Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-161-AC3, 2018

Geoscientific
Model Development
Discussions

Interactive comment on "Simulating migration in dynamic vegetation models efficiently with LPJ-GM" *by* Veiko Lehsten et al.

Veiko Lehsten et al.

veiko.lehsten@nateko.lu.se

Received and published: 9 October 2018

Dear reviewer, we want to thank for the time spend for this thorough review. Please find a step by step response below. All line numbers will be entered once the final version of the ms is ready.

Interactive comment on "Simulating migration in dynamic vegetation models efficiently with LPJ-GM" By Veiko Lehsten et al. Anonymous Referee #2 Received and published: 21 September 2018 Comment: The paper presents two methods for simulating tree species migration, newly imple- mented in the dynamic global vegetation model LPJ-Guess. I find the paper mostly well written and generally an interesting scientific contribution. Response: Thank you for this summary.

Comment: What I, based on the presented material, cannot consent to is the reoccurring state- ment that the model can be used for continental simulations of multiple interacting species, nor that it is suitable for DGVMs beyond special cases (i.e. species simulations in Europe).

Response: We agree that the performance of the dispersal algorithm as presented creates a doubt that a continental simulation is possible. Given that we clearly make this claim should require for us to present an algorithm that would have a suitable performance. The main aim of this paper was to introduce the two algorithms into the DGVM and to perform continental scale simulations in a upcoming work. Hence we did not invest a lot of time into optimization, except for a paragraph in the discussion for simplicity of the paper. We have now used two more options to optimize the speed that we evaluate in the Matlab script which allowed us to increase the performance by more than an order of magnitude of the FFTM versus the explicit simulation (one of the improvements was already part of the LPJ-GM simulations and was not included in the Matlab script for readability),while the other (which only improves by 30% to 50%) is currently not implemented. We will rephrase all statements where we state that our implementation can be used for continental applications and write that it has the potential to be used for large areas and also has a lot of potential for performance improvements. Again the intention was to present the two algorithms for seed dispersal, while any continental scale simulation experiments would require to also present a completely new parameterization of some of the trees, as well as many other aspects hence we would like to refrain from performing continental scale simulations for now. We completely agree with the second statement. Given that we worked a lot in the past with species simulations we completely forgot that the currently most common application of DGVMs uses PFTs and there it is not necessary to include seed dispersal.

Comment: Most DGVMs use plant functional types with mixed dispersal/reproduction traits, particularly when used for large spatial applications. The example application deals with only two species, and only the dominant late successional tree species Fagus Sylvatica is tracked, which has a quite narrow dispersal kernel. Furthermore, the application deals with a homogenous landscape. From what the authors show and write I am not convinced that/how a continental simulation with multiple interacting and dispersing species would/can be possible. Response: Again we completely agree that only for simulations on the species level, an inclusion of seed dispersal is useful. Currently all dispersal simulation is performed at the master node, if the dispersal simulations for each species are performed at one node per species there should be (theoretically) no reasonable reduction in performance. We also repeat that we have not proven that we can simulate at a continental scale, but only that we are 2-3 orders of magnitude faster in simulating seed dispersal compared to the explicit simulation. Whether this is sufficient for a continental simulation will be shown in successive work and we will rephrase this claim.

Comment: From the paper I understood that using FFTM with widely spread transects would not be appropriate in heterogonous areas. SMSM with terrain, on the other hand, would not save enough computation time to be applicable on continental area. Is the plan to use FFTM with transects in homogeneous areas and SMSM in heterogeneous areas? But if so, how would these algorithms then communicate with each other in a continental simulation? Response: Though not formulated in the paper (for simplicity) yes this is the main idea. Depending on the parameters of the species specific dispersal kernel, there is a maximum distance that the seeds are transported (theoretically there is no such limit, but given the strong decrease of the tail, this assumption has no influence on the final result). One option is to define a certain area as heterogeneous. The seeds produced in this area are dispersed by the SMSM algorithm, while the seeds of the remaining area are dispersed by the FFTM. Though for both the seed production is only taken into account for the assigned areas the seed fall will be calculated for the area plus an edge surrounding the area with the width of half a maximum kernel width. In a last step the dispersed seeds of the two methods are added to a final distribution of seeds. This way there is no complicated communication of the two algorithms, but the edges of the areas that are simulated overlap while the areas of the

seed production that go into the algorithm do not. Since the reviewer asks for it we will present this reasoning in the Discussion.

