Supplemental materials: A simple explanation for declining temperature sensitivity with warming

E. M. Wolkovich J. Auerbach, C. J. Chamberlain, D. M. Buonaiuto,
A. K. Ettinger, I. Morales-Castilla & A. Gelman

1 A first-hitting-time model of leafout

Our model follows the general understanding of how warm temperatures (forcing) trigger leafout in temperate deciduous trees (Chuine, 2000). We use a first-hitting-time model, which describes the first time a random process hits a threshold, because of its broad applicability and conceptual simplicity. We define leafout day, n_β, as the day, n, that cumulative daily temperature, S^n, hits the threshold, β.

We derive the relationship between daily temperature and leafout in two common scenarios. In the first, we take the average daily temperature up until the leafout date. In the second, we take the average daily temperature over a fixed window, such as March 1st to April 30th. In both cases, we discretize time since, although many biological processes depend continuously on time, research typically measures time in discretized units, such as days, weeks, or months.

1.1 Scenario 1: Using average daily temperature until the leafout date

We use the following notation:

$$
\begin{aligned}
n &= \text{day since temperatures start to accumulate, } n = 0, 1, ..., N \\
X_n &= \text{observed temperature on day } n \\
S_0^n &= \sum_{i=0}^{n} X_i, \text{the cumulative daily temperature from day 0 to day } n \\
M_0^n &= \frac{S_0^n}{n}, \text{the average daily temperature from day 0 to day } n \\
\beta &= \text{the threshold of interest, } \beta > 0, \text{ (for example, F* or required GDD)} \\
n_\beta &= \underset{n}{\text{argmin}}\ S_n > \beta, \text{the first day the cumulative daily temperature passes the threshold} \\
&\quad\ \text{(for example, day of year (doy) of budburst).}
\end{aligned}
$$

We model X_n as a Gaussian random walk, $X_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(\alpha_0 + \alpha_1 n, \sigma)$, where $\alpha_0 > 0$ is the average temperature on day $n = 0$, $\alpha_1 > 0$ is the day-over-day increase in average temperatures, and σ is the standard deviation. Note that this model differs from the traditional Gaussian random walk because of the factor n.

This model has two important consequences:

(1) Leafout time is inversely related to average temperature at leafout time.

Under this model, $M_0^{n_\beta}$ and n_β are inversely proportional. To see why, assume for the moment that the cumulative daily temperature hits the threshold exactly on leafout day. That is, $S_0^{n_\beta} = \beta$. Then

$$M_0^{n_\beta} = \frac{S_0^{n_\beta}}{n_\beta} = \frac{\beta}{n_\beta}$$

rearranging yields

$$n_\beta = \frac{\beta}{M_0^{n_\beta}}$$

Much research uses linear regression to quantify the relationship between n_β and $M_0^{n_\beta}$. Regressing n_β on $M_0^{n_\beta}$ finds a best fit line to the inverse curve, $n_\beta = \frac{\beta}{M_0^{n_\beta}}$. The relationship is linearized with the logarithm transformation: $\log(n_\beta) = \log(\beta) - \log(M_0^{n_\beta})$. That is, $\log(n_\beta)$ is linear in log-average daily temperature with slope -1 and intercept $\log(\beta)$.

(2) The variance of the average temperature may decreases as temperatures rise.

Under the model, the mean and variance of M_0^n is $\mathrm{E}(M_0^n|\alpha_0, \alpha_1) = \frac{1}{n}\sum_{i=0}^n(\alpha_0 + \alpha_1 i) = \alpha_0 + \alpha_1\frac{(n+1)}{2}$ and $\mathrm{Var}(M_0^n|\alpha_0, \alpha_1) = \frac{\sigma^2}{n}$.

