Dear Editor:

We are pleased to resubmit our manuscript "Warmer spring temperatures in temperate deciduous forests advance the timing of tree growth but have little effect on annual woody productivity", by Cameron Dow and co-authors, for consideration for publication in *Nature*. The manuscript demonstrates that although warm spring temperatures shift the timing of temperate deciduous tree growth earlier, they do not augment annual growth. The implication is that, contrary to the expectations imbued in most climate models, warming spring temperatures are unlikely to increase woody productivity enough to strengthen the long-term carbon sink of temperate deciduous forests.

We have received and responded to the comments of four reviewers. The largest changes changes included the following:

- We redid dendrometer band analyses to provide more stable parameter estimates, a refined data cleaning process, and a more conservative approach to discarding outliers. This did not substantively alter the results (although there were some modest changes) or affect the conclusions.
- We reworked the introduction and discussion to incorporate content and citation suggestions of the reviewers, as well as discussion of relevant papers that have been published since our initial submission. This includes discussion of the literature on wood formation and reworking of the discussion on C cycling.

We believe that the revised manuscript is significantly improved and thank the reviewers for their constructive feedback.

We believe our manuscript to be an appropriate length for *Nature*, with 2,687 words, 3 display items, and 51 references in the main article. The article is accompanied by an Extended Data file (3 tables, 7 figures) and a supplementary table.

Thank you for considering our submission for *Nature*.

Sincerely,

Kristina Anderson-Teixeira, on behalf of all co-authors.

Responses to referees' comments

Referee #1 (Remarks to the Author):

GENERAL COMMENTS

The manuscript entitled: "Warmer spring temperatures in temperate deciduous forests advance the timing of tree growth but have little effect on annual woody productivity" by Cameron Dow, Albert Y. Kim, Loïc D'Orangeville, Erika B. Gonzalez-Akre, Ryan Helcoski, Valentine Herrmann, Grant L. Harley, Justin T. Maxwell, Ian R. McGregor, William J. McShea, Sean McMahon, Neil Pederson, Alan J. Tepley, Kristina J. Anderson-Teixeira, is well written, well-structured and easy to read. The main message is clear, important and timely.

Thank you.

Key results: Please summarise what you consider to be the outstanding features of the work.
Using dendrometer band measurements from about 500 trees, over 10 years and across two forest plots, the authors wanted to show that warmer spring temperatures shifted the stem growth of deciduous trees earlier but had no effect on peak growing season length, maximum growth rates, or annual radial growth. Using tree-ring data from more than 200 chronologies from 100 forest plots across eastern North America, the authors wanted to confirm that, in this region, ring width is not sensitive to spring but to summer temperatures. The authors point out that these findings imply that extra CO2 uptake in years with warmer springs is not sequestered in long-lived woody biomass as expected by vegetation models.

We agree with this summary.

Validity: Does the manuscript have flaws which should prohibit its publication? If so, please provide details.
I like the manuscript and the approach presented. I Think this approach provide some important elements, like the fact that despite increasing temperature the length and rate of peak growing season are not changing, but I don't think the results provided are sufficient to establish the full story from climate change to leaves phenology, carbon capture and woody biomass production.

We have revamped the introduction and discussion to provide a clearer presentation of the broader context and role of our study. In particular, we have revamped our discussion of wood formation (adding citations, no longer referring to this as "phenology") and carbon cycling (clarifying logic, adding literature, etc.). Most of these improvements are detailed in response to comments below.

Originality and significance: If the conclusions are not original, please provide relevant references. On a more subjective note, do you feel that the results presented are of immediate interest to many people in your own discipline, and/or to people from several disciplines?
I believe the subject is important and relevant for a large audience including scientists working on vegetation models and carbon cycle. However, contrary to what is said this manuscript is not unique in trying to relate climate, growth phenology and wood production. A couple of recent papers mentioned already for example that growth rate contribute more than phenology to wood production (Delpierre et al. 2016; Etzold et al. 2021).

Thank you for pointing out these recent studies, both of which are highly relevant and consistent with our results. We have added citations throughout the manuscript to these studies and others focused on climate, growth phenology, and wood production. We have also revised the introduction to include a discussion of the current state of knowledge on how spring temperatures affect growth phenology and wood production (3rd paragraph).

Data & methodology: Please comment on the validity of the approach, quality of the data and quality of presentation. Please note that we expect our reviewers to review all data, including any extended data and supplementary information. Is the reporting of data and methodology sufficiently detailed and transparent to enable reproducing the results?
To my point of view the overall quality of the paper and the data are excellent. however, I don't believe that 25 and 75% thresholds computed from a logistic model fitted on dendrometer bands can provide an estimate of stem growth phenology (onset and cessation of cambial activity). This is more than a technical dispute, this a conceptual misunderstanding. To my point of view growth phenology and growth dynamics may have quite different drivers. Onset of growth may be driven by temperature and day length, while the rate of growth may be driven by carbon and water availability. Using a simple logistic model to compute phenology and dynamics parameter means that this parameters won't be independent and this prevent a proper analysis.

We appreciate this point, and have adjusted language throughout the manuscript to clearly communicate that we are getting at the timing/ seasonality of stem growth, as opposed to growth phenology (defined based on detectable phase transitions). Specifically, we have removed use of the term "phenology" in association with stem growth (but continue to use it for leaf phenology). We have also been careful not to imply that our results are able to define the start or end of stem growth, or total growing season length. These wording changes do not alter our conclusions.

In addition, we have revised the introduction to include a paragraph on the current state of knowledge on how spring temperatures affect growth phenology and wood production (3rd paragraph). A couple sentences in this paragraph that are particularly relevant to setting up a proper understanding of the relationship between our results and the seasonality of wood formation are as follows: "Close coordination of budburst and initiation of xylogenesis[11] suggests that warm springs should shift the onset of growth earlier alongside observed advances in leaf phenology[1–3,4]. However, earlier initiation of growth would not necessarily translate to earlier, faster, or greater stem growth; rather, stem growth is dependent upon environmental conditions on hourly to daily time scales[28,29], and annual

growth is more closely linked to conditions within the growing season than to growing season length[30], GPP[23], or NEE[23].“

Appropriate use of statistics and treatment of uncertainties: All error bars should be defined in the corresponding figure legends; please comment if that’s not the case. Please include in your report a specific comment on the appropriateness of any statistical tests, and the accuracy of the description of any error bars and probability values.
Error bars are not presented in figure 2.
Otherwise statistics are fine to me.

Error bars are not shown on this plot because there are no error bars for the leaf phenology (extracted for a single pixel at the center of each forest plot), and those for tree growth would overlap in a way that they are indistinguishable from one another, making the plot too messy. The purpose of this figure is to summarize and clearly visualize results. To clarify where readers may view the spread of these timing variables, we have added the following statement to the Figure 2 caption: “Mean DOY_{25}, DOY_{50}, and DOY_{75} were estimated using the Bayesian model visualized, with confidence intervals, in Extended Data Figure 4.”