Comment: Given that the paper, the presented ideas and the LPJ-GM implementation are already a substantial contribution, I recommend reconsidering the (over?) statements regarding continental applications and DGVMs e.g. in the last sentence of the abstract and particularly the first sentence in the discussion and talk about DVMs with species and spatial extents exceeding applications of a few ha, which is a good and sound contribution. Response: We will tune down our statements and clearly show what level of areal coverage we show the algorithms being able to simulate.

Comment: Another way could be closing the explanatory gaps, i.e. (1) discussing issues with DGVMs and how DGVMs, which usually use plant functional types (PFTs) for large scale applications, could be parameterised for the algorithms, (2) discuss the costs/difficulties of an application with a realistic number of interacting species, with differing dispersal traits and (3) explaining how a realistic continental simulation could be assembled with the FFTM and/or the SMSM simulation, given spatial fragmentation and spatial heterogeneity. Response: We will do this in the Discussion. General comments 1. In many places in the text the authors state that transect simulations lead to similar/only slightly reduced migration speed. However, from the figures/table it seems to be quite a significant underestimation, and the less transects the worse the underestimation of the migration speed (>20%; i.e. in the 3000y application >600y delay). I recommend stating this clearly and to discuss the consequences. Response: An underestimation of 20% might look great but given that most current approaches completely ignore seed dispersal and also that the parameterization of the seed dispersal kernel comes with quite an uncertainty puts the 20% in perspective. The currently reached migration speeds are very likely to be too low as described in the Discussion, which is probably caused by uncertainties in the parameterization of the seed dispersal kernel. However as the aim was to implement the two methods in a DGVM we aimed to keep the kernel similar to the parameterization within TREEMIG to be able to compare results.

Comment: 2. The authors claim that they fulfil the stochastic requirements because they have 200 or so 1km2 grid-cells when comparing to the usual 0.5 x0.5 grid cell. However,this only holds if the spatial heterogeneity caused by the stochastic disturbances and stochastic mortality does not affect tree species migration. In the example application the authors choose Fagus Sylvatica, a dominant late successional tree species, and I can imagine that for this species the stochasticity might indeed play a minor role. However, what in case of e.g. pioneer, less dominant/more specialised species? These might depend on disturbed areas for establishment – is the transect approach valid for such species? I would find it very helpful to see how the stochasticity and the few available transect cells might affect the spread of such species. Response: Pioneer species are typically fast migratory species, hence they would typically be able to colonize the area before the late successional species arrives (if both occupy the same climate space). If on the other hand the late successional species has a larger climate range you are correct that the early successional species will be hindered in its migration into the few spots that are available over a short time. This is true for both a full simulation as well as a simulation along transects. The fact that the seed survival of early successional species is typically higher due to the lower seed mass, should allow the spread still in the small temporal successional gaps. Again the focus here is to introduce the methods. A parameterization of the species that results in a migration speed comparable to observed values is outside the range of this study. However we agree that this is an important point and will discuss it in the limitations and further work section. Comment: It is correct that if applied globally DGVMs usually use 0.5 grid-cells, however when applied as DVMs on continents or regionally the resolution is usually much smaller. See e.g. the dispersal experiments by Snell (2014), and the simulation of European potential natural vegetation with LPJ-Guess (Hickler et al. 2012). 4. I would recommend referring a bit more to relevant literature in some parts of the text, since several of the ideas/methods have already been discussed/used elsewhere. (I mentioned some references in the specific comments list below). Response: The

maximum resolution is dictated by the climate data available. Given that LPJ-GUESS is parameterized using CRU climate data the simulations at coarse scale are typically performed at 0.5 degree or sometimes using the CRU climatology to bias correct the CRU timesseries it is performed at 0.25 or 0.1 degree. Even when run with 10 patches and at 0.1 degree would result in 250 patches per 0.5 degree which is comparable to the 200 simulations that we perform at any 50 by 50 km cell. However we will mention that non-global simulations and especially regional simulations use finer grid cells.

Comment: What I miss in the current introduction is a bit more on why migration is missing in DGVMs. The authors state that one reason is the '1D' property, i.e. that cells are not interacting and thus the computation costs of making them interacting. But what should also be mentioned is the problem of parameterisation: DGVMs usually use PFTs, often compiled of species with various different traits with respect to migration (dispersal vectors, competitiveness, generation times, ...) (e.g. Snell et al., 2014). Response: We will mention this.

