Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-31-RC2, 2018


Geoscientific Model Development Discussions

# *Interactive comment on* "Implementing spatially explicit seed and pollen dispersal in the individual-based larch simulation model: LAVESI-WIND 1.0" *by* Stefan Kruse et al.

**Anonymous Referee #2**

Received and published: 19 April 2018

The paper presents the wind driven seed and pollen dispersal algorithms newly implemented in the existing model LAVESI which simulates Larix gmelini vegetation dynamics in north-central Siberia. Language wise the paper is very well written. Content-wise, however, I feel that the manuscript is currently a bit thin. I do agree that seed dispersal and particularly pollination are neglected processes worth to be further investigated. New contributions are the insertion of wind speed and distance in the already existing seed dispersal kernel and the new pollination probability function. Beside this, however, it is only shown that the implementation of the functions work as intended and that two unspecified parallelisations lead to a computation time reduction. Else, the authors mainly write what potentially could be done with the model, but unfortunately



miss to give sufficient instructions for how this could be done, nor do they show example simulations for the potential applications.

I do not understand why the authors decided to put so much emphasis on describing the verification with all its sophisticated applied statistics. In my opinion the verification does not contribute to a knowledge gain and could be dealt with a few sentences, referring the main share to the supplementary. However, I do believe that a transparent description of the parameterisation and applied sensitivity analyses could be of large value for the community.

I recommend a thorough revision of the manuscript targeting that in the end only those applications are emphasised in the abstract, introduction and conclusion that are sufficiently prepared in the methods and results and appropriately discussed.

In the following I give examples for recommended revisions in the different sections followed by a list of small comments, questions and remarks.

(1) Abstract: In my opinion the abstract does currently not fit to the content of the paper. Reading it, I developed different expectations not addressed in the manuscript, such as model applications (at least exemplary) dealing with the advocated passing of genetic information and/or diversity and/or identifying important drivers of migration dynamics. Also global vegetation models are mentioned twice prominently in the abstract, e.g. last sentence: '... substantially help in unveiling the important drivers ...', however, unfortunately, I did not get an idea from the manuscript how this could be done and how global vegetation models can be improved. I recommend rewriting the abstract more closely describing the manuscript.

(2) Introduction: The authors refer to DGVMs and list processes not implemented in them. However, they do not mention why these processes are not included – e.g. issues of parameterisation (see e.g. Snell et al. 2014), particularly also due to the high variation in seed dispersal rates (and vectors) within the plant functional types currently used (see e.g. Huntley et al. 2010, Lee 2011, Snell et al. 2014, Sven-

ning et al. 2014), issues of coarse spatial resolution and of computational complexity arising with inter-grid cell communication via seed dispersal or other spatially linked processes (e.g. Nabel 2015) – yet these issues that cause problems for larger scale models seem to even get visible in this manuscript describing a local single species model: when discussing the huge increase in computation cost and that the model is not properly parameterised for the suggested studies (indicated by sentences such as "Before applying this new model version, however, a proper parameterisation is necessary." "If this can be efficiently parameterised, the model could further be used to track genetic lineages in time."). I recommend mentioning these prominent problems of seed dispersal implementation in the introduction and relate to them in the discussion.

(3) Methods: the methods seem largely appropriate. Below I list some specific questions/recommendations:

(a) According to 2.1 the presented simulations assume homogeneous forest stands with periodic boundary conditions. However, applications to study tree line dynamics, or migration in general, or the suggested "impact of migration processes on the genetics", require dispersal among inhomogeneous plots/gridcells and I would appreciate at least a small application showing how the model can be used for such conditions, showing e.g. what migration speed results from the model, which processes have largest influences etc.

(b) "2.2.4 parameterisation to fit field data": even if the authors only visually compared the resulting number of trees, it might be helpful to have a table listing the number of trees for different parameter combinations (at least in the supplementary). What sensitivities were detected, which parameters were particularly important/ influential in tuning the number of trees? The authors state that the model parameters required a revision after implementing the extensions - what can we learn from this? It might be worth to discuss which parameters had to be revised and how, which are sensitive and influential - these things could be really helpful for the community.



