

# A simple explanation for declining temperature sensitivity with warming

E. M. Wolkovich<sup>1,a</sup>, J. L. Auerbach<sup>2</sup>, C. J. Chamberlain<sup>3</sup>, D. M. Buonaiuto<sup>3</sup>,  
A. K. Ettinger<sup>4</sup>, I. Morales-Castilla<sup>5</sup> & A. Gelman<sup>2</sup>

<sup>1</sup>Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, British Columbia, Canada

<sup>2</sup>Department of Statistics, Columbia University, New York, NY 10027, USA

<sup>3</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

<sup>4</sup>The Nature Conservancy in Washington, 74 Wall Street, Seattle, WA, USA

<sup>5</sup>Department of Life Sciences, University of Alcalà CTRA N-II, KM., 33,600, 28802, Alcalà de Henares, Spain

<sup>a</sup>Corresponding author (ORCID: 0000-0001-7653-893X)

## Abstract

Temperature sensitivity—the magnitude of a biological response per °C—is a fundamental concept across scientific disciplines, especially biology, where temperature determines the rate of many plant, animal and ecosystem processes. Recently, a growing body of literature in global change biology has found temperature sensitivities decline as temperatures rise (Fu et al., 2015; Güsewell et al., 2017; Piao et al., 2017; Dai et al., 2019). Such observations have been used to suggest climate change is reshaping biological processes, with major implications for forecasts of future change. Here we present a simple alternative explanation for observed declining sensitivities: the use of linear models to estimate non-linear temperature responses. Corrections for the non-linearity of temperature response in simulated data and long-term phenological data from Europe remove the apparent decline. Our results show that rising temperatures combined with linear estimates based on calendar time produce observations of declining sensitivity—without any shift in the underlying biology. Current methods may thus undermine efforts to identify when and how warming will reshape biological processes.

# 1 Main text

Climate change has reshaped biological processes around the globe, with shifts in the timing of major life history events (phenology), carbon dynamics and other ecosystem processes (IPCC, 2014). With rising temperatures, a growing body of literature has documented changes in temperature sensitivity—the magnitude of a biological response scaled per °C. Many studies have found declining responses to temperature in recent decades (Fu et al., 2015; Güsewell et al., 2017; Piao et al., 2017; Dai et al., 2019), and some have reported more uniform sensitivities across elevation (Vitasse et al., 2018), or lower sensitivities in warmer, urban areas (Meng et al., 2020).

Most studies attribute changes in temperature sensitivity to shifts in underlying biological processes. For example, researchers have suggested weaker temperature sensitivities are evidence of increased light limitation in the tundra (Piao et al., 2017), or a decline in the relative importance of warm spring temperatures for spring phenological events (e.g., leafout, insect emergence) in the temperate zone (Fu et al., 2015; Meng et al., 2020), as other environmental triggers (e.g., winter temperatures that determine ‘chilling’) play a larger role. Yet, despite an increase in studies reporting declining or shifting temperature sensitivities, none have provided strong evidence of the biological mechanisms underlying these changes (e.g., Fu et al., 2015; Meng et al., 2020). The missing mechanisms may be hidden in the data: environmental factors moderate biological processes in complex ways (Chuine et al., 2016; Güsewell et al., 2017), are strongly correlated in nature (e.g., Fu et al., 2015), and temperature variance shifts over time and space (Keenan et al., 2020).

Here we propose a simpler alternative explanation: the use of linear models for non-linear responses to temperature. Researchers generally use methods with assumptions of linearity to calculate temperature sensitivities, often relying on some form of linear regression to compute a change in a quantity—days to leafout or carbon sequestered over a fixed time, for example—per °C, thus ignoring that many biological responses to temperature are non-linear. We show, theoretically then with simulated and empirical data, how the use of linear methods for non-

linear responses can produce an illusion that the mechanisms underlying biological processes are changing.

Many observed biological responses are the result of continuous non-linear processes that depend on temperature, which are discretized into temporal units for measurement. For example, a biological response, such as leafout, occurs when a certain thermal sum is reached (Dijkhuis, 1956; Lindsey and Newman, 1956), and plants will reach this threshold more quickly—in calendar time—when average daily temperatures are warmer (Valentine, 1983; Lechowicz, 1984; Kramer, 2012). Biologically, however, the plants may require the same temperature sum. Indeed any process observed or measured as the time until reaching a threshold is inversely proportional to the speed at which that threshold is approached. Temperature determines the speed of many biological processes (Bonan and Sirois, 1992; Hinrichsen, 2009; Hofmann and Todgham, 2010). Thus, at very low temperatures plants would never leaf out and at higher temperatures they could leaf out in only a matter of days—yet sensitivities estimated from linear regression at higher (warmer) temperatures would appear much lower than those observed at lower temperatures. Warming acts to step on the biological accelerator, and thus may produce the illusion of shifting processes when non-linear responses are modeled as linear.

We show this by deriving the relationship between a biological response and temperature using a simple stochastic model, which describes the first time a random process hits a threshold (see ‘A first-hitting-time model of leafout’ in Supplementary Information). Our model holds the temperature threshold for leafout constant. Even though the mechanism by which temperature leads to leafout does not change, the model produces declining sensitivity—as measured in days per °C—with warming. Indeed, under this model constant temperature sensitivity would be evidence that the temperature threshold is not constant and the mechanisms underlying the leafout process have changed.