By the law of total variance,

$$\begin{aligned}
\mathrm{Var}(M_0^n) &= \mathrm{E}(\mathrm{Var}(M_0^n|\alpha_0, \alpha_1)) + \mathrm{Var}(\mathrm{E}(M_0^n|\alpha_0, \alpha_1)) \\
&= \frac{\sigma^2}{n} + \mathrm{Var}(\alpha_0 + \alpha_1\frac{n+1}{2}) \\
&= \frac{\sigma^2}{n} + \mathrm{Var}(\alpha_0) + \frac{(n+1)^2}{4}\mathrm{Var}(\alpha_1) + (n+1)\mathrm{Cov}(\alpha_0, \alpha_1)
\end{aligned}$$

As temperatures rise and leafout date becomes earlier, the variance of the average temperature will decline—provided the variation in temperatures, σ^2, is sufficiently small.

1.2 Scenario 2: Using average daily temperature over a fixed window

We slightly modify the notation:

$$
\begin{aligned}
n &= \text{day since temperatures start to accumulate, } n = 0, \ldots, a, \ldots, b \\
X_n &= \text{observed temperature on day } n \\
S_a^n &= \sum_{i=a}^{n} X_i, \text{the cumulative daily temperature from day } a \text{ to day } n \\
M_a^n &= \frac{S_a^n}{n-a}, \text{the average daily temperature from day } a \text{ to day } n \\
\beta &= \text{the threshold of interest, } \beta > 0, \text{ (for example, F* or required GDD)} \\
n_\beta &= \underset{n}{\operatorname{argmin}}\ S_0^n > \beta, \text{the first day the cumulative daily temperature passes the threshold} \\
&\quad \text{(for example, day of year (doy) of budburst).}
\end{aligned}
$$

As before, we model X_n as a Gaussian random walk, $X_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(\alpha_0 + \alpha_1 n, \sigma)$, where $\alpha_0 > 0$ is the average temperature on day $n = 0$, $\alpha_1 > 0$ is the day-over-day increase in average temperatures, and σ is the standard deviation. We make the additional assumption that $X_n > 0$ for all n and $a < n_\beta < b$. That is, the cumulative temperature acquired by the tree always increases.

Note that

$$S_a^b \sim \mathcal{N}\left(\alpha_0(b-a) + \tfrac{\alpha_1}{2}(b-a)(b+a+1), \sigma\sqrt{b-a}\right)$$

$$M_a^b \sim \mathcal{N}\left(\alpha_0 + \tfrac{\alpha_1}{2}(b+a+1), \tfrac{\sigma}{\sqrt{b-a}}\right)$$

$$S_n^b - S_a^b \sim \mathcal{N}\left(\alpha_0(b-a-n) + \tfrac{\alpha_1}{2}((b-n)(b+n+1) - a(a+1)), \sigma\sqrt{b+a-n}\right)$$

so that

$$
\begin{aligned}
Pr\left(n_\beta \le n \mid M_a^b = m\right) &= Pr\left(n_\beta \le n \mid S_a^b = (b-a)m\right) \\
&= Pr\left(S_0^n \ge \beta \mid S_a^b = (b-a)m\right) \\
&= Pr\left(S_n^b \le (b-a)m + S_0^a - \beta\right) \\
&= Pr\left(S_n^b - S_0^a \le (b-a)m - \beta\right) \\
&= \Phi\left(\frac{(b-a)m - \beta - [\alpha_0(b-a-n) + \frac{\alpha_1}{2}((b-n)(b+n+1) - a(a+1))]}{\sigma\sqrt{b+a-n}}\right)
\end{aligned}
$$

The distribution of M_a^b shows that consequence (2) above still holds with this model. Consequence (1) no longer holds directly, but will in many situations where average daily temperature until an event correlates strongly with average daily temperature calculated over a window (Figs. S1-S2; this correlation is likely common because researchers generally select temporal windows over which to calculate average temperatures to be highly relevant to the event or process). We note two additional consequences:

