Dear referee,

We would like to extend our sincere thanks to the referee for carefully reading and reviewing our manuscript. Your insightful comments and valuable suggestions have greatly improved the quality of our work. We deeply appreciate the time and effort you have put into providing such detailed feedback. We have incorporated your recommendations to enhance the clarity and robustness of our study. Thank you for your dedication and expertise

Referee 1:

Please include Ips typographus in the title and abstract (see my comments on the first version about lack of applicability to other bark beetle systems).

Ips typographus has been added to the title and the introduction. Also the other section where screened for a better use of the specific species, i.e., *Ips typographus,* and the more general wording “bark beetles”.

The abstract is still too suggestive of model accuracy, which was not evaluated (comparison of specific situations against observations). L 37-38 isn’t very helpful; I suspect “any” model could produce a wide range of observed dynamics, depending on what parameters were used. Similar to “accurately mirroring a wide array of observed…” I think it’s fair to say that by using different sets of parameters, model results spanned the range of observations, and that the model results indicated that including beetle outbreaks had a major effect on carbon dynamics.

The abstract has been adjusted to address this comment. The sentences on model evaluations now read :

“*Simulation experiments demonstrated the capability of ORCHIDEE to simulate a variety of post-disturbance forest dynamics observed in empirical studies. Through an array of simulation experiments across various climatic conditions and disturbance intensities, the model was tested for its sensitivity to climate, wind disturbance, and selected parameter values. The results of these tests indicated that with a single set of parameters, the ORCHIDEE outputs spanned the range of observed dynamics*.”

L 43: Can the profound insights be described briefly here?

I changed the sentence to explain why it was the main insight. I removed “profound” because it was too strong.

“Notably, the study revealed that modeling abrupt mortality events, as opposed to a continuous mortality framework, provides new insights into the short-term carbon sequestration potential of forests under disturbance regimes by showing that the continuous mortality framework tends to

overestimate the carbon sink capacity of forests in the 20 to 50 year range in ecosystems under high disturbance pressure, compared to scenarios with abrupt mortality event"

L 117: Which is used here, coupled or uncoupled?

I added a sentence that clearly says we use the uncoupled version. The text now reads:

"Unlike the coupled setup, which needs to run on the global scale, the stand-alone configuration can cover any area ranging from a single grid point to the global domain. In this study ORCHIDEE was run as a stand-alone land surface model."

At the beginning of Section 2.4, please add a paragraph summarizing which mechanisms are included and which are not. For instance, I don't see beetle survival in winter (as indicated on L 494) (also, it appears that ibeetlessurvival in Table 1 is different from the description in Section 2.7, which appears to calculate ibeetlessurvival as a function of host availability, not winter temperature…or, if I don't understand, please add more description about how winter T is included).

Adding such a paragraph would be redundant with section 2.4, 2.5, 2.6 and 2.7 in which the mechanisms are already described. Concerning $\mathbf{i_{beetles\ survival}}$ a dedicated section 2.7 explains our approach. We now explain why winter temperature is not explicitly used in eq. 11. The relevant text now reads:

"The capability of the bark beetles to survive the winter in between two breeding seasons is critical in simulating epidemic outbreaks. During regular winters, winter mortality for bark beetles is around 40% for the adults and 100% for the juveniles (Jönsson et al. 2012). In our scheme, this mortality rate is implicitly accounted for in the calculation of the bark beetle survival index (ibeetles survival). A lack of data linking bark beetle survival to anomalous winter temperatures, justifies the implicit approach and prevented including this information as a modulator of ibeetles survival. The latter explains why winter temperatures do not appear in eq. 11. Instead the model simulates the survival as a function of the abundance of suitable tree hosts which decreases the competition for shelter and food "

Providing one figure per variable that shows the different assumptions of the logistic curves (as well as the input variables?) would be helpful for readers.

It was our initial thought to but given the numbers of variable x assumption, we choose to show only variables that are relevant to understand the effect of the assumption on the behavior of the model. Figures in S2 have been added to the manuscript that show the parameter sensitivity effect on the tree stand densities, relative density index and total above ground biomass.

Section 3.3: Italicize variable names to be consistent?

Done. We agree it enhances readability.

L 570: Is the figure reference correct?

Thanks for noticing. The reference was not correct and has been corrected to “figure 4”.

L 617: Please add panel labels (e.g., “(a)”) to figures and refer to them in the main text.

Done. We agree it enhances readability.

L 620: Please add text noting that there is a difference between the duration and severity of bark beetle-caused tree mortality at the landscape scale versus at the stand/pixel scale. The landscape scale may have a longer duration and lower severity than the stand scale because beetles disperse across a landscape and cause mortality at different times. Such a distinction will be important in Table 3 and here in the interpretation of the model results.

