**Warner College of Natural Resources**

**Forest and Rangeland Stewardship**

123 Forestry Building

Fort Collins, Colorado 80523-1472

<https://warnercnr.colostate.edu/frs>

June 16, 2025

Dear editor,

Thank you for giving us the opportunity to submit a revised draft our manuscript, “Drought may initiate western spruce budworm outbreaks, but multi-year periods of increased moisture availability promote widespread defoliation” for publication in the journal *Forest Ecology and Management.* We appreciate the time and effort that you and reviewers put into providing comments on manuscript.

We have carefully considered each comment and done our best to incorporated nearly all of the suggestions. Notably, in response to a large portion of Reviewer 1’s comments, we have clarified that here we are testing for a generalized Moran effect (sensu Engen and Saether 2005) not the Moran effect, which has strict assumptions about population dynamics. In response to Reviewer 2’s comments on the presentation of the manuscript, we have revised the text to be more concise in several areas.

Below we provide the point-by-point responses. All modifications in the manuscript have been highlighted in *blue italicized text*.

We believe these revisions have greatly improved our manuscript and we hope that the manuscript is now suitable for publication.

Sincerely,

Sarah Hart

Assistant professor of Applied Forest Ecology

Forest and Rangeland Stewardship

Email: [Sarah.Hart@colostate.edu](mailto:Sarah.Hart@colostate.edu)

Tel: 720-980-9264

Review of FORECO-D-25-00360

Reviewer comments:

Reviewer #1: Review of FORECO-D-25-00360

Summary:

This is an important spatial analysis of an important insect - the western spruce budworm - using tree-ring data as a proxy for insect impact over deep time - hundreds of years. This is a topic first addressed substantively by Swetnam and Lynch (1993) in "nearby" Arizona, and subsequently revisited by Ryerson et al. (2003) in Colorado. The question of the role of cycles of drought and moisture in promoting outbreaks of high intensity is also important. The approach is also novel because it separates triggering/initiation from intensity post-triggering. This is not just novel, it's appropriate given emerging insights about the eruptive behaviour of the budworm complex. [That matter is addressed in greater depth in the detailed review.]

The data are a high quality, even if the study area is limited in extent, to a part of just one state. The analyses are appropriate. Readers of this journal would welcome a contribution such as this. The manuscript was a pleasure to read because the writing style is clear and concise with perfect grammar.

The problems with the manuscript are, in my view, not in the Methods or Results, but in how the study is framed, regarding theoretical ecology, insect population dynamics, and vis a vis the discussion of closely related systems, such as eastern spruce budworm, as well as underlying models of eastern and western spruce budworm, both modern and historical.

I'm going to sign my review, because I think the substance of this manuscript has a lot of merit, but the manuscript overall could be even more impactful if it can be framed more robustly - and this is a matter of opinion, albeit considered opinion. It's obviously up to the authors how they want to frame (or re-frame) the work, but I hope my copious criticisms are taken as constructive and are seen as worth acting on. If not, at least they will know who to blame. This is also a case where knowing the reviewer's name could be helpful in formulating a reply. It's ok to dismiss points that I make if a robust counter can be presented.

This analysis needs to be published. I think it's a question of how the analysis ought to be interpreted, and therefore how it ought to be framed. The detailed review will clarify what I mean by this.

I'm not going to cover the Methods in depth because I find them appropriate and convincing enough for the purpose. I'm hoping my review will be complemented by additional reviews that I suspect will tend to focus more on the detailed methods on far, far less on the framing in the introduction and discussion.

Note that I insert a number of graphs as part of this review - roughly a dozen or so. That's because I think there is ground uncovered that might usefully be covered. This is opinion. Not everyone will agree. This review will read more smoothly if the file containing graphics is read instead of this text-only window.

Signed Barry J. Cooke

Detailed review:

Abstract.

Lines 17-18: “regionally synchronous” – I’m curious to see what is meant later on by “region”. Line 14 refers to “12 sites across central to northern Colorado”, and Colorado spans only 400km in latitude. “northern Colorado” is a relatively small “region”. Meanwhile, Figure 3 suggests synchrony (covariance) is never actually greater than zero (when the width of the confidence interval is considered), and that the mean covariance degrades to 0.2 at distance as small as~40km. First, this parameter – 40km - is not what I would consider to be a “region”; it’s not large enough. The Fig. 3 covariogram truncates at 100km, which is the smallest “region” I can imagine that would be relevant for questions of spatial synchrony of pest dynamics. Second, it’s not clear the Fig. 3 covariance is high enough even at its maximum (~0.3), to be considered “synchronous”. For example, contrast with Peltonen et al. (2002) who illustrate in their Figure 2 (below) western spruce budworm outbreaks that are highly correlated (r ~ 0.7) at short distance classes of ~100km and don’t decline to ~0.2 until 300km. That is a process most would consider “regionally synchronized”. And it's the same for five other insect species in their 20-year old graph. Not even their mountain pine beetle series shows as low synchrony as these WSBW data in Fig. 3. And note their Fig. 2 correlograms all \*begin\* at 100km, where this Fig. 3 covariogram \*ends\* at 100km.

For comparison:

A graph of different weather conditions

AI-generated content may be incorrect.

The fact that one study (this one, WSBW in CO) uses tree ring data and the other (Peltonen, WSBW in WA+OR) defoliation data is not relevant to the question of what defines a “region” and what constitutes a “significant” level of “synchrony”, although that may well explain the different correlogram/covariogram shapes/heights. I’ll look forward to that point being addressed in the Discussion.

