Results and Discussion

The influence of the length of the period of temperature integration

Examining the temperature sensitivity of phenology via S_T (Eqn 1) requires specifying a period of time over which to

integrate temperatures. Some studies relate observed variability in phenology to the mean annual temperature (e.g. Wolkovich *et al.*, 2012) or mean meteorological spring temperature (i.e. March, April, May in the northern hemisphere) (e.g. Keenan *et al.*, 2014), whilst others use an optimisation approach to identify the period of time that is most correlated to the observed variability in phenology (e.g. Fu *et al.*, 2015). The period over which temperatures are integrated invariably has a defined length. The length of the integration period is of crucial importance, due to the fact that longer integration periods tend to have lower interannual variability in aggregated temperature. For example, year-to-year variability on a particular day or week of the year is typically much higher than year-to-year variability of monthly, seasonal or annual temperatures. In addition, the covariance between P and T is dependent on the length of the integration period used, and inversely related to the relevance of the integration period to the phenological event. The timing and length of the chosen integration period affect the relevance of the integration period to the phenological event, and could therefore affect the derived apparent S_T, through its influence on both var(T) and cov(T, P) (Eqn 2).

To illustrate the potential impact of varying integration lengths on the derived S_T, we use 20 yr of phenological and meteorological observations from the Hubbard Brook Experimental Forest, for three deciduous forest species. For each species, we use temperatures from all periods of length ≥ 2 wk within the first 150 d of each year that show a significant correlation with observed bud-burst ($P < 0.01$), and examine how the derived S_T varies in dependence of the effect of integration length on var (T). As expected, var(T) declined with increasing integration period length, from 9°C^2 for integration periods of 2 wk (as used in, e.g. Gunderson *et al.*, 2012; Shen *et al.*, 2014; Fu *et al.*, 2015; Zhang *et al.*, 2015b; Güsewell *et al.*, 2017), to between 1–2°C^2 for integration periods of over 2 months (Fig. 1a). Over the same range of period lengths, the covariance between temperature and bud-burst also declined, but much less than the variance in temperature (Fig. 1a). The combined changes in temperature variance and temperature−phenology covariance resulted in a large change in the derived apparent S_T, which increased from *c.* −1.5 d per °C when using a temperature integration period of 2 wk, to a range of −3.2 to −6 d °C^{-1} over longer time periods (Fig. 1a,b), representing a greater than three-fold difference in the apparent sensitivity of spring bud-burst to temperature change. Although the derived S_T was least sensitive to period length for longer period lengths, the correlation between phenology dates and integrated temperature was distributed across a range of period lengths, start dates and S_T values (Figs 1c, S2). Studies often use the period with the highest correlation between phenology and integrated temperature as reference period, but these results suggest that a high correlation (Fig. 1c) can be obtained across a range of period start date–length combination, and therefore S_T values (Fig. 1b). The influence of the integration period on the derived S_T complicates the interpretation of results across studies, across sites, or even within a site for species that

