Interactive comment on "Dispersal distances and migration rates at the arctic treeline in Siberia – a genetic and simulation based study" by Stefan Kruse et al.

Anonymous Referee #4
Received and published: 22 November 2018

We thank the reviewer for reviewing our manuscript and especially for a closer look on the supplementary material. The comments helped to improve the first version of our manuscript. This was revised at the corresponding positions for each specific comment below.
Our response are placed in bold font below each of the reviewer's comments in italics. Followed by a citation of changed text with a line statement that refers to the version of the manuscript with tracked changes.

General comments:

This study utilizes genotyping and parentage analysis of individual trees to improve larch seed dispersal simulation within an individual-based, spatially-explicit forest model. The study is carried out at a single 100 m x 100 m site in the Taymyr Peninsula in northern Siberia. LAVESI, the forest model used, is specifically designed for individual larch growth, mortality, and regeneration, and the updated model is used to simulate northward migration of the larch treeline and forestline under two different climate scenarios. The updated model performed well when compared to observation data, though it slightly overestimated the number of recruits close to the parent tree as well as an overestimation of very long dispersal. The south-north migration simulation under static climate resulted in a migration rate of 0.6 m/year and 1.6 m/year for the forest- and treelines, respectively. Under a climate scenario of decreasing temperature and slightly increasing temperature from south to north, the south-north migration rate was slower. They also found an accelerating rate of dispersal over the simulation time under the static climate scenario.
The study is important for field ecologists as well as the ecological modeling community. Currently, northward tree migration across the circumpolar boreal region is of crucial importance due to its potential impact and feedback to climate. However, most forest models do not adequately represent dispersal mechanisms. This study showcases an innovative way to determine in situ effective seed dispersal and incorporate such data into a forest model for calibration and application.

While the study is effective and well-structured, and shows how well the LAVESI model can perform at a local-scale, the model was tuned quite heavily to the small study area (only 100 m2), and the model output was compared only to data that was used in the tuning process. Before this model can be utilized at a larger scale I believe it will require more generalized parameter values. In particular, because the model produced fairly slow migration rates compared to other studies, I feel it may be overfitted to this study site and data, though only additional comparisons and simulations with the model will be able to determine if this is the case. It would be nice to see a sentence or two acknowledging this in the Conclusions. It would be nice in future studies to see this model compared to independent data at a separate site as

well. I would also be interested to see how the migration would play out under a climate change scenario, though this is likely planned for future work.

Overall, I think this paper is well-written and the manuscript should be accepted with only a few minor revisions. This study is a great starting point for future work with this model and the equations developed within it. It should be of interest to other ecologists working on similar problems across the boreal region.

Response to the the centre part of the general comment in starting with "While […]". A similar comment came from R3. We added a short discussion about the "small" study area that is already challenging for such an analysis to the discussion in section 4.1. Nevertheless, it would be worth to undergo this work at more sites to compare the findings of this study to other treeline locations.

Line 284ff:
"Unfortunately, the labour-intensive sample collection and genetic analyses restricted the analysis to a rather small area in comparison to the large area of the treeline transition zone. Assessing the parentage across a broader scale and for different positions in the treeline ecotone would further help to understand dispersal dynamics at the treeline but the additional knowledge gain does not scale with effort."

Additionally, we extended our conclusion covering the comment on further studies that would help unravelling if our slow migration rate estimate is flawed by overfitting to only one study site or not, as requested by the reviewer.

Line 400ff:
"To find out if the estimated slow migration is an outlier coming from overfitting to only one study site or the general response rate under current warming, further similar studies at other treeline positions would be necessary."

Below are some minor comments and edit suggestions for consideration by the authors:

Line 92: Change "Subsequent" to "Subsequently,"

Response: Done

Line 120: Change "larch species" to "larch individuals"

Response: Done

Line 129: You say here and in the Supplementary Material that active layer depth influences tree mortality (which I am guessing is based on growth rate). However, it seems based on the information in the Supplementary Material that active layer depth directly influences tree growth, which in turn would also influence mortality (and potentially seed dispersal?).

We use the actual tree growth in comparison to the maximum potential growth of the same tree as currency for productivity and mortality.

The given information was not sufficient to explain how active layer depth influences trees (growth/mortality). In consequence, we edited the text in the Methods section for clarification.

Line 137ff:
"The original model of Kruse *et al*. (2016) was updated with the following processes (details in Suppelement 2): (i) seed dispersal distances now depend on species-specific traits (tree height, seed properties) and wind speed and direction (Kruse et al., 2018b), (ii) the tree diameter growth function is newly calibrated to the climate forcing (Epp et al., 2018), and (iii) the active-layer thaw depth directly influences the tree's growth that is used to estimate it's seed production and mortality."

Lines 135-140: I'm not sure why some of these parameter descriptions are in quotes and some aren't. In general this sentence is difficult to get through. You may want to consider just publishing a table instead of listing them in the text.

For clarification we decided to remove the listing of only some of the varied parameters and refer readers to the complete information in the supplementary material. A complete list and further detailed information on each parameter combination and the process can be found there.

Line 139: I'm not sure what "different modes to compute the competition" are

We tested the impact of several implementations of influence areas and strengths of competition on the trees diameter growth. The actual growth of an individual is the currency in the model by which other functionalities are based on (seed production/mortality).

We refer now the reader to the supplementary material, as there is the information on modified parameters/modes and tested model variables.