*Thanks for this feedback. To avoid confusion with other research conducted at different spatial extents and scales, such as the work by Peltonen et al. 2002, we’ve revised our text to refer to landscape-wide outbreaks, rather than regional outbreaks.*

*We note that in our analyses, as in those presented by Peltonen, significant of synchrony is defined using a bootstrapping approach. We agree that interpreting the ecological importance of any statistically significant association is a little subjective.*

I’m frankly inclined to think the horizontal scale differences – the breaking at 100km - may be a “feature, not a bug”. If the WSBW system in this Fig. 3 exhibits rates of synchrony as low as MPB (see their Fig. 2), then maybe the WSBW system is not as periodic, and is more eruptive, than most tree-ring studies would have us believe (Alfaro, Swetnam, etc). Maybe WSBW is more like Campbell (1993) described (his Fig. 29, below), with larval survival rates \*rising\* with larval density – an “Allee effect” that leads to highly eruptive and non-synchronizable intrinsic dynamics.

A diagram of a graph

AI-generated content may be incorrect.

This eruptive dynamic is reflected in Campbell’s (1993) synthetic Figure 53 (which he compares to Morris & Watt’s (1963) similar model for eSBW):

A graph on a white background

AI-generated content may be incorrect.

Note there is no hope for such a nonlinear first-order growth process to exhibit periodic and synchronizable intrinsic oscillations. Mass dispersal will lead to an unceasing sequence of cascading eruptions – a domino effect. The only way such a nonlinear eruptive population process can be synchronized is through the forcing effect of correlated environmental triggers – a powerful exogenous effect that (a) has nothing to do with the weak “Moran Effect” (1953) observed in near-linear second-order systems where cycle phase is undetermined (“phase forgetting”: Nisbet & Gurney 1976) and therefore free to be shaped by the sum history of stochastic perturbations, but (b) is the actual subject of this paper!

Scanning forward in the manuscript, I note the authors cite Moran (1953) three times in the introduction and three times in the discussion. So the points I’m making here regarding the Abstract will become relevant later in the review.

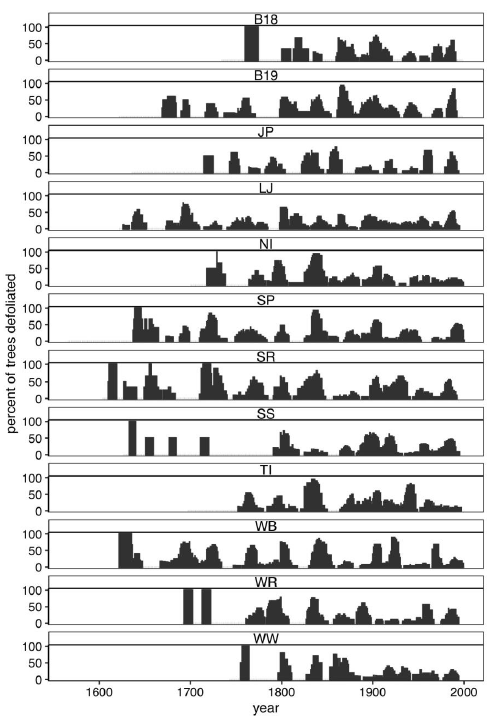
At this point I anticipate the authors will have a difficult choice to make during a major revision. Because Moran’s theorem is simply not applicable to nonlinear first-order systems. Flower (2016, p.14 ) gets this right: “Strictly speaking, the Moran theorem requires that the populations in question have homogeneous, linear density dependence structures.” She also writes: “The range of outbreak durations and time between outbreaks seen in the records I analyzed suggest locally varying population dynamics. Broadly similar, but not identical, periodicities have been reported for WSB outbreak records in different regions [20]. Differences in non-climatic conditions, such as predator characteristics, habitat quality, disturbance regimes, or land management practices, may be responsible for this geographic variability [1,45,55]. Local variations in WSB population dynamics violate a key assumption behind the classic Moran theorem”. This represents a huge caveat to the oversold manuscript title: “Three Centuries of Synchronous Forest Defoliator Outbreaks in Western North America”.

This is hardly the epitome of “synchronized” “cycling”:

A black and white image of a city skyline

AI-generated content may be incorrect.

And yet the current manuscript shows even LOWER levels of synchronization, at even SMALLER scales!



A graph of different colored squares

AI-generated content may be incorrect.

So (1) the proposed body of theory doesn’t apply (to either nonlinear systems or to heterogeneous systems). Moreover, (2) this Colorado WSBW system also doesn’t appear to be synchronized to the level expected either under Moran’s hypothesis or to the levels reported in Peltonen et al. (2002) for WSBW or Flower (2016) (who used her own custom “similarity” index).

Moreover, the methodological split of dynamics into the pre-eruptive environmental response function versus the post-eruptive response phase is inherently more aligned with Campbell and Morris’s multiple equilibrium hypothesis (nonlinear first order) than with Royama’s and Moran’s single equilibrium hypothesis (near-linear second-order), where there would be no material basis for having separate environmental responses. For this reason, a major re-write is going to be required on the Introduction and Discussion, even if the Methods and Results were to stand as is. The difficult part is that the authors are going to have to choose which framing they prefer.

Similarly, scanning ahead on “Campbell” I notice Campbell 1993 is not cited.

Campbell, R.W., 1993. *Population dynamics of the major North American needle-eating budworms* (Vol. 463). USDA Forest Service, Pacific Northwest Research Station.