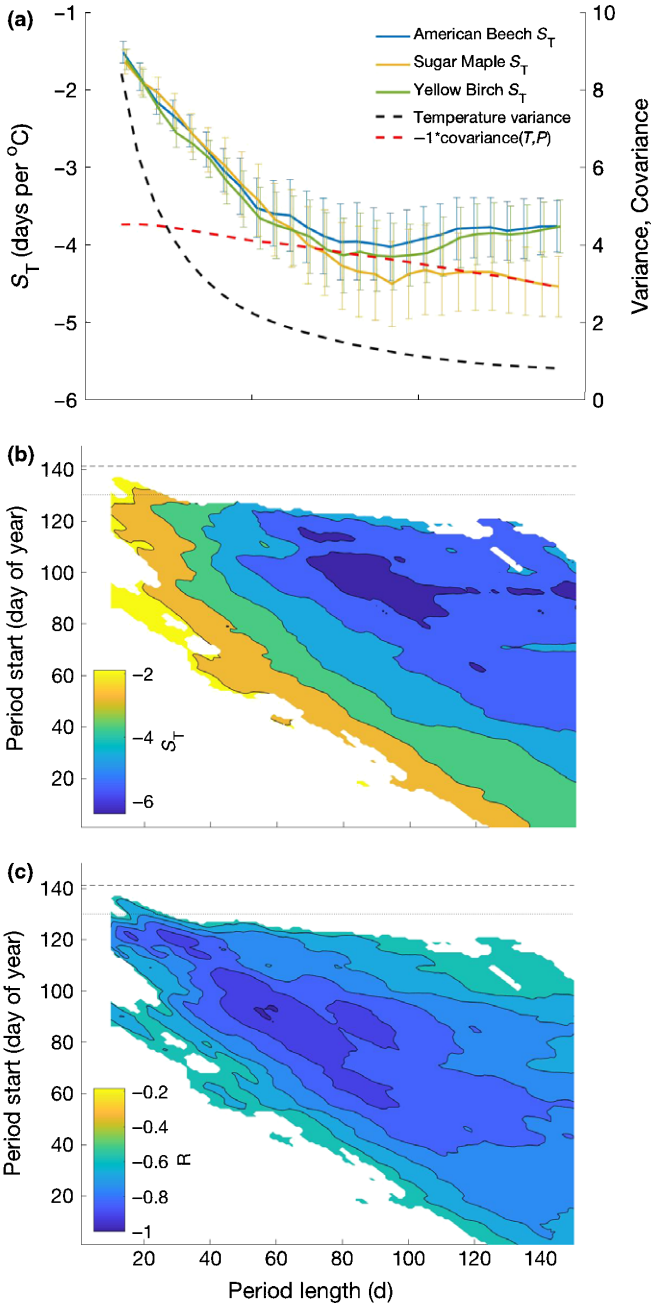

Fig. 1 The effect of temperature integration length on the derived temperature sensitivity (S_T) of spring bud-burst at the Hubbard Brook Experimental Forest. (a) Temperature sensitivities calculated for a range of temperature integration periods (period length), binned by the period start day of year, for three tree species (American Beech, Sugar Maple, and Yellow Birch), along with the associated temperature variance (°C^2) and the covariance (d°C^{-1}) between integrated temperature (T) and observed phenology dates (P). (b) A map of the derived apparent temperature sensitivity for Sugar Maple, and (c) the associated correlation (R) between observed phenological dates for this species and interannual changes in integrated temperature. Error bars in *a* represent the standard deviation within a period length bin. Horizontal lines on (b, c) represent the earliest (dotted) and mean (dashed) of observed phenology dates across years. See Supporting Information Fig. S1 for comparable figures (b, c) for American Beech and Yellow Birch. S_T is derived from Eqn 2 using linear regression for 22 yr (1992–2013) of observed phenological records. Only significant S_T values ($P < 0.01$) are shown.

may be sensitive to temperatures during different periods (Friedl *et al.*, 2014).

The potential for spurious correlations due to changes in temperature variance

The S_T metric is often used to examine spatial and temporal changes in the temperature sensitivity of phenology (e.g. Shen *et al.*, 2014; Fu *et al.*, 2015; Zhang *et al.*, 2015a). There are, however, potential spatial and temporal changes in var(T) that could lead to the identification of spurious relationships. For example, spring (March, April, May) temperature variance is a strong function of latitude, and varies from 4.3 to 2.2°C^2 over the extent of the deciduous forests of northern Europe (Fig. S3). Examining changes in S_T in space or time therefore potentially constitutes a special case of a well studied statistical phenomenon known as 'ratio correlations' (Pearson, 1897; Chayes, 1971). Ratio correlations arise when a ratio (in this case Eqn 2) is compared with the denominator of the ratio, or to some quantity of which the denominator is a function. As we have shown above, ∂var(T) can be much larger than ∂cov(T, P). In addition, as phenology is not responding exactly to the integrated temperature being used, there can be variance in T that is unrelated to P. Although the covariance will also be influenced, the fact that P is not an exact function of T implies that ∂var(T) > ∂cov(T,P). Strong correlations (e.g. between S_T and var(T) (Wang *et al.*, 2014)) could therefore potentially emerge due to the large changes in temperature variance even in the absence of any causal relationship. Latitudinal changes in var(T) could also result in changes in cov(T, P), however. This implies that changes in all terms of Eqn 3 need to be assessed in order to interpret derived changes in S_T.