Following the comment by the referee, we understood that the initial text in section 2.5 was not clear enough and we used the wording of the referee to clarify this issue. The new text now reads:

“At the landscape scale, which can cover areas up to 2500 km², the duration of mortality may be longer and the severity lower because beetles disperse across the landscape and cause mortality at different times. This distinction is important for interpreting model results, particularly when considering parameters like ird limit in the ORCHIDEE model. ird limit describes the proportion of trees surviving after an outbreak and should therefore be adjusted for the spatial scale of a gridcell in ORCHIDEE. In model set-up where a gridcell represents a single stand (~1 ha), ird limit should be close to 0, indicating that nearly all trees may be killed. However, in a simulation with grid cells representing 2500 km², not all trees will be killed, which is reflected in setting ird limit to 0.4“

L 622: What is “favorable”? For what objectives? If the authors mean comparison with Table 3, then I’m not sure I’m convinced: outbreaks can exhibit a range of characteristics. The authors may be able to eliminate some combinations of parameters as too extreme, but I’m not sure how parameters can be identified as most favorable.

The sentence was reworded as “By comparing the outcomes of the sensitivity tests (section 4.1) to a compilation of observations (Table 3), a first estimate for several parameters was proposed (Table 4).”

I’m less convinced of the selection of parameters in Table 4 based on comparison with Table 3. Table 3 should focus on Ips typographus, not other bark beetle species, and allow for a range in outcomes. Adding columns to Table 3 that describe 1) the parameter set yielding the best comparison, and values, and 2) the range of outcomes for the other parameter sets may help the reader understand the selection process better.

We agree that the link between table 3 and table 4 is weak and unclear. In order the make the method clearer and reinforce the link between the outbreak criterias describe in table 3 and the parameter set chosen in table 4, we made a series of adjustment in the manuscript:

- We add a new column in table 3 call “reference range” which give a credible range use in the calculation of the 5 score (one for each criteria)
- We add a paragraph explaining how the Credible score was calculated :

*“Based on Table S1 and the reference range in Table 3, a score is calculated for each parameter set. The Credibility Score (CS) is the sum of four scores, indicating that the result falls within the four reference ranges described above and no outbreak is triggered when DRwindthrow = 0.1%. The CS is computed as follows: CS = (score1 + score2 + score3 + score4) * score5. Only parameter sets achieving a CS of 4 will be selected. If multiple values are possible for a given parameter, the most frequently selected value will be preferred.”*

- We add a supplementary table (S1) which gives all the parameter sets tested and their respective credibility score according to table 4

- In table 3, we now stress the fact that our data come from different beetle species, different host species and different climate zones in the caption of table 3, i.e., “”. We agree this is suboptimal but it currently is the only approach to compile some data. Despite its shortcomings, lumping data from different species and climate regions is, however, common in literature reviews.

We hope that with these changes. Our choice seems less subject to interpretation while referring clearly to a protocole.

Figure 1. “host susceptibility” instead of “host weakness” will increase reading comprehension. Same with text in the main body.

Done.

Why does host weakness decline? Isn’t this an external factor associated with climate and age/density? Does this result from killed hosts that relieve density competition?

Yes, when the density is declining from the outbreak itself the competition between trees for light and nutrients declines too. As a consequence the host weakness decreases**.** In our model it is the main reason why an outbreak stops. This explanation has now been added to the caption of Fig 1.

“Figure 1 : Dynamic interplay of the different host and beetle characteristics during a bark beetle (Ips typographus) outbreak. The time window spans four outbreak development stages: build-up, epidemic, post-epidemic, and endemic. The curves represent key characteristics,

showing the growth in beetle population and subsequent decline in host population. Ihosts dead characterizes the presence of defenseless uprooted or cut spruce trees; ihosts alive, characterizes living spruce trees that could become hosts for the bark beetles; ihosts susceptibility, susceptibility of spruce trees to bark beetle attack; ibeetles mass attack, quantifies the capability of the bark beetles to mass attack; ibeetles survival, characterizes the survival of bark beetles. Host and bark beetle characteristics are detailed in the subsequent text. When the density of the host trees is declining due to an increased host mortality from the bark beetle outbreak itself, the competition between trees for light and nutrients declines as well. As a consequence, the host susceptibility decreases which in ORCHIDEE is the main pathway for an outbreak to move back to the endemic phase. After 1 year the wood from a storm is not fresh enough for bark beetles to breed in. In ORCHIDEE, the bark beetle population needs to be capable of mass attacking living trees within a year to make the transition from the build-up to the epidemic phase. "

Why does hosts dead decrease during build-up and appear to be zero during the outbreak?