But Campbell1987 is cited, in line 116, and it is cited as though it is in alignment with Royama (1984), re: the role of natural enemies, even though it is not. [And in the next four lines, frequent citation is made of Nealis and Regniere (2009, 2021). But Regniere and Nealis (2007) is also not aligned with Royama.] It’s important to appreciate, as I outlined above, how Campbell did not agree with Royama on the role of natural enemies. Cambell (1993) reports Allee effects occurring from predator escape at low to moderate densities. This is the same function described by Regniere & Nealis 2007 (their Figure 7) for eastern spruce budworm:

A graph of a function

AI-generated content may be incorrect.

… and this nonlinear response agrees with Watt and Morris - a point Campbell makes explicit in his Fig. 47 - but disagrees with Royama. This is the problem with using both empirical population ecology and theoretical ecology as a context for a study: there’s a lot of disagreement, and the limited areas of narrow agreement are just starting to be mapped out. It takes a lot of background reading to make sense of one species, let alone two. One can choose (as most do) to avoid this “hairball” or one can choose (as few do) to tackle it head on. Problems arise when the deep, murky topic is skimmed introductorily or in discussion. It leads to an incoherent introduction that can only lead to confusion later on , in the discussion. A skilled writer could perhaps avoid the “hairball” and get a faster publication, but this only leaves it to another occasion to set the work in its full and proper context. So why not do that now?

*Thank for you for this very detailed critique. In response to this, we have revised our framing to highlight that we are looking for generalized Moran effects.*

Lines 18-19: “In the first several decades following Euro-American colonization, outbreaks were shorter, more severe, and less synchronous across the region”. I’ll be looking for a quantification of duration and synchrony before and after this timepoint. At the moment all I’ve noticed is a single covariogram spanning the full time frame Figure 3.

*Thank you for this comment, unfortunately we are limited in how much detail we can present in our abstract.*

Line 20: “changes” – Be explicit about rising or falling host abundance

*Changed to “decreases”.*

Line 21: “land-use practices” – As in Line 20, be explicit about what “practices”. By Line 21 it should be clear whether you’re talking about increased burning reducing host availability and thereby decreasing outbreak duration and synchrony.

*We have revised this sentence to read: “In addition, we found outbreaks were often initiated by drought events and sustained by periods of above average moisture availability, consistent with the pulsed stress hypothesis.”*

Line 23 “pulsed stress hypothesis” – at this point it seems you have two co-existing hypotheses, the resource pulse hypothesis (at slow time scales) and the “pulsed stress hypothesis” (at fast time scales). The two ideas are compatible, and are even separable and additive, if they operate at distinct time scales. It seems to me this an opportunity to talk about “slow-fast” systems (Holling) that exhibit “cross-scale dynamics” in time, not just space (Raffa).

*Thank you for the suggestion. To keep our abstract to a sensible length, we were unable to include discussion of this point here.*

Main text:

Line 31: “Euro-American colonization was initially linked with lower susceptibility to outbreak” – It’s not clear what is meant by the word “initially”. You would want the putative mechanism mentioned here, i.e. what land use practice and whether it putatively raises or lowers host abundance.

*To improve clarity, we have removed this highlight.*

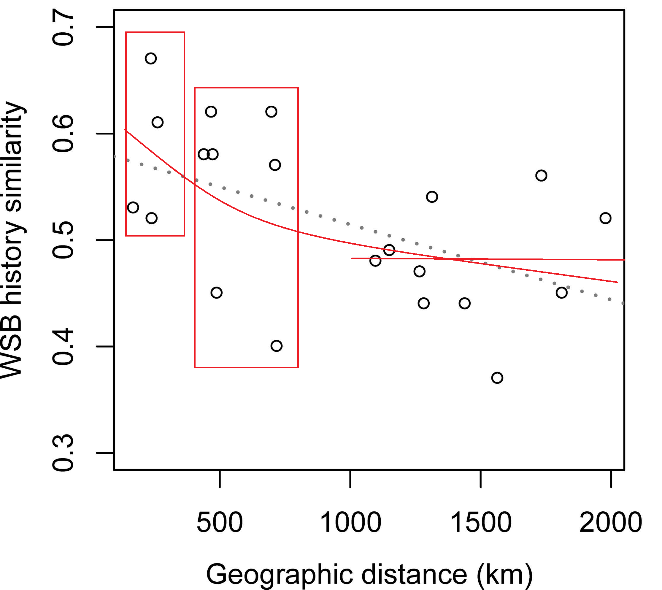
Line 42: “WSB”. I always prefer WSBW, to align with “SBW” = spruce budworm. [SB tends to get used for “spruce beetle”.]

*We have considered this suggestion, but have decided to stay with WSB as our abbreviation because it is widely used in the literature by forest managers in the region.*

Lines 48-51: no mention of the potential role of positive feedback in generating eruptions, just the failure of negative feedback. See Campbell 1993.

Thanks. We have revised this text to now read: “These outbreaks occur when several thresholds in the host-WSB system are crossed and feedbacks among the WSB populations, host trees, and natural enemies lead to rapid growth of WSB populations.” We have also and included reference to Campbell 1993.

Lines 55-56: “Importantly, outbreaks often occur synchronously across broad spatial extents (*i.e.*, 1000s of kilometers) (Flower, 2016)”. Flower’s caveat above is an important caveat. They often do. But they often don’t. First, her metric of “synchrony” actually drops off quickly at 400-500 km. But she fit a linear model instead of a nonlinear one. Second, her metric is questionable, as are the baseline outbreak reconstructions used. If these are done to a crude specification, counting only the most intense growth departures, synchrony appears to rise. This makes it difficult to compare across studies, which is what the community desperately needs.