To illustrate this point, we consider the scenario in which the covariance between phenology and spring temperature is independent of latitude, but where spring temperature variance increases with increasing latitude (Fig. 2), based on observed spatial patterns for winter deciduous forests over a European latitudinal gradient from 42.5° to 55°N (Figs S2, S3). We generate 10 000 covariance values from a normal distribution, randomly distributed along the European latitudinal gradient (Fig. 2a). The increasing temperature variance (Fig. 2b) and constant covariance (Fig. 2a) across the latitudinal gradient inevitably leads (Eqn 3) to a decline in the apparent S_T with increasing latitude (Fig. 2c). A logical conclusion would be that forest phenology is less sensitive to temperature where temperature variance is high, as has been reported (Wang *et al.*, 2014). S_T is a function of corr(T, P) and sd(P), however (Eqn 3), so S_T could be declining due to a change in either, with important implications for the interpretation of the apparent change in S_T. For instance, in this example, we know that var(T) increases with latitude, and cov(T, P) is constant (Figs 2, S3). As cov(T, P) = corr(T, P)×(var(T).var(P))$^{1/2}$, a constant cov(T, P) and increasing var(T) implies a decrease in var(P) and/or a decrease in corr(T, P). This is important, as a latitudinal gradient in corr(T, P) could be indicative that the chosen integrated temperature is not equally relevant to the observed phenology dates across latitudes, and that the change in S_T might be an artifact of a spatial gradient in the appropriateness of the chosen reference