Comment : Furthermore, if I understood it correctly, the example simulation is for 3000y and the tracked species migrates 100km in that time. Several of the criticized studies with DGVMs (e.g. "land use change on vegetation and ecosystem properties") would use well below 3000y; mostly around 100/200y – given the comparable cheapness of 1D simulations and the mentioned constraints due to parameterisation: wouldn't a 'no dispersal' simulation be sufficient for many simulations with large spatial extent and coarse resolution? Response: As already mentioned before (and highlighted in the Discussion) the migration speed that we are calculating are way too small compared to measured values. The aim of this paper was to introduce the method and here we choose to measure our success in response to the migration speed simulated by TREEMIG since we used a similar dispersal kernel. Any real world application will require a new parameterization of the kernel to gain a realistic speed. Currently most simulations are not 'no dispersal' simulations but 'extremely fast dispersal' simulations (given that they use free establishment). Otherwise they could not show any response of vegetation on climate. For short term studies where the time horizon is well below the generation time of the species there is a limited use of a dispersal kernel. Especially for Europe where land use is dominating the vast majority of the landscape and at least in the northern part also plants (often alien) species in forest any kind of simulations assuming semi-natural conditions are questionable. However to understand current tree distribution in those parts which are still semi-natural and especially to understand forest species history taking seed dispersal into account might be important. We will mention this reasoning in the Discussion.

Comment: I would appreciate a more detailed description of the SMSM method. Maybe an illustration? Would this method work with a species with a more pronounced long distance dispersal tail than Fagus Sylvatica? What would this mean regarding computation costs? How to parameterise the SMSM? Could a setting like Fig 5 be simulated with transect at all? Looking at the supplementary figure it seems that the matrix shift method with a terrain has a very small computational gain? Response: Thank you for suggesting to add a figure that will illustrate the SMSM, we will do so in the revised version. In general the SMSM method is a relatively direct implementation of seed dispersal, by moving seeds from one cell to another with a certain probability. Comment: I would appreciate more discussion of the limitations and a clearer directive how to apply the algorithms for a continental simulation, if possible. When reading the text I got the feeling that the remedy for the FFTM limitations (heterogeneity/fragmentation, wind directions) is to use the SMSM, but that this method, particularly if used with terrain, is not performant enough for continental applications. Some more buzzwords for the limitation section: parameterisation of SMSM; species parameterisation; fragmentation when using the FFTM; what about ecosystems with many species (i.e.. tropics). You read correctly that FFTM is not able to handle landscapes resulting in heterogeneous seed dispersal while SMSM has strong performance constrains. As a matter of fact, the FFTM is still applicable if the barrier is larger than the kernel width since it will place seeds there but if the cells are not suitable then the seeds will not germinate (Baltic sea, alps). The SMSM is only required in areas where we have (or rather know) different

dispersal tail lengths depending on the terrain. This might be the case in some valleys in the alps, where seed dispersal acts mainly along the valleys, but not the mountain, but given the typical resolution of the output for continental studies this might not be necessary to apply the SMSM at all, while for finer scale studies SMSM might be the best choice. We will discuss this in the Discussion section

Comment: Reduction of migration speed by >20%, i.e. in the 3000y simulation > 600y delay. 8. The editor provided me with the model code. Unfortunately I was not able to understand how the simulations were done. There are no hints on how the simulations were conducted, nor was I able to identify the configuration file (instruction script (ins)?) used for the simulations or to find out how/where the transects were defined. I know that it is cumbersome but in the spirit of "good scientific practice" it might be nice to provide and mark the configurations files? Response: The transects are defined in the gridlist. Basically while a typical LPJ-GUESS gridlist contains only columns one for the longitude and one for the latitude, in LPJ-GM there are additional columns in which for each species a time is given in which free establishment is allowed see below for the start of a gridlist. 23 50 TeBS,101.0,IBS,101.0 23.01 50 TeBS,100.0,IBS,10000.0 23.02 50 TeBS,100.0,IBS,10000.0 The first line indicates that at position 23 degree longitude, 50 degree latitude both species TeBS (temperate broadleaved Summer green tree or beech and IBS Intermediate shade tolerant broadleaved Summer green tree; or birch) are allowed free establishment (hence no seed limitation) at the year 101, which is one year after the initialization phase for nitrogen initialization. Hence this cell would form a refugia for beech. The next cell is located at 23.01 degree longitude and 50 degree latitude and beech is only allowed free establishment after the year 10000, hence not within the simulated time of 3000 years, it can only establish at this site if seeds arrive there. Birch is allowed to establish at this site without seed limitation. The transects are defined in a way that only the cells that are on the transects are listed in the gridlist.txt. This way of defining them might not be the most elegant one, but since the current setup of LPJ-GUESS simply cuts the gridlist into as many pieces and distributes them into different directories in which the simulation is performed, this way