A group of graphs showing different types of data

Description automatically generated

So this notion of strong, spatially homogenous periodicity and strong, temporally unwavering synchrony has, I think, reached, or even passed, its zenith, in both the eastern and western spruce budworm systems. What people \*say\* a paper shows versus what a paper actually shows can be two somewhat different things. Flower (2016) is a case in point. I think the interesting question is WHY does the periodicity (which is pretty weak) vary latitudinally, and WHY does synchrony break down for a century? If you don’t allow for the significant spatial and temporal variation that clearly exists, then there is no incentive to explain that variation. Instead of doing science, we’re honoring a tradition.

*This critique of Flower (2016) was helpful for thinking about revising our manuscript. We have revised our text to highlight that we are really testing for generalized Moran effects not a true Moran effect. This change in our framing allows for outbreaks to some areas, but not all sites to experience synchronous outbreaks because generalized Moral effects are only one of several potential drivers of outbreak dynamics.*

Lines 73-75: “the forcible displacement of native peoples by Euro-American settlers and the ensuing changes in land management practices altered forest composition, structure, and disturbance regimes across much of the Western US”

I wonder specifically what the role was of the Spanish gold miners of New Mexico, for example, in altering forest landscape structure. Do all types of clearing not lead to a rise in weedy Douglas fir? I’m unclear on the hypothesized mechanisms of land use change and specific changes in forest landscape structure. Increased clearing lowers the abundance of everything. Increased fire favors some species over others, but only in the long-run post-fire.

*We discuss in more detail the effects of Euro-American settlement on forests in the subsequent paragraph.*

Lines 89-91: “Collectively, similar changes in effects of Euro-American land-use practices have been hypothesized to make stands more susceptible to WSB outbreaks and lead to increases in the severity and synchrony of WSB outbreaks” … specifically because of the rise of Douglas-fir? If so, then in replace of what?

*It’s not actually that Douglas-fir is necessarily replacing other species, but more that active fire suppression in this system can increase overall stand density, often due to recruitment of Douglas fir.*

Lines 92-94: “evidence for this effect appears to vary regionally” How so?

*We have revised this text to highlight that different studies have come to different conclusions. Specifically, the text now reads “However, prior research has found mixed support for this effect”*

Line 102: “This apparent contradiction may arise for several reasons.” Not the least of which would be that factors affecting eruption might occur on fast time-scales (Larsson’s stress hypothesis), and factors affecting duration might operate on decadal time-scales (White’s nutrition hypothesis). Finally, factors affecting regime trends (forest succession) might operate on centennial time scales (Holling’s resource pulse hypothesis). And all these independent effects could easily be operating simultaneously. i.e. They’re not mutually exclusive “hypotheses”. They’re all postulated effects.

*Thanks for this comment. We agree.*

Figure 1: So we are dealing with a strip of a box 50km x 300km, in CO. The goal in discussion is going to have to be to navigate what we get from this relatively small -scale study to the larger scale studies of Peltonen (WA+OR) and Flower (range-wide, including Canada). Because otherwise we have this apparently weird contradiction that the smaller the scale of study the lower the observed synchrony. This won’t make sense to the community, so it needs to be reconciled in this paper, and that means it needs to be introduced as an object of study.

Full disclosure that I analyzed the data of Swetnam and Lynch (1993) back in 1994, and found a result not dissimilar from yours, in terms of these lesser localized outbreak pulses that did not align all that well with the regional outbreak average. There was a fair amount of asynchronous pulse eruption between watersheds, which is not consistent with the Royama-Moran theory of budworm outbreak cycle indication and synchronization. But it is consistent with the Watt/Morris/Campbell/Nealis view of the budworm “outbreak” “cycle”, where the degree of “synchrony” of an outbreak varies depending on the triggering and amplifying perturbations on each eruption, including the extent of dispersal. (That some eruptions may be extensive is not evidence of a “Moran Effect”, which pertains the phase of oscillation in a second-order cycling system (where phase is forgotten), and not the intensity and scale of eruption in a nonlinear first order system.) The scale of individual eruptions and the scale of cycle synchronization though time are different questions.

*It would be great if any analyses of the Swetnam data were published. It’s hard to interpret these findings without understanding how the data were analyzed or being presented with actual results, rather than just a discussion.*

*We note that the response variables analyzed in Peltonen et al. 2002 and Flower 2016 are different then response variables than we examine here. We’ve also clarified that we are testing for generalized Moran effects. Taken together, we agree our results are not directly comparable with either of these studies. In our discussion, we interpret the relatively low magnitude and spatial scale of synchrony as an indicative of the role of dispersal processes.*

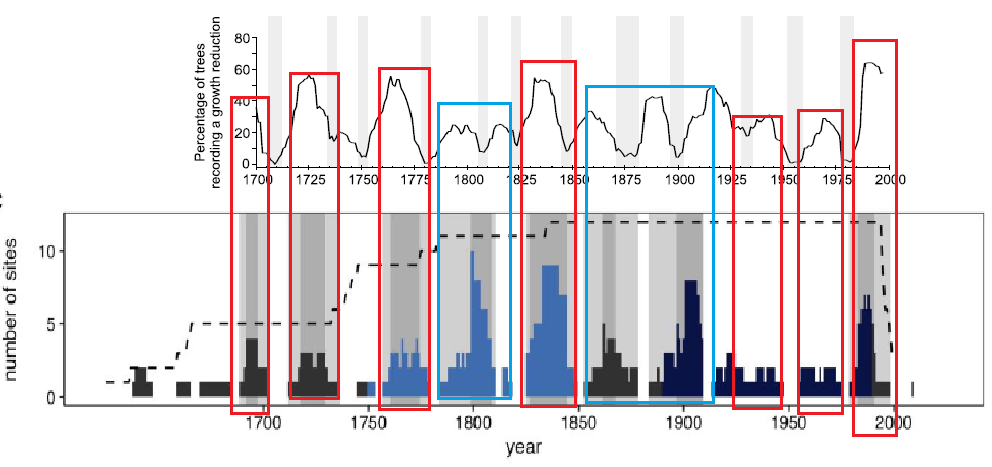
Lines 154-155: “Ring-width data for reconstructing periods of WSB outbreak were collected from Douglas fir in the 1990s, but unpublished (Table S1).” Ok, this explains a lot for me. Because I’m trying to make sense what new things we can learn from a smaller-scale study. Since Swetnam & Lynch (1993) in New Mexico and Ryerson et al in Colorado (2003) everyone has been trying to address larger and larger scale issues. So why go smaller? I personally see value in getting more data, and it doesn’t bother me that the data from 30 years old. But what I want to see 30 years after the fact is how these independent data bear on old data, old analyses, old interpretations.