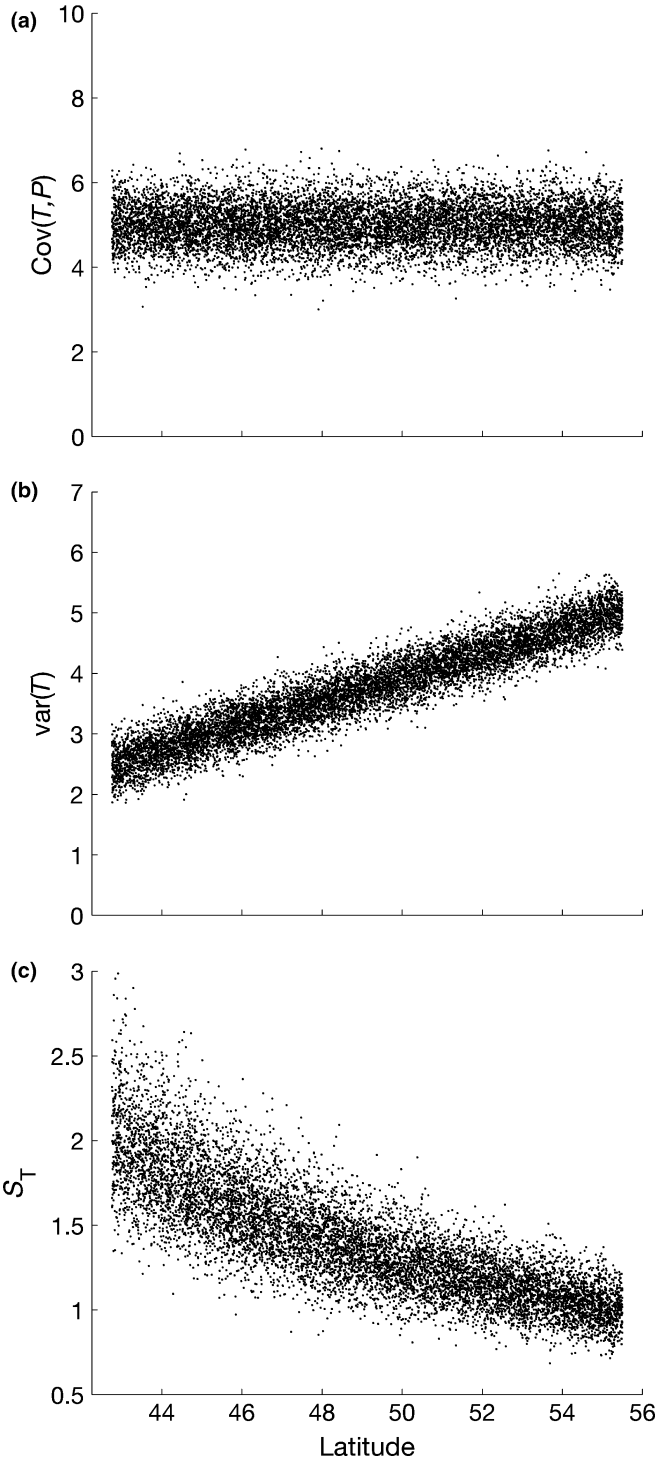

Fig. 2 An illustration of the potential effect of spatial changes in the variance of an integrated temperature metric (T) on the derived temperature sensitivity (S_T). Given (a) a synthetic covariance distribution, cov(X_Y,T) ~ N (5,0.5), that is uncorrelated with latitude, and (b) a temperature variance, var(T), (°C^2), that is correlated with latitude, leads to (c) an apparent S_T (days °C^{-1}) that has a strong relationship with latitude. The relationship of temperature variance to latitude (var(T) = N(−6 + 0.19×Latitude, 10%)), Supporting Information Figs S2, S3), and the relationship of the covariance of temperature and phenology to latitude, were extracted from the ERA global meteorological observations and MODIS satellite observations (see the 'Materials and Methods' section) over the range of European winter deciduous broadleaf forests (42.5°:55°N).

temperature. The relationship between S_T and var(T) should only be considered valid if there is no spatial gradient in corr(T, P). Spatial and temporal changes in S_T are commonly reported in the literature (Shen *et al.*, 2014; Fu *et al.*, 2015; Wang *et al.*, 2018), without consideration for how changes in the different terms of Eqn 3 could affect the derived sensitivity of phenology to temperature independently of changes in phenology. This illustration does not aim to claim that all reported latitudinal differences in S_T are due to independent changes in var(T), as sd(P) and cov(T, P) may also vary, depending on the study region in question, reflecting real changes in phenological sensitivity. It does however show that the relative change in each term of S_T (Eqn 3) needs to be assessed to fully understand the implied spatial changes in the response of phenology to temperature.

The temperature on which phenology depends is not known with accuracy

Although it is a widely accepted fact that phenology responds to temperature, there is no broad consensus how exactly that dependence manifests. For example, the timing of warming matters (Clark *et al.*, 2014; Friedl *et al.*, 2014), as do other moderating factors such as winter chilling, dormancy requirements, photoperiod, humidity and leaf longevity (Murray *et al.*, 1989; Myking & Heide, 1995; Morin *et al.*, 2009; Basler & Korner, 2012; Laube *et al.*, 2014). The representativeness of the chosen temperature metric is therefore difficult to quantify. In addition, studies commonly include temperatures that happen after the phenological event, such as when using annual or even spring integrals, or when trends or interannual variability lead to an encroachment of phenological dates into the period of temperature integration. Co-located meteorological and phenological observations are also often lacking, with researchers forced to use gridded meteorological data or observations from the nearest station (e.g. Olsson & Jönsson, 2014). There are also potential nonlinear relationships between temperature and phenology. The integrated temperature used is therefore always only a proxy to the real integrated temperature to which plants respond.

The fact that the integrated temperature being used is often a rough estimate is important, as it means that the predictor used to define S_T inherently has an associated but unquantified uncertainty. Unacknowledged uncertainty in predictors represents a common 'error in variables' scenario that unequivocally leads to a phenomenon known as regression dilution or attenuation bias (Pindyck & Rubinfeld, 1991). This implies a likely underestimation of S_T in the presence of unaccounted for error in T. We demonstrate this by considering the hypothetical case of a 1 : 1 relationship between the date of leaf phenology (P) and an unknown temperature metric (T), with an arbitrary 'true' S_T of 5 d per degree change in temperature. To this relationship we add varying degrees of random error (e ~ N(0)). Adding random error to the integrated temperature metric leads to an underestimation of the true S_T, with the bias increasing as random error in T increases (Fig. 3). The effect of the unknown error in T can be reduced by estimating S_T using regression approaches that account for unknown errors in both axes, such as reduced major axis (or Type-2) regression (e.g. White *et al.*, 2009; Fu *et al.*, 2015; Yang *et al.*, 2018). Even with such approaches, however, sensitivity to the error in variables problem remains. The fact that the true temperature signal to which phenology is sensitive is not known therefore complicates the interpretation of between-species differences in S_T values at a given site, or between individuals of the same species at different sites.

