Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-31
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 7 March 2018

Geoscientific Model Development
Discussions
Open Access

Implementing spatially explicit seed and pollen dispersal in the individual-based larch simulation model: LAVESI-WIND 1.0

Stefan Kruse[1], Alexander Gerdes[1,2], Nadja J. Kath[3], Ulrike Herzschuh[1,3,4]

[1]Polar Terrestrial Environmental Systems Research Group, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, 14473 Potsdam, Germany

[2]Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany

[3]Institute of Biology and Biochemistry, University of Potsdam, 14476 Potsdam, Germany

[4]Institute of Earth and Environmental Science, University of Potsdam, 14476 Potsdam, Germany

Correspondence to: Stefan Kruse (stefan.kruse@awi.de)

Abstract. It is of major interest to estimate the feedback of arctic ecosystems to the global warming we expect in upcoming decades. The speed of this response is driven by the potential of species to migrate, tracking their climate optimum. For this, sessile plants have to produce and disperse seeds to newly available habitats, and pollination is needed for the seeds to be viable. These two processes are also the vectors that pass genetic information through a population. A restricted exchange among subpopulations might lead to a maladapted population due to diversity losses. Hence, a realistic implementation of these dispersal processes into a simulation model would allow an assessment of the importance of diversity for the migration of plant species in various environments worldwide. To date, dynamic global vegetation models have been optimised for a global application and overestimate the migration of biome shifts in currently warming temperatures. We hypothesise that this is caused by neglecting important fine-scale processes, which are necessary to estimate realistic vegetation trajectories. Recently, we built and parameterised a simulation model LAVESI for larches that dominate the latitudinal treelines in the northernmost areas of Siberia. In this study, we updated the vegetation model by including seed and pollen dispersal driven by wind speed and direction. The seed dispersal is modelled as a ballistic flight, and for the pollination of seeds produced, we implemented a wind-determined and distance-dependent probability distribution function using a von Mises distribution to select the potential pollen donor. This individual-based and spatially explicit implementation of both dispersal processes makes it easily feasible to inherit plant traits and genetic information to assess the impact of migration processes on the genetics. The final model can substantially help in unveiling the important drivers of migration dynamics and, with this, guide the improvement of recent global vegetation models.

1. Introduction

How fast vegetation communities can follow their shifting climate envelope in a changing environment is determined by their ability to migrate. This is exceptionally challenging under current global change and plants might strongly lag behind their moving climate envelope (Harsch et al., 2009; Loarie et al., 2009; Moran and Clark, 2012). Temperatures are increasing most strongly in the Arctic. Accordingly, forests in the tundra-taiga transition zone are expected to respond by migration into the tundra (Bader, 2014; Holtmeier and Broll, 2005; MacDonald et al., 2008). However, empirical studies show diverse responses to the warming, including treelines being stable, advancing or even retreating (Harsch et al., 2009). A taiga range expansion though, might positively feedback to a global temperature increase due to albedo reduction (Bonan, 2008; Piao et al., 2007; Shuman et al., 2011).

To predict forest stand responses to climate, computer models were designed with different scopes of complexity, between highly general or very specific (Grimm and Railsback, 2005; Thuiller et al., 2008). Among these, simulation studies with dynamic global vegetation models (DGVMs) tend to overestimate the turnover of treeless tundra into forests (Brazhnik and Shugart, 2015, 2016; Frost and Epstein, 2014; Kaplan and New, 2006; Roberts and Hamann, 2016; Sitch et al., 2008; Snell, 2014; Yu et al., 2009; Zhang et al., 2013). On the other hand, small-scale models (forest-gap or individual-based) provide sufficient detail to realistically represent the responses at a stand level, but need much effort for parameterisation and are typically not applied over large areas (Martínez et al., 2011; Pacala et al., 1996; Pacala and Deutschman, 1995; Zhang et al., 2011). Nonetheless, the ability to form a closed canopy forest depends mainly on species traits acting at a fine-scale level such as (1) time needed to mature (life-cycle, high generation time) and produce viable seeds, (2) dispersal distance and the chance for long-distance seed dispersal and (3) germination and establishment of new individuals (Svenning et al., 2014). The overestimation of DGVMs is mainly caused by unconstrained seed availability when climate variables allow a vegetation type to establish, which was recently pointed out by using a time lagged response function between the grid points in simulations with a DGVM (Snell, 2014; Snell and Cowling, 2015). In addition, the structure of a tree stand, and its response to changes in external forcing, is determined by further local processes, such as spatially explicit competition among individuals of all ages and their interactions. Of special interest is the local adaptation of populations at the molecular level, which is constrained by a high gene flow through seed or pollen distribution across populations, but also prevents negative consequences due to founder effects (Austerlitz et al., 1997; Burczyk et al., 2004; Fayard et al., 2009; Nishimura and Setoguchi, 2011; Ray and Excoffier, 2010). This has, so far, not been implemented in such models.