Ryerson et al (2003) took place in southern Colorado, in an area sized similarly to this one:

A map of the state of the united states

AI-generated content may be incorrect.

The two series line up pretty well (red boxes), Ryerson top, current ms bottom:



This is encouraging. Not surprising. But there are a couple of windows (blue boxes) where the outbreak bounces back and forth between north (bottom) vs south (top) Colorado.

*We agree that comparing with previous research is important. We present a comparison between our reconstruction and existing reconstructions in the discussion, along with presenting the data in Table S5.*

Drought (PDSI) is not surely not bouncing back and forth at this limited spatial scale of ~300km. But the outbreaks are. The two outbreak series are going to yield slightly different results when subjected to the same analyses. What would we make of this? Of course I’m not expecting the authors to answer this question. By asking it I’m pointing out the dire need for a synthesis, and the problem created when we report a new result without comparing it directly to existing results. I’m betting the level of synchrony in the Ryerson data are comparable to this ms. But nobody has reported on it yet. Same for Swetnam in NM. Flower (2016), in contrast, examines synchrony between states without reporting on within states.

*We agree! In fact, we have addressed the dire need for people to share data and code in the discussion. We are not so certain that outbreaks are “bouncing back and forth” between northern and southern Colorado. While an interesting idea, we’re don’t see this pattern in the figure.*

Line 204: “ponderosa pine” – what assurances can you give that ponderosa pine are a pest free climatic control? Would they not be the recipient of “jolts” (Alfaro reference) of positive growth resulting from release during past MPB outbreaks?

*We removed any interdecadal patterns in each ponderosa pine tree-ring series that may emerge due to disturbance by detrending the ponderosa pine data using a 30-year 50% frequency response cubic smoothing spline. This approach is explained in the section on our dendroecological approach.*

Lines 245-247: “1.28 standard deviations below the mean (Harvey et al., 2018; Swetnam,1985). We then defined stand-level periods of outbreak as periods where: (1) at least 40% of the host trees recorded a defoliation event for 4 or more years”

There are three parameters referred to here in data pre-processing - 1.28sd, 40%, and 4y –, and one is obliged to ask how the results change if greater or lesser stringency is used in defining what one considers to be an “outbreak”. Given that the data were processed in R – a scripted language - it should be relatively easy to show what happens if lower and higher values (e.g. 0.56sd, 20%, and 2y vs 2.56 sd, 60%, and 6y) are used on the three reconstruction sensitivity parameters. This would be a valuable addition to the SI.

Given the focus of our paper was not on methods for reconstructing WSB outbreaks, we elected to follow established methods. Nonetheless, our data and code are all publicly available so we would encourage other scientists to explore this topic.

Lines 291-293: “To determine if changes in land-use practices associated with Euro-American colonization influenced the dynamics of WSB outbreaks, we split our dataset into two 100-yr periods: (1) 1750-1849 and (2) 1890-1989, or approximately the centuries before and after intensive colonization”

But this does not control for the cooler climate through the end of the Little Ice Age. Why is this not a major caveat for Colorado? Worth noting that ALL defoliator outbreak reconstructions across North America go through a period of reduced outbreak activity through this period 1750-1850.

Here is a CWT for eSBW in northeastern Quebec (data in Boulanger & Arseneault 2004).

A screen shot of a graph

Description automatically generated

Here is a CWT for 2BW in northern BC (data in Zhang & Alfaro 2002).

A diagram of a waveform

Description automatically generated

Both series show reduced intensity and synchrony of outbreaks in that interval 1700-1900. How can colonization be affecting forest-pest dynamics in areas this far North, where no forest change is suspected until the invention of the chainsaw and the feller-buncher - the only change in land use through the Holocene? And yet it’s the same pattern, coast to coast. (It might help to use the CWT function in R package DefoliatR on these Colorado WSBW data.)

We tested for the potential confounding effect of differences in climate between the two time periods and found no significant effect (see Figure S6). In our discussion, we address how concurrent changes in land-use practices and climate can limit our ability to detect the effects of changing forest management practices. We’ve also included a wavelet analysis in the supplemental figures, which confirms that the dampening in synchronicity occurs in the later part of our record and does not support a widespread dampening from 1700-1900 (see Figure S5).

Line 351:”one site experienced only 3 outbreaks, and two sites experienced 9 outbreaks”

Now we are into the results and they are as variable as I was guessing. 3 versus 9 is a big difference. It’s a three-fold difference.

We agree this variability is interesting. With our revised framing that tests for generalized Moran effects as one driver of spatial synchrony that this type of variability would be expected.