Simulations show that correcting for non-linearity using a log transformation removes apparent declines in temperature sensitivity (Fig. 1, S2, code link). In empirical long-term leafout data from Europe, correcting for non-linearity in responses produces little evidence for declining

sensitivities with warming (Figs. 1, S6, S7). An apparent decline in sensitivity for silver birch (*Betula pendula*) from -4.3 days/°C to -3.6 days/°C from 1950-1960 compared to 2000-2010 disappears using a log-log regression (-0.17 versus -0.22). We see similar corrections using 20-year windows, and a potential increase in sensitivity for European beech (*Fagus sylvatica*, see Tables S1-S2). Moreover, the variance of the leafout dates of both species declines as temperatures rise—(declines of roughly 50%, see Tables S1-S2), which is expected under our model as warming accelerates towards the thermal threshold that triggers leafout (and in contrast to predictions from changing mechanisms, see Ford et al., 2016).

Fundamentally rising temperature should alter many biological processes, making robust methods for identifying these changes critical. In spring plant phenology, where declining sensitivities are often reported (Fu et al., 2015; Piao et al., 2017; Dai et al., 2019), warming may increase the role of ‘chilling’ (determined mainly by winter temperatures) and daylength (Laube et al., 2014; Zohner et al., 2016)—potentially increasing the thermal sum required for leafout at lower values of these cues (Polgar et al., 2014; Zohner et al., 2017; Flynn and Wolkovich, 2018). Adjusting our simulations to match this model yielded shifts in sensitivities with warming. Unlike a model with no underlying biological change, however, after correcting for non-linearity the shifts in sensitivities remained and they occurred in step with the biological change (Fig. S4a, c). In contrast, sensitivities estimated from a linear model showed shifts across the entire range of warming, well before the simulated biological change (Fig. S4a, c). Further, we found that an increase in the thermal sum required for leafout should yield larger temperature sensitivities, not smaller, as often expected (e.g., Fu et al., 2015), thus highlighting the complexity of what trends to expect in sensitivities with warming (see ‘Common hypotheses for declining sensitivity’ in Supplementary Information for an extended discussion).

Our theoretical model and empirical results show that rising temperatures are sufficient to explain declining temperature sensitivity. It is not necessary to invoke changes to the mechanisms that underlie the biological processes themselves. Our results provide a simpler explanation for observations of declining temperature sensitivities, but do not rule out that important changes in biological processes may underlie such declines. Instead, our results highlight how the use

of linear models may make identifying when—and why—warming alters underlying biology far more difficult.

Inferring biological processes from statistical artifacts is not a new problem (e.g., Nee et al., 2005), but climate change provides a new challenge in discerning mechanism from measurements because it affects biological time, while researchers continue to use calendar time. Other fields focused on temperature sensitivity often use approaches that acknowledge the non-linearity of responses (e.g., Yuste et al., 2004). Researchers have called for greater use of process-based models (Keenan et al., 2020), which often include non-linear responses to temperature, but rely themselves on exploratory methods and descriptive analyses for progress (Chuine et al., 2016). The challenge, then, is to interrogate the implicit and explicit models we use to interpret data summaries, and to develop null expectations that apply across biological and calendar time.

## References

- Bonan, G. B., and L. Sirois. 1992. Air-temperature, tree growth, and the northern and southern range limits to *Picea-mariana*. *Journal of Vegetation Science* 3:495–506.
- Chuine, I., M. Bonhomme, J.-M. Legave, I. García de Cortázar-Atauri, G. Charrier, A. Lacointe, and T. Améglio. 2016. Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break. *Global Change Biology* 22:3444–3460.
- Cook, B. I., E. M. Wolkovich, and C. Parmesan. 2012. Divergent responses to spring and winter warming drive community level flowering trends. *Proceedings of the National Academy of Sciences of the United States of America* 109:9000–9005. Cook, Benjamin I. Wolkovich, Elizabeth M. Parmesan, Camille.
- Cornes, R. C., G. van der Schrier, E. J. van den Besselaar, and P. D. Jones. 2018. An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres* 123:9391–9409.

131 Dai, W. J., H. Y. Jin, Y. H. Zhang, T. Liu, and Z. Q. Zhou. 2019. Detecting temporal changes in  
 132 the temperature sensitivity of spring phenology with global warming: Application of machine  
 133 learning in phenological model. *Agricultural and Forest Meteorology* 279.

134 Dijkhuis, F. J. 1956. Computation of heat unit accumulations in maize for practical application.  
 135 *Euphytica* 5:267–275.

136 Flynn, D. F. B., and E. M. Wolkovich. 2018. Temperature and photoperiod drive spring phe-  
 137 nology across all species in a temperate forest community. *New Phytologist* 219:1353–1362.

138 Ford, K. R., C. A. Harrington, S. Bansal, J. Gould, Peter, and J. B. St. Clair. 2016. Will changes  
 139 in phenology track climate change? A study of growth initiation timing in coast Douglas-fir.  
 140 *Global Change Biology* 22:3712–3723.