I did not had to read in a separate file for the refugia definition, and I am sure that the information is linked to the gridcell. The configuration script (the ins file) is similar to the one used in Hickler et al. except that it contains these additional entries at the global level. ! migration INSTRUCTION years_total 3000 ! How many years the dispersal simulation is performed domain 23 50 0.01 0.01 ! which domain is simulated and with what resolution param "size_lat" (num 100) ! how many cells are in the domain along the latitude param "size_lon" (num 100) ! how many cells are in the domain along the longitude dispersal_patchsize 0.99 ! How big a single patch is. if_dispersal_fft 1 ! whether FFTM dispersal is performed if_dispersal_float 0 ! whether SMSM dispersal is performed if_dispersal_ext_fft 0 ! whether another variant of FFTM (not described in the paper) is performed stochastic_seed_est_scaler 0.01 ! scaler for the patch size output_interval 10 ! in years save space since not all years are needed in the output

Each species contains the following extra parameters which are taken from TREEMIG(here are the values for beech): max_fecundity 29. ! maximum fecundity min_height_for_maturity 14.4 ! minimum height for maturity germination_rate 0.3 ! rate of seeds germinating per year max_seed_age 3.3 ! maximal survival times for seeds in seed bank short_range_disp_frac 0.99 ! fraction of seeds that go into short seed dispersal short_disp_alpha 25 ! parameter for short distance dispersal long_disp_alpha 200 ! parameter for long distance dispersal

I also would prefer to make the whole model code publicly available. However current policies within the modelling consortium only allows to give access to model code after individual contact with the author. I decided that my unit containing the implementation of the code for the actual migration will be made publicly available (as a supplement to this paper), but there are of course some other small bits and technical issues, like for example the MIP related code that is located in other units.

Specific comments: l.1: Maybe consider to adapt the title, since LPJ-GM does not necessarily lead to a more efficient simulation of migration in dynamic vegetation models per se – e.g.: "Simulating migration in dynamic vegetation models efficiently on the

example of LPJ- GM" or maybe better "Simulating migration efficiently in the dynamic vegetation model LPJ-GM" Response: Yes we will change the title to “Simulating migration efficiently in the dynamic vegetation model LPJ-GM 1.0” Comment: l.21: Most DGVMs do not use species but plant functional types (PFTs) Response: We will mention this. Line… Comment: l.31: From the last Figure in the supplementary, SMSM with terrain seems to be much slower than FFTM? Response: Yes it is especially now that we have optimized the Matlab code (though at the expense of readability) it is. We will delete the word ‘marginally’. However it is still faster than an explicit seed exchange. Line : … Comment: l.40: “Furthermore, with the transect method both methods”? Response: We will introduce the ‘with the transect methods’ . We will replace continents with large regions, since we have not really shown that continents can be simulated with our method. Line:… Comment: l.49: DGVMs assume that some instance (i.e. species) of the PFT can establish

Response: We will add a sentence before stating that while most DGVM applications use PFTs which are not suitable for the seed dispersal simulations due to different seed dispersal mechanics within the same PFT, we are concentrating here on applications simulating explicit tree species. Line: …

l.51: Something seems not correct with the embedded sentence – maybe that instead of the? Response: No here that would give a different meaning. We inserted an ‘a’ and hope the sentence is now easier to read. Line:…. Comment: l.53 & 63: Anyway DGVMs usually do not simulate species but only PFTs Response: Since we already write in line …. that we are only considering species simulating DGVMs we consider this covered.

Comment: l.60: When considering ecosystem properties in the future hardly any study would make projections Âż100y, maybe 200y, but the example in this study uses 3000y. Wouldn’t – based on what is shown in this paper – a “no migration between large gridcells (0.5 âŮę) assumption” be appropriate for studies with âĹij 100-200y? Response: We mention this limitation at line :….. “However, given that most studies using future climate simulate only 50 or 100 years ahead, which is way below the generation time of trees, and because of human activities which plant many tree species outside its native range, the use of explicit modelling of seed dispersal in DGVMs might be limited for studies of future tree distribution." Comment: l.76: Another example of a model even accounting for wind speed and direction: LAVESI-WIND (Kruse et al. 2018) Response: Thanks for pointing us to the paper, we are citing it now in line .... Comment: l.82: What does the spatial heterogeneity refers to in this context – soil and climate? If I understood the set-up right LPJ-GM also does not account for such heterogeneities within the grid cell, only to such with regards to species interactions and stochasticity? Response: Yes: Soil and climate, mountains blocking seed transport as well. This sentence should simply highlight that a simple transfer of migration speeds calculated with models at fine scale into models at coarse scale is challenging.