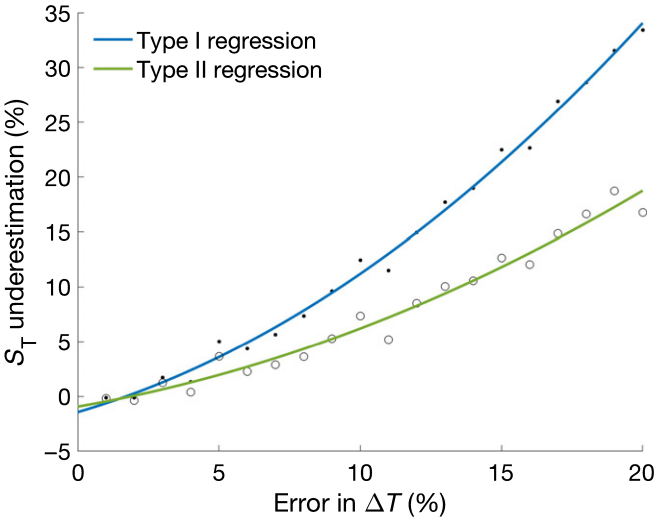

Fig. 3 The implicit underestimation of the apparent temperature sensitivity due to unaccounted for uncertainty in integrated temperature. This synthetic example shows the effect of increasing error in integrated temperature (T) on the derived apparent temperature sensitivity (S_T, days °C^{-1}), using both linear regression (Type-I) and reduced major axis (Type-II) regression. Both the phenology dates and the temperature metric used are obtained from the hypothetical relationship of $\Delta P = 5 \times \Delta T$ (representing a temperature sensitivity of 5 d per °C), with random error drawn from a normal distribution with zero mean and standard deviation of 0–20%.

Implications and solutions

The issues raised above have implications for the detection and attribution of spatial and temporal changes in S_T. Temporally, S_T has been reported to have changed over the past few decades (Fu *et al.*, 2015). Spatially, S_T has been reported to change in dependence of latitude (Shen *et al.*, 2014), altitude (Piao *et al.*, 2011; Zohner & Renner, 2014), and in response to the temperature variance of a region (Wang *et al.*, 2014). Differences between species are commonly reported (Marchin *et al.*, 2015), between populations (Parmesan, 2007), and between experiments and natural observations (Wolkovich *et al.*, 2012). But such changes in S_T are difficult to assess without consideration of the issues highlighted above. For example, a recent study reported that experiments underpredict the sensitivity of spring phenology to warming, when compared with natural observations (Wolkovich *et al.*, 2012). But the observation-based sensitivity was calculated using annual temperatures, which have a much lower variance than, say, spring temperatures. Due to this lower variance, the temperature sensitivity derived from observations using annual temperatures would be over-estimated, and also poorly estimated because annual ΔT is a poor proxy for the actual temperature signal to which plant phenology responds. Using spring temperature

instead could lead to a lower apparent S_T in observations, but no change in the apparent S_T of experiments (due to the fact that the experimental change in temperature is typically applied evenly throughout the year), and greatly reduce the difference between the two. Similarly, studies using S_T to compare differences in the response of co-located species to temperature would have difficulty discerning whether the differences between the species' S_T values were due to an inherently different temperature sensitivity, or a difference in the representativeness of the chosen integration period for a particular species. Unfortunately, analyses rarely control for such effects, or test the statistical characteristics and representativeness of the integrated temperature metric used.

Fundamentally, the above issues exist because the representativeness of an integrated temperature metric to the observed phenological event can rarely be quantified (e.g. integration lengths from 2 wk to 2 months can have a statistically equivalent relationship with the observed dates (Fig. 1)). In reality, T is a noisy proxy for the true temperature signal to which P responds, and can contain significant variation that is not associated with changes in P. As the true integrated temperature signal is unknown, the derived S_T can be assumed to always be inaccurate, and can lead to results that are influenced by statistical artifacts.

We advocate for improvements in the application and reporting of the S_T metric. Statistical methods such as Type-II regression should be used (as in, e.g. White *et al.*, 2009; Fu *et al.*, 2015) to help minimise the influence of unaccounted for uncertainty in T (Fig. 3). Studies should also incorporate analyses of the potential effect of changes in temperature variance due to integration length, or spatial/temporal differences. This will help attribute the apparent differences in S_T, and inform the interpretation of results. In addition, clearly reporting integration periods would improve reproducibility, and using multiple integration periods would help assess the robustness of results.

