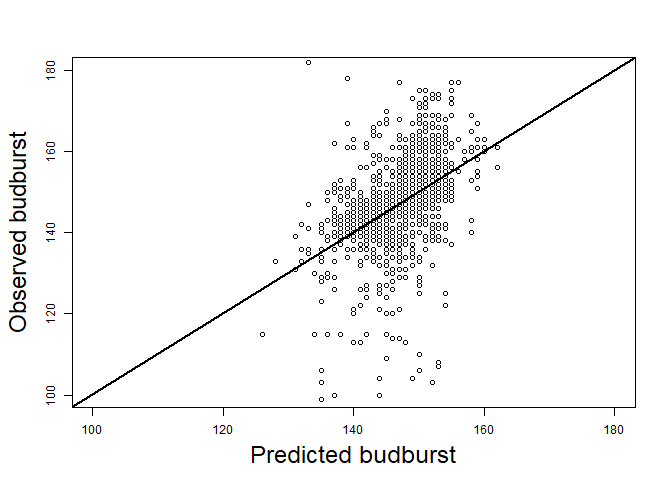
ASSOCIATE EDITOR COMMENTS TO THE AUTHORS   
Associate Editor   
Associate Editor Comments for Authors:   
Two reviewers have now provided comments on this revision. Based on their comments and my own reading of the paper, I see the following issues that remain with the manuscript. First, the specific methods are not well explained. Although there are certainly more details than in the previous draft, there are still several instances where information is lacking. Both reviewers note missing details on methods, validation, and model assumptions. Second, one reviewer notes that there is a well established body of literature on predicting phenological shifts that is not mentioned or discussed in the paper. The reviewer provides a number of citations to support this claim. The current work should be placed in the context of previous work, specifically addressing any improvements and potential disadvantages related to assumptions or otherwise.   
  
REFEREES' COMMENTS TO AUTHORS   
  
(NB. If there is no comment from a Reviewer listed below, this probably means that they have uploaded a separate 'file for author' to the Central Site. You can see these comments in your Author's Centre by clicking ‘manuscripts with decisions’ and then using the 'files attached' link at the bottom of the decision letter).   
  
Reviewer: 1   
  
CONFIDENTIAL COMMENTS TO AUTHORS   
The presented manuscript covers a key interest of modern times, namely how natural populations will cope with future global warming. Here, the authors focus on the mismatch between a consumer and resource species (spruce budworm and balsam fir model system). The manuscript aims to project future phenology of both species and the resulting mismatch using a modeling approach based on rate accumulation functions. The approach is interesting and promising, but the manuscript lacks detailed information on the data used, analyses performed and results presented. Furthermore, I am worried about the insufficient validation of the proposed modeling approach with observed data on phenology.   
  
  
Major comments   
1. Missing validation   
The validation of the SWB and Uniforc model to project phenology in response to temperature is insufficient. The observed data on phenology used to validate the models only spans two years across two sites (no details given on data). Considering that there is year-to-year variation, I don’t see how the models can be validated using two years of observed data?

Ideally, observed data of species-specific phenologies and observed temperature data across the entire historical period (1996-2016) should be used for validating the models. If no long-term data on phenology of the chosen study system is available, maybe the choice of study system is not ideal?

Do we need to validate the SBW model since it has been used for year within BioSim (and all related papers based on it)?

For the tree model, we can do the regression of Observed versus Predicted data. If the slope is non-significantly different from 1 and the intercept from 0, then the model is unbiased.



Here, neither the slope (p = 0.42), nor the intercept (p = 0.38) are significantly different from 1 and 0 respectively.

The only issue is that the dataset used to fit the model is the same as the one used to test it (unless we do a cross-validation).

The authors also make comparisons across latitudes/sites (but see ‘3. Comparisons across latitudes/sites’ below) and assume that the models are valid across all sampled sites. It is, however, well possible that populations are adapted to local conditions and hence that the same parameterizations of the model or the model per se does not hold true across sites.

Maybe I am wrong, but I would assume that the trees do not need to adapt too much: if they rely on temperature as a signal for the “good season”, then this is how they adapt.