141 Fu, Y. S. H., H. F. Zhao, S. L. Piao, M. Peaucelle, S. S. Peng, G. Y. Zhou, P. Ciais, M. T. Huang,  
 142 A. Menzel, J. P. Uelas, Y. Song, Y. Vitasse, Z. Z. Zeng, and I. A. Janssens. 2015. Declining  
 143 global warming effects on the phenology of spring leaf unfolding. *Nature* 526:104–107.

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

147 Hinrichsen, H. H. 2009. Biological processes and links to the physics. *Deep-Sea Research Part*  
 148 *II-Topical Studies in Oceanography* 56:1968–1983.

149 Hofmann, G. E., and A. E. Todgham. 2010. Living in the now: Physiological mechanisms to  
 150 tolerate a rapidly changing environment. *Annual Review of Physiology* 72:127–145.

151 IPCC. 2014. *Climate Change 2014: Impacts, Adaptation, and Vulnerability*. Cambridge Uni-  
 152 versity Press, Cambridge, United Kingdom and New York, NY, USA.

153 Keenan, T. F., A. D. Richardson, and K. Hufkens. 2020. On quantifying the apparent temper-  
 154 ature sensitivity of plant phenology. *New Phytologist* 225:1033–1040.

155 Kopp, C. W., B. M. Neto-Bradley, L. P. J. Lipsen, J. Sandhar, and S. Smith. 2020. Herbarium  
 156 records indicate variation in bloom-time sensitivity to temperature across a geographically  
 157 diverse region. *International Journal of Biometeorology* 64:873–880.

- 158 Kramer, P. 2012. *Physiology of woody plants*. Elsevier, New York.
- 159 Laube, J., T. H. Sparks, N. Estrella, J. Höfler, D. P. Ankerst, and A. Menzel. 2014. Chilling  
160 outweighs photoperiod in preventing precocious spring development. *Global Change Biology*  
161 20:170–182.
- 162 Lechowicz, M. J. 1984. Why do temperate deciduous trees leaf out at different times - adaptation  
163 and ecology of forest communities. *American Naturalist* 124:821–842. Lechowicz, mj.
- 164 Lindsey, A. A., and J. E. Newman. 1956. Use of official weather data in spring time - temperature  
165 analysis of an Indiana phenological record. *Ecology* 37:812–823.
- 166 Linkosalo, T., H. K. Lappalainen, and P. Hari. 2008. A comparison of phenological models of  
167 leaf bud burst and flowering of boreal trees using independent observations. *Tree Physiology*  
168 28:1873–1882.
- 169 Lundell, R., H. Hanninen, T. Saarinen, H. Astrom, and R. Zhang. 2020. Beyond rest and qui-  
170 escence (endodormancy and ecodormancy): A novel model for quantifying plant-environment  
171 interaction in bud dormancy release. *Plant Cell and Environment* 43:40–54.
- 172 Meng, L., J. Mao, Y. Zhou, A. D. Richardson, X. Lee, P. E. Thornton, D. M. Ricciuto, X. Li,  
173 Y. Dai, X. Shi, and G. Jia. 2020. Urban warming advances spring phenology but reduces the  
174 response of phenology to temperature in the conterminous United States. *Proceedings of the*  
175 *National Academy of Sciences* 117:4228.
- 176 Nee, S., N. Colegrave, S. A. West, and A. Grafen. 2005. The illusion of invariant quantities in  
177 life histories. *Science* 309:1236–1239.
- 178 Park, D. S., I. Breckheimer, A. C. Williams, E. Law, A. M. Ellison, and C. C. Davis. 2019.  
179 Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity  
180 across the eastern united states. *Philosophical Transactions of the Royal Society B: Biological*  
181 *Sciences* 374:20170394.
- 182 Park, H., S. J. Jeong, C. H. Ho, C. E. Park, and J. Kim. 2018. Slowdown of spring green-up  
183 advancements in boreal forests. *Remote Sensing of Environment* 217:191–202.



184 Piao, S., Z. Liu, T. Wang, S. Peng, P. Ciais, M. Huang, A. Ahlstrom, J. F. Burkhardt, F. Cheval-  
185 lier, I. A. Janssens, et al. 2017. Weakening temperature control on the interannual variations  
186 of spring carbon uptake across northern lands. *Nature Climate Change* 7:359.

187 Polgar, C., A. Gallinat, and R. B. Primack. 2014. Drivers of leaf-out phenology and their  
188 implications for species invasions: insights from Thoreau’s Concord. *New Phytologist* 202:106–  
189 15.

190 Prevey, J., M. Vellend, N. Ruger, R. D. Hollister, A. D. Bjorkman, I. H. Myers-Smith, S. C.  
191 Elmendorf, K. Clark, E. J. Cooper, B. Elberling, A. M. Fosaa, G. H. R. Henry, T. T. Hoyer,  
192 I. S. Jonsdottir, K. Klanderud, E. Levesque, M. Mauritz, U. Molau, S. M. Natali, S. F.  
193 Oberbauer, Z. A. Panchen, E. Post, S. B. Rumpf, N. M. Schmidt, E. A. G. Schuur, P. R.  
194 Semenchuk, T. Troxler, J. M. Welker, and C. Rixen. 2017. Greater temperature sensitivity of  
195 plant phenology at colder sites: implications for convergence across northern latitudes. *Global*  
196 *Change Biology* 23:2660–2671.