Comment: l.94: why every time-step? LPJ-GM only does it once per year? Response: Yes of course we mean annually. We change this at line ....

Comment: l.100: If I understood it correctly the presented simulations only simulate two species. Response: Yes in our example simulation only two species are simulated but the method can simulate more species in a real application case.

Comment: l.101: What would a simulation with several species look like, does each need one FFT/SMSM? What are the resulting costs? Response: In the current simulation time that we present in the table we are actually simulating the seed dispersal of both species independently (though it would of course be faster to only simulate one species). Yes each species needs its own FFTM or SMSM to be performed if the migration of several species is to be evaluated, however they could be potentially performed at separate nodes which would decrease calculation time again. We are mentioning this in the section where we discuss the performance (Line 453 to 455 in the old ms). Comment: l.110-111: Please list a few key references describing LPJ-GUESS 4.0 Response: We included Smith 2014 and Lindeskog 2013 which are the main references describing the 4.0 version. Comment: l.119: Above and below this

node is called master Response: We now call it master here as well. Comment: .123: "no seed dispersal"-> "no seed limitation"? Response: Yes we added this as well. Line ... l.130: There are species producing seeds throughout the year (see e.g. Owens 1994, Brokaw, 1998) Response: Yes this is one of the discretization errors that we have to make. Given unlimited computing power and knowledge of weather conditions and plants behavior, we would perform the FFTM or SMSM daily over the time when seeds are produced. However, as a first improvement of the situation in which most models do not consider seed dispersal at all, we suggest to simulate at an annual time step. Comment: l.140: Here or generally in LPJ-Guess? Response: It is variable but this is the recommended size. Comment: l.153-157: How is this similar to Lischke et al., 2006? Lischke et al. (2006) do not men- tion LAI but state: "The number of seeds S produced per year by each tree depends on its height, species and mast seeding period."? Response: We also use the height of maturity, but no mast seeding period, while Lischke et al. scaled the seeds with height we did scale them with LAI, you are correct this is not the same and we have taken that sentence away as it was meant introduce into the chapter but it is necessary. Comment: l.175: For Fagus Sylvatica?! Response: Added. Actually we also simulated seed dispersal for Birch but since Birch is set to no seed limitation that does not matter. Comment: l.181-182: But wouldn't the implementation of wind direction lead to anisotropy and therefore make the FFTM not applicable anymore (E.g. Neupane (2015))? Response: The FFTM can apply any shape of seed dispersal kernel, it can just not change it with the landscape. Hence certain wind directions are possible like the kernel used for illustration, which is also skewed, for example by wind. Neupane simualates effects of the landscape on fruit dispersing birds. Such an effect would have to be modelled by the SMSM. If the different wind directions in different parts of the domain (e.g. caused by a certain terrain) is to be taken into account, this also needs to be done by SMSM.

l. 185: maybe use θ and add the θ = 1 in the text below? We will adjust this equation. Comment: l.186: long term -> long distance? Response: Changed. Comment: l.197: but how is the number of seeds defined in this case, since in the next para it is stated that the establishment of seedlings depends on the number of available seeds? LPJ-GUESS calculates the number of established individuals per species depending on the light reaching the forest floor. LPJ-GM takes this value and either and sets it to zero depending on the presence (or rather absence) of seeds. In case of establishment free from seed limitation (in our case the birch), this step is not performed. Hence the species can always establish depending only on the light reaching the forest floor. Comment: l.202: "depending stochastically depending" Response: Thanks for spotting this repetition. It is fixed. Comment: l.206: "seed bank per and the germination" remove the per? Response: Done Comment: l.216: The authors should definitely mention that the method has also already been broadly applied in simulating dispersal. E.g. have a look at Powel (2001) + shortly googling I e.g. found Pueyo et al. (2008) and Prasad et al. (2013) and I assume there are more. Response: We are now mentioning that there are a number of applications which already use ffts to simulate dispersal and cite a few of them. Line . . ..

Comment l.235: "different wind distributions" -> only if they are valid for the whole simulated area, or? Response: We added this remark in line: . . .. l.242: How is this proportion determined? A few lines later we point to a derivation of the parameters (in this case this proportion) in the supplementary material S.1.