(c) Most of the parameters in Table 1 are stated as "estimated" but there at least needs to be a starting point! If the study aims to help parameterising DGVMs it would be helpful to at least know on what the parameters are originally based, even if its from another species etc.

(d) 2.3.2: I find this section difficult to read and understand. If I am not mistaken the dispersal function was once implemented inside and once implemented outside the model (which I think is not explicitly stated somewhere?). I would move most of this and the according results section to the supplementary and only retain that the model was successfully verified.

(e) 2.3.3: The header of this section suggests an evaluation. The text of the section suggests that the simulation results will be quantitatively compared to observed values from literature. Yet I did not find such a comparison in the results section.

(4) Results: thin (see remarks above)

(a) As stated above, I would move most of 3.1 to the supplementary (all comparison to 'direcly' estimated/calculated).

(b) I do not get the "fraction of significant differences" in Table 2

(c) The second parallelisation variant outruns the first when using four cores, however, there seems to be a saturation in that 8 cores are not doing better but even worse than 4, why is this? It would be interesting to see what happens with 16/more cores!

(5) Discussion:

(a) The authors start with stating that unlimited seedbeds cause high uncertainty in DGVM predictions. While I agree that the assumption of unlimited seedbeds contributes to the uncertainty in DGVM studies/projections I think that there are heaps of at least equally important uncertainty sources, like modelling of mortality, competition, canopy structures, etc. Furthermore, I do not see that the manuscript present a remedy for the uncertainty connected to seedbeds. In my opinion a min/max assumption

with no and full dispersal currently gives a better uncertainty frame than difficult to parametrise dispersal kernels. If one of the main messages of the manuscript stays that LAVESI-WIND can be used to enhance DGVMs, e.g. by parameterising dispersal kernels, I strongly recommend to discuss the issue and difficulties of (PFT) parameterisation (see e.g. Huntley et al. 2010, Lee 2011, Snell et al. 2014).

(b) LAVESI is a single species model parameterised for a certain Larix species. I wonder how not represented processes, such as competition with other species/understory, trait variability, or others, could influence migration processes, particularly in the currently tree-less tundra? Is this implicitly covered by the parameterisation? If it's implicitly covered, how might this change with climate change?

(c) Large parts of the model are realised as stochastic processes. How many repetitions are required to get a sound result, e.g. how is a targeted output variable such as migration speed influenced by the stochasticity?

(d) The last sentence of 4.2. stating that the model is not yet parameterisation to field data sounds as if it would only be a small step, but I assume that this is rather difficult and would like to read a bit more about how this can be done and what problems are anticipated.

(e) The increase in computation costs associated with seed dispersal and the like have been mentioned in several places (e.g. Snell et al. 2014, Svenning et al. 2014, Nabel 2015) and are one of the reasons why seed dispersal is usually not accounted for on larger simulation extents. I think it would be worth to mention that this is a common finding when dealing with seed dispersal. Furthermore, the authors mention hundreds of CPUS here but did only conduct simulations with 1,4 and 8 CPUs (and with one of the parallelisation the avg sim time was worse with 8 than with 4 CPUs). In order to better support their statements and in order to assess the merit of the parallelisations another set of simulations with e.g. 16 or 32 CPUs would be helpful.

(f) In my opinion section 4.4 is not well supported by the rest of the manuscript. I for



example miss a link from simulations on the homogeneous 200 x 200 m plot with periodic boundaries to a study investigating migration rates. How are borders of simulation plots connected? How is the long distance dispersal dealt with and pollination which potentially reach over several plots? How does this increase the computational complexity, is such a simulation currently computationally feasible? Especially, when you a certain set of repetitions is required to account for the stochasticity? Furthermore, how can pollen influx be studied when the amount of pollen is not simulated by the model?