(3) The conditional median is quadratic in n:

$$
\begin{aligned}
\frac{1}{2} &\overset{\text{set}}{=} Pr\left(n_\beta \leq n \mid M_a^b = m\right) \\
&\Rightarrow 0 = (b-a)m - \beta - \left[\alpha_0(b-a-n) + \frac{\alpha_1}{2}((b-n)(b+n+1) - a(a+1))\right] \\
&\Rightarrow m = \frac{1}{(a-b)}\left[-\beta - \alpha_0(b-a-n) - \frac{\alpha_1}{2}((b-n)(b+n+1) - a(a+1))\right] \\
&\quad = \frac{1}{(a-b)}\left[-\beta - \alpha_0(b-a) - \frac{\alpha_1}{2}(b-a)(b+a+1)\right] + \frac{\alpha_0 + \frac{\alpha_1}{2}}{(a-b)}n + \frac{\frac{\alpha_1}{2}}{(a-b)}n^2 \\
&\quad := \gamma_0 + \gamma_1 n + \gamma_2 n^2
\end{aligned}
$$

(4) The conditional mean and variance are sums of negative sigmoids, according to the following identities

$E\left(n_\beta \mid M_a^b = m\right) = \sum_{n=0}^{\infty} Pr\left(n_\beta \geq n \mid M_a^b = m\right)$

$E\left(n_\beta^2 \mid M_a^b = m\right) = \sum_{n=0}^{\infty} n \; Pr\left(n_\beta \geq n \mid M_a^b = m\right)$

2 Results using long-term empirical data from PEP725

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 European Phenology Project (PEP725, Templ et al., 2018). 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. 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).

Our estimates of temperature sensitivity from a linear model using untransformed variables shows a decline in sensitivity with recent warming for *Betula pendula* over 10 and 20-year windows, but no decline for *Fagus sylvatica*; using logged variables estimates appeared more similar over time or sometimes suggested an increase sensitivity (see Figs. S4-S5, Tables S1-S2). This shift in estimated sensitivity when regressing with untransformed versus logged variables suggests these declining estimates from untransformed variables may not be caused by changes in the underlying mechanisms of leafout (i.e., reduced winter chilling) and driven instead by using linear regression for a non-linear process. This hypothesis is supported further by large declines in variance of leafout in recent decades.

Shifts in variance provide another hurdle to robust estimates of temperature sensitivity. Previous work has highlighted how shifting temperature variance (over space and/or time) could lead to shifting estimates of temperature sensitivities (Keenan et al., 2020), but our results stress that variance in both leafout and temperature are shifting. If both shift in step, estimates would not be impacted by changes in temperature variance, but our results suggest variance in temperature—for these data—has declined more than variance in leafout, though both have declined substantially in recent decades (Tables S1-S2).

These results highlight how the acceleration of biological time due to climate change requires researchers to clarify their assumptions. Expecting temperature sensitivity to remain constant as temperatures rise assumes the relationship between response and temperature is proportional. But the underlying biological processes suggest this relationship is seldom proportional, or even linear.

References

Chuine, I. 2000. A unified model for budburst of trees. Journal of Theoretical Biology 207:337–347.

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.

Keenan, T. F., A. D. Richardson, and K. Hufkens. 2020. On quantifying the apparent temperature sensitivity of plant phenology. New Phytologist 225:1033–1040.

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

3 Tables

Table S1: Climate and phenology statistics for two species (*Betula pendula, Fagus sylvatica*, across 45 and 47 sites respectively) from the PEP725 data across all sites with continuous data from 1950-1960 and 2000-2010. ST is spring temperature from March 1 to April 30, ST.leafout is temperature 30 days before leafout, and GDD is growing degree days 30 days before leafout. Slope represents the estimated sensitivity using untransformed leafout and ST, while log-slope represents the estimated sensitivity using log(leafout) and log(ST). We calculated all metrics for for each species x site x 10 year period before taking mean or variance estimates. See also Fig. S4.