Comment: l.242: 1km2 cell? Response: Yes in our application all cells have one kmˆ2 extent. Comment: l.249: How often is this done/ needs to be done to account for long distance dispersal? Response: Currently this is done 10 times hence we are reaching a maximum of 10 km. Comment: What happens with the seeds at the boundary of a simulation area? Response: For both the FFT as well as the SMSM simulation we extend the area by one kernel width to avoid / minimize edge effects. Basically all seeds that land of the seed domain are lost.

Comment: l.255: Figure 2 is not cited in the main text (only in Fig. 5). Response: Thanks for spotting this. This sentence must have gotten lost in one of the internal revisions. We now refer to Figure 2 in the description of the simulations. Line . . ..

Comment: l. 266: Wouldn't the heterogeneous landscape be much more crucial to test the appli- cability of the methods? Response: No the idea is to only use the corridors in homogenous landscapes and to speed up the simulation there. In heterogeneous landscapes this simplification is not suitable. Hence we only test the corridors in homogenous landscapes. Comment: l.274-275: And? But? Response: And we do not want to strongly increase the migration speed. We have spent a lot of time trying to come up with a better solution like some kind of weighted average, however so far we have not found one. Hence we prefer to have a reduction by 20% hence a conservative estimate rather than a strong increase which also was heterogeneous within the simulated area depending on the arrangement of the corridors. We have not given up the hope to come up with a better solution in a real world application. Comment: l.276: Out of interest: how many CPUs were used/ what computing environment? Response: We used 200 nodes (with 20 nodes per CPU) at the LUNARC computing facilities. Comment: Would it be possible to add a 'no dispersal and no communication' (i.e. a 1D) simulation for comparison?

Response: Yes we are working on it and will add it to the final version of the paper

Comment: l.303: 1km2 grid cells Response: We added this. Comment: l.306: somewhat? >20%! Response: We have removed the somewhat. Comment: l.309: Maybe add the numbers for the variability Response: What we mean here is visual realization that the distance of the points increases. Since we do not use a mean value to estimate the migration front, it is hard to quantify this variability since we have a different variability above compared to below the line Comment: l.312: Which probably also explains the patterns in the migration front?! Response: Yes that is the reason. Comment: l.314: When FFT when FFTM? Response It should always be FFTM. Thanks for spotting this. Comment: l.318-319: How do the simulations compare to a simulation without communication between grid cells, i.e. 1D simulations? Response: We are currently running this simulation and will add this comparison in the table in the final paper. Comment: l.323: How to specify this parameter when not having a FFTM simulation at hand? Response: Here the aim was to parametrize the SMSM in a way that we have a similar migration speed compared to FFTM. In a practical application one would have a certain dispersal kernel and the derivation in Supplement 2 allow to estimate the parameter to fit a Gaussian kernel. It is also possible to transform the Gaussian kernel to any other shape by adding several Gaussian kernel. If would wanted to do this we would have increased the calculation time for the SMSM. Hence we opted for a more practical approach to get comparable results with the two and still keep the kernel and parameterization from TREEMIG. Comment: How for SMSM with terrain, does this require a simulation without terrain before? Response: Well as stated before, one can mimic the function used in the FFTM or one can use a Gaussian dispersal function to start with and calculate the parameter for the SMSM from the distribution. However in our case we wanted to be comparable to TREEMIG, so we choose their function and parameterization. And to avoid to use several Gaussian to approximate the function used in TREEMIG we simply tested in an homogenous area. We are now mentioning this in the text on line: . . . Comment: What are the cost reductions then? Response: When the final dispersal kernel is approximated by stacking several Gaussian dispersal kernel the SMSM has to be performed several times.

Comment: l.335-337: I would find it valuable to have the simulation times for the terrain simulation in the table, too! Response: All SMSM calculations are with terrain, though the terrain is a homogenous grid or 1s. In the Matlab script we have differentiated between simulation of SMSM with terrain (one extra multiplication) and without, however since we are not planning to use any SMSM without terrain and since the LPJ-GM code always does a terrain, we decided to remove the SMSM without terrain from the Matlab script. Comment: l.341: This would probably not work with transects?! Response: Yes, the transects have to be chosen in a way that they are not disrupted by barriers that are larger than the dispersal kernel. The main idea behind using the transects is to use them only in heterogeneous areas where you would simulate the whole area. However some parts of the typically squared domain might be homogenous so one might to choose to use transects there as well. Comment: l.358: K is the number of iterations?!