197 Tansey, C. J., J. D. Hadfield, and A. B. Phillimore. 2017. Estimating the ability of plants  
198 to plastically track temperature-mediated shifts in the spring phenological optimum. *Global*  
199 *Change Biology* 23:3321–3334.

200 Templ, B., E. Koch, K. Bolmgren, M. Ungersböck, A. Paul, H. Scheifinger, T. Rutishauser,  
201 M. Busto, F.-M. Chmielewski, L. Hájková, S. Hodzić, F. Kaspar, B. Pietragalla, R. Romero-  
202 Fresneda, A. Tolvanen, V. Vučetič, K. Zimmermann, and A. Züst. 2018. Pan European  
203 Phenological database (PEP725): a single point of access for European data. *International*  
204 *Journal of Biometeorology* 62:1109–1113.

205 Valentine, H. T. 1983. Budbreak and leaf growth functions for modeling herbivory in some  
206 gypsy-moth hosts. *Forest Science* 29:607–617.

207 Vitasse, Y., C. Signarbieux, and Y. H. Fu. 2018. Global warming leads to more uniform spring  
208 phenology across elevations. *Proceedings of the National Academy of Sciences* 115:1004–1008.

209 Wang, C., R. Y. Cao, J. Chen, Y. H. Rao, and Y. H. Tang. 2015. Temperature sensitivity  
210 of spring vegetation phenology correlates to within-spring warming speed over the northern  
211 hemisphere. *Ecological Indicators* 50:62–68.

212 Xu, Y. J., H. J. Wang, Q. S. Ge, C. Y. Wu, and J. H. Dai. 2018. The strength of flowering-  
 213 temperature relationship and pre-season length affect temperature sensitivity of first flowering  
 214 date across space. *International Journal of Climatology* 38:5030–5036.

215 Yuste, J., I. A. Janssens, A. Carrara, and R. Ceulemans. 2004. Annual  $Q_{10}$  of soil respiration  
 216 reflects plant phenological patterns as well as temperature sensitivity. *Global Change Biology*  
 217 10:161–169.

218 Zhang, H. C., W. P. Yuan, S. G. Liu, W. J. Dong, and Y. Fu. 2015. Sensitivity of flowering  
 219 phenology to changing temperature in China. *Journal of Geophysical Research-Biogeosciences*  
 220 120:1658–1665.

221 Zohner, C. M., B. M. Benito, J. D. Fridley, J. C. Svenning, and S. S. Renner. 2017. Spring  
 222 predictability explains different leaf-out strategies in the woody floras of North America,  
 223 Europe and East Asia. *Ecology Letters* 20:452–460.

224 Zohner, C. M., B. M. Benito, J. C. Svenning, and S. S. Renner. 2016. Day length unlikely to  
 225 constrain climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate*  
 226 *Change* 6:1120–1123.

227 *Acknowledgements:* Thanks to TJ Davies, A Donnelly, TM Giants, D. Lipson, C. Rollinson and  
 228 C. Zohner.

229

230 *Data & Code Availability:* Code for simulations and plots is provided here. For empirical ex-  
 231 amples, we used PEP 725 phenological data and E-OBS climate data, both of which are freely  
 232 available via the links.