Treeline stands in the Siberian Arctic were densifying, but only rather slowly colonising the tundra during the past decades (Frost et al., 2014; Kharuk et al., 2006; Montesano et al., 2016), which could be attributed to seed limitation (Wieczorek et al., 2017). We developed the ***La****rix* **ve**getation **si**mulator LAVESI to simulate tree stand dynamics at the Siberian treeline on the southern Taymyr Peninsula and use it as a framework to explore impacts of climate change on larch forests (Kruse et al., 2016). In the first version, the dispersal function randomly dispersed seeds by a probability density function describing a Gaussian term with a fat-tail. This ensured the most realistic implementation and was parameterised to fit observed stand patterns. The model simulates tree stands on plots, representing a homogeneous forest, which can easily be enlarged to simulate wider areas. However, for simulations on larger transects passing from forests to treeless areas, wind direction and strength become more important for seed dispersal and needed to be included in the model. Seed dispersal processes are well studied (Nathan et al., 2011a; Nathan and Muller-Landau, 2000) and commonly implemented in vegetation models but are not coupled with wind speed and direction (Levin et al., 2003; e.g. Snell, 2014). Among others, Student's 2Dt dispersal kernel depicts best the leptokurtic behaviour of dispersal kernels (Clark et al., 1999; Clark, 1998; Nathan and Muller-Landau, 2000; Petrovskii and Morozov, 2009). Also wind patterns might change over time, as the pressure levels vary in a changing climate (Trenberth, 1990), or are directed (Lisitzin, 2012) so that an implementation of wind-dependent dispersal would enable a more realistic simulation of migration (cf. Nathan et al., 2011b).

Pollen was not represented in the former LAVESI version, but is needed to independently track gene flow by seeds and pollen through time. Pollen dispersal functions are frequently used to reconstruct vegetation composition from palaeo archives, for example in the Landscape Reconstruction Algorithm by Sugita *et al.* (2010), whereas other models have been used to track pollen clouds in tree stands (review in Jackson and Lyford, 1999; Prentice, 1985). Calculating every pollen dispersal event for each tree and seed is computationally challenging, but it can be simplified following the assumptions of Kuparinen *et al.* (2007). Accordingly, an individually based pollination for each seed could be implemented using a wind-determined and distance-dependent probability distribution function for pollen dispersal (similar to Gregory, 1961). It would make use of the von Mises distribution, which is an angular equivalent to the Gaussian normal distribution, for the two-dimensional representation (Abramowitz and Stegun, 2012). Besides tracking the full genealogy of a simulated tree stand, this pollination function allows the inheritance of individually varying traits of each tree, rather than randomly drawing the actual trait value from the pool of available traits (cf. Scheiter et al., 2013). Additionally, the implementation of spatially explicit seed dispersal and pollination would enable us to align the model to detailed biogeographical knowledge gained from molecular methods (e.g. Navascués et al., 2010; Polezhaeva et al., 2010; Semerikov et al., 2007, 2013; Sjögren et al., 2017).

We aim with this study to enable the simulation of spatially explicit and wind-dependent seed dispersal and pollination in the individual-based model LAVESI. The coupling of the seed dispersal kernel to prevailing winds and the incorporation of the pollination is described in Section 2. Results of the validation of the updated model by comparisons of simulation experiments to observed wind patterns are presented in Section 3, which are discussed in Section 4 and followed by conclusions in Section 5.