		mean(ST)	mean(ST.leafout)	var(ST)	var(leafout)	mean(GDD)	slope	log-slope
1950-1960	*Betula pendula*	5.6	7.0	3.4	110.5	71.7	-4.3	-0.17
2000-2010	*Betula pendula*	6.6	6.8	1.2	47.0	64.6	-3.6	-0.22
1950-1960	*Fagus sylvatica*	5.6	7.5	3.3	71.9	83.8	-2.8	-0.11
2000-2010	*Fagus sylvatica*	6.7	7.7	1.2	38.3	86.7	-3.4	-0.20

Table S2: Climate and phenology statistics for two species (*Betula pendula, Fagus sylvatica*, across 17 and 24 sites respectively) from the PEP725 data across all sites with continuous data from 1950-2010. ST is spring temperature from March 1 to April 30, ST.leafout is temperature 30 days before leafout, and GDD is growing degree days 30 days before leafout. Slope represents the estimated sensitivity using untransformed leafout and ST, while log-slope represents the estimated sensitivity using log(leafout) and log(ST). We calculated all metrics for for each species x site x 20 year period before taking mean or variance estimates. See also Fig. S5.

		mean(ST)	mean(ST.leafout)	var(ST)	var(leafout)	mean(GDD)	slope	log-slope
1950-1970	*Betula pendula*	5.8	7.1	2.6	79.9	72.5	-4.3	-0.19
1970-1990	*Betula pendula*	5.9	7.2	1.3	104.8	72.2	-6.1	-0.33
1990-2010	*Betula pendula*	6.8	6.7	0.9	36.2	60.0	-3.3	-0.21
1950-1970	*Fagus sylvatica*	5.6	7.6	2.7	63.4	86.0	-3.1	-0.12
1970-1990	*Fagus sylvatica*	5.6	7.5	1.3	56.2	81.3	-2.5	-0.12
1990-2010	*Fagus sylvatica*	6.7	7.3	1.2	31.4	76.0	-3.4	-0.19