Line 353: “average quiescent period ranged from 12.7 to 61.8 years”

As above, this is extremely variable. It is a five-fold difference. And it is aligned with the high degree of variability emphasized by Blais (1968, 1983) for eastern spruce budworm – a perspective that changed immediately in the literature after the publication of Royama (1984), when suddenly the conversation was about the high degree of periodicity and synchrony and the cause of it all : the “Moran Effect”.

Thank you for the historical perspective here. We note that with our revised framing that tests for generalized Moran effects as one driver of spatial synchrony that this type of variability would be expected.

Line 356: “range:6-18 years”. Again, a three-fold variation in estimated duration. This is consistent with Blais, not Royama. It’s also not consistent with the homogeneity required by Moran’s (1953) logic – not when all three parameters are varying 3-5 fold.

*With our revised framing that tests for generalized Moran effects as one driver of spatial synchrony that this type of variability would be expected.*

Line 359: “duration was on average 2.1 times greater (mean:23.2 years; range:14-27 years)”. Exactly, and this is the LEAST variable parameter.

*With our revised framing that tests for generalized Moran effects as one driver of spatial synchrony that this type of variability would be expected.*

Table 1. The bottom row should summarize the range across sites in all outbreak characteristic parameters, including the ratio of min to max (e.g. two-fold vs five-fold). Part of your opening hypothesis is that Moran’s logic is relevant. This extreme range suggests maybe it is not. In which case, what do we replace it with? Campbell gives us a clue.

*We have added the landscape-level means and standard deviations where appropriate to the bottom of Table 1.*

Lines 377-379: “moderate agreement among time series of the percent of trees defoliated at a site (W = 0.26; p <0.001; mean inter-site rs = 0.19; mean inter-site C= 34%)

What values of W, r, C, would the community consider “low” and “high” if these are “moderate”? I think these are moderate to moderately low, especially compared to the high number of papers reporting relatively high levels of synchrony. I also think they’re representative. They’re not anomalously low because of, say, biased sampling. I think the community should see, say, a whisker plot of the distributions on these parameters. This is going to emphasize the rather variable levels of synchrony – a fact the literature has been down-playing for 30 years. And all three indices point in the same direction, also something worth highlighting.

*We revised our text to read “low to moderate” rather than moderate. Additionally, we note that we present the range of inter-site concurrency and Spearman’s rank correlation in the following sentence. While not quite as informative or easy to read as a box and whisker plot, given the number of figures already in present in the manuscript and the minimal gain in information we have retained this information in sentence format.*

Line 379: “relationships between some pairs of sites were neutral or even negative”

Exactly! And how does this square with Moran’s requirement for spatial homogeneity of regulatory ecology?

Again, with our revised framing that tests for generalized Moran effects we believe this variability would be expected.

Line 382: “significant synchrony among time series of the percent of trees defoliated existed at distances up to 50 km”

In some cases yes. But in many cases, no. The confidence interval in the spline of Fig. 3 is so wide that it always includes zero at all distances. That’s not a bug. It’s a feature. It’s telling us something that the literature continually downplays.

We disagree with this interpretation. In figure 3, the confidence interval only overlaps with zero for distances greater than 50 km. Thus, significant synchrony exists when sites are within 50 km one another, but not beyond. This analysis does not test for individual cases where outbreaks may or may not be occurring at the same time, but tests across all sites over a combined time period if outbreaks occurred synchronously.

Figure 4. It’s not clear to me what the index of “temporal clustering” is actually measuring. The black line with its envelope and the blue line – it’s not clear what each implies and what the deviation between them means. What should we have been expecting under a null hypothesis?

We have revised this figure to make it clearer what is displayed on the y-axis and how to interpret the figure. Specifically, we changed the y-axis to read “L(t) function” and added the following to the caption “The dashed gray lines indicate the 95% confidence envelopes. Gray shading indicates significant synchrony, where L(t) values are greater than the upper bound on the 95% confidence interval”.

Figure 5. There are no differences before and after, as stated in line 405. Do I take this is a concrete proof that the colonization->land use->forest composition-> pest response causal model is wrong? If so, then which part of it is wrong? In my view the negative result is suggestive at best. But what it suggests is not at all clear.

We disagree with this interpretation. In figure 5, significant differences between groups are represented by different colored boxes/bars (as stated in the figure caption). These results are also presented in Table S4. Our statistical analyses have led us to conclude that relative to the 1890-1989 period, outbreaks were longer and more severe during the 1750-1849 period.

Lines 417-418: “We found synchrony in outbreak dynamics was lower in the 1890-1989 period than the 1750-1849 period”.

First, this pattern is counter to what I showed above for eSBW in Quebec and 2BW in BC. Second, what do we make of the possibility that climate is warming and cycle amplitude is attenuating? This is exactly what is happening to the geometrids in Europe, where climate warming is driving outbreaks upslope, right of the tops of the mountains – a process that was complete by 1989 but started decades earlier. Why would this not happen in Colorado?

This may occur because our paper focuses on a different insect and/or a different host tree species and/or different land use histories.

Figure 6: the darkest shade is nearly indistinguishable from the intermediate shade. I can distinguish, but it’s not easy.

We have changed the colors here to make the easier to read.

Lines 446-455: This is consistent with the idea that drought triggers eruption on a fast time scale, but moisture shapes outbreak intensity or duration on a slower time scale. Coupled with your argument that land use change shapes intensity and/or synchrony on a centennial time scale you have positive support for all three hypotheses operating at three distinct time scales: resource stress hypothesis (Larsson), environmental adequacy hypothesis (White) resource pulse hypothesis (Holling). Why not organize the hypotheses this way, and present them in a more structured, orderly manner? It seems to me you’ve gone a long ways to fleshing out Holling/Raffa “cross scale hypothesis” as a sort of overarching hypothesis for these limiting cases, which you find to be non mutually exclusive.