Conclusions: Do you find that the conclusions and data interpretation are robust, valid and reliable?
In fact, I found a gap between the results (results section), which I found accurate and robust, and the interpretations (discussion section), which lack internal logic and extrapolate too much. Again, the strength of the paper is the characterization of the stem peak growth (rate and duration) using large number of dendrometer bands and appropriated methods of data analysis. However, the same method is not appropriated to describe wood formation phenology, and there is no data in the paper about carbon fluxes.

We have removed use of the term “phenology” in association with stem growth (but continue to use it for leaf phenology). We have revised the introduction to summarize the relevant literature on wood formation and clarify the relationship of our results to this body of knowledge (3rd paragraph; also see response two comments above). We have also been careful not to imply that our results are able to define the start or end of stem growth, or total growing season length. Our original conclusions were not dependent upon characterization of phenology (defined based on detectable phase transitions), and therefore this rewording clarifies the message but does not alter conclusions.

We have also reworked the discussion of carbon fluxes to better clarify concepts and to improve the internal logic. Specific changes include explicitly clarifying the relationship between stem growth and ANPP_woody, providing a clear test of potential 1-yr lag effects of spring temperatures on growth in the results section, and limiting discussion of the “extra C” in years with warm springs to two modest paragraphs (2nd - 3rd paragraphs of discussion).

We have reviewed the logic of our discussion on C cycling and believe it to be sound, with key claims backed either by our own results or by multiple other studies. For instance, while we unfortunately lack the appropriate eddy flux data needed to demonstrate increased CO2 sequestration in years with warm springs at both of our focal study sites, there are widespread observations that this occurs, including at one of our sites (Harvard Forest).

Suggested improvements: Please list additional experiments or data that could help strengthening the work in a revision.
Flux tower data on the carbon fixation side and wood formation monitoring data on the carbon sequestration side would be perfect to bridge the gap, but I am aware that adding this kind of data to this already large study would be a huge work.

We carefully considered bringing in flux tower and/or wood formation data, but ultimately decided that neither would add substantially to this paper or warrant the extra effort (particularly considering that the lead author is an MS student and no longer at the institution where he performed the work). We detail our logic below.

We had previously considered including eddy flux data, and reconsidered based upon this comment, but both times decided that it would not make sense. Both Harvard Forest and SCBI have eddy flux towers. Harvard Forest has the longest eddy flux record that exists and would be fantastic for this type of analysis, and we heavily cite a recent paper analyzing that record (Finzi et al. 2020). However, SCBI flux data (from NEON) only start at the end of 2016, which would give us only 4 years overlap with the dendrometer band data. Moreover, the tower is on complex terrain, which violates the assumptions of eddy flux measurements and means that we couldn't trust the carbon balance from this site. It would probably be possible to extract some parameter of interest– for example, timing of when daytime C uptake passes a certain threshold in the spring. This could nicely supplement the analysis, but (a) wouldn't cover enough years, and (b) wouldn't give us the C balance (what we ultimately want). Even if we could relate spring temperatures to timing of the start of C uptake at SCBI, this wouldn't be a novel finding of resolve where the extra carbon may be going. Thus, there would be little to be gained from bringing in eddy covariance data. We believe that an analysis of this question would be very interesting, but would make sense as a separate paper.

We also considered whether it would make sense to pursue bringing in automated point dendrometer data from TreeNet (e.g., used in Etzold et al. 2022). This is a very nice data record that would be highly relevant to the question at hand, and we were not yet aware of it in the early stages of this project. The main reason that we decided against seeking to bring it in now is that, while it would expand our analysis and allow a more detailed analysis of the early stages of annual wood formation, it wouldn't fundamentally enrich the conclusions. The downside of bringing in these data would be simply a matter of the work involved: not only would we want to bring in dendrometer data, but also parallel tree-ring data for Europe. This is all theoretically possible and would be quite interesting, but is beyond what we can do with current time and resources.

References: Does this manuscript reference previous literature appropriately? If not, what references should be included or excluded?
The manuscript discusses about stem growth phenology, but lack literature on wood formation phenology (ex: Delpierre et al. 2019; Huang et al. 2020; Rossi et al. 2008; Savage & Chuine 2021).

Three of the four references given as examples here are for conifers, and the exception (Savage & Chuine 2021) is a review saying we know little about the drivers of the timing of stem growth in angiosperms. We have incorporated 3/4 of these references into the third paragraph of the introduction (limited to 2 conifer citations because of length limitations): "Most studies on tree growth responses to warmer spring temperatures have focused on boreal or temperate conifers, which tend to respond to warmer spring temperatures with an earlier start to growth[24, 25] and increased annual growth in mesic climates[26, 27]. In contrast, we have little evidence as to how stem growth and $ANPP_{woody}$ respond to warmer spring temperatures in deciduous forests[10,11]. …"

Clarity and context: Is the abstract clear, accessible? Are abstract, introduction and conclusions appropriate?
The abstract is excellent in presenting Author's ideas, which I don't completely agree with, as explained earlier.

We have adjusted the entire manuscript – including the abstract – in response to the concerns raised above, as detailed below each concern. We believe that the current version of the abstract is now entirely substantiated by our analysis. One small but substantive change was revision of the final sentence of the abstract to read: "Rather, contradicting current projections from global C cycle models[2,14], our empirical results imply that warming spring temperatures are unlikely to increase woody productivity ~~or~~ *enough to* strengthen the long-term CO_2 sink of temperate deciduous forests."

Please indicate any particular part of the manuscript, data, or analyses that you feel is outside the scope of your expertise, or that you were unable to assess fully.
My background in dendroecology, dendrometers measurements and wood formation monitoring allow me to understand the paper well. However, I learned a lot reviewing this manuscript since the authors are nicely using up-to-date methods, which were not familiar to me.

Thank you.

DETAILED COMMENTS

L39. Well, there is a quite remarkable bunch of papers in the field of wood formation now (Ex: Delpierre et al. 2019, Huang et al. 2020).

We have added Huang et al. (2020) and several other references on wood formation (noting, however, that there are more studies on conifers–including Delpierre et al. 2019– than broadleaf species). We have also removed this statement and restructured the introduction to include a paragraph on the

current state of knowledge on how spring temperatures affect the timing of growth and wood production (3rd paragraph).

L42. I am not convinced by the 25% – 75% criteria to define the beginning and the end of stem growth. Usually 5% and 95% are chosen. But anyway, my main concern is that logistic curve is very appropriate to simply describe peak growth (as it is done in the manuscript), but not phenology (as it is extrapolated in the manuscript). The logistic curve is fitted to all the data for a tree-year, this means that the computed parameters are not independent, what happened in the middle of the growing season will impact on the parameters of the beginning. A drought, which will stop growth in summer for example, will delay DOY25, this is not what we want! Moreover, phenology for me is a change in quality, from bud to leaf or flowers for example, not a change in quantity of the same process (from 24% of growth to 25).