233

234 *List of Supplementary Information:*

235 A first-hitting-time model of leafout

236 Simulations of common hypotheses for declining sensitivity

237 Methods & results using long-term empirical data (PEP725)

238 Table S1-S2

239 Fig S1-S7

240

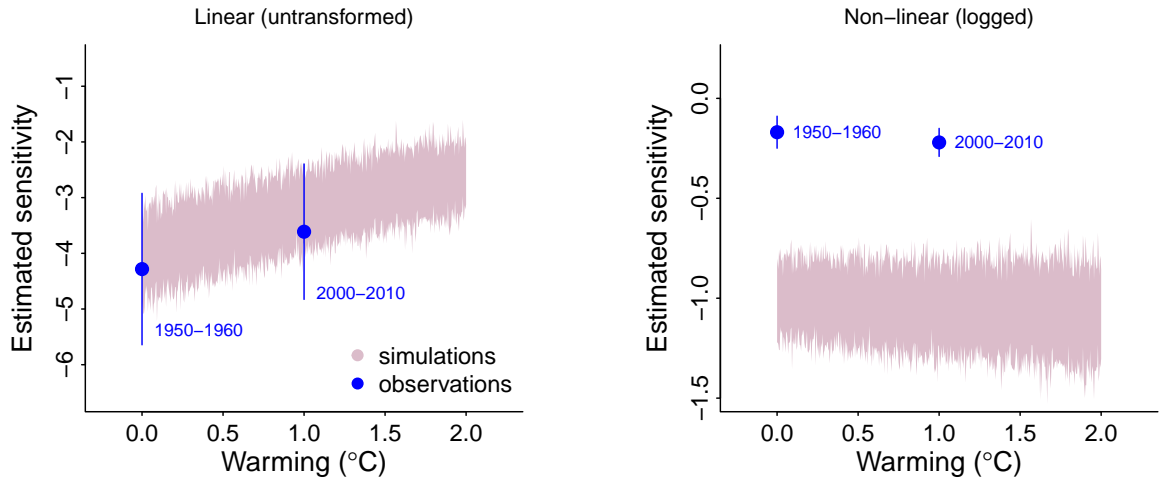


Figure 1: **Shifts in temperature sensitivities (response per °C) with warming occur when using linear models for non-linear processes.** Estimated sensitivities decline with warming in simulations (shading, estimated across 45 sites with a base temperature of normal(6,4), variation comes from fluctuation in the Monte Carlo simulations) with no underlying change in the biological process when sensitivities were estimated with linear regression (left). This decline disappears when performing the regression on logged predictor and response variables (right). Such issues may underlie declining sensitivities calculated from observational data, including long-term observations of leafout across Europe (for *Betula pendula* from PEP725 from for the 45 sites that had complete data for 1950-1960 and 2000-2010), which show a lower sensitivity with warming when calculated on raw data, but no change in sensitivity using logged data. Shading, symbols and lines represent means  $\pm$  standard deviations of regressions across sites. See Supplementary Information for a discussion of why estimated sensitivities are -1 or lower in non-linear models.

## 2 Letter

Editor and reviewer comments are in *italics*, while our responses are in regular text.

### Editor's comments:

*You'll see that while two of the reviewers (1 and 3) are convinced by your study, the other two still need more clarity and detail to persuade them.*

We thank the editor for the opportunity to revise our manuscript. We found the reviewers comments very helpful, with many overlapping requests for clarity or a more careful message. We provide detailed point-by-point responses below.

### Reviewer 1 (A Donnelly) comments:

*The authors present a very interesting and important analysis to demonstrate how a decline in temperature sensitivity with warming is obtained using linear regression but the decline disappears when both variables (temperature and leafout dates) are log transformed. This is a very important and timely finding which has implications for future predictions of the impact of climate change on phenology. I would expect this report to motivate researchers to re-examine trends in phenology using this method. The manuscript is extremely well written, explained and presented. Therefore, I recommend this short report for publications without revision.*

We thank the reviewer for the positive comments on the manuscript's topic and writing style. We have worked to maintain these good aspects of the manuscript while addressing other reviewer comments.

### Reviewer 2 comments:

*This MS provides a simple explanation for the observed non-linearity in biological temperature sensitivity (the observation that at higher temperatures the biological temperature sensitivity per degree Centigrade is decreasing): non-linearity in the biological temperature response. The MS develops a model to prove this point and it is very nicely written. Nevertheless, I found this somewhat unconvincing, because that the explanation appears to be the same as the observation (e.g. crows are black, because crows are black). The main problem appears to be a criticism in how words are used in temporal ecology literature: "sensitivity" versus "response". Indeed, many authors have used the word "(in)sensitivity" rather than the word "response" and this could be criticized, because "(in)sensitivity" implies some fundamental shift in the temperature response mechanism, even though what we are really seeing is an unchanged response. However, taking this into account the MS is reduced to a semantic discussion, in that not the sensitivity but the response is non-linear.*

We understand the reviewer's concerns to be that our finding may seem obvious to some, but we don't see it as a purely semantics debate for several reasons. First, there is a growing use of

shifting or declining sensitivities as evidence of shifts in the underlying biology in the climate change literature, without a robust interrogation of whether this evidence supports the underlying hypothesis. Second, other reviewers of this manuscript here, and additional reviews we gathered from researchers in the field suggest the main crux of our manuscript (that temperature responses are non-linear and thus linear models used to estimate those responses will find declining sensitivities with warming temperatures) was surprising to them. We ourselves were surprised by this result and did not find it intuitive. Third, our simulations show that many models expected to produce declining sensitivities do not (we discuss this more in the following comment), and thus show how critical developing more robust expectations in this area of research may be to progress. Thus, while we understand our findings may not be surprising to some in the field we believe they are novel to many and that highlighting the statistical issue and providing a theoretical underpinning to explain it is an important and useful advance.

*Either way, if temperature becomes a non-limiting factor with warming, other environmental factors will inadvertently take over, changing selective pressures on organisms. Therefore, I also do not see how this simple explanation for non-linearity of temperature sensitivity could discredit the conclusion that fundamental aspects of biological processes are being reshaped by climate warming, as the abstract appears to suggest.*

We agree with the reviewer that our model does not refute that shifting biological processes could underlie observations of declining sensitivities, and we did not mean to suggest otherwise. We fully agree that with continued warming, we should expect a greater role of factors beyond spring temperatures, but we believe our results highlight that most current methods may not capture this well. To address this we have: (1) adjusted our language in the abstract, which we now close with “Current methods may thus undermine efforts to identify when and how warming will reshape biological processes” and in the main text, line 101-line 105:

Our results do not rule out that important shifts in biological processes may underlie observations of declining temperature sensitivities, but provide a simpler explanation for them. Instead, our results highlight how the use of linear models may make identifying when (and potentially why) warming alters underlying biology far more difficult.

(2) We have provided a longer discussion of the expectation that other environmental factors should play a larger role with warming in the main text (line 83-line 97). (3) We have included an extended discussion and analyses of alternative models in the supplement (see new sections ‘Effect of shifting  $\beta$  on sensitivity’ and ‘Simulations of common hypotheses for declining sensitivity’). In our previous draft we provided a simulation where reduced overwinter chilling led to a greater thermal sum requirement with warming, we now explain this simulation in more detail. Interestingly, this simulation does not always produce declining sensitivities—for example, we now show in ‘Effect of shifting  $\beta$  on sensitivity’ that increasing the thermal sum for leafout would produce *increasing* (i.e., more negative) sensitivities estimated from linear models and the exact predictions for how sensitivities should shift with climate change depend strongly on the underlying model of chilling, forcing and how warming co-varies with each. We now also provide one photoperiod simulation as well, which also highlights the complexity of teasing out

other cues. While the focus of our manuscript remains on the statistical issue of using linear models for non-linear temperature processes with warming—which extends beyond spring plant phenology—we hope these examples better highlight the challenges of modeling these processes with current data and show that our assumptions from verbal models may not always be accurate.