2. Methods

2.1. General model description of the *Larix* vegetation simulator LAVESI

LAVESI is an individual-based spatially explicit model that currently simulates the life cycle of larch species as completely as possible from seeds to mature trees (Kruse et al., 2016). It was set up to improve our understanding of past and future treeline displacements under changing climates, focusing on the open larch forest ecosystem in northern Siberia, which is underlain by permafrost. The relevant processes (growth, seed production and dispersal, establishment and mortality) are incorporated as submodules, which are parameterised on the basis of field evidence and complemented with data from literature. Simulation runs proceed in yearly time steps and are forced by monthly temperature and precipitation time series. The area simulated represents homogeneous forest plots of 100 x 100 m and seeds dispersed beyond the plot borders are deleted. The model is programmed in C++ using standard template libraries. This and its modular structure allow a straightforward implementation of further extensions.

The model was successfully applied to conduct temperature-forcing experiments, where simulations revealed that the responses of the larch tree stands in Siberia lag the hypothetical warming by several decades, until the end of 21st century (Kruse et al., 2016; Wieczorek et al., 2017).

Here we present the implementation of wind-dependent seed dispersal as well as the newly introduced pollination. The strict boundary condition had to be revised to allow the simulation of larger areas. Hence, we introduce a new modes of periodic boundary conditions for seeds leaving the simulated area, so that the borders of a simulation plot are connected along all borders, which we used in all the simulations for this manuscript. This mimics a tree stand within a homogeneous forest, similar to forest gap models (e.g. Brazhnik and Shugart, 2016; Pacala et al., 1996; Pacala and Deutschman, 1995; Zhang et al., 2011).

2.2. Implementing dispersal processes coupled to wind speed and direction

2.2.1. Seed dispersal

In the initial version of LAVESI, seeds are dispersed in random directions and at a distance r in m, estimated by a Gaussian and negative exponential (fat-tailed) dispersal function (Eq. 5, Kruse et al., 2016):

$$r = \sqrt{2E_0^2(-log(rand)) + 1/2distanceratiorand^{-1.5}} \qquad (1)$$

where E_0, originally named *width*, is the Gaussian distribution's standard deviation in m, $rand$ stands for a random number $\in [0,1]$ and $distanceratio$is a weighing factor for the fat tail in m^2. Parameter estimates were based on a sensitivity analysis and numerical experiments.

The wind-dependent distance estimation was implemented as a ballistic flight following the assumptions of Matlack (1987). Accordingly, seed dispersal distances depend on the height of the releasing tree H_t in m, currently estimated as 75% of H_t, and are modified by wind speed V_W in m s^{-1} and a species-specific fall speed of propagules (seed plus wing) V_d = 0.86 m s^{-1} for *L. gmelinii* and tuned by the parameter λ (set to 0.01):

$$E_0 = 0.75 \cdot \lambda \cdot H_t \cdot \frac{V_W}{V_d} \qquad (2)$$

Finally, the direction for the seed dispersal is determined by the randomly selected wind observation (speed and direction).

2.2.2. Pollination probability

A pollen dispersal function was newly implemented as a distance-dependent probability function for pollination of each individual seed, rather than simulating the large amount of pollen released by each tree (Gregory, 1961; Kuparinen et al., 2007). The pollination probability of a tree is proportional to the amount of pollen in the air column around it; subsequently, the following function is used here as the distance-dependent probability distribution of arriving pollen:

$$p_r = exp\left(\frac{-2p_e r^{1-0.5m}}{\sqrt{\pi}C(1-0.5m)}\right) \qquad (3)$$

where p_r is the distance-dependent distribution, r is the distance in m, p_e is the ratio of wind speed and pollen descending velocity estimated for *Larix gmelinii* (Eisenhut, 1961) and parameters *C* and *m* are set to $C = 0.6$ cm and $m = 1.25$ (Gregory, 1961).

The distribution described in Eq. 3 is multiplied by the von Mises distribution (Eq. 4), a continuous probability distribution on the circle, to include pollen distribution over a certain area and couple the process to the wind direction (illustration in Fig. 1; Abramowitz and Stegun, 2012).

$$p_v = \frac{exp\left(\kappa cos\left(\theta - \bar{\theta}\right)\right)}{2\pi I_0(\kappa)} \qquad (4)$$

where κ is the inverse of the von Mises distribution's variance, and $I_o(\kappa)$ is the modified Bessel function of order 0 as a function of κ, θ is the angle between trees and $\bar{\theta}$ the actual wind direction. The modified Bessel function in the von Mises distribution is programmed in its integral representation using the Simpson integration scheme (Abramowitz and Stegun, 2012).