Part of this comment has to do with the use of the term "phenology" throughout the manuscript. We have reworded mentions of stem phenology to timing of peak stem growth.

This comment also relates to the independence of our DOY timing variables. While we agree that the timing variables (DOY_{25}, DOY_{50}, and DOY_{75}) are interdependent, variables quantifying the rate or extent of growth (g_{max} and ΔDBH) are independent of these. If a drought were to curtail g_{max} and/or ΔDBH, this could cause DOY_{25} to occur earlier or later, depending on the timing of the drought. The fact that we see little correlation between spring temperatures and annual growth (ΔDBH or RWI) implies that, while summer drought may have influenced timing variables in some years (e.g., we believe a modest drought at HF in 1999 reduced summer growth such that DOY_{25} occurred earlier), they are unlikely to have had much influence overall.

L43. Daily is too much her, since there is not daily data…

We have renamed this variable as "maximum growth rate".

L44. The conclusion about tree ring is not direct, it is because there are more correlated with summer T, but there is no test to check that there is not an increasing trend in tree-ring width…

Our study focuses on interannual variation, which is much greater than long-term trends over the time frames considered here. The standard dendrochronological methods applied here, including removal of long-term trends (e.g., those associated with changing tree size) through detrending, focus on retaining the interannual variation. Because interannual variation in temperature has a much greater magnitude than the long-term trend over the time period considered, this analysis should be appropriate for determining the dominant temperature controls over annual growth. We agree with the reviewer that, in theory, significant long-term trends could bias this analysis. Disentangling long-term trends in tree growth and their drivers is a challenging problem. However, our recent study including three of the sites analyzed here (SCBI, Harvard Forest, + one site in Indiana) showed that, after correcting for climate and tree size, long-term growth trends were negative (probably in large part due to increasing competition as these secondary stands mature; Anderson-Teixeira et al. 2022, GCB: https://doi.org/10.1111/gcb.15934). Thus, it is extremely unlikely that warming springs are

causing a long-term trend of increasing growth in the predominantly secondary forests in our study region.

L46. We can still imagine that tree rings are more sensitive to summer T, but benefit from spring T.

We tested for this (3rd paragraph under "Tree-ring analysis" section) and found significant (at $p=0.05$) positive correlations with April T for only 4% of chronologies (less than expected at random).

L47-48. This is a very interesting though, and an important question, but I found no demonstration in the manuscript that C capture increase at Harvard and Symp. Forests.

This has been shown at Harvard Forest, and a paper demonstrating this (Finzi et al. 2020) is one of the citations here. Unfortunately, the eddy flux record at SCBI (from NEON) cannot be used to address this question, as it only began in late 2016, and because the tower is located on complex topography, making the C balance unreliable.

Because we do not confirm extra C sequestration in years with warm springs in our analysis, we softened this statement to read: "These findings imply that <u>any</u> extra CO_2 uptake in years with warmer springs[5–7] does not significantly contribute to increased sequestration in long-lived woody stem biomass."

L49-51. Again, the idea is interesting, but I found it is over interpretation here since the process is not understood and there are only two sites in the study…

While the dendrometer band data that allow us to describe how growth timing shifts in years with warm springs are limited to two sites, it is important to note that the tree-ring study confirms a lack of extra stem growth in years with warm springs across the temperate deciduous biome of the eastern US. The latter is the more relevant result to the sentence in question.

We have reworded this final sentence of the abstract to read: "Rather, contradicting current projections from global C cycle models[2,14], our empirical results imply that warming spring temperatures are unlikely to increase woody productivity ~~or~~ *enough to* strengthen the long-term CO_2 sink of temperate deciduous forests."

L52. I read 30% in Fredlingstein.

While Freidlingstein's numbers would give ~30% (S_LAND/ E_FOS), they're not quite appropriate because S_LAND includes all terrestrial biomes, not just forests. The ~20% figure was derived using an estimate of the global forest C sink from Harris et al. 2021, but this was confusing (as evidenced by this comment). We've decided that it's important to be more specific here and have revised the sentences related to the global C budget to read as follows:

"In recent decades, tree growth in Earth's forests has more than offset losses from deforestation and other disturbances, such that a net forest CO_2 sink of ~1.6 Gt C yr^{-1} offsets ~20% of anthropogenic emissions[15], dramatically slowing the pace of atmospheric CO_2 accumulation and climate change. Of this important C sink, ~ 47% occurs in temperate forests[15], with temperate deciduous forests sequestering >0.3 Gt C yr^{-1}.[16] The future behavior of this C sink will play an important yet uncertain role in influencing atmospheric CO_2 and climate change[17–19]."

L68-70. This sentence is misleading. There is a large corpus of dendrometer studies over time (see Salomón, R.L., Peters, R.L., Zweifel, R. et al. The 2018 European heatwave led to stem dehydration but not to consistent growth reductions in forests. Nat Commun 13, 28 (2022). https://doi.org/10.1038/s41467-021-27579-9) but to my point of view this is not phenology. There is also studies on wood formation phenology, but mainly on conifers, this is true (Ex Delpierre et al. 2019, Huang et al. 2020). But then why deciduous forests are of such importance to test this idea?

We now clarify the motivation for better understanding the seasonality of temperate deciduous tree growth in the introduction. First, in the first paragraph of the introduction, we outline the importance of temperate deciduous forests for the global C cycle: "....Of this important C sink, ~ 47% occurs in temperate forests[15], with temperate deciduous forests sequestering >0.3 Gt C yr^{-1}.[16] The future behavior of this C sink will play an important yet uncertain role in influencing atmospheric CO_2 and climate change[17–19]."

Second, we explain the current understanding of how spring temperatures affect the seasonality of growth, including the contrasting states of knowledge on conifers vs broadleaf trees: "Most studies on tree growth responses to warmer spring temperatures have focused on boreal or temperate conifers, which tend to respond to warmer spring temperatures with an earlier start to growth[24, 25] and increased annual growth in mesic climates[26, 27]. In contrast, we have little evidence as to how stem growth and $ANPP_{woody}$ respond to warmer spring temperatures in deciduous forests[10,11]. …"

L73. Growth phenology should be defined, and to my point of view it is a weak notion.

As noted above, we have removed use of the word "phenology" to refer to stem growth throughout the manuscript. We have reworded this particular sentence to read, "Tree-ring records, which can be used to examine relationships of annual growth to temperature but not to understand growth seasonality,.."

L75-76. Typos: additional spaces

Fixed.

L81. Works from Delpierre et al. 2016 and Etzold et al. 2021 are missing in the introduction. They already show that productivity is more related to peak growing season than to length of the growing season, using similar, dendrometer approach…

We now cite both these studies, for example in the third paragraph of introduction: "...annual growth is more closely linked to conditions within the growing season than to growing season length[30], GPP[23], or NEE[23]."

L88-94. This is perfectly fine to me, there is no 'phenology' in here!

Thank you. As noted above, we have removed use of the word "phenology" to refer to stem growth throughout the manuscript.

L95-96. I also like this wording.