4 Figures

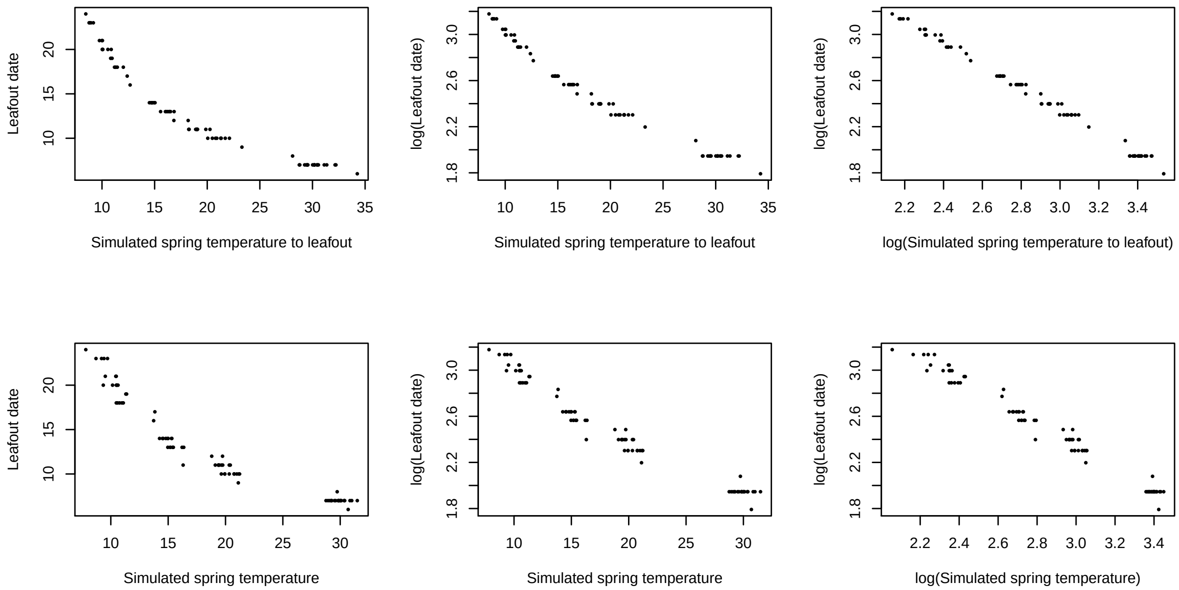

Figure S1: **Simulated leafout as a function of temperature across different temperatures highlights non-linearity of process.** Here we simulated sets of data where leafout constantly occurs at 200 growing degree days across mean temperatures of 0, 5, 10 and 20C (constant SD of 4), we calculated estimated mean temperature or until leafout date (top row) or across a fixed window (bottom row, similar to estimates of 'spring temperature'). While within any small temperature range the relationship may appear linear, its non-linear relationship becomes clear across the greater range shown here (left). Taking the log of leafout (middle) reduces this some, but taking the log of both leafout and temperature (right) further linearizes the relationship.

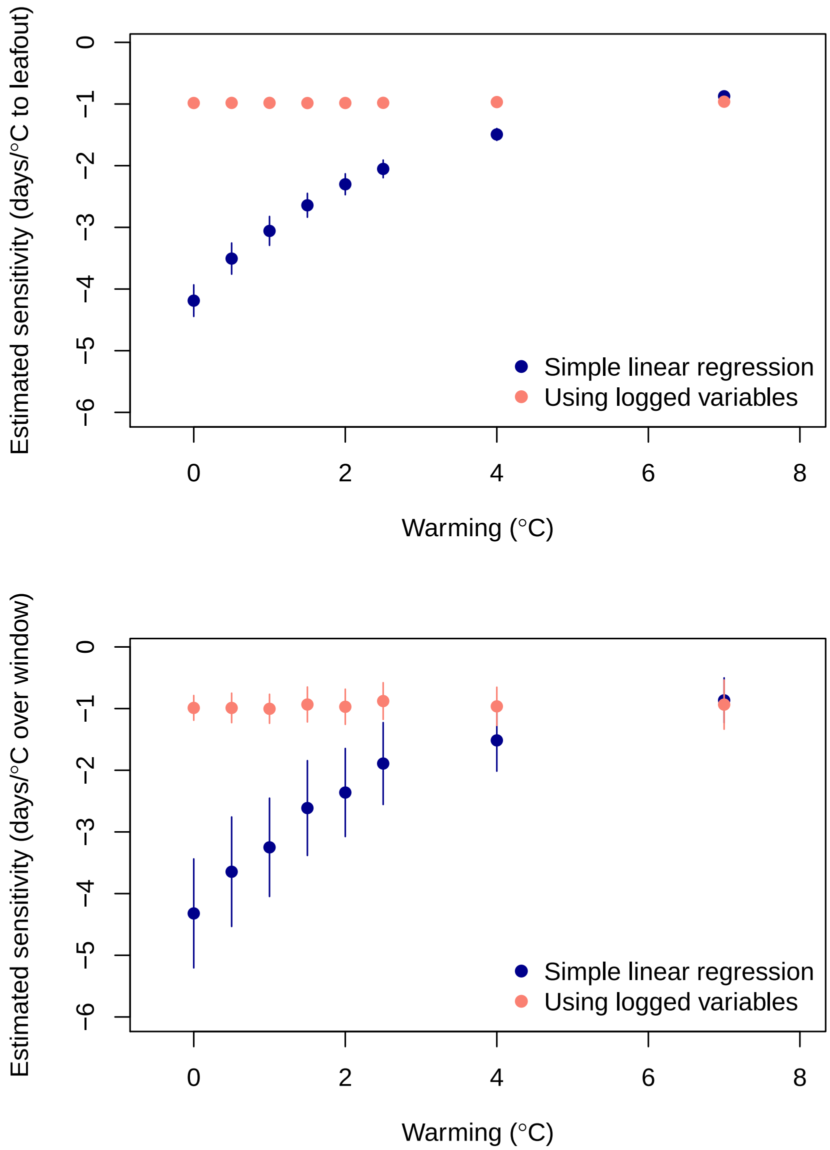

Figure S2: **A simple model generates declining sensitivities with warming.** We show declines in estimated sensitivities with warming from simulations (top: using average temperature until leafout, bottom: using a fixed window) with no underlying change in the biological process when sensitivities were estimated with simple linear regression ("Simple linear regressions"). This decline disappears using regression on logged predictor and response variables ("Using logged variables").