*We agree there are a lot of drivers here and the hypotheses are not mutually exclusive and because of this there are several ways that these ideas could be presented. To compare with prior work, we tried to structure our text based on some of the more recent work on WSB outbreaks (e.g., Ellis et al. 2017).*

Figure 8. It’s my preference to have the y axes labelled on the left, not the right, but I can live with these “as is”.

We have added a y-axis title “value” to this figure, and left the facet labels on the right, as is convention with faceted plots (Wickham 2016)

Line 465: “WSB outbreaks occur synchronously across sites, suggestive of the combined effects of density-dependent processes and the Moran effect”

I think you showed that the Moran effect is quite possibly not 100% relevant for a system that is this spatially variable in the parameters describing its intrinsic dynamics. Moreover, if the environmental factors governing initiation and duration are different, then the system may not be linear in its dynamics, and this would further undermine application of Moran’s theorem to the case. WSBW might be behaving more like, say, a mountain pine beetle system, where drought triggers eruption possibility, but overall warmth drives seasonality of developmental phenology and the probability of successful synchronized mass attack. This is precisely how Campbell (1993) viewed the system. He was sharply critical of Royama’s style of rebutting Watt & Morris, and rejected the blind application of Royama’s theory to the WSBW system. The nonlinearity in the WSBW system doesn’t come from co-operative defeat of resin-based tree defenses, as in MPB, but in satiating the response of generalist predators, such as spiders and birds. Also, the co-operative effect of rising density on mating success, which has been shown in eSBW (Regniere et al 2013) but possibly (I would say probably) also exists in wSBW – because it also exists in FTC (Evenden et al 2015).

Régnière, J., Delisle, J., Pureswaran, D.S. and Trudel, R., 2013. Mate‐finding allee effect in spruce budworm population dynamics.*Entomologia Experimentalis et Applicata*,*146*(1), pp.112-122.

Evenden, M.L., Mori, B.A., Sjostrom, K.D. and Roland, J., 2015. Forest tent caterpillar, Malacosoma disstria (Lepidoptera: Lasiocampidae), mate-finding behavior is greatest at intermediate population densities: implications for interpretation of moth capture in pheromone-baited traps. *Frontiers in Ecology and Evolution*, *3*, p.78.

We have clarified here and throughout that we are referencing a generalized Moran effect, which we believe is applicable.

Lines 500-502: “This is consistent with previous research on spatiotemporal patterns of contemporary WSB outbreak in British Columbia, which has suggested that dispersal is key in driving the synchronization of WBS population dynamics”

If dispersal is key in driving the extent of outbreaks this is not the same thing as arguing that dispersal is driving the synchronization of cycles. Don’t argue things that don’t need to be argued, and don’t copy the logical errors of others who have treated the subject far too casually. Your results are not entirely consistent with Flower (2016), and there has to be a good reason for this. Scale is probably the answer, but why? Your series is also not identical to Ryerson’s just 100km to the South – even though it’s close – and there has to be a reason why the series differ, even if slightly. Similarly, I showed above that synchrony between BC, CO, NM declines in the window 1700-1900. Why? Why the different result? I would not be so keen to highlight superficial similarities when there are such meaningful differences to be explained/explored.

*We have changed our text here to indicate that we are talking about spatiotemporal patterns of outbreaks. We also caution against making direct comparison with other studies because of differences in the data and methods.*

Lines 502-506 “Our records also show that outbreaks can develop concurrently in disjunct populations. For instance, evidence of the 1820s outbreak first appears in the tree-ring record in 1825 at the TI and JP sites, the furthest apart of our sites. While budworms are strong fliers that can disperse hundreds of kilometers in the right weather conditions (Sturtevant et al., 2013), most dispersal occurs at much shorter distances (i.e, <15 km) (Senf et al., 2017). Thus both dispersal …”

So dispersal fails to synchronize the 1820s outbreak across the study area, because it was not long-range enough. But then how did it do it during other outbreaks, earlier and later?

Our point here is that insect populations can develop at multiple disjunct locations due to common variance in environmental drivers.

Lines 506-507: “… both dispersal and the Moran effect are likely important in driving the synchrony of local populations of WSBs”

But if Moran effect drives “local” synchrony (what is “local”? <50km?), then what drives larger-scale synchrony (e.g. during most outbreaks, but not the 1820s outbreak)? If the hypothesis is that plant stress and food quality drive eruption probability and outbreak duration respectively, this is not at all a “Moran Effect”. This is about environmental triggering of a nonlinear process, much as for bark beetles. It has nothing to do with phaseless oscillations that are generated by a nebulously linear and lagged predator-prey interaction, where natural enemies are somehow miraculously expected to keep pace with WSBW though similar long-range dispersal. How are tiny wasps and flies going to to disperse at such rates? It makes sense maybe for the lynx and hare that Moran studied. But for WSBW and their parasites?

I think your analysis points in a different direction than where you took it.

Again, we apologize for the confusion here, but we were from the onset examining generalized Moran effects.

Lines 519-521: Greater regional host availability can increase probability and severity of local WSB outbreak, independent of stand susceptibility, by influencing the potential for dispersing WSBs to arrive at the focal stand (Howe et al., 2024”

This is the “resource pulse hypothesis” that I believe Howe cites by name. So why not state it as such in the Introduction? i.e. Start with an orderly processing of multiple hypotheses operating at multiple time scales?