Thank you.

L131. Correct for 'temperature'

Fixed.

L131. Why Tmax? It is difficult to interpret from and ecophysiological point of view?
Tmax = day T, while growth occurs mainly during the night (Zweifel 2021).

We don't believe that there is any conflict here, for two reasons, both relating to the fact that there can be lags between temperature and physiological responses. First, the critical temperature window (i.e., the period over which temperature is most relevant to spring growth phenology) occurs ahead of most growth – temperature is acting as a cue for growth to start, as opposed to being directly linked to daily growth. Second, even for time scales over which temperature is more directly controlling growth (e.g., during the peak growing season), we'd expect some lagged response. The proximate cause of stem expansion is turgor pressure, which is more readily achieved on nights following days with low Tmax, when lower transpiration and soil evaporation result in higher water storage in both the stem and soil. Thus, we see no reason why Tmax shouldn't be more strongly correlated with growth than Tmin.

L133. I prefer xylem structure / architecture groups as referred elsewhere.

Changed to xylem architecture. Related: we also changed mention of "wood-type" to xylem architecture as well.

L149-150. OK, but Tmax and CWT were chosen to maximize the relation with timing, not with rate… Same fitting should be done for rate to compare. By the way, timing of growth is also a good term.

We're not looking to find the critical window during which temperature most strongly influences g_max or ∆DBH (Fig. 3 clearly shows that ∆DBH is most strongly influenced by June-July temperatures, and we'd expect similar results for g_max), but rather to understand how these are variables are influenced by an earlier start to growth. DOY25 represents the beginning of peak stem growth, and was the most sensitive to spring temperatures among the variables we tested. By selecting a CTW based on DOY25, we thereby capture the most influential window of spring temperatures. Note that we also ran a version of the analysis using monthly CTWs (Extended Data Figure 5), with similar results, demonstrating that our results are not dependent upon selection of the CTW with maximum influence on DOY25.

L177-178. When only temperatures are tested, but summer Tmax point to water stress…

We agree, and we now devote a paragraph in the discussion (2nd-to-last) to water stress associated with high summer temperatures. The particular sentence in question has been reworded to remove the statement "... summer temperatures were the more important driver of annual stem growth…".

L192-193. I found the discussion disconnected from results, with several slippery changes that distort the logic. Ex: Stem growth timing become stem growth phenology; tree-ring climate sensitivity become annual wood productivity…

We have reworked the discussion, paying careful attention to the logic. In particular, addressing the two concerns raised here, we have (1) removed use of the term "phenology" when referring to the timing of stem growth, while enriching the discussion through reference to literature on wood formation, and (2) clarified the relationship between stem growth and woody productivity. We now define ANPP_woody in the introduction (2nd paragraph) and clarify its relationship to stem growth in the discussion (2nd paragraph): "It remains theoretically possible that warm springs could augment $ANPP_{woody}$, which, although routinely calculated based on stem growth, can be partially decoupled from it through differences in wood density or C content[23]. "

L196-197 + L204. I didn't find the evidence in that results that 'warmer springs hasten the cessation of stem expansion'

Throughout, we have reworded to ensure that we don't imply that DOY75 translates to cessation of stem expansion. For example, we have reworded the sentence previously at L196-7 to read as follows: "However, inconsistent with the concept that an earlier start to growth would increase $ANPP_{woody}$, we demonstrate that warmer springs either hasten the deceleration of stem expansion or otherwise fail to translate extended growing seasons into biologically significant increases in stem growth (Fig. 1), and thereby have negligible effect on total annual growth for most species and locations (Fig. 3). "

L204-209. I don't understand this very long sentence. Moreover, SPEI was not reported in the result section.

This sentence described a non-result (lack of correlation between spring temperatures and summer drought stress within our analysis). As it was not essential and not reported in the results section, we have therefore removed it here, but retain text on this in the methods section.

L295-296. But this means that tree are partially changing from one year to another, this should be taken into account in the analysis.

This comment is a bit confusing. Answering what seems to be the concern: The timing of the first census should not affect the analysis so long as the record always starts prior to the start of growth. The initial spring survey always began prior to canopy green-up, being shifted forward when we were concerned that leaf-up would occur before the normal census start time.

L316 (actually 314). I don't think 'phenological dates' could be applied here…

We have reworded this sentence to read: "We fit a five-parameter logistic growth model[35] to dendrometer band data from each tree-year to define stem growth milestones and growth rates (Fig. 1)."

L318. What is the justification for the use of this logistic model instead of a Gompertz model for example?

The logistic model applied here was carefully developed/ documented in a publication led by one of our coauthors (McMahon & Parker, 2015; doi: 10.1002/ece3.1117), and we believe it to be the best function available. Although the Gompertz is a good function, it lacks the lower asymptote and some of the flexibility of the Richards function (5-parameter logistic), and for some inference, a very specific fit is needed. The five parameters have reasonable interpretations: high and low asymptotes, rate and inflection point along with an additional parameter, theta, which skews the curve so that it can be asymmetric. Although it has five parameters, it was fit easily using a series of maximum likelihood methods, balancing breadth of search with precision in a pipeline. Again, McMahon and Parker 2015 expand on this.

We did revisit the model fits and revised our fitting method to provide more robust and stable estimates of starting and ending diameter. Briefly, this entails a shift from use of "L" and "K" to "a" and "b" parameters, as described in McMahon and Parker (2015; this is the paper describing the seasonal growth model that we employed). While "L" and "K" describe asymptotes of the fit curves, these are not always realistic, and the parameters "a" and "b" provide constraints based on observations. For example, the spring (initial / minimum) diameter is now constrained by spring measurements, whereas it was previously free to be significantly lower than the initial spring measurements. This analysis improvement did not substantially alter our results or affect our conclusions.

L326-328. If the authors want to say that they used the fitted parameters to compute the variable of interest the sentence is quite disturbing?

The variables of interest were computed based on the fitted parameters, following the fitting of the growth model on an individual basis for each tree-year. Given that the dendrometer bands do not provide a continuous record, this model fitting is necessary to compute meaningful response variables (in our case, DOY_{25}, DOY_{50}, DOY_{75}, g_{max}, and ∆DBH). This model has been carefully tested and, for trees that are growing at a healthy pace, generally produces an excellent fit (McMahon & Parker 2015). It also acts as a smoothing function that offsets short-term variability associated with stem hydration status. We now explain this advantage in the methods section of the manuscript, "An advantage of this approach is that short-term shrinkage and swelling associated with rain events[35,57] and measurement errors show up as residual variation and do not unduly influence the parameters of interest."

L332. In total what proportion of data have been removed? from my calculations 500/2500 = 1/5 = 20%, this is far above the 10% limit… Without any explanation about why the proposed model do not adjust to 1/5 of the trees.

We've revisited our data cleaning methods in response to this concern. This – together with an improvement to the method of estimating parameters describing the starting and ending DBH – resulted in a reworking of the dendroband analysis. Results remained similar.