We only use the woody biomass of hosts that are dead during the current years by other sources of mortality than bark beetles. After 1 year the wood from the storm is not fresh enough to host bark beetles which tend to favor freshly dead trees.

This is now explained in 2.7 at line 399-400 "However, after a year, this substrate becomes unsuitable for breeding and is excluded from the ihosts dead calculation." and has been added to the caption of Fig. 1 (see previous comment for the exact wording)

Please define the variables in the caption.

Done

Figure 2. Having a flow diagram is a good idea. This diagram is very hard for me to understand. The reader cannot discern the meaning of the variable names. Please add them to the caption. Add boxes/ovals/circles that indicate (separately) actions/state variables/etc.

Figure 2 has been improved.

What is the difference between the red and black lines?

Black lines denote intermediate variables such as indices. For the red lines the source is always a variable that describes forest structure such as RDI. The flow diagram was redesigned to better distinguish between variables internal to the ips typographus model and variables that come from pre-existing ORCHIDEE functionality.

It would be helpful to have the processes (mechanisms) grouped together with a surrounding labeled box.

The flow diagram was redesigned and mechanisms have been grouped (shown in the same box) to strengthen the link with the text.

Table 3: Please use a different word than "measure" for model output, which is not a measurement.

Done

Also, because this study is no longer an evaluation against measurements, please use a different term than "how to check", such as "how to calculate".

Done

Figure 3. I assume that DRwindthrow is a parameter specified before a model run…please be explicit in the caption. Is this a relationship from an equation or an emergent property that results from a model run? Also, it would be helpful to readers to provide brief explanations of that parameter and ihostsdead in the caption.

As now explained in section 3.5 line 560-562 windthrow damage has been forced in order to focus our study on bark beetles. A paper in which storm damage is an emergent property that results from a model run in under preparation and will be submitted in the coming months. The text now reads "In this simulation experiment, the amount of fresh dead tree hosts (Nwood) used by the bark beetles to breed was controlled by modifying the maximum damage rate of a windthrow event (DRwindthrow) in ORCHIDEE. Seven DRwindthrow were simulated (i.e, 0.1%, 5%, 7.5%, 10%, 15%, 20%, 35%)."

A short explanation has been added to the caption of Figs X. and Y to better explain the meaning and values of these parameters.

Figure 8: What are the asterisks?

I add the Wilcoxon test explanation into the caption**.**

Also, please provide a brief explanation of the three versions (or frameworks? pick one term).

Done

Adding panel labels (e.g., "(a)") to all figures and referencing them in the captions and main text will minimize reader confusion.

Done

There are multiple typos and writing and citation style errors in the manuscript.

We used this revision to carefully check the spelling, grammar and terminology used in the manuscript. The only native English speaker among the authors has left research which means that we have to rely on software tools to check grammar.

Referee 2

Substantive comments

• Figs. 1 and 2 are very low-resolution and text is hard to read. Fig. 3 also has weird text but it's larger so not as much of a problem.

New higher quality figures were prepared.

• Line 265: How is this limited to between 0.5 and 1? Looking at the subsequent equations I don't understand.

Thanks for noticing. This sentence was right in the previous version but given the equation evolved any value from 0 to 1 is possible. I change the sentence on line 265 to reflect this change.

• Line 281 (Eq. 5a): The "1 –" in the exponential seems to make this function point the incorrect way, with i lower (i.e., stand is less attractive to beetles) when PWS*+, is high (i.e., a strong drought happened recently) and vice versa.

In orchidee PWS=1 means no stress, this is the reason why with a 1- but I understand your concern. To make it clear and to be consistent with the code, I decided to add a sentence explaining this around line 281.

• Line 285 (Eq. 5b):
-What is "nb age class"?
-Why is age class 1 the maximum considered?
-(The above two questions also apply to Eq. 6b.)

I clarified this issue at line 285. The text now reads:

"$$PWS_{max} = \sum_{nb\ age\ class=3}^{age\ class=1} max(PWS_{spruce,\,n},\ PWS_{spruce,\,n-1},\ PWS_{spruce,\,n-2}) \times \frac{F_{spruce\ class}}{F_{spruce}} \qquad (5b)$$"

$$\text{“}i_{rd\ spruce} = \sum_{nb\ age\ class=3}^{age\ class=1} \frac{D_{age\ class}}{D_{max}} \times \frac{F_{age\ class}}{F_{spruce}} \quad (6b)\text{”}$$

-Are the PWS values here averages? Over what time periods?
-If PWS#-./0' is this year's plant water stress, PWS#-./0',)23 is last year's, etc., then PWS#-./0',)24 being included means that it's actually looking over the past four years, not three as mentioned at line 279.