*The MS could narrow its focus to just being about the difference between “sensitivity” and “response” to temperature, and that it is important to provide accurate definitions. Moreover, it could urge the field not to overreach in their conclusions on biological temperature responses.*

We agree and have worked to balance the manuscript better to contrast our findings with biological expectations given warming, and how well we can detect them. We have also reviewed our uses of the terms, ‘sensitivity’ and ‘response’ for accuracy and clarity; for example, see changes on line 44, line 47, line 73 and in main text figure caption.

*On a different note, the number of citations is inadequate. For example, on p4 paragraph 1 is the first mention of “thermal sum” without any citation. This happens more often, and should be revised thoroughly.*

Our updated manuscript includes 13 additional citations (our original draft included 12), see in particular line 49, line 51, line 65.

*I have no small textual comment, as the MS is very well written.*

### **Reviewer 3 (CM Zohner) comments:**

*I really like this study. It is an important warning to phenologists that the commonly-used temperature sensitivity metric can be dangerously misleading. The modeling part makes sense to me but I wonder whether more information is required at some parts to allow readers to fully grasp what was done here. For instance, it is not totally clear to me how the PEP data was analysed. How did you determine the preseason lengths (which of course are also problematic per se) on which temperature sensitivities are ultimately based on? I recognize that this is a short report that does not allow for too many details but I think that the Supplementary part should still contain all necessary information to fully understand what has been done here.*

We are glad to hear the reviewer appreciated our manuscript and agree we could do more to make our methods clearer. We have updated our supplemental section, ‘Methods & results using long-term empirical data (PEP725)’ to better explain our methods. We used pre-defined pre-seasons; we are now more explicit about this and for transparent provided two shorter pre-season windows (see updated tables S1-S2). The relevant text now reads:

To examine how estimated sensitivities shift over time, we selected sites of two common European tree species (silver birch, *Betula pendula*, and European beech, *Fagus sylvatica*) that have long-term observational data of leafout, through the Pan Euro-

pean Phenology Project (PEP725, Templ et al., 2018). We selected these two species given that they were best represented for consistent data at the same sites over our study years for an early-leafout (*Betula pendula*) and a late-leafout (*Fagus sylvatica*) species (e.g., *Betula pendula* had 17 sites with leafout data from 1950-1960 and 2000-2010, while the next best option for an early-leafout species, *Alnus glutinosa*, had data for only five sites). We used sites with complete leafout data across both our 10-year (and 20-year) windows to avoid possible confounding effects of shifting sites over time (see Tables S1-S2 for numbers of sites per species x window).

To calculate temperature sensitivities, we used a European-wide gridded climate dataset (E-OBS, *Cornes et al., 2018*) to extract daily minimum and maximum temperature for the grid cells where observations of leafout for these two species were available (Fig. S5 shows a subset of the climate data for 14 sites used). Determining the appropriate window over which to estimate a temperature sensitivity for spring plant phenology is an area of active research (*Güsewell et al., 2017; Xu et al., 2018*). Ideally researchers wish to separate windows over which chilling and forcing apply (if they can be cleanly separated, *Linkosalo et al., 2008; Lundell et al., 2020*), but this is generally impossible given our limited understanding of the two processes (*Chuine et al., 2016*). Researchers thus either use a pre-defined spring window (e.g., *Zhang et al., 2015; Park et al., 2018, 2019; Kopp et al., 2020*) or use a statistical search to determine window attributes; for example, some use a set period then search for a start date (e.g., *Cook et al., 2012; Wang et al., 2015*), while others search for both a start date and window length (e.g., *Fu et al., 2015; Tansey et al., 2017*). Given the non-identifiability of the simple chilling + forcing model described above (‘Simulations of common hypotheses for declining sensitivity’), we do not feel there is sufficient evidence that statistical searches for pre-season windows will select biologically relevant periods related to forcing, and more easily could add additional layers of non-identifiability (to date, most pre-season windows are fit as separate analyses making such issues harder for researchers to identify). Thus, we used pre-determined windows from 1 March to 30 April (60 d; we also present windows of 45 d, from 1 Mar to 15 Apr, and for 31 d, from 15 Mar to 15 Apr, for comparison). Our window represents a period in the spring when forcing is likely the dominant cue, and is similar to many other studies of temperature plants using pre-defined windows (e.g., *Prevey et al., 2017; Park et al., 2019*). We then estimated the sensitivity as a simple linear regression of leafout day of year versus the mean daily temperature over the window (‘slope’ in Tables S1-S2), or a simple linear regression using logged versions of these predictors (‘log-slope’ in Tables S1-S2).