One promising alternative to the S_T metric lies in the use of data-informed process-oriented models to characterise the response of phenology to temperature change along with other potential modifiers such as photoperiod (e.g. Migliavacca *et al.*, 2012). Such an approach could be used, based on both natural and experimental observations (Hänninen *et al.*, 2019), in tandem with a range of climate scenarios to rigorously characterise the potential response to climate variability and long-term change. A key strength of a model-based analysis is that model projection uncertainty can be directly quantified, and used to inform the interpretation of results. A model, parameterised from the data, could also be used in 'experiment mode', in which warming is applied (either uniformly or following climate projections) throughout the year to quantify the model's implicit temperature sensitivity. Not only can the combination of models with observations aid researchers better interpret the observations, it will also inevitably lead to the development of new state-of-the-art models and fundamental theory. Although models too have their pitfalls (Hänninen *et al.*, 2019), and can be subject to structural error, ensembles can be used to quantify uncertainties and test competing hypotheses (Hufkens *et al.*, 2018). A full consideration and propagation of measurements and their associated uncertainties, for example through using model–data fusion techniques, would both allow for the identification of weak model components and the design of more effective experimental strategies (Keenan *et al.*, 2011).

The S_T metric is subject to multiple issues that undermine its credibility, in particular for natural observations, and can generate misleading results. Although we focus here on the temperature sensitivity of phenology, some of the issues raised could also apply to the temperature sensitivity of other ecological phenomena, or indeed the sensitivity of phenology to other factors, such as precipitation in water limited regions (Moore *et al.*, 2016). We therefore advocate for more rigorous statistical assessments of potential underlying biases, and the use of data-informed modelling approaches for the interpretation and projection of phenological changes.

Acknowledgements

TFK acknowledges funding support from NASA Terrestrial Ecology Program IDS Award NNH17AE86I. ADR acknowledges support from the National Science Foundation's Macrosystems Biology program (award EF-1702697). The authors thank Amey Bailey for providing phenological observations at Hubbard Brook Experimental Forest. Research at Hubbard Brook (DEB-1637685) is partially supported through NSF's LTER program. The European Center for Mid-Range Forecasting is acknowledged for providing the ERA-Interim gridded weather data. KH acknowledges support from the BELSPO Brain programme (project BR/175/A3/COBECORE). All data used in this manuscript are publicly available.

Author contributions

TFK designed the study, detailed and performed the analysis, and is responsible for the integrity of the manuscript. ADR aided in the interpretation and framing of the results. KH contributed to the analysis and commented on the results. TFK led the writing, with input from ADR and KH.

ORCID

Koen Hufkens https://orcid.org/0000-0002-5070-8109
Trevor F. Keenan https://orcid.org/0000-0002-3347-0258
Andrew D. Richardson https://orcid.org/0000-0002-0148-6714

References

Basler D, Korner C. 2012. Photoperiod sensitivity of bud burst in 14 temperate forest tree species. *Agricultural and Forest Meteorology* **165**: 73–81.

Chapman DS. 2013. Greater phenological sensitivity to temperature on higher Scottish mountains: new insights from remote sensing. *Global Change Biology* **19**: 3463–3471.

Chayes F. 1971. *Ratio correlation.* Chicago, IL, USA: Chicago University Press.

Chuine I, Cour P, Rousseau DD. 1999. Selecting models to predict the timing of flowering of temperate trees: implications for tree phenology modelling. *Plant, Cell & Environment* **22**: 1–13.

Clark JS, Melillo J, Mohan J, Salk C. 2014. The seasonal timing of warming that controls onset of the growing season. *Global Change Biology* **20**: 1136–1145.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P *et al.* 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**: 553–597.

Friedl MA, Gray JM, Melaas EK, Richardson AD, Hufkens K, Keenan TF, Bailey A, O'Keefe J. 2014. A tale of two springs: using recent climate anomalies to characterize the sensitivity of temperate forest phenology to climate change. *Environmental Research Letters* **9**: 054006.

Fu YH, Zhao H, Piao S, Peaucelle M, Peng S, Zhou G, Ciais P, Huang M, Menzel A, Peñuelas J *et al.* 2015. Declining global warming effects on the phenology of spring leaf unfolding. *Nature* **526**: 104–107.

Ganguly S, Friedl MA, Tan B, Zhang X, Verma M. 2010. Land surface phenology from MODIS: characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment* **114**: 1805–1816.