We correct for $PWS_{spruce,n-2}$ because ORCHIDEE only checks 3 years. We also add the description of PWS_{spruce} at line 289.

• Lines 358-361: This text suggests that Fig. 1 shows both a return to endemic stage and an evolution to epidemic, but it looks like it only shows the latter.

The referee is correct. I removed the reference figure 1.

• Line 393: Is Bwood live biomass only?

Yes, it is used as a reference in order to reflect that when Nwood = Bwood the susceptibility reaches its maximum. That is arbitrary but we think it is reasonable to think that above 50% of the living trees kill, there is too much substrate than an endemic bark beetle population can eat in 1 year.

• Lines 451-452: Why is this index calculated separately for each age class whereas the indices contributed to by Eqs. 5b and 6b are calculated as cross-age-class averages?

In order to consider different mortality rates between dominant, subdominant and understorey trees (here represented by the 3 age classes) the indices are calculated for each age class. In equations 5b and 6b indices are calculated across age-classes to account for stand susceptibility. Eq 5b for exemple is used as a proxy of stand healthiness for beetle attraction.

• Lines 490-491: What does “acclimation” refer to in this sentence?

It is now clarified that “acclimation” means “change their behavior or their resistance to external stressor such as winter temperature, of the bark beetle population to each location.”

• Line 517 and in Results: Ranges don't make sense in this context. What exact values were tested? Line 516 says that only three values were chosen.

We reformulate that paragraph in order to clean the method. it is now read :

"For each of the four variables, three distinct values were assigned to two parameters labeled "Shape" and "Limit". The Shape parameter determines the shape of the logistic relationship, with three values tested: (a) Shape=-1, yielding a linear relationship, (b) -5<Shape<-30, resulting in a logistic curve, and (c) Shape=-500, turning the logistic relationship into a step function. For the logistic curve, the exact Shape value between -30 and -5 is chosen according to each index under study: (1) Ssusceptibility = -5; (2) Sactivity= -5; (3) Smass attack= -30; and (4) Sgeneration=-5 . For Smass attack, a higher value has been chosen because the slope of the logistic curve has a significant impact in order to trigger an outbreak."

• Lines 559-560: Is this 20% number referring to max65""& ?

Yes, I have clarified in the text.

• Lines 566-571: How do you calculate the number of generations per year?

It is the exact same equation as in the original paper from Temperli et al. 2013. The equation is now given on line:

$$DD_{eff} = \sum_{i=1}^{n_{diapause}} (T_{opt} - T_{min}) * e^{(0.0288* T_{bark,i})} - e^{(0.0288* b_{eff} - (40.99 - T_{bark,i})/3.59)} - 1.25 \quad (10)$$

Where i is a day, $n_{diapause}$ is the number of days between the 1st of january and the day of the diapause. T_{opt} *(30.3°C)* is the optimal bark temperature for beetles development and T_{min} *(8.3°C)* is the temperature below which the beetles developpement stop. $T_{bark,\ i}$ is the average daily bark temperature. $T_{bark,\ i}$ is calculated as the daily average air temperature minus 2°C. All parameters values are taken from Temperli et al. 2013

• Lines 573-581 and subsequently: "Control" and "climate experiment" feels weird. For the sites, it's more of a climate gradient that starts with HYY and continues through HES/FON. There's nothing a priori different about HYY that makes it a "control" and the other sites "experimental." Similarly for the windthrow damage rates: Maybe if the lowest value was 0 that would be a control, but it's not. (Note that I'm not saying true control treatments are needed—I don't!)

I understand your concern, HYY should be considered as a reference location for no outbreak and 0.1% damage has a small enough damage rate that should never trigger an outbreak. This comment was addressed by changing the wording throughout the text.

• Lines 588-593: Note that (1) is the default or previous ORCHIDEE setup.

Yes, I make it clearer

Sect. 4.1: I'm unclear on how the sensitivity tests indicate anything about "false positives" due to the calculation of any specific parameter. Weren't all experiments done with all parameters?

False positives are directly deducted from the control (now renamed as reference for no outbreak). If an outbreak is observed for HYY or 0.1% damage rate we can consider it as false positive but not in its statistical meaning but more as a boundary that ORCHIDEE should not cross. So given the confusion surrounding the term "false positives" in this context, the wording was replaced by "improbable outbreaks" throughout the manuscript.

