Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author
This manuscript evaluates the proximate factors associated with late spring frosts and freezes ("false springs") using a spatially and temporally extensive dataset for six species of European trees. In contrast with previous studies that have primarily examined one proximate factor at a time (e.g., lines 78-83), the current manuscript simultaneously models multiple proximate correlates—climatic and geographic—of false spring, some of which are changing with climate change. The scope of this study is impressive. However, one critical factor seems to be missing to me, which is the extent to which false springs reduce tree growth, survival and fecundity. Have the fitness effects of false springs remained the same through time? What are the ultimate ramifications of these shifting drivers of false spring for population growth rates and persistence under climate change? I recognize the challenge in collecting fitness data, but the current manuscript seems incomplete without a link to tree performance.

Here are some comments I had as I read the manuscript:

Line 18 (second sentence of the abstract): The second sentence of the abstract starts by saying: "Research to date has generated conflicting results," but the text has not yet stated what the problem or research question is. It seems pre-mature and out of place to indicate here that there have been conflicting results when previous text has not yet defined the scope of the research. Some restructuring could be useful.

Lines 20-23: The sentence starts " Here, we assessed the effects of...," but the text never indicates what the response variable is. I presume that the response variable here is leafing phenology, but the sentence could be strengthened by making the connection between climate and phenology more explicit. For example, instead of "using PEP725 leafout data," the text could read "on leafout phenology by using PEP725 data..."

Lines 55-57: "Temperate plants are exposed to freezing temperatures numerous times throughout the year, however, individuals are most at risk to damage in the spring, when freeze tolerance is lowest (Sakai & Larcher, 1987)." I got stuck on this sentence for a minute and wonder whether it would be worth rephrasing to indicate the timing of the year when freezing temperatures are likely to happen (presumably, fall and winter) and when freeze tolerance is the highest (presumably fall and winter). As stated, it led me to question the frequency of freezing temperatures during the growing season in the summer, and the timing of peak freeze tolerance. Are plants the least tolerant of freezing temperatures early in the spring specifically because leaf tissue is so young? Does that pattern also hold for non-deciduous species? [[Some of these questions are answered in lines 121-133, but this earlier sentence could be rephrase for clarity. In fact, I might recommend moving the text from 121-125 to this part of the manuscript because it provides biological details that are otherwise somewhat absent from the introduction.]]

Line 109: The URL listed here is incorrect. It should be www.pep725.eu, not .edu

Lines 113-114: Please provide family names for all six species.

Lines 125-128: How variable is the rate of budburst? Given the scope of this study, it seems incredibly important to analyze accurate budburst dates—inaccuracies in those dates could bias model results. I would imagine that there is intra and interspecific variation in the rate

of budburst. The text indicates that budburst data were not available for âĂIJthe majorityâĂİ of individuals. Are there sufficient data from the minority of individuals (given then extensive dataset) to account for intra- and interspecific variation in budburst rate? I appreciate the alternative model articulated that uses different rates for different species. Are these growth chamber data translatable to what happens in nature?

Lines 138-139: The authors defined false springs at temperatures of -2.2 degrees C and reference Schwartz (1993) or -5 degrees C, referencing (Lenz et al. 2013 and Sakai and Larcher 1987). A one sentence description of why these temperatures are relevant would be helpful. I appreciate that the authors provide citations, but the readers of this manuscript will want a more descriptive justification for these temperatures.

Lines 188-193, and later: There is a simple transformation between logistic regression beta values and odds ratios, where odds ratio= exp(beta). For example, using your data from Table S6 (also reported in lines 219-220), for every 2 degree C increase in men spring temperature, there would be a exp(-0.48)= 0.619 or $\tilde{3}8\%$ reduction in the odds of a false spring. This is perfectly interpretable, and IâĂŹm not quite sure what the value is in using an approximation for the probability that y=1 (which the divide by 4 rule is). Of course, my calculation depends on the values reported in Table S6 being the beta coefficients, I donâĂŹt quite understand how the authors calculated a 7.64% reduction in the probability of a false spring from a coefficient of -0.48 using the divide by 4 rule (as that would result in a 12% reduction). Are the values reported in Table S6 the coefficients?