Response: Yes thanks for spotting that we did not explain this. We now added it to the text at line . . . Comment:

l.361: “a very similar migration pattern” I would delete the “very” Response: Done. Comment: l.362: it is slower by 20%! Response: Yes but given the differences in the literature of migration speed within and between measured and simulated migration speed as well as the uncertainty in the parameters of the seed dispersal kernel this is still relatively similar. Comment: l.362-363: in l. 310-312 the authors state that its slower because of the migration path? How do these two different explanations contribute? The stochasticity leads to an increase in migration speed if there are surrounding cells right and left that can contribute via diagonal seed exchange to the cells along the transect. We are currently testing the effect of transects being wider than a single cell, but the results of this would make the ms more complex and we will present them in the next application.

l.364: how to parameterize “explicit considerations of wind directions” Basically one could calculate different Gaussian distributions in different directions and according to the wind distribution in one area using the considerations in Supplement 2. Comment: l.376: Something is missing in this sentence Response: we added an ‘or the other’ to make it clearer Comment: l.379: Maybe in a DVM? But not in ecology and not to simulate dispersal; the authors should mention some applications - as mentioned above: have a look at Powell (2001) + other references such as Pueyo et al. 2008; Prasad et al. 2013 + I imagine there are much more. Response: Yes we agree, we meant DGVMs, this is certainly misleading and we are now relating to some other applications in the introduction. Comment l.384: “DGVMSs” Response: Thanks for spotting this. Comment: l.410-412: 63-85% instead if 85%? I would maybe remove this quantitative compar- ison. In my understanding the size of the reduction will be dependent on the model and the set-up of the simulations, i.e. on a variety of factors, such as the number of simulated and dispersing species, the resolution, settings of the applied algorithms, etc., and since its two different models and probably very different simulation set-ups, it seems to me to be comparing apples and oranges? Response: Yes we certainly agree that there are a variety of factors influencing this and therefore it might be more suitable to not quantify it here. Instead we write that our method leads to a reduction in a similar range depending on the configuration of the corridors (Line #CONFIGURATION) Comment: l.413: more pronounced than what or where?

Response: Thanks for spotting this, the sentence that this was referring to was lost in an internal revision. Line ... Comment: l.416: maybe 0.5 and 1.0? Response: No here we actually mean 0.1 There are some applications at 1 degree and even 2.5 degree, but when vegetation or even species are in the focus the finer scales are more common. Comment: l.449: From the last Figure in the supplementary it seems that SMSM simulations with terrain are comparably much slower. Is it possible to speed them up with transects? Response: The idea is to have only those parts where the area is very complex or in which we are actually able to define different seed dispersal kernel to be used with the SMSM, all other areas should use the FFTM. The SMSM did also speed up if used with corridors. The values in table 1 are actually calculated with the extra one multiplication required for the SMSM with terrain. We therefore decided to remove the SMSM-without-terrain from the figure in Supp.2. Comment: l.462: I would not call 20% slightly Response: With respect to the uncertainty both in the parameters available for the seed dispersal kernel as well as the estimates of migration speed in the literature from pollen analysis 20% is still a low uncertainty. However we are removing the word slightly. Comment: l.457: “FTTM” -> “FFTM” Response: Thanks for spotting this.

l.465: unfortunate – I think this would be really interesting, especially when simulating fragmented landscape or non-dominant species Response: We absolutely agree and we are already performing test simulations for a further study. Comment: 481: missing “) Response Thanks for spotting this. Comment: l.486: What was the tested set-up? I assume FFTM? I.e. no ‘terrain’? Transects with 50km distance? How many competing/migrating species? All grid cells homogeneous? How many years? Response: We