Our new methods remove tree-years where (1) total growth for the year was < 0.05mm. These trees did not grow enough for an appropriate modeled curve to be fit to the data. (2) Tree-years where the first manual dendrometer band survey happened late in the spring, so the beginning of growth was missed. By missing the start of growth, the predicted starting dbh value was skewed, so subsequent predictions of DOY% variables was unreliable. (3) Tree-years where timing variables DOY_gmax, DOY25, DOY50, and/or DOY75 were >2.5 standard deviations away from the means for their xylem architecture group, year, and site. (4) Tree-years where single day growth rates were > 2.5 SD away from the mean for each xylem architecture group at each site. Categories 3 and 4 represent the tree-years which were outliers due to potential poor model fits or unreliable raw dendrometer band measurements.

Categories 1 and 2 do not have to do with the model failing to fit, but rather are cases where data was not suitable to build modeled growth curves in the first place. Removing these two categories is necessary to ensure our results are trustworthy. We therefore have moved the description of removal of tree-years falling into these categories in the the paragraph describing pre-model cleaning: "We also removed tree-years with small or negligible total growth (∆DBH ≤ 0.05 mm; SCBI = 26, Harvard Forest = 253), and tree-years where first intraannual measurement was later than the first spring survey (trees that were missed in the initial census; SCBI = 22, Harvard Forest = 8)."

Under these new criteria and clarifications, we remove 189 tree-years (~8% of tree-years) due to poor model fits instead of the 500+ previously reported.

L370. Why not testing Tmoy?

We did not test for an effect of mean temperature because Tmax and Tmin provide more biologically meaningful information (potential peak daily temperatures/ potential heat stress and chilling cues, respectively). From related analyses done in the past, we've found that Tmean tends to be more weakly correlated with growth metrics than Tmax.

L379. The duration of CTW is not clear to me. Why not starting the analysis at monthly level and then downscaling progressively to the week level?

We experimented with different time scales for defining the CTW. We expected that monthly resolution may be too coarse for the analysis of interest, and therefore experimented with daily and weekly resolutions in the R package *climwin*. We ruled out the use of daily resolution because this sometimes produced spurious correlations with temperatures over very short time periods. We therefore focused on defining CTW with weekly resolution. We also ran the analysis with a monthly resolution, yielding similar results (Extended Data Figure 5). We note also that there was a negative correlation between Tmax and DOY25 over a wide range of potential spring time windows (Extended Data Figure 3).

To clarify our decision to use a weekly time frame in the *climwin* analysis, we have added the following statement to the paragraph describing the identification of CTWs: "To avoid spurious correlations that could occur using temperature data at the daily resolution, we ran this analysis with weekly resolution, using temperatures averaged over weekly time periods."

L383 (378 in our version). Remove the '}'

Done.

L384. What is the relation between this analysis and the climwin analysis? What mean exactly spring temperatures here?

We have clarified this. The sentence in question now reads: "Correlation between the dendrometer band-derived growth parameters (DOY_{25}, DOY_{50}, DOY_{75}, L_{PGS}, g_{max}, and $\triangle$DBH, Fig. 1) and CTW (at weekly or monthly resolution, as described above) were assessed using a linear mixed model in a hierarchical Bayesian framework. "

L396. Add (SPEI).

This line number doesn't match our version; we think it refers to line 391. We have added this abbreviation: "we examined the relationship between spring temperatures and summer Standardized Precipitation Evapotranspiration Index (SPEI)".

L417. The point is not at the correct place.

Fixed.

L429. Despite I have a background in dendroecology, this paragraph was very difficult to follow for me.

We have made various revisions in this paragraph to improve its clarity.

L503-504. Remove the capitals.

Done.

L505-506. Journal name is missing.

Fixed.

L556-558. Remove the capitals.

Done.

L578. n/a?

Fixed.

L580. gcb?

Reference fixed.

L590-605. Check problems in most of the references!

We have checked and fixed problems in the references.

Thank you for the review!

Referee #2 (Remarks to the Author):

Dow et al investigated whether warmer spring temperatures result in earlier and/or increased wood growth of temperate trees. This is an important question as (1) rising temperatures are causing earlier leaf-out of many plants but it is not known whether wood growth is responding similarly, and (2) carbon sequestered wood has a long residence time in the ecosystem and hence if warmer spring temperatures enhanced wood growth this could be an important negative feedback to climate change. Their primary data source consists of band dendrometer data from almost 500 deciduous trees at two forests in the eastern United States. Although these data show that warmer spring temperatures did shift the timing of

wood growth earlier, there was no consistent effect found on other measures, including growing season length or total annual growth. Overall, this is an interesting and well-executed study, and the writing is clear and direct (in particular, L189-191 give a very nice summary). I
particularly liked the use of the dendrometer data, which enables quantification of the **timing** of wood growth, in addition to the overall amount of annual wood growth in terms of changes in a tree's diameter. In that regard this is an important advancement over some related papers that have tried to investigate this question using tree ring chronologies, which do not indicate anything about the timing of wood growth.

Thank you.

However, dendrometer data provide an imperfect measure of the mass of wood growth (what I think the authors are referring to with their term "woody productivity") because they are measuring diameter—which may be subject to shrinking and swelling of tissues as water relations change—and not biomass; the relationship between biomass and diameter is reasonably well predicted by allometric equations but these do not capture variations in wood density that may occur across years or even within years (early vs. latewood). Finally, dendrometer band estimates of tree diameter tend to have negatively autocorrelated errors; an over-estimate of diameter in period x will lead to an over-estimate of growth between period x-1 and x, and an under-estimate of growth between x and x+1.

We are very aware that dendrometer bands capture hydraulic shrinkage and swelling (e.g., Herrmann et al. 2015: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0169020). Short-term diameter changes caused by hydraulic shrinkage and swelling appear as residuals around the fit of the logistic growth model that we apply to estimate relevant parameters (McMahon & Parker 2015; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4314258/). McMahon & Parker (2015) explicitly demonstrate this. To clarify this in the manuscript, we have added the following statement in the methods section where we describe the McMahon & Parker model: "An advantage of this approach is that short-term shrinkage and swelling associated with rain events (McMahon et al. 2015, Herrmann et al. 2016) and measurement errors show up as residual variation and do not unduly influence the parameters of interest."

We agree that variation in wood density across years is possible and has the potential to decouple stem growth from woody productivity. We have revised the discussion around this issue (2nd paragraph of discussion). The most relevant text on this now reads: "It remains theoretically possible that warm springs could augment $ANPP_{woody}$, which, although routinely calculated based on stem growth, can be partially decoupled from it through differences in wood density or C content[23]. Extra C fixed in years with warm springs could potentially be allocated to formation of more C-dense wood, either through enhanced cell wall thickening (a process that lags behind stem expansion[39]) or to a higher ratio of high-density latewood to lower-density earlywood. However, existing evidence indicates that vessel features are most strongly controlled by summer drought stress in the previous (earlywood) or current (latewood) year, while warm springs have a neutral or negative effect on

latewood width[40–42]. Thus, it is unlikely that warm springs have a positive effect on total C content of annual rings or $ANPP_{woody}$."