For the insect, it may adapt at some point, but since the SBW model is widely used to predict SBW dynamics across several provinces, people certainly have reason to do so, haven’t they?  
  
I also wonder how well the scenarios are able to predict temperature especially in regards to cold and warm spells and the frequency of their occurrence? In general, information on the frequency of such spells at the tested locations during the historical period would be interesting and helpful for understanding the relevance of such spells for the study system. This could easily be reported for observed temperature data.   
Yes, it is possible. Should we do it?

2. Details on analysis missing   
The authors did a great job in outlining the general model (Section 2) and how the models can be adjusted to the two species studied here. It is however, very unclear what analyses were performed and underly the presented results.   
  
For example, it is not clear how the results for Figure 6 are derived; are the data plotted derived over the entire period or show a certain year (i.e. last year within time period)? If plotted across the entire period, it might it be more useful to present the data as trends over time. It is also not clear whether the temperature data constitute a constant increase in temperature only (or also includes spells and what the frequency of such spells is). Hence, it would be good to give details on the temperature data and the models used for the results.   
  
Overall, a clear description of the data used (whether observed or projected) as well as all the models used in the analyses would be very useful.   
To discuss.

3. Comparisons across latitudes/sites   
If I understand correctly (based on R output provided in Supplementary Information; 2 Analysis of variance), a difference between sites was tested. This not necessarily corresponds to a difference in latitude. If the authors want to show an effect of latitude on emergence date, budburst date and/or mismatch, the authors have to explicitly test that.   
I can clarify that, but it is clearly stated that the sites are the points shown on the map (figure 4), and that the distance between each point is 0.5° latitude (and that we tried to correct for differences in altitude as much as possible). So, I can replace the word “site” by “latitude” in the appendix, if it can avoid any misleading interpretation.

Also here, details on the models fitted for the respective results would improve the clarity of the manuscript.   
  
  
Other comments:   
- Are the temperature data for historical period based on recorded data? Based on lines 304-309 it seems like they were projected, too. Assuming that temperature data differ between sites (for both historical and future temperatures), it would be good to present the data and how it differs between sites/latitudes.

That is feasible (although it might be less obvious for future data).  
- Details on data used for parameterization for the Uniforc model are lacking (e.g. “in the 1980s and 1990s” [line 292; please give specific range] or “different sites during growing season” [line 294]).

OK, we can give more details.  
- Lines 108-110 imply that there is a known effect of temperatures on the two species studied here. Are there any data/studies available that can be used to back this up? Especially as the study focuses on accumulated temperatures it would be good to clearly point out that accumulated temperature effects resting phase of these species.

I was wondering if that was a joke. We cite several paper in the manuscript about the SBW (OK, we may find one for balsam fir).  
- “The emergence of SBW generally precedes balsam fir budburst by several days” [Lines 260-261]. But in Figure 2 the species are displayed the other way around which is a bit confusing. Does “generally” mean that this direction of trend was observed across latitudes for the historical period?

I wonder if the reviewer made effort to understand the paper, Figure 2 is about sensitivity, not about emergence time.

- Supplementary Information; 2 Analysis of variance: I don’t think that the R output constitutes a good way to present model output. The R code used would be very useful though!   
- “Consecutive spells will have additive effects” [Line 389 and following]. This was not tested, was it?   
- SWB has several host species [Lines 245-248]. How important is the balsam fir relative to the other host species? Does this differ between sites/latitudes?   
- Figure S1: not clear which temperature values are used for predicted values (dotted line) to validate linear approximation.   
  
  
Reviewer: 2   
  
CONFIDENTIAL COMMENTS TO AUTHORS   
The authors present a mathematical model for quantifying the phenological mismatch between ectothermic resource and consumer species.  They develop an explicit mathematical formulae for calculating the length of the resting period for each species, using a rate accumulation function for either a constant temperature difference or a warm/cold spell of short duration.  The differences between the resting periods of the resource and consumer species constitutes the phenological mismatch.  The authors validate their model using data for the balsam fir – spruce budworm system. Their main finding is that a cold/warm spell during the resting period could increase or decrease phenological phenological synchrony.  The authors claim that their method is important for modelling the effects of climate change on consumer-resource systems.   
  