*Additionally, I think it is important to emphasize in the text that, although the days per C metric does not seem to be adequate to detect changes in the phenological responses to temperature, it is still likely that these changes are happening at the moment. You could for example cite some experimental studies, including your own meta-analysis of these studies, that clearly show that the required warming to leaf-out is increasing as winters become shorter, e.g., refs(1-4). The linear dashed line in Figure 3b in ref(5) also demonstrates that, while winter chilling and day*



length seem to have large effects on the amount of warming (degree-days) to leaf-out, the apparent temperature sensitivity doesn't really change, otherwise one should see a flattening curve. So, while I totally agree that the days per C metric can be totally misleading, I think the paper would benefit from a short discussion on the importance of spring phenology drivers apart from spring temperature to emphasize that they are important, although probably not directly visible in the days per C metric.

This is a good point and somewhat related to points raised by Reviewer #2. To address it we have, (1) changed the ending the abstract to "Current methods may thus undermine efforts to identify when and how warming will reshape biological processes," (2) provided a longer section in the manuscript, line 83-line 97, which reads (and now cites most of the suggested citations):

Fundamentally rising temperature should alter many biological processes, making robust methods for identifying these changes critical. In spring plant phenology, where declining sensitivities are often reported (Fu et al., 2015; Piao et al., 2017; Dai et al., 2019), warming may increase the role of 'chilling' (determined mainly by winter temperatures) and daylength (Laube et al., 2014; Zohner et al., 2016)—potentially increasing the thermal sum required for leafout at lower values of these cues (Polgar et al., 2014; Zohner et al., 2017; Flynn and Wolkovich, 2018). Adjusting our simulations to match this model yielded shifts in sensitivities with warming. Unlike a model with no underlying biological change, however, after correcting for non-linearity the shifts in sensitivities remained and they occurred in step with the biological change (Fig. S4a, c). In contrast, sensitivities estimated from a linear model showed shifts across the entire range of warming, well before the simulated biological change (Fig. S4a, c). Further, we found that an increase in the thermal sum required for leafout should yield larger temperature sensitivities, not smaller, as often expected (e.g., Fu et al., 2015), thus highlighting the complexity of what trends to expect in sensitivities with warming (see 'Common hypotheses for declining sensitivity' in Supplementary Information for an extended discussion).

## References

1. J. Laube, T. H. Sparks, N. Estrella, J. Hfler, D. P. Ankerst, A. Menzel, *Chilling outweighs photoperiod in preventing precocious spring development. Glob. Chang. Biol. 20, 170-182 (2014).*
2. C. Polgar, A. Gallinat, R. B. Primack, *Drivers of leaf-out phenology and their implications for species invasions: Insights from Thoreau's Concord. New Phytol. 202, 106-115 (2014).*
3. C. M. Zohner, B. M. Benito, J. D. Fridley, J. C. Svenning, S. S. Renner, *Spring predictability explains different leaf-out strategies in the woody floras of North America, Europe and East Asia. Ecol. Lett. 20, 452-460 (2017).*
4. D. F. B. Flynn, E. M. Wolkovich, *Temperature and photoperiod drive spring phenology across all species in a temperate forest community. New Phytol. 219, 1353-1362 (2018).*
5. C. M. Zohner, L. Mo, T. A. M. Pugh, J. F. Bastin, T. W. Crowther, *Interactive climate factors restrict future increases in spring productivity of temperate and boreal trees. Glob. Chang. Biol. 26, 4042-4055 (2020).*

#### Reviewer 4 comments:

*In this short manuscript, the authors intended to show that rising temperatures combined with linear estimates based on calendar time are expected to produce observations of declining sensitivity without any shift in the underlying biology. While the topic is very interesting and the conclusion is eye-catching, I found that the reasoning and presentation of the manuscript deserved much more reconsideration and the conclusions were not convincing.*

*My general feeling is that the manuscript was presented in a form that was very hard to follow. For example, the authors presented a stochastic experiment, whose details were buried in the supplementary information. Many critical details of the experiments were missing, making it difficult to assess. The authors later argued a “correction” for non-linearity, but the theorem of this correction and its caveats were not presented. Coming after that, the authors intended to test their theory with the observation dataset (PEP725), but critical details on how they calculated the temperature sensitivity was not clear, which, as Keenan et al. (2020) mention, is a fundamental issue in deriving temperature sensitivity. I have to read the short manuscript for several times, tracing backward and forward between the main-text and SI, and taking some guesses to what they have done. Thus, I strongly recommend the authors to make a full-length paper with clear presentation on their methods and reasoning. In its current form, the manuscript is not publishable in my opinion.*

We appreciate the reviewers concerns that we should have provided more clear methods for our PEP725 analyses, and this was also a concern of Reviewer 3. To address this we have greatly extended the section in the supplement on the PEP725 analysis, providing full methods (see also the response above to Reviewer 3’s first point, where we show much of the relevant text).

Our stochastic experiment is designed to be simple—a model where an event is triggered based on a thermal sum threshold, as we describe in the main text. We fully describe the stochastic model both mathematically and in a longer text format in the supplement. The correction we suggest (log transformation) is the natural transformation for an inverse relationship (by definition a log linearizes this relationship) so we are unsure how to further describe this. While we could move the mathematical description of the model into the main text, we feel that will narrow the readership unnecessarily—as the model is fundamentally very simple and already described in the main text. We provide the supplement math as a more developed proof for those interested. Based on several reviews by those in this research area, including those of the other three reviewers here, we believe this format of the manuscript allows readers to understand our main point, analyses and findings.