Table 4: Missing *S*0"*-'$7$7") .

Done

Lines 683-684: This sentence is confusing. "Turning back" from what? If the "tipping point" is 9 kgC/m2, what is the "threshold"?

Line 690-694: I have rewritten this part because some of the statement was wrong; It is now more clear since 9.0 is not considered as the tipping point that led to an epidemic.

Lines 684-685: REN clearly reaches some "tipping point"—i.e., it reaches a minimum and then recovers. What's so special about 9 kg/m2?

At line 691 : "at a biomass level around 9.000 gC.m-2 equivalent in our simulation to an $i_{rd\ spruce} = i_{rd\ limit}$=0.4, there's no turning back until that threshold is passed."

Fig. 7: Add shading or some other indicator of where the outbreak (as defined according to Table 3) begins and ends.

We added a grey area that represents the epidemic phase in each panel/site.

Lines 721-722: "a more resilient recovery" doesn't seem right for the continuous case, where there's no specific event to "recover" from.

I agree that by rereading the sentence I realize that a mistake was made because it makes sense for the "abrupt" configuration but not for the "continuous" one. I decided to remove this sentence.

Lines 752-753: Unclear how this sentence is related to the rest of the paragraph (snags).

I have removed this sentence because none of our figures suggest the mechanism even if it is indeed happening.

Line 767-775: Terms should be defined in this paper. It's not really helpful to cite a paper for readers to look up the definitions, especially after the terms are used. Ideally definitions should be mentioned in the introduction or methods so that they're defined before they're used—note that "resilience" is used in the results section.

The definitions were added in the text.

Lines 777-799 (Sect. 5.3): Discussion (and also Conclusion) should mention that this depends on spatial scale and model use. If one is looking at global-scale NBP, there might not be as much of an impact of including abrupt disturbances because they're happening constantly somewhere.

In this study, we only check the temporal aspect of the abrupt versus continuous mortality. Increasing the spatial domain is one of the main objectives of the second paper. This is the reason why we did not cover this aspect in this study.

Technical corrections, etc.

• Line 257: "analogue the the" should be "analogous to the". done
• Line 279: "average" is sort of confusing here. It looks like it's the (weighted) average across different age classes of the maximum water stress in each age class. Maybe just delete "the average" from this sentence to make it work. Changed to maximum
• Line 288: A paragraph break before "In addition to drought" would help break this section up a bit. Done
• Line 308: "mean quadratic" should be "quadratic mean". done
• Line 323 (Eq. 7b): "none" should be "non". done
• Line 372: Is "excess" the right word? replace by stimulation
• Line 409: Incomplete sentence: "The amount of suitable tree hosts." done
• Line 464: "bettles" typo in subscript. done
• Line 501 (Table 2):
o I thought the spelling looked weird for "Wetstein," and indeed, it doesn't actually seem to be a place in Germany. The listed coordinates point to Třeboň, Czechia, about 400 km southeast of the similarly-named German mountain Wetzstein. The listed coordinates do, however, look similar to those of the CZ-Wet (CZECHWET) site near Třeboň. done
o I can't find the HES site in the FLUXNET site table, or any site with similar coordinates. It's possible it's just missing from that table, but given the above issue it warrants a double-check. HES was removed from FLUXNET in 2015 but we still use the forcing from 1997-2014.
o The FON latitude appears incorrect according to the FLUXNET site table. done

o REN site latitude should round up to 46.6°N. (Check rounding of other sites as well.) done
Line 570: Reference to Fig. 3 should be to Fig. 4. done

Figs. 3 and 4: These figures refer to “Control” and “Climate experiment” which isn’t explained until later in the text. Done
Line 606: Delete “)”. done
Line 607: Delete “)”. done
Line 617 and elsewhere: Instead of having to say “4th row, 2nd column,” add subplot labels a-h. done
Lines 678, 683: Thousands separator should be comma, not period. So 9,000 instead of 9.000. Or to avoid cross-cultural confusion, just use 9 kg instead of 9,000 g. Done
Line 718: “Abrupts” typo. done
Line 731: “maining”? replace by impact
Line 761: First comma should be a semicolon. done
Line 767:
o “resistance” should be defined. It’s not really helpful to cite a paper for readers to look up the definitions, especially after the term is used. done
o “locations” should be “location” (or just deleted) ??
Line 810: “Ips Typographus” should be “Ips typographus” done
Lines 811-812: Latin name should be italicized done
Line 860: Same DOI repeated twice. Maybe these both will be replaced with the new DOI the authors mentioned in their cover letter? data and script have been pulled together for an unique DOI