Consequently, following Gregory (1961) the pollen distribution's mathematical form is:

$$p = p_r p_v = exp\left(\frac{-2p_e r^{1-0.5m}}{\sqrt{\pi}C(1-0.5m)}\right)\frac{exp\left(\kappa cos\left(\theta - \bar{\theta}\right)\right)}{2\pi I_0(\kappa)} . \qquad (5)$$

For each seed produced in simulation runs, every tree that bears cones is taken into account as a potential pollen donor. From these, one tree is randomly chosen based on the probability density to become the father of the produced seed.

2.2.3. Temperature and precipitation

Simulations are forced with monthly climate series from the CRU TS 3.22 database (Harris et al. 2014). We selected a grid box intersecting a location with a known northern taiga tree stand (CH06 at 70.66° N; 97.71° E, site CF in Wieczorek et al., 2017) and a northern forest tundra stand (TY04, 72.41 °N; 105.45 °E, site FTe in Wieczorek et al., 2017). From the available data we excluded years before 1934, because of missing climate station data and hence unreliable extrapolations in the data set (Mitchell et al., 2004). Furthermore, the final year was set to 2013, which is the latest year of field work. The climate at these sites either allows strong tree growth with mean July temperatures of 13.50 °C, coldest temperatures during January of -33.24 °C and a precipitation sum of ~328 mm per year or only sparse stands to emerge with 13.11 and 36.07 °C in July and January, respectively, and ~247 mm annual precipitation (cf. Kruse et al., 2016).

2.2.4. Parameterisation to fit field data

The model's parameters had to be revised after implementing the model extensions to achieve simulated tree densities comparable to field data. We conducted simulations on 100 x 100 m areas with closed boundaries initialised by introducing 1,000 seeds in the first 100 years of a stabilisation period of 1,000 years, with forcing climate data randomly sampled from the available data. For the final 80 years of each simulation we used the climate series from one of the field sites (TY04, 72.41 °N and 105.45 °E, site FTe in Wieczorek et al., 2017). We visually compared the number of trees at year 2011 to the field survey data, which was the first year of fieldwork. The parameters were manually tuned and we iteratively performed simulation runs to improve the simulation results until finally achieving a good fit to the observed pattern (data not shown; parameter values in Table 1).

Climate forcing data

2.2.5. Wind speed and direction

The model is driven with pairs of wind speed in ms^{-1} and wind direction in degrees [°]. The winds at 10 m above the surface for the years 1979–2012 at 6 hourly resolution were extracted from the ERA-Interim reanalysis data set (Fig. 3; Balsamo et al., 2015). Because of the coarse spatial resolution (80 x 80 km), we considered only the grid box over the climate station Khatanga, which is situated roughly in the centre of the treeline ecotone on the southern Taymyr Peninsula (71.9° N; 102.5° E; Wieczorek et al., 2017). During simulation runs, values are randomly drawn from the year's vegetation period (May to August; Abaimov, 2010) for each seed dispersal event and for the determination of pollination. For simulated years in which climate data are available but no corresponding wind data, a year is randomly selected.

2.3. Model evaluation protocol

2.3.1. Overall simulation setup

The simulation experiments were conducted on 200 x 200 m plots using the model with the new processes. Populations were initiated on empty areas by randomly distributing a fixed number of seeds during the first 100 years of a 1,000 year long stabilisation period. The simulation model randomly drew weather conditions for each year from the complete available period 1934–2013 during the stabilisation period. In the final 80 simulation years, the actual weather data were used.

First, we performed simulation experiments with constant wind conditions to verify the implemented dispersal processes. Wind forcing was from the north or from the south, both with a constant wind speed of 10 km h^{-1}. Simulations were repeated 50 times with an input of 250 seeds per year during initialisation.

Second, we evaluated the functionality of the seed dispersal function by forcing the model with wind data from the reanalysis data set ERA-Interim (Balsamo et al., 2015). Simulations were repeated 10 times and population growth initiated by introducing 100 seeds per year during initialisation.