*A major argument made by the authors was a stochastic experiment: when leafout date was determined by a constant cumulative daily temperature, a declining apparent temperature sensitivity could be derived. However, this argument has two limitations: (1) the assumption of constant threshold of cumulative heat is a very strong one, and not supported by observations*

*(e.g. Fu et al., 2014); (2) even we accepted the strong assumption, it does not prove that the observed decline in apparent temperature sensitivity comes from unchanged heat accumulation. As the authors showed, increasing heat requirements could indeed yield declining apparent temperature sensitivity. As a reader, I do not see the value of this stochastic experiment, which did not successfully prove the points made by the author.*

We did not mean to suggest that our model proves declining sensitivities are driven by the model we present and we should have made this clearer. As noted above in our reply to Reviewer 2, to address this we have: (1) adjusted our language in the abstract, which we now close with “Current methods may thus undermine efforts to identify when and how warming will reshape biological processes” and in the main text, line 101-line 105:

Our results do not rule out that important shifts in biological processes may underlie observations of declining temperature sensitivities, but provide a simpler explanation for them. Importantly, our results highlight how the use of linear models may make identifying when (and potentially why) warming alters underlying biology far more difficult.

Regarding the assumption of the constant threshold of cumulative heat, we have added a new section (‘Effect of shifting  $\beta$  on sensitivity’) that examines the effects of relaxing this assumption and find that, perhaps contrary to expectations, an shifting thermal sum would produce increasing, not declining sensitivities. We did not include this before as we aimed to test a simple model so that we could understand expected shifts in sensitivities without this effect. We are not aware of strong observational evidence that this threshold is varying (Fu et al. 2014 in *PLOS One* and Fu et al. 2015 in *Nature* both use methods similar to the we critique here to suggests shifting thresholds, but we may be reviewing the wrong papers), though we now clearly state and cite experimental evidence that these thresholds can shift (line 83-line 88):

Fundamentally rising temperature should alter many biological processes, making robust methods for identifying these changes critical. In spring plant phenology, where declining sensitivities are often reported (Fu et al., 2015; Piao et al., 2017; Dai et al., 2019), warming may increase the role of ‘chilling’ (determined mainly by winter temperatures) and daylength (Laube et al., 2014; Zohner et al., 2016)—potentially increasing the thermal sum required for leafout at lower values of these cues (Polgar et al., 2014; Zohner et al., 2017; Flynn and Wolkovich, 2018).

In addressing concerns from Reviewers 2-3, we also provide an extended discussion of the complexity of outcomes from models that predicting shifting thermal sums based on other cues for spring phenological events. While many of the authors ourselves assumed that sensitivities to spring temperatures should decline as other cues come into play, this is not always the case given the co-variation of many of these cues, and highlights the value of explicit models to challenge assumptions.

*In the next paragraph, the authors argued that a log transformation could make the declining temperature sensitivity disappear. While this may be true, an issue may arise as the log transform would make the regression sensitive to a few outliers and it will necessarily hide change in*

*temperature sensitivity, if it happens. So I hardly see how the disappearance of linearity after log-transformation proves that non-linearity did not exist in the observations.*

Log transformations make regressions less (not more) sensitive to outliers, which we assume is what the reviewer means, though we're not clear if the concern here is regarding outliers in the predictor or response, or both. Our argument is that a log is the correct transformation for the data given the model that plants leafout after reaching a threshold thermal sum. As mentioned above, we do not mean to suggest that our model in any way proves that there are not biological shifts underlying the PEP 725 data, but that our current statistical methods may not be robust to detecting to these changes, and we have edited the abstract and main text to clarify this.

*Also about the analyses on the observation dataset: the PEP dataset includes much more species than the two used here, why selectively analyzing these two species?*

Good catch, we now clarify our methods regarding species selection in the supplement ('Methods & results using long-term empirical data (PEP725)'). We write:

We selected these two species given that they were best represented for consistent data at the same sites over our study years for an early-leafout (*Betula pendula*) and a late-leafout (*Fagus sylvatica*) species (e.g., *Betula pendula* had 17 sites with leafout data from 1950-1960 and 2000-2010, while the next best option for an early-leafout species, *Alnus glutinosa*, had data for only five sites).

For clarity we provide on the following pages the number of sites for other major species in the PEP725 data, including our analyses of the species with the third most data for our analyses, *Quercus robur*.

*Overall, the stochastic experiment showed that declining apparent temperature sensitivity could have some alternative explanation, but it is far from sufficient to conclude that the observed declining temperature sensitivity was from the authors' assumption. The empirical analyses were not performed in a transparent and convincing way. Therefore, I could not draw similar conclusions as the authors did.*

Our goal with the stochastic experiment was to show, as this reviewer writes, "that declining apparent temperature sensitivity could have some alternative explanation," and thus we believe our manuscript, on some level, did communicate our main finding. We have worked to adjust our language to make it clear we do not mean to state that we have shown this is the definitive explanation, but instead that it is an explanation that should be better considered in analyses using declining sensitivities with warming as evidence of shifting biology with warming. We believe our revisions have further clarified our methods, analyses and conclusions and thank the reviewer for this.