Line 151-152: Could you expand on the 20mx20m vs. the surrounding 100mx100m section? I'm not sure I follow where the spatial differences are coming from.

We needed to make the simulated data comparable to the inferred effective seed dispersal distances. Therefore, we followed directly our sampling scheme as described in Section 2.1 sample collection "[…] We sampled all individuals >0.4 m in height in a 20 x 20 m area as well as all trees >2 m high or bearing cones from the surrounding 100 x 100 m area (Fig. 3). Additionally, in the central 12 x 12 m area individuals <0.4 m were collected."

Here we added a reference to the sampling scheme description in section 2.1 in the regarding sentence.

Line 166ff:
"We resampled these simulated distances to consider the same frequency of observed parenthoods in the central 20 x 20 m as in the surrounding 100 x 100 m area (sampling scheme details in section 2.1 sample collection)."

Line 224: Add "for this model" after "Mean dispersal" Line 229: change "have the smallest" to "has the smallest"

Response: Done

Supplement S2:

Line 74: Change "correspondingly" to "corresponding" and delete "roughly"

Response: Done

Line 76: Change "of Matlack" to "from Matlack"

Response: Done

Lines 76-79: I'm confused by what 0.86 m/s is referring to. I this Vd? Or w? Additionally this sentence is some- what awkward and I would recommend breaking it up into two sentences and clarifying.

The value 0.86 m/s is referring to the descent rate for seeds, which is abbreviated by Vd. We separated the sentences as suggested and edited it for clarification.

The corrected part of the text is now in line 74ff:
"The release height H_t is estimated at 75% of the individual's height. V_d is the descent rate for seeds and is estimated for *Larix gmelinii* by a linear regression using species data from Matlack (1987). For species having wing-scales attached to the seeds, this rate can be calculated by $V_d = 0.0032 * \sqrt{w} + 0.4807$ and is 0.86 m s^{-1}, with the wing loading w (Matlack 1987) for *L. gmelinii*. The variable w is calculated by dividing the average seed weight (in microdyne) of 3.5 mg (Heit and Eliason, 1940; Lukkarinen et al., 2009) by the propagule area of 0.2 cm^2 (Fu et al., 1999)."

Line 84: how did you obtain the sdist and the scaling parameter? I see that you tuned them variously but did you have initial starting values based on literature or data?

When implementing the seed dispersal kernel into the model (Kruse et al., 2016), we made a first guess for the resulting dispersal kernel based on literature values and tuned those values to observed patterns.

Line 88: Where did you obtain the data for the study showing no significant influence of temperature? Was it at the same study site? I am concerned about this growth function modification as it further "tunes" the model to a specific area, and may need to be re-tuned if the model is moved elsewhere

We used a tree ring series from Yamal of the National Climatic Data Center data bank for *Larix sibirica* and own data for *Larix gmelinii* from Khatanga near the study site and for both data from the nearest weather station. For further information, please see the supplement of Epp et al. (2018) published in Scientific Reports.

Regarding the second part of the comment. The modelled tree diameter growth in the current version of the model is adapted to weather in Taymyr and Yamal. Therefore, it has to be tuned for each species and region when using it for further applications.

Lines 93-98: See my above comment on permafrost-tree growth influence. It seems ALT impacts tree growth directly and mortality indirectly, though I may be wrong.

Yes, answered in the other comment above.

Line 97: What is the parameter fe?

It is a soil property parameter, see definition in Hinkel and Nicholas (1995).

Table S4: I would suggest also adding variable symbols next to the parameter descriptions, especially if they are mentioned in this text or other published works.

We added for the model parameters the corresponding symbols. Corresponding changes were made in Table S5.

Line 105: Why do you need to shift the dispersal peak by 2-3 m? Is this based on comparisons with the observation data? I would mention this here.

We tried to explore potential setting to align the modelled effective dispersal distances to the observations. For clarification, we edited the sentence and refer to the results presented in the main article.

The text now line 105 reads:
"To fit the simulated seed effective dispersal distance to observations (Fig. 5) we explored potential settings …"

Line 120: What is the reference simulation? Additionally please expand on what you mean by "general performance."

We extended the statement of the reference simulation, which is the baseline simulation with the original model. Furthermore, we added for clarification of the "general performance" a reference to the correlation coefficients in Table S5.

This sentences in line 122ff changed to:
"This was improved by other simulations (qt-wJ) but their general performance (lower correlation coefficients, Table S5) was weaker than the reference simulation without parameter changes or adaptations of the model (a)."

Line 123: I'm not sure what you mean by "In parts"

We deleted the confusing beginning of the first sentence of the regarding paragraph. In the following sentence we briefly state the achievements, but also at which results the best fitting model version deviated from the observed pattern.

Text now in line 126:
"We achieved a good fit when increasing the peak of the dispersal function in the model to longer distances."

Line 127: What is the ecological basis for changing the density competition to improve the on-site recruitment ratio?

Similar to Janzen and Connell's findings, recruits have the highest chance to survive at intermediate distance to the producing tree, not directly at it. They are "pushed back" by the mother tree for a variety of reasons (shadow of the tree's crown, high pest pressure/seed predators,

exhausted nutrients in the active layer, insulating accumulation of needles and other litter, etc.).

This is implicitly implemented in the model and can be manipulated by varying the competition density, e.g. by increasing the influence on smaller trees. With this, seedlings from farther distances could have a likely higher chance to establish.

Line 128: Delete “were” in between results and strongly

Done