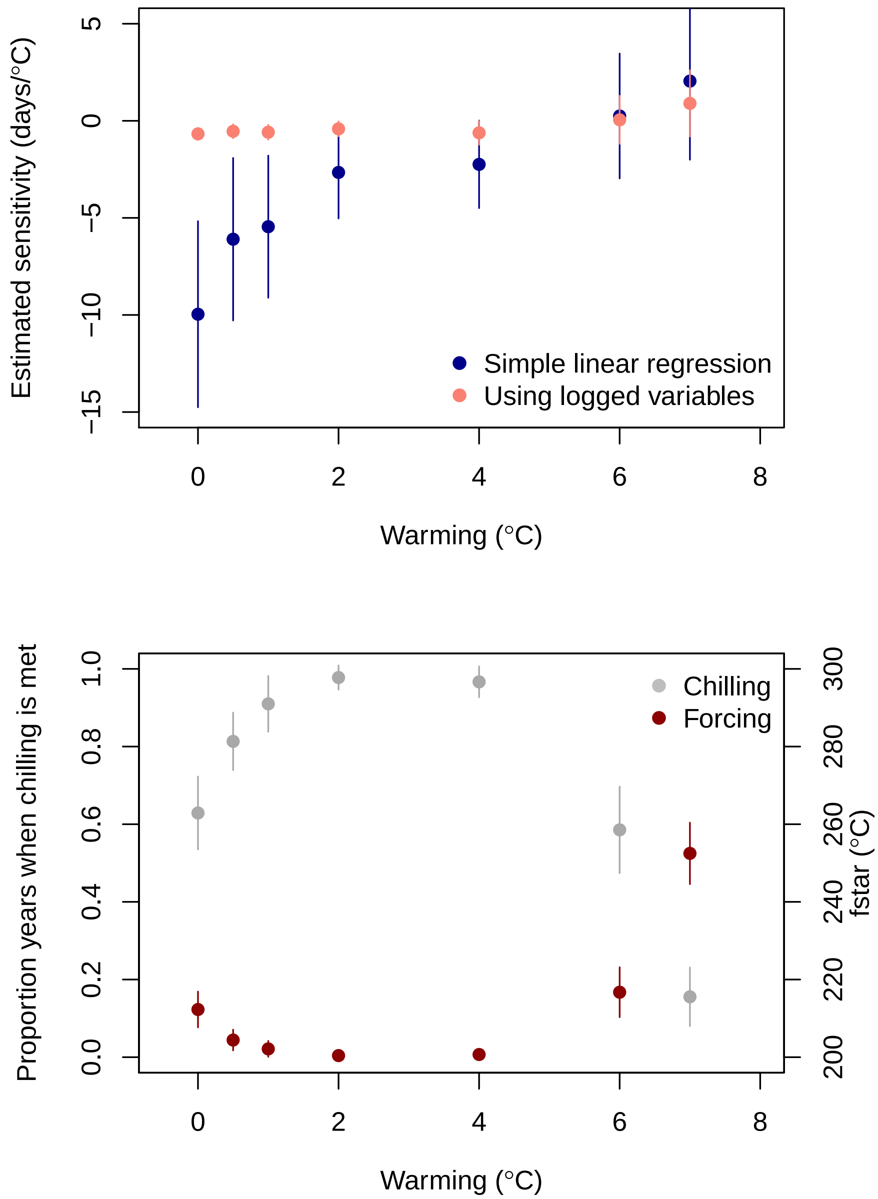

Figure S3: **Simulated leafout as a function of temperature across different temperatures with shifts in underlying cues.** Here we simulated sets of data where leafout occurs at 200 growing degree days ('fstar') when chilling is met, and requires additional growing degree days when chilling is not met. We show estimates sensitivities in the top panel, and the shifting cues on the bottom panel.