Gunderson CA, Edwards NT, Walker AV, O'Hara KH, Campion CM, Hanson PJ. 2012. Forest phenology and a warmer climate – growing season extension in relation to climatic provenance. *Global Change Biology* **18**: 2008–2025.

Güsewell S, Furrer R, Gehrig R, Pietragalla B. 2017. Changes in temperature sensitivity of spring phenology with recent climate warming in Switzerland are related to shifts of the preseason. *Global Change Biology* **23**: 5189–5202.

Hänninen H, Kramer K, Tanino K, Zhang R, Wu J, Fu YH. 2019. Experiments are necessary in process-based tree phenology modelling. *Trends in Plant Science* **24**: 199–209.

Hänninen H, Tanino K, Ha H, Tanino K, Hänninen H, Tanino K, Hänninen H, Tanino K. 2011. Tree seasonality in a warming climate. *Trends in Plant Science* **16**: 412–416.

Hufkens K, Basler D, Milliman T, Melaas EK, Richardson AD. 2018. An integrated phenology modelling framework in R. *Methods in Ecology and Evolution* **9**: 1276–1285.

Keenan TF. 2015. Phenology: spring greening in a warming world. *Nature* **526**: 48–49.

Keenan TF, Carbone MS, Reichstein M, Richardson AD. 2011. The model–data fusion pitfall: assuming certainty in an uncertain world. *Oecologia* **167**: 587–597.

Keenan TF, Gray J, Friedl MA, Toomey MP, Bohrer G, Hollinger DY, Munger JW, O'Keefe J, Schmid HP, Wing IS *et al.* 2014. Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Climate Change* **4**: 598–604.

Laube J, Sparks T, Estrella N, Menzel A. 2014. Does humidity trigger tree phenology? Proposal for an air humidity based framework for bud development in spring. *New Phytologist* **202**: 350–355.

Leith H. 1974. *Phenology and seasonality modeling*. New York, NY, USA: Springer-Verlag.

Leopold A, Jones S. 1947. A phenological record for Sauk and Dane Counties, Wisconsin, 1935-1945. *Ecological Monographs* **17**: 81–122.

Linderholm H. 2006. Growing season changes in the last century. *Agricultural and Forest Meteorology* **137**: 1–14.

Marchin RM, Salk CF, Hoffmann WA, Dunn RR. 2015. Temperature alone does not explain phenological variation of diverse temperate plants under experimental warming. *Global Change Biology* **21**: 3138–3151.

Melaas EK, Richardson AD, Friedl MA, Dragoni D, Gough CM, Herbst M, Montagnani L, Moors E. 2013. Using FLUXNET data to improve models of springtime vegetation activity onset in forest ecosystems. *Agricultural and Forest Meteorology* **171–172**: 46–56.

Migliavacca M, Sonnentag O, Keenan TFF, Cescatti A, O'Keefe J, Richardson ADD. 2012. On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model. *Biogeosciences* **9**: 2063–2083.

Miller-Rushing AJ, Primack RB. 2008. Global warming and flowering times in Thoreau's Concord: a community perspective. *Ecology* **89**: 332–341.

Moore CE, Brown T, Keenan TF, Duursma RA, van Dijk AIJM, Beringer J, Culvenor D, Evans B, Huete A, Hutley LB *et al.* 2016. Reviews and syntheses: Australian vegetation phenology: new insights from satellite remote sensing and digital repeat photography. *Biogeosciences* **13**: 5085–5102.

Morin X, Lechowicz MJ, Augspurger C, O'Keefe J, Viner D, Chuine I. 2009. Leaf phenology in 22 North American tree species during the 21st century. *Global Change Biology* **15**: 961–975.

Morin X, Roy J, Sonié L, Chuine I. 2010. Changes in leaf phenology of three European oak species in response to experimental climate change. *New Phytologist* **186**: 900–910.

Murray M, Cannell M, Smith R. 1989. Date of budburst of 15 tree species in Britain following climatic warming. *Journal of Applied Ecology* **26**: 693–700.

Myking T, Heide OM. 1995. Dormancy release and chilling requirement of buds of latitudinal ecotypes of *Betula pendula* and *B. pubescens*. *Tree Physiology* 697–704.

Olsson C, Jönsson AM. 2014. Process-based models not always better than empirical models for simulating budburst of Norway spruce and birch in Europe. *Global Change Biology* **20**: 3492–3507.

Parmesan C. 2007. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology* **13**: 1860–1872.