Regarding autocorrelated errors, we agree that these exist but do not believe they would bias our results. On an intraanual basis, the fit of a logistic growth model to estimate relevant parameters minimizes the influence of erroneous measurements, which would appear as residual error. While erroneous measurements are inevitable (despite screening of the data pre- and post- model fitting), we expect them to be random and not to directionally bias our results. To clarify this in the manuscript, we have added the following statement in the methods section where we describe the McMahon & Parker model: "An advantage of this approach is that short-term shrinkage and swelling associated with rain events (McMahon et al. 2015, Herrmann et al. 2016) and measurement errors show up as residual variation and do not unduly influence the parameters of interest."

On a year-to-year basis, data for each year on a given tree are independent, as we make the first measurement in the late winter/early spring before leaf-out and continue measurements past the cessation of growth in the fall. We do not include data from the preceding or following year when fitting the growth model. To clarify this in the manuscript, we have added the following statement in the methods section: "In our analysis, each band-year was treated independently, with no data overlap from one year to the next."

The authors note that their results are in contrast to studies based on CO2 flux measurements which show that earlier spring leaf-out increases both springtime CO2 uptake as well as annual CO2 uptake. While different explanations are proposed for this result, these explanations are successively shot down—leaving this contradiction apparently unresolved. I found this to be highly unsatisfying. I could not follow the argument presented in the paragraph beginning L248, and am unsure how this would explain the discrepancy between tower and dendrometer data. I feel that the authors need a stronger explanation for where the "extra" CO2 taken up in years with warm springs ends up, if it's not in wood. (Note that if it is allocated to a fast-turnover pool then we would see it respired back to the atmosphere in late summer or fall, resulting in no "extra" net CO2 uptake in warm spring years—which is not what the tower data show).

We have reworked the discussion, including the following improvement (which address the concerns listed here):

- **We have significantly shortened and revised our discussion of the fate of the "extra" CO2. While our previous version of the discussion sought to explore this in some depth, the fact remains that the current study cannot answer the question as to where this extra C goes. We now present a more concise discussion with improved logic in the 2nd-3rd paragraphs of the discussion.**
- **We now present the analysis ruling out a 1-yr lag effect in the results section, which allows us to be more concise in the discussion.**
- **We believe the most likely explanation of the discrepancy between tower and dendrometer / tree-ring data is that C is allocated to shorter-lived pools, but still remains in the ecosystem beyond the end of the calendar year. This is explained in the 3rd paragraph of the discussion:**

"The fate of any additional C fixed during years with warm springs remains unresolved, but possible destinations – including respiration, NSC storage, and production of foliage, reproductive structures, roots[32], or root exudates – generally have shorter residence times than woody growth. Indeed, when GPP of a mature forest was increased through experimental CO_2 enrichment, $ANPP_{woody}$ remained unchanged, while additional C was released back to the atmosphere on relatively short time scales through enhanced respiration[21]. Consistent with this, it has been observed that C gains from an earlier spring can be offset through fall or winter respiration[22], although even the C in shorter-lived pools would often be carried over into the following year[43]. Thus, observed augmentation of NEE by warm springs[2,8,9] is likely to be compensated by increased respiration in subsequent years."

- **The paragraph previously started on L248 has been completely changed. The intended point was that long-term effects of continuously warming spring temperatures may be different than the intra-annual effects characterized here. That possibility is now discussed in the 4th paragraph of the discussion: "It is possible that as spring warming continues, forests will adjust to directional changes in growing season length with an enhancement of $ANPP_{woody}$. …"**
-

Specific comments

1. It would be helpful for the readers if the authors could define exactly what they mean by "tree growth" and distinguish it from other terms they use including "woody productivity".

Our first mention of tree growth (besides in the title) occurs in the 2nd sentence of the summary paragraph, where it is clear that we are talking about stem growth: "However, less is known about how spring temperatures affect tree stem growth[10,11], …"

We now define woody productivity in the 2nd paragraph of the introduction:
Model representation of C allocation to stem growth – or woody aboveground net primary productivity ($ANPP_{woody}$) on the ecosystem level – does not capture known decoupling of stem growth from photosynthate production[11,19,23]. As a result, the consequences of rising spring temperatures on stem growth may not be accurately represented in models[11,19].

We also clarify that there is potential decoupling of stem growth and $ANPP_{woody}$ in the 2nd paragraph of the discussion: ". It remains theoretically possible that warm springs could augment $ANPP_{woody}$, which, although routinely calculated based on stem growth, can be partially decoupled from it through differences in wood density or C content[23]. "

2. How old are the trees? L270 makes it sound like SCBI trees would be 170 y (since mid-19th C) and HF 80 y (since 1938 hurricane). It would be helpful for the reader if explicit ages were reported. I wonder how the relatively advanced age of the SCBI trees might influence the observed patterns.

We have added a sentence addressing this to the 3rd paragraph of methods: "While we do not estimate the ages of trees in our sample, bands at both sites were placed on individuals of differing sizes in an attempt to measure trees across a range of ages."

3. I found that Figure 2 was not very easy to interpret or understand.

We have revised the caption of Figure 2 to better explain the figure, including (1) revising the initial summary sentence, (2) adding a better description of canopy foliage phenology data source, and (3) adding an explanation of how the results in panels c-d were obtained.

4. L303+ The data cleaning procedure sounds somewhat arbitrary and I wonder if "inconvenient" data points were excluded. It would be nice to see how the analysis yields different results as the criteria for exclusion were adjusted.

We've revisited our data cleaning methods in response to this concern. This – together with an improvement to the method of estimating parameters describing the starting and ending DBH – resulted in a reworking of the dendroband analysis. Results remained similar.

The nature of our dendrometer band data (and presumably dendrometer band data in general) was such that significant data cleaning is required. Erroneous measurements get recorded, bands can temporarily stick and then rapidly release or get damaged / bumped by squirrels, trees may grow very little, etc. The revised data screening procedures are described in the methods. First, we describe pre-model cleaning in the paragraph starting with the following sentence: "The raw dendrometer band data were screened before analysis to remove records or entire tree-years that were either unreliable or inappropriate for our analysis." This included removal of anomalous individual records that were likely measurement or data entry errors, tree-years where there had been a problem with the band (e.g., sticking and subsequent slippage), tree-years with very little growth (accurate data, but insufficient to resolve temporal patterns in growth), and tree-years with a late start to measurements (again, accurate data but failing to capture the start of growth). Even following this screening, some anomalous records remained. To avoid potential skewing of our analysis by such records, we removed tree-years for which the model fits produced outlier estimates, as described in the paragraph beginning with the sentence "After fitting the growth model, we removed tree-years with poor fits.". We tried the analysis using outlier definitions of 2, 2.5, and 3 SD away from the group means, all of which yielded similar results. We ended up using the 2.5SD cutoff for the analysis presented here.