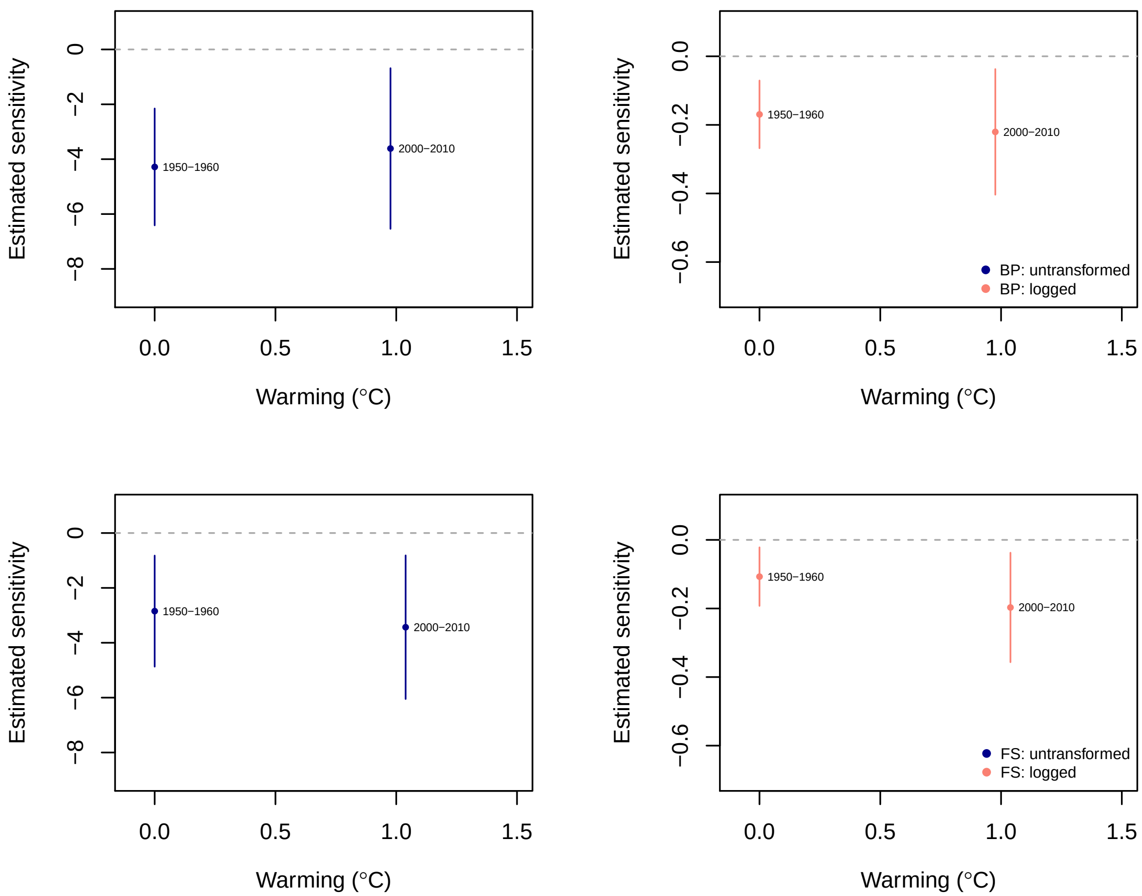

Figure S4: **Sensitivities from PEP725 data using 10 year windows of data** for two species (top – *Betula pendula*, bottom – *Fagus sylvatica*; all lines show 78% confidence intervals from linear regressions). Amounts of warming are calculated relative to 1950-1960 and we used only sites with leafout data in all years shown here. Both approaches show variation in sensitivity across time. See Table S1 for further details.