Pearson KS. 1897. On a form of spurious correlation which may arise when indices are used in the measurements of organs. *Proceedings of the Royal Society of London* **60**: 489–502.

Piao S, Cui M, Chen A, Wang X, Ciais P, Liu J, Tang Y. 2011. Altitude and temperature dependence of change in the spring vegetation green-up date from 1982 to 2006 in the Qinghai-Xizang Plateau. *Agricultural and Forest Meteorology* **151**: 1599–1608.

Pindyck RS, Rubinfeld LR. 1991. *Econometric models and economic forecasts, 3rd edn.* Boston, MA, USA: McGraw-Hill/Irwin.

Primack RB, Higuchi H, Miller-Rushing AJ. 2009. The impact of climate change on cherry trees and other species in Japan. *Biological Conservation* **142**: 1943–1949.

de Reaumur RAF. 1735. Observations du thermometre, faites a Paris pendant l'annee 1735 comparees avec celles qui ont ete faites sous la ligne a l'Ile de France, a Alger et en quelques-unes de nosıles de l'Am erique. *Memoires de l'Academie Royale des Sciences de Paris* 545–576.

Richardson AD, Keenan TF, Migliavacca M, Ryu Y, Sonnentag O, Toomey M. 2013. Climate change, phenology and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology* **169**: 156–173.

Schaaf CB, Gao F, Strahler AH, Lucht W, Li X, Tsang T, Strugnell NC, Zhang X, Jin Y, Muller J-P *et al.* 2002. First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment* **83**: 135–148.

Shen M, Tang Y, Chen J, Yang X, Wang C, Cui X, Yang Y, Han L, Li L, Du J *et al.* 2014. Earlier-season vegetation has greater temperature sensitivity of spring phenology in northern hemisphere. *PLoS ONE* **9**: e88178.

Thackeray SJ, Henrys PA, Hemming D, Bell JR, Botham MS, Burthe S, Helaouet P, Johns DG, Jones ID, Leech DI *et al.* 2016. Phenological sensitivity to climate across taxa and trophic levels. *Nature* **535**: 241–245.

Wang H, Dai J, Rutishauser T, Gonsamo A, Wu C, Ge Q. 2018. Trends and variability in temperature sensitivity of lilac flowering phenology. *Journal of Geophysical Research: Biogeosciences* **123**: 807–817.

Wang T, Ottlé C, Peng S, Janssens IA, Lin X, Poulter B, Yue C, Ciais P. 2014. The influence of local spring temperature variance on temperature sensitivity of spring phenology. *Global Change Biology* **20**: 1473–1480.

White MA, de Beurs KM, Didan K, Inouye DW, Richardson AD, Jensen OP, O'Keefe J, Zhang G, Nemani RR, van Leeuwen WJD *et al.* 2009. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006. *Global Change Biology* **15**: 2335–2359.

Wolkovich EM, Cook BI, Allen JM, Crimmins TM, Betancourt JL, Travers SE, Pau S, Regetz J, Davies TJ, Kraft NJB *et al.* 2012. Warming experiments underpredict plant phenological responses to climate change. *Nature* **485**: 494–497.

Yang J-X, He W-M, Peng Y, Zhou X-H, Peng P-H, Li J-J. 2018. Warming delays the phenological sequences of an autumn-flowering invader. *Ecology and Evolution* **8**: 6299–6307.

Zhang H, Yuan W, Liu S, Dong W. 2015a. Divergent responses of leaf phenology to changing temperature among plant species and geographical regions. *Ecosphere* **6**: 2009–2016.

Zhang H, Yuan W, Liu S, Dong W, Fu Y. 2015b. Sensitivity of flowering phenology to changing temperature in China. *Journal of Geophysical Research G: Biogeosciences* **120**: 1658–1665.

Zohner CM, Renner SS. 2014. Common garden comparison of the leaf-out phenology of woody species from different native climates, combined with herbarium records, forecasts long-term change. *Ecology Letters* **17**: 1016–1025.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 The distribution of deciduous broadleaved forest sites for which climate and remote sensing data were used in this study.

Fig. S2 The effect of temperature integration length on the derived temperature sensitivity of spring bud-burst at the Hubbard Brook Experimental Forest.

Fig. S3 The relationship between the variance of spring temperature, and covariance of temperature and phenology, with latitude.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.