tested using the FFTM (hence without terrain) using 4000 by 4000 grid cells, running for a few years only. Looking at the numbers in table 1 shows that running a full scale simulation with 21000 years over the 3463 $\frac{1}{2}$ degree cells that we typically use for European runs would take a long time: 1800 (CPUh per 100000 cells and 3000 years) /100000(cells in the MS)*3463(half degree cells in Europe) *50*50(rough estimation of how many 1km cells are in a half degee cell)) *21000 years in LGM simulation / 3000 years in testsimulation gives us roughly 10 mill CPUh. Given that my current account allows me 45000 CPUh a month that is currently not feasible and that is why we suggest the transect method. (Actually there might be even more time needed given that there are more than 2 species in the final runs). Comment: l.488: “considerable computation costs” – what were they in the tests (e.g. CPU h per simulated y)? Are continental applications possible, or are they not possible? Response: See above. If the corridors are clever placed yes they are possible and if a more efficient parallelization of the FFT is implemented. In this ms we are not providing a proof for this (we will do in the next where we aim to perform a European simulation). Therefore we do not refer to continental runs anymore. The statement at this point is meant to say that from a memory requirement there is no problem performing the FFTM over large areas. Comment: l.488: plural and singular mix: “a high Response: Thanks for spotting this. ... amounts” + what does “of the FFT as the local simulations” mean? Response: amounts : see calculations above. FFT as local simulation: Currently only one node is taking care of the calculation of the FFT. One could theoretically perform the FFT at each node and use one master node only for collecting the amount of dispersed seeds and performing the communication. Hence there is still some untapped optimization potential. Comment: l.498: what do you mean with “truly mechanistic”? I recommend deleting this statement Response: We meant that the migration rates are a result of the dispersal kernel and establishment in a mechanistic way. We agree that the term might be misleading and have removed it. Comment: l.503: “related estimates the Conclusion section”? Response: Thanks for spotting that there is an ‘in’ missing.

Comment l.613: Again “10” + something with the formatting Response: Thanks for spotting this. It is changed. Comment: l.627: Please provide a legend – even if the figure is only schematic Response: We will add a legend. Comment: l.635: When looking at the Figure and reading 2.4 I wondered where the 5*10-7 came from and how this parameter is determined? – Finally I found some information in 3.3 Response: See responses above. It is a fitted parameter to make the two methods result in comparable migration speed. Fig.3 and 4: I would appreciate if the y-axis of the distance plots on the right would have similar scales, this would really help for comparison Response: We will change the axes. Comment: l.647-648: difficult sentence – maybe: "only taken into account for grid cells ..."? Response: Thanks for the suggestion.

Comment: l.661: Comparing the dark blue spots in Fig . 2 and the white ones in Fig. 5 the Figures seem to be mirrored along the diagonal? Comment:

l .664: cpu*h = CPU h? Response: Changed. Comment: l.671: FFTM with 10: shouldn't this be 64% instead of 67%? Response: Thanks for spotting this typo.

References Brokaw, N. V. (1998). Cecropia schreberiana in the Luquillo mountains of Puerto Rico. The Botanical Review, 64(2), 91-120. Fisher, R., McDowell, N., Purves, D., Moorcroft, P., Sitch, S., Cox, P., Huntingford, C., Meir, P., and Woodward, F. I. (2010). Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations, New Phytol., 187, 666–681. Hickler, T. , Vohland, K. , Feehan, J. , Miller, P. A., Smith, B. , Costa, L. , Giesecke, T. , Fronzek, S. , Carter, T. R., Cramer, W. , Kühn, I. and Sykes, M. T. (2012). Projecting the future distribution of European potential natural vegetation zones with a generalized, tree speciesâ ÌĘ A ĚĞ Rbased dynamic vegetation model. Global Ecology and Biogeography, 21: 50-63. Kruse, S., Gerdes, A., Kath, N. J., and Herzschuh, U. (2018). Implementing spatially explicit seed and pollen dispersal in the individual-based larch simulation model: LAVESI-WIND 1.0, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-31, in review. Neupane, R. C. (2015). Modeling Seed Dispersal and Population Migration Given a Distribution of Seed Handling Times and Variable Dispersal Motility: Case Study for Pinyon and Juniper in Utah. Owens, J. N. (1995). Constraints to seed

production: temperate and tropical forest trees. Tree Physiology, 15(7-8), 477-484. Powell, James (2001). Spatio-temporal models in ecology; an introduction to integro-difference equations. Utah State University, Logan (UT). Prasad, Anantha M., et al. (2013). Exploring tree species colonization potentials using a spatially explicit simulation model: implications for four oaks under climate change. Global Change Biology 19.7, 2196-2208. Pueyo, Y., et al. (2008). Dispersal strategies and spatial organization of vegetation in arid ecosystems. Oikos 117.10, 1522-1532. Snell, R. S. (2014), Simulating long-distance seed dispersal in a dynamic vegetation model, Glob. Ecol. Biogeogr., 23(1), 89–98, doi:10.1111/geb.12106. Snell, R. S., Huth, A., Nabel, J. E. M. S., Bocedi, G., Travis, J. M. J., Gravel, D., Bug- mann, 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 (Cop.), 37(12), 1184–1197, doi:10.1111/ecog.00580. Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-161, 2018.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-161, 2018.