==========

small comments, questions and remarks:

l.1 - title: Since LAVESI did already contain spatially explicit seed dispersal, maybe add wind driven "Implementing spatially explicit wind driven seed and pollen dispersal ..."

l.23 - is the selected pollen donor still potential? maybe: "... select the pollen donor."

l.46 - "is mainly caused"

l.47 - "using a time lagged response function"? Snell 2014 uses a dispersal function for dispersal between grid cells and a logistic function to limit within grid cell dispersal"

l.53 - "this" refers to what? local adaptation?

l.59 - "this ensured the most realistic implementation" I would remove this part of the sentence or rephrase the paragraph, because of the previously mentioned dynamics at the treeline. I had to read the sentence several times before I realised that this refers to the homogeneous forest only.

l.63 - "are not coupled" – There is a phd thesis integrating wind dispersal in the DGVM CLM, see https://globalchange.mit.edu/sites/default/files/Lee_PhD_2011.pdf

l.64 - 'Among others, Student's 2Dt...' This sentence seems lost - delete it?

l.69-77: move to methods or discussion

l.84: "Results of the validation" -> verification?

l.83-86: I would delete these sentences, particularly the second part informing that there is a discussion and conclusion section.

l.93: I would use past here (i.e. "which were parameterised"), since you describe the parameterisation it in Kruse et al. (2016) and, if I understood it correctly, re-parametrised it for this study?

l.95: What exactly are "homogeneous forest plots of 100 x 100 m" - 1. does a simulation runs on one plot or several plots? 2. Does homogeneous refer to the environmental conditions? What is the resolution and what is the extent of the simulation runs?

l.95: "seed dispersed beyond the plot borders" - thus out of the 100 x 100m plot?

l.99: "... lag the hypothetical warming ..." – maybe "... densification and northwards migration might lag the applied hypothetical warming ..."

l.111: for better readability maybe insert a space or comma or other separator between distanceratio and rand

l.113: Does this refer to (1) the sensitivity analysis in Kruse et al. 2016 or (2) a new analysis? If (1) maybe point again to the reference, if (2) can you show results of the sensitivity analysis in your supplementary?

l.116+117: do you have references for the 75%, for Vd and lambda? Is there any biological interpretation for lambda?

l.120: a reference to 2.2.5 would be helpful here

l.124: "subsequently" to what?

l.128: could you shortly write what C and m are? Why do you use m=1.25? Gregory (1961) seems to recommend different values for m for different conditions.

l.133: the symbols for the angle between trees and the actual wind direction are very



difficult to visibly distinguish, because the bar is hardly noticeable. Maybe make the bar more visible, lower, or use another symbol for the wind direction

l.143: "monthly climate series" - reading that the model has an annual temporal resolution I had to went back to the LAVESI description. I think it might be helpful to very briefly state which CRU variables are used to derive which annual driving variables

l.147: in this sentence field work is mentioned but it is neither explained what field work this was, nor a reference given. In l.156 I can infer that the reference probably is Wieczorek et al. (2017) and from this and other places I infer that the field work took place from 2011-2013. Maybe it would be easier to swap 2.2.3 and 2.2.4 and shortly properly introduce the field work.

l.149: should probably be "-36°C"

l.153: which field data? (see above). 100 x 100 areas -> is this the one 100x100m plot? How many repetitions did you simulate?

l.155: I do not get which climate you use. Is this the climate described in 2.2.3 or not? If so: do not repeat, if not: why not?

l.156: even if the number pf trees was only visually compared it would be nice to see some combinations of number of trees and parameters in a table

l.158: what is a good fit?

l.160: "Climate forcing data" – leftover?

l.171: do you refer with "plots" to the 50 repetitions of the one simulated plot?

l.182: from all pollen sources or from the selected donors?

l.184: "implemented simulations" -> "conducted simulations"

l.184-: repetition with 225

l.187: How do these two parallelisations differ?

l.203: "north between 345° and 15°" - how does this compare to the directions in Table 2?

l.219: plural vs singular: "computation time increases"

l.220: readability: "produced seeds"

l.225-: repetition with 184 (however without 2 CPUs)

l.239: uniform ? -> 2.2.2: "based on the probability density"

l.242: "the higher values" - which?