The full data screening process is described in the methods, and was not adjusted to produce desired results. Indeed, we have now tried multiple iterations of the analysis, always with qualitatively similar results.

5. L326. I'm not sure why tree years "with poor fits" were removed - what does this indicate, what are the implications?

Please see response above.

6. L343 - why focus on the NEON tower when the biometric data predate NEON? Why not use the center of the biometric plot?

We have redone the analysis to use the coordinates for the center of the ForestGEO plots. Results are not substantially altered.

7. L344. Why use satellite data when, at least for Harvard Forest, ground observations of phenology are available?

Unfortunately, SCBI does not have ground-based phenology observations for a long time period (although NEON has been collecting phenology data since 2016). We used satellite data for the longer record and for consistency across sites.

8. L404 – It's not clear why Juglans nigra was ruled out. Because it is semi ring porous? This could be stated explicitly.

We have changed the sentence here to read "We analyzed records for the ring- and diffuse- porous species at each site (Extended Data Table 1), but excluded semi-ring porous species (e.g., *Juglans nigra* L. at SCBI) and conifers (e.g., *Tsuga canadensis* at Harvard Forest)."

Thank you for the review!

Referee #3 (Remarks to the Author):

An excellent study with important implications for forest modeling and management in the face of temperature increases, warmer springs, and lengthening of the growing season. However, while realizing that stringent text limitations are being observed, I do feel that the conclusions were sometimes obscured by rather wordy descriptions. For instance, in the abstract on lines 47-49, the sentence "…findings imply that extra CO_2 uptake in years with warmer springs10–12 is not allocated to long-lived woody biomass, where it could have a substantial and lasting impact on the forest C balance.…" is the key finding of the paper but it is a rather awkward sentence. Perhaps you could just go with something more straightforward such as " … findings imply that any extra CO_2 uptake in years with warmer springs does not significantly contribute to increased sequestration in long-lived woody stem biomass…" and leave it at that.

Thank you for the review. We have implemented the abstract wording suggestion.

Similar wordiness somewhat obscuring the importance of the findings is also reflected in the discussion section (lines 189-198) and perhaps this section could also be streamlined…

We have reworded the entire discussion, and believe the updated version to be significantly improved/ more concise.

One limitation of this study is in the description of the MODIS satellite derived phenology data and an explanation of how the data are actually used. While the authors are to be commended for providing the reprocessing version information (version .006 identified in the legend for Figure 2), it is not at all clear if the phenology dates used are from a single pixel (a gridded 500m quantity) from each forest, or an average of selected pixels scattered over each forest, or indeed larger spatial areas encompassing portions of each forest. How much forest do these MODIS samples represent?? Also there is no mention of the use of the Quality Flags associated with this MODIS phenology product (this is an common but unfortunate drawback of using Google Earth Engine which does not encourage users to access Quality Flags).

We clarify that MODIS satellite data are extracted from a single pixel for each plot in our methods section. We've added a statement to this paragraph explaining that we only used data flagged as "good" or "best" in our description of the leaf phenology variables.

However, just a few more lines of description would really clarify Figure 2 and emphasize how it captures canopy foliage phenology.

We have revised the caption of Figure 2 to better explain the figure, including (1) revising the initial summary sentence, (2) adding a better description of canopy foliage phenology data source, and (3) adding an explanation of how the results in panels c-d were obtained.

Also note there is a typo (temperature) on line 130.

Fixed.

Figures and tables are appropriate.

Thank you.

Referee #4 (Remarks to the Author):

Warmer spring temperatures in temperate deciduous forests advance the timing of tree growth but have little effect on annual woody productivity

Cameron Dow et al.

General comments
This paper addresses the question of whether warmer spring temperatures, in addition to lengthening the growing season, lead to increased stem growth in North American deciduous trees. The results very clearly indicate no.

Crown and trunk growth phenology were found to occur 6-10 days earlier for ring and diffuse species in years with a warmer spring, resulting in generally earlier onset of phenological milestones. However, increased temperature had no consistent effect on stem growth at the annual scale. Only 5 of the 207 trees had a significant positive relationship between spring temperature and stem growth. However, annual stem growth was negatively correlated with maximum summer temperature in approximately 50% of the trees. The authors conclude that summer temperatures are the more important factor in annual stem growth than spring temperatures.

A minor correction of understanding here: each of these chronologies is a composite of multiple trees from one species-location combination.

The analyses are based on a very solid database of (i) band dendrometer measurements at two sites with an average of 463 trees, and for tree ring data from 207 trees, at 108 sites with a total of 16 species. Data and analyses were done for ring and diffuse-porous species; conifers were not studied.

Both the statistical approaches and their application appear robust. The results are unambiguous with respect to answering the questions posed.

The significance of the results is of utmost importance for understanding climate change on forest dynamics and specifically on its carbon sink performance. The results are also likely to attract very broad interest.

Thank you.

In my view, there would be potential for improvement in the discussion of the results if the recent papers by Etzold et al. (2021) and Zweifel et al. (2021) could be included. These two papers were probably published after this work was completed, but could contribute significantly to the interpretation of the data presented here for the following reasons:
o Both papers are based on high-resolution stem radius change data and are probably the first of their kind with such a large number of sites and tree species.
o Etzold et al came to the same conclusion as the authors here regarding vegetation length and its only small effect on annual tree growth. However, Etzold's data are based on automatic dendrometers at daily resolution and include coniferous and deciduous trees. The higher temporal data resolution offered additional information about the course of growth, e.g., that only a relatively small number of growing days lead to annual growth. This also explains, at least in part, the result found here that it is not so much the length of the growing season, but rather the quality of a few growing days that plays a role in annual growth.
o Zweifel et al showed that trees grow mainly at night because negative water potentials are then more moderate, allowing cell division and cell expansion, whereas during the day transpiration lowers water potentials in the stem to the point that the necessary turgor threshold in the cambium is usually not reached. In addition, Zweifel et al demonstrate that growth is suppressed more rapidly (by high VPD and low soil water potential) than stomata are known to close. Picking up on this finding in this paper here, leads directly to the explanation of why stem growth is limited long before photosynthesis is stopped.

What all three papers cannot answer is what happens to this excess assimilated carbon when it obviously does not end up in the wood. However, the discussion in this regard seems balanced to me.

We were excited to see the Etzold et al. 2022 study when it came out after submission of this manuscript, and added citations to it throughout the manuscript even before receiving this review. We have also incorporated references to Zwiefel et al (2021) and other references on wood formation in reworked versions of the introduction and discussion. This includes discussion of sink-limitation of growth, including drawing parallels with CO2-enrichment studies where GPP is enhanced but stem growth unchanged. (We believe that the most likely explanation is that any extra C goes into short-lived pools and is released as respiration over a relatively short time frame.) We feel that these changes have greatly improved the manuscript.

Some details
Consider using the term 'band dendrometer' instead of 'dendrometer band'.