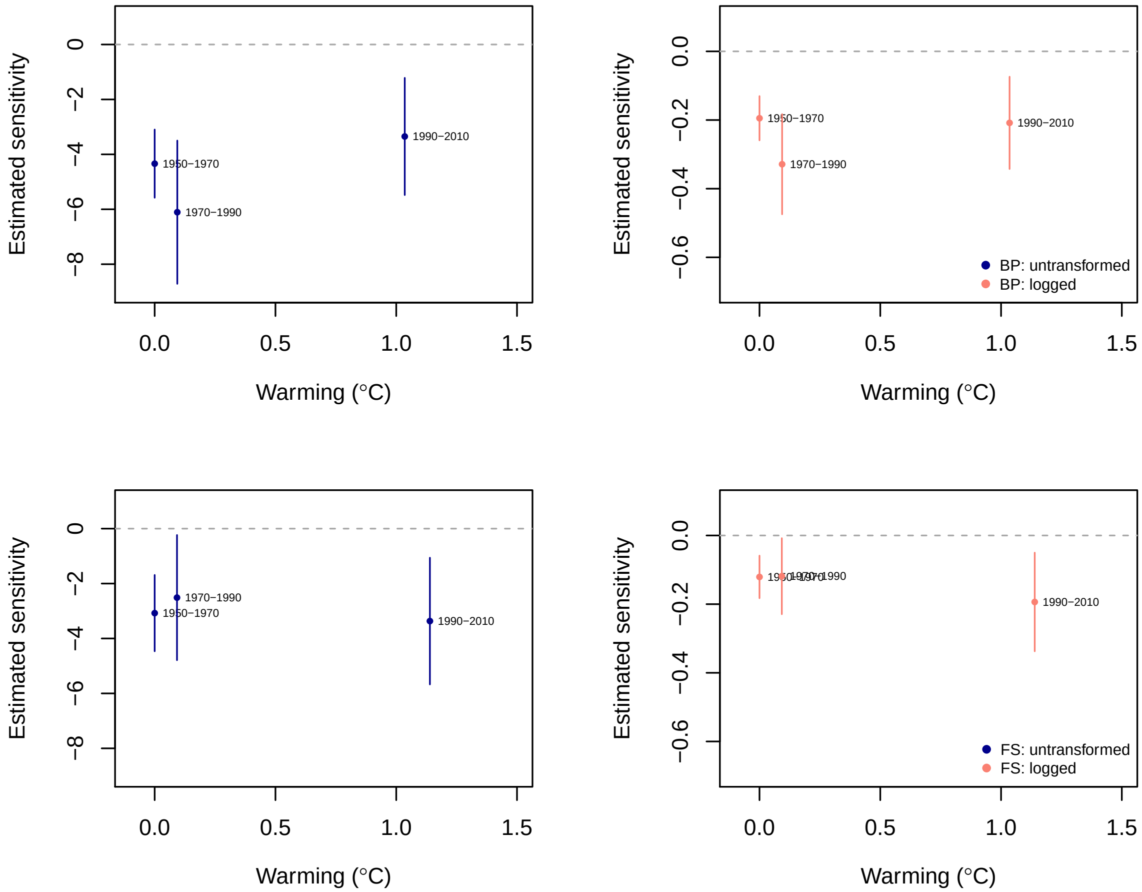

Figure S5: **Sensitivities from PEP725 data using 20 year windows of data** for two species (top – *Betula pendula*, bottom – *Fagus sylvatica*; all lines show 78% confidence intervals from linear regressions). Amounts of warming are calculated relative to 1950-1970 and we used only sites with leafout data in all years shown here. Both approaches show variation in sensitivity across time. See Table S2 for further details.