l.242-: I do not get: "4-6% of test show significant differences ... no evidence that the differences are significant"?

l.257: Maybe output storage is not in the Table S1?

l.269: maybe replace inferring by calculating/tracking

l.270: maybe replace estimating by calculating

l.273: the second variant outruns the first when using four cores, however, there seems to be saturation in that 8 cores are not doing better than 4 but slightly worse, why is this?

l.289-: I would delete these three last sentences

l.300: "∼ 12 m metres"

l.303: -> Duncan (1954)

l.312: how did the winds change, such that the potential migration rate slowed down and not e.g. increased?

l.317: maybe add 'pollination' and 'this', i.e.: "overestimation of pollination distances, but this seems unlikely"



l.323: LAVESI-WIND?

l.324: maybe "evaluate with field-based data"?

l.336: from the perspective of a DGVM grid cell (typically > 0.5°) 5,000 x 5,000 m is still very small scale

l.337: LAVESI-WIND?

l.620: LAVESI-WIND?

l.621: "for each adult tree (potential pollen source) and for each seed source."

l.622 maybe write: "the shaded areas ... pollination probability for the labelled source ..." -> this is the probability for the one labelled source, or?

Fig 3.: I am surprised that 150-200m does not occur for pollination. Is this due to the perspective taken for the plot? I.e. when I understand it correctly the seeding is from the perspective away from a centre and the pollination is from the perspective of the centre? A short sentence on this would be great for a quick understanding. When looking solely at the figure it would also be nice to have some more information in the caption on what the figure shows, e.g. a short note if this is an average of 50 runs with the full LAVESI-WIND. Table 1: For a discussion of the re-parameterisation it would be nice if the original values of the parameters could be listed here, too. This would also help to identify the new parameters. Table 2: The different notations -15 and 345 are a bit irritating.

Supplementary

- l.8: note 5?

- "D.f." / "d.f." ? SS? F?

- Table S3: comma for thousand misplaced

- Table S3: why do the distance percentiles differ as compared to Table S2?

- Table S4: notation mixing initially confuses (-15 vs 345)

Additional References:

Huntley, B. , Barnard, P. , Altwegg, R. , Chambers, L. , Coetzee, B. W., Gibson, L. , Hockey, P. A., Hole, D. G., Midgley, G. F., Underhill, L. G. and Willis, S. G. (2010), Beyond bioclimatic envelopes: dynamic species' range and abundance modelling in the context of climatic change. Ecography, 33: 621-626. doi:10.1111/j.1600-0587.2009.06023.x

Lee, E. (2011). Impacts of Meteorology-Driven Seed Dispersal on Plant Migration: Implications for Future Vegetation Structure under Changing Climates. PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Nabel, J.E.M.S. (2015). Upscaling with the dynamic two-layer classification concept (D2C): TreeMig-2L, an efficient implementation of the forest-landscape model TreeMig. Geosci. Model Dev., 8: 3563-3577. https://doi.org/10.5194/gmd-8-3563-2015

Snell, R. S., Huth, A. , Nabel, J. E., Bocedi, G. , Travis, J. M., Gravel, D. , Bugmann, H. , Gutiérrez, A. G., Hickler, T. , Higgins, S. I., Reineking, B. , Scherstjanoi, M. , Zurbriggen, N. and Lischke, H. (2014). Using dynamic vegetation models to simulate plant range shifts. Ecography, 37: 1184-1197. doi:10.1111/ecog.00580

Svenning, J. , Gravel, D. , Holt, R. D., Schurr, F. M., Thuiller, W. , Münkemüller, T. , Schiffers, K. H., Dullinger, S. , Edwards, T. C., Hickler, T. , Higgins, S. I., Nabel, J. E., Pagel, J. and Normand, S. (2014), The influence of interspecific interactions on species range expansion rates. Ecography, 37: 1198-1209. doi:10.1111/j.1600-0587.2013.00574.x

---

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-31, 2018.