Both “dendrometer band” and “band dendrometer” are broadly used terms. Because ForestGEO, in which our author team is rooted, has traditionally used the term “dendrometer band”, we prefer to stick with this.

Lines 67-69: See Etzold et al 2021 and Zweifel et al 2021. Both papers include analyses of stem-growth phenology of temperate deciduous forests over multiple years.

This sentence has been removed, but we now cite both papers multiple times throughout the introduction and discussion.

Line 87ff: The method neglects rehydration processes in spring, i.e. the stem swells, but this swelling has nothing to do with growth yet, but with the replenishment of winter tree water deficits. However, I agree that the method presented does not have a significant impact on the results when comparing different groups of data, because any systematic error will cancel out when comparing the groups.

We agree that our dendrometer band analysis approach captures only stem expansion, which is driven by both wood formation and hydraulic changes. We also agree that this should not bias our results. We also note that the tree-ring portion of the analysis is not influenced by stem hydration.

Related to this comment, we have revised our fitting method to provide more robust and stable estimates of starting and ending diameter. Briefly, this entails a shift from use of “L” and “K” to “a” and “b” parameters, as described in McMahon and Parker (2015; this is the paper describing the seasonal growth model that we employed). While “L” and “K” describe asymptotes of the fit curves, these are not always realistic, and the parameters “a” and “b” provide constraints based on observations. For example, the spring (initial / minimum) diameter is now constrained by spring measurements, whereas it was previously free to be significantly lower than the initial spring measurements. This analysis improvement did not substantially alter our results or affect our conclusions.

Lines 111-112: How is the length of a peak quantified? An explanatory sentence would be helpful or then the use of a different word for 'peak'.

We define the calculation for peak growing season length in the first paragraph under the section titled "Dendrometer band analysis".

L198ff: See Zweifel et al 2021 for the importance of various environmental factors on stem growth and see Etzold et al 2021 for a discussion of the importance of day length on growth.

We have added citations to both studies here.

L204ff: See Etzold et al 2021 who found that growing season length was not correlated with annual stem growth.

We now cite this study (and Zwiefel et al. 2021) in our discussion of drought, which now occurs in a different part of the discussion (2nd to last paragraph) following our revisions: "Regardless of the influence of spring temperatures on C cycling within the ecosystem, our results clearly demonstrate that the dominant effects of temperature on deciduous tree growth occur not in the spring, but during the peak growing season of the current or sometimes previous year (Fig.3, Extended Data Figure 7), when increased atmospheric demand associated with high temperatures can limit both leaf-level gas exchange and stem growth[23,28,30,49]. "

L228ff: However, see all work showing legacy effects for stem growth. Growth in the current year is strongly influenced by conditions that prevailed in the months and years prior. In addition, the concept of a balance between leaf and sapwood area implies that crown size strongly determines target stem growth. In summary, there is much evidence that a good (or bad) production year has its effect on growth with some delay (up to several years). Although no significant results were found here, a statement to this effect might lead to what I consider a more balanced discussion.

We're aware that legacy effects of the previous year(s) on stem growth are widespread, and our analysis including the previous year (Extended Data Figure 7) reveals that these are common within our data set (mainly for summer months). Here, however, we are specifically interested in whether warmer spring temperatures would have a lag effect, and find that this is not the case.

We have improved the presentation of this analysis by shifting it from a sentence in the discussion to a short paragraph in the results section: "To test whether warm springs might result in storage of non-structural carbohydrates (NSC) that would augment growth the following year[37], we extended the analysis to examine correlations between RWI and T_{max} in the previous year (Extended Data Figure 7). This revealed little effect of previous spring temperatures on annual growth, with significant positive correlations of RWI to previous March or April T_{max} for 5/142 ring porous chronologies and to previous April or May T_{max} for 7/66 diffuse porous chronologies." We then make brief mention of this in the discussion (1st sentence of second paragraph): "...our findings imply a lengthening of the period from

peak stem growth to the cessation of CO_2 uptake by the ecosystem and an increase in C allocated to functions other than stem expansion in the current or following year."

We have added brief mention of lag effects from the previous summer to the 2nd-to-last paragraph of the discussion: "Regardless of the influence of spring temperatures on C cycling within the ecosystem, our results clearly demonstrate that the dominant effects of temperature on deciduous tree growth occur not in the spring, but during the peak growing season of the current or sometimes previous year (Fig.3, Extended Data Figure 7), when increased atmospheric demand associated with high temperatures can limit both leaf-level gas exchange and stem growth[23,28,30,49]. " However, given space limitations and the fact that these lag effects are at most tangentially related to the influence of spring temperatures on annual growth, we did not expand our discussion of them beyond this.

L257 delete one "are"

Done.

L402: Why were conifer species excluded from the analysis?

We excluded conifers primarily for the practical reason that they are rare at our focal study sites. In addition, they are subject to different mechanisms and have been more broadly studied. We now include a statement on the contrasting states of knowledge on conifers vs broadleaf trees in the introduction (2nd paragraph): "Most studies on tree growth responses to warmer spring temperatures have focused on boreal or temperate conifers, which tend to respond to warmer spring temperatures with an earlier start to growth[24, 25] and increased annual growth in mesic climates[26, 27]. In contrast, we have little evidence as to how stem growth and $ANPP_{woody}$ respond to warmer spring temperatures in deciduous forests[10,11]. …"

Data statement, field conditions (subtitle field work, collection and transport): It is claimed that weather conditions are not relevant to the collection of band dendrometer data. This is not true, as stems can shrink in the range of up to one year's increment during dry periods (depending on location and species). If the authors are aware of this fact, it would be helpful to clarify the statement accordingly. In addition, a sentence should be added explaining why the authors believe this fact is not relevant to this study.

We are well aware of this and have fixed the error (caused by our misinterpretation of the information being requested). The statement now reads as follows:

"Dendrometer band measurements were made during the growing season from spring to autumn. Weather conditions have a modest effect on these readings. For example, stems shrink and swell depending upon hydration status (which, in turn, is affected by soil moisture and vapor pressure deficit), and metal bands may be modestly affected by temperatures (through thermal expansion of metal). These effects should be largely removed through the fitting of a seasonal growth model (detailed in methods section) and are not expected to bias our analysis."

Further, we address the effect of weather on dendrometer measurements with the following statement in the methods section (paragraph describing model fitting): "An advantage of this approach is that short-term shrinkage and swelling associated with rain events[35,57] and measurement errors show up as residual variation and do not unduly influence the parameters of interest. "

References

Etzold S, Sterck F, Bose AK, Braun S, Buchmann N, Eugster W, Gessler A, Kahmen A, Peters RL, Vitasse Y, et al. 2021. Number of growth days and not length of the growth period determines radial stem growth of temperate trees. Ecology Letters 00: 1-13.

Zweifel R, Sterck F, Braun S, Buchmann N, Eugster W, Gessler A, Haeni M, Peters RL, Walthert L, Wilhelm M, et al. 2021. Why trees grow at night. New Phytologist 231: 2174–2185.

Thank you for the review!