Lines 146-193: This section indicates that the statistical models use "climate change" as a predictor (crossed with species in model 1, as well as being crossed with environmental variables in model 2). I believe âĂIJclimate changeâĂİ is a simple binary variable, with data coded as 0 (1951-1983) vs. 1 (1984-2016). For me, the wording at lines 146-149 could be revised to indicate that the authors created a binary climate change parameter, coding the data by whether it was collected before temperature trends increased (1951-1983) or after (1984-2016). The current wording ("we split the data:...") did not clearly signal to me that this was the "climate change" parameter referred to in the data analysis section.

Could you please define what the "NAO index" is? What data go into the calculations of the NAO index? Surely, I am not the only reader unfamiliar with this index.

Lines 200-201: The average values reported here for acceleration in budburst date and minimum temperature are not apparent in Tables S3 and S4, which seem to report statistical results from simple linear models, but not overall means. Please be clear about how these averages were generated, especially since it seems that there were climate change by species interactions for both variables. Perhaps additional tables could be provided with the species-level average changes, which would facilitate future meta-analyses. I know those data are provided in Fig. 2, but it would also be useful to have the associated sample sizes in tabular form for future meta-analyses.

Fig. 3: Should the X axis read "Change in the probability of false springs"?

Fig. 4: Please figure out a way to include the data points in the manuscript. I recognize that you have a very large dataset, but it is not appropriate to show predicted regression lines without some indication of the underlying data. The authors could provide histograms of the climatic and geographic factors for each species under false springs and not in the supplemental file. But, some effort needs to be taken to allow readers to evaluate the fit of

these models to the data.

Fig. 4E: I recommend recoding the X axis to 1951-1983 vs. 1984-2016 instead of 0 vs. 1. This recoding would facilitate reader comprehension.

Lines 239-240: This sentence refers to Figure 3 to show that the effects of climate and geography depend on the climate change variable. However, Fig. S2 is a much more powerful demonstration of these changes. I strongly recommend that Figure S2 is moved to the main document and Fig. 4 is placed in the supplemental file. Panel 4E could be retained in the main text.

Reviewer: 2

Comments to the Author
This was a highly robust study evaluating late spring freezing events following budburst (ie, false springs) across a 1951 to 2016 timeframe for six tree species across 11,648 sites in Europe, to determine the effect of mean spring temperature, distance from the coast, elevation, and the North Atlantic Oscillation using PEP725 leafout data. I thought the study was nicely done. I had few comments or issues with the manuscript.

The introduction does a nice job of introducing the topic and potential issues and impacts associated with false springs. Methods, results, and discussion are clear with the few exceptions I list below.

Strengths of the manuscript are definitely the robustness of the dataset. The largest weakness is the writing style. Although technically correct the authors could do a better job of making it interesting to the reader. It is a tough manuscript to read and didn't keep my interest.

Specific comments by line numbers follow.

56. Add semi-colon before "however".

70. Add semi-colon before ";however," and a comma after it.

99-106. Too late now but I would have liked to have seen more of a hypothesis here rather than just throwing everything in a barrel and seeing what falls out.

100. Do all of the tree species have the same risk (-2.2)?

107. I may be old-fashioned but I really like study area descriptions to put a study into the appropriate context but maybe GCB does not require one.

116-120. Impressive data set!

137-138. This sentence is repetitive with an earlier sentence.

192. Why 98% uncertainty intervals?

200. This is somewhat unclear. Why is it "advanced -5.22 days". So it didnâĂŹt advance?

278-287. You indicate your data has "reshaped the driving factors" twice in this paragraph yet your results seem to be in line with what previous researchers have found. Please clarify.

The graphics are informative.

I wish you well with the revisions.