There is no reference to “resource pulse hypothesis” in Howe et al. 2024.

Lines 523-524: “Many, if not most, of the Douglas-fir dominated stands in our study area originated after widespread burning in the second half of the 19th century (Sherriff and Veblen, 2007).”

That was my understanding. The introduction needs to start with this point, explaining exactly the causal chain of events whereby:

colonization->land use->forest composition-> pest response

We describe this in paragraph four of the introduction.

At this point I’m not going to have too much to say about the “Implications” section. I would rework the implications so that they are better aligned with what I think is the more accurate interpretation of this analysis, as described above.

We have re-read through the implications section and revised the text to be clear that we are referring to generalized Moran effects.

After all this, it would be understandable if the authors chose to reject me as a choice in any re-review. But I would pleased to see any revision that addressed these points in a material way.

Reviewer #2: Summary

This article examines interesting questions around the potential linkages between long-term WSB outbreak periodicity and regional synchrony and external variable such as drought and forest management/disturbance history. I quite liked this article and think the authors did a good job weaving together several different issues into a coherent study. There are a variety of data sets tied together here and appropriate dendrochronological methods/controls used. The data presentation is clean and thorough, with the addition of several supplementary figures and tables. The writing was generally good though I have some comments below recommending a bit of refinement to tighten the story. I will also note that I found few typos or odd grammatical errors, which indicates some careful editing prior to submission (which I appreciate as a reviewer). While the results were not necessarily groundbreaking or unexpected, I think this is a quality article that will, with a bit of minor revision, make a nice contribution to WSB outbreak science and FEM.

Major comments:

None.

Minor comments:

Introduction:

In general, the introduction hits on the major points to set up the questions. I did, however, feel that it was a bit unfocussed, moving from several key hypotheses describing insect-plant interactions to WSB outbreak dynamics without really giving me a sense of how it was building to the three main questions of the article. The hypotheses are useful for understanding insect-plant interactions but maybe don't need to be covered so explicitly to set up the questions addressed in the article, i.e., are outbreaks regionally synchronous, does drought explain the timing of when and where outbreaks begin and end, do these dynamics differ in pre vs. post colonization. This point seem is reinforced after reading the discussion where many of these elements are never revisited (i.e., they may be interesting but they were nonessential to the final thrust of the article). I suggest a bit of tightening up the focus of the introduction to trim out the extra non-essential details to the questions, perhaps using the major points of the discussion as a basis to clear out non-essential parts.

Thanks for this suggestion. We have revised our introduction to be more concise. For example, we combined paragraphs on stand characteristics that promote outbreaks and also the effects of historical land use practices on stand characteristics.

Ln. 76-77. Reads a bit awkwardly, maybe "Early Euro-American settlers carried out extensive logging operations near settlements, which occasional ignited large-scale forest fires."

We have revised this statement to read “Early Euro-American settlers carried out extensive logging operations near settlements.” We note that some fires were lit intentionally and not necessarily because of logging.

Ln. 80. Maybe "Selective harvest of large trees coupled with severe wildfires can reduce host…"

Changed as recommended.

Ln. 115. "…response is lagged and density-dependent."

Changed as recommended.

Methods:

Ln 207 - Nice use of a non-host to tease out the impacts of WSBW.

Thank you.

Ln. 333-334 - Is this a standard threshold for drought vs. non-drought conditions? It might be good to provide a couple citations where this was either validated or at least used enough to qualify as precedent.

Values less than –1 correspond with mild or more severe drought, whereas values greater than 1 correspond with mild or more extreme wet spells. This follows Palmer’s (1965) original presentation of PDSI. We have added this reference to this sentence so that it now reads: “First, we classified time series of SC-PDSI as periods of either drought or above average moisture availability, which we defined as consecutive years with SC-PDSI values below -1.0 (drought) or above 1.0 (wet periods) following Palmer (1965).”

Results:

The results look good though I think they could be trimmed down to just describing trends. There are a few sidelines into what would better fit in results (see below and other instances). Trimming these sorts of statements and just describing the significance (and nuances) in the discussion would make for a tighter results section.

Thanks for this comment. We understand it to mean that some of the details of the results could be presented in the discussion. We appreciate concise writing and have read back through the text to look for places to simplify the presentation of the results. For instance, we have removed much of the text describing the characteristics of site and landscape-level outbreaks because it largely repeated information presented in Table 1 and Figure X.

Ln. 405-406. This is interesting but would probably be a more impactful in the statement where you can more effectively draw attention to the significance of the results.

Thanks for this suggestion. We have moved this text to the paragraph on Euro-American settlement in the discussion.

Ln. 446. Also, here the results are framed in a way that would fit better in the discussion. This first sentence (and others like it) could be cut to make the results more to the point.

Thanks for this suggestion. We have retained this text because we felt that including a topic sentence made the paragraph easier to follow.

Discussion

A good summary of the study and important cautions on study limitations. I don't think the subtitles help, rather they seem to more break up the discussion than add to it. Later parts of the discussion seem to abandon the convention to make more general points so I think it best to remove them altogether.

We have removed the subtitles.

462-464. I think these first two sentences could be cut without losing anything from the paragraph. The methods are already described so it seems redundant to explicitly repeat them here.

We considered this recommendation, but elected to follow advice from Schimel (2012), who recommends that discussions form self-contained stories.

References

Schimel, J. 2012. Writing science: how to write papers that get cited and proposals that get funded. Oxford University Press, Oxford.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. Available online: https://ggplot2.tidyverse.org