The two main issues are the the biological realism of the mathematical approach and its utility in predicting mismatches in real systems.   The authors claim that science is currently unable to predict phenological mismatches.  This is not correct.  The mathematical approach of delay differential (DDE) equations with temperature- and time-dependent developmental delays provide a mechanistic and biologically realistic approach for predicting both phenological shifts and mismatches (Gurney 1983, Nisbet 1983, Nisbet 1997, Amarasekare and Coutinho 2014, Scranton and Amarasekare 2017, Amarasekare 2019).  The authors use a phenomenological rate accumulation function.  The DDE approach provides a mechanistic rate accumulation function based on temperature effects on maturation and juvenile mortality.  This function has the advantage that it can incorporate mechanistic descriptions of maturation and mortality rates, based on first principles of thermodynamics, that can be parameterized using empirical data on growth/maturation rates.  Moreover, temperature effects on ectotherm developmental delays, which arise from the multiplicative effects of the temperature responses of are what drives phenology, the seasonal timing of life history events.  Importantly, any computation of the resting period should explicitly incorporate temperature effects on maturation and mortality. The point is that there are alternative approaches that are both biologically realistic and can be easily parameterized using data that are readily available (there is a plethora of empirical studies that have quantified the temperature responses of maturation and mortality for many multicellular taxa).     
  
The authors calculate the end of the resting period for two scenarios: (i) temperature change is constant throughout the period, independent of time, and (ii) there is a warm or cold spell of relatively short duration at a particular time during the resting phase.  While these cases allow for the analytical approximations the authors use to quantify phenological shifts and mismatches, they do not capture the realities of climate warming, which is a nonstationary process.  Climate warming involves a change in the mean temperature as well as the amplitude of temperature fluctuations (diurnal and seasonal).  When cold/hot extremes occur, it is against the backdrop of this nonstationary process.  This raises the question of how accurate the mismatch predictions given the authors also use a linear approximation in the interests of analytical tractability.  The DDE approach with variable time delays has the advantage that it can accommodate any nonstationary climate change scenario.   
  
The authors’ approach provides for a way to quantify phenological mismatches, but not the dynamical consequences of such mismatches for consumer-resource interactions.  Given the complex non-linearities inherent in consumer-resource dynamics, combined with the non-linearities in the temperature response functions of life history traits and the climate change scenarios, it is not possible to make any predictions about whether and how a phenological mismatch quantified solely in terms of the end of the winter resting period about the effects of such a mismatch on consumer-resource dynamics.  The variable delay models have a significant advantage over the authors’ approach in that they can both predict the mismatches AND their dynamical consequences on consumer-resource interactions.   
  
None of the above is made clear in the manuscript, which gives the misleading impression that there are no alternative approaches to quantifying mismatches and that the author’s approach provides the means for predicting the effects of warming-induced phenological mismatches on consumer-resource interactions.  What the authors’ model does it to provide tractable means for quantifying the length of the resting period, albeit under strong assumptions about both the biology and the nature of warming.  The DDE approach, while more biologically realistic and able to predict both the mismatch and its consequences for consumer-resource dynamics, requires the kind of mathematical and computational expertise that is out of reach for most biologists.  The authors’ approach has the advantage of tractability, especially in practical situations such as biological pest control.  Having some knowledge of the length of the mismatch can be useful in implementing integrated pest management practices to counter or ameliorate the effects of climate warming on pest control, which is a serious problem with potentially dire consequences for global food security.  I think the paper’s value should be highlighted using these practical aspects.     
  
It is important to cast the Introduction and Discussion in the context of the DDE models with time delays which has a long history in addressing these same questions, and pointing out why a more tractable approach is needed.  Acknowledging the previous body of work in no way diminishes the authors’ work.

I have the feeling that this reviewer is involve in these DDE approaches and want to do an “advertising campaign” here. He/she did not discuss anything but DDE approach. So, I am OK with adding a paragraph about that, but as an editor, I would be unhappy with such a review.

I acknowledge that the idea of putting our results in the perspective of pest control might be interesting.

End of Comments