The dispersal distance and direction as well as the release height of every 100th seed dispersal event were recorded during each year of the complete simulation run. Pollination was assessed by recording the distance and direction from the pollen sources to the seed positions prior to seed release of all tree individuals present at the final year of the simulation.

Finally, we parallelised the code of the model using the OMP-library and implemented simulations using 1, 2, 4 and 8 CPUs. The performance of the model was evaluated by recording the computation time of each single simulation year for complete simulation runs (1,080 years). We conducted four different runs, one with only wind dispersal of seeds (SEED), one with seed and pollination (+POLL), and two different parallelised pollination computations (+POLL_PAR-A and B).

2.3.2. Verification of wind-dependency

We verified the wind-dependent seed dispersal by comparing the simulation results with the simulated results from the implemented dispersal function and refer to these values as ‘expected’. Therefore, we estimated the dispersal distances with the implemented dispersal functions (seeds Eq.1 and pollen Eq.5) under the same conditions as used for driving the model, namely winds from north or south with a constant speed of 10 km h^{-1}. The observed mean log-transformed (to achieve normal distribution) dispersal distances were compared to each other with a Welch two-sample t-test. We compared simulated to expected dispersal distances in each wind direction in height classes of trees in 10% steps of the sorted values, excluding the minimum and maximum values where only few observations could be recorded. The simulated data values in the range of ±5% at the determined height classes were first tested for an interaction between the release height and the dispersal distance with a Spearman’s rank-based measure of association (rho) and a t-test. Then we compared the simulated distances for each height class to the expected distances, both log-transformed prior to analysis, with a t-test. Furthermore, we compared the dispersal in both directions to each other using the same procedure.

The simulated pollination distances were compared to estimates from the implemented function under the same conditions, namely the adult trees’ position and constant north or south winds with a speed of 10 km h^{-1}. The resulting three-dimensional pollination distribution probabilities at simulation year 2011 (first year of field work) were binned to cardinal directions in steps of 30°, starting with North between 345° and 15°. The simulated and directly estimated log-transformed pollination distances were tested for differences in each individual direction with a Welch two-sample t-test for cases with >10 observations. We tested whether the mean of the p-values is significantly greater than 0.01 with a Student t-test.

2.3.3. Evaluation of dispersal processes

The seed dispersal and pollination distances were evaluated by calculating the mean dispersal distances for both processes in bins of 30° cardinal directions over the complete simulation run for years with available wind data (1979–2012). In addition, we selected two years with contrasting wind patterns, the year 1990 with predominant winds from the East, and 1998 from the West. The results were qualitatively compared to observed values from the literature.

During the simulations, we recorded only one pollination event per tree per year reaching the central area of 20 x 20 m of the plot and every 100th seed dispersal event.

2.3.4. Model performance experiments

The memory load was estimated by adding up the size of all data types within each handled structure simulating a plot of one hectare (Table S1). These were multiplied by the actual number of elements in each of the structures. We calculated mean values of the number of handled items of the final 80 years of the simulations for the evaluation of dispersal processes to estimate the total memory needed for the arrays of trees and seeds and the grid representing the environment (Kruse et al., 2016).

The computation times increases with the actual number of trees and seeds present in simulations. We analysed the dependency between the time needed for each simulated year and the number of trees and, additionally, the number of seeds produced by generalised nonparametric regression (using the "gam"-function in R-package "gam"; Hastie, 2017). The dependent variable time t was log-transformed prior to analysis. The explanatory variables – number of trees Nt and seeds Ns – were non-parametrically fitted and tested for non-linearity by comparing the deviance of a model that fits the terms linearly with a chi-squared test. In the initial model formula, we also included the interaction between the explanatory variables and excluded non-significant terms from the linear model (p>0.05) until yielding the final best model. The code for estimating pollination probabilities, seed dispersal and tree density computation was parallelised using the OMP-library for C++ and we compared the reduction in computation time of simulations with 1, 4 and 8 CPUs.

3. Results

3.1. Verification of wind-dependency

The simulated seeds were solely dispersed in a north or south direction in coherence to the forcing winds (Fig. 2). The median seed dispersal distances were ~12.2 m with a north wind and ~12.0 m with a south wind with a majority of 95% falling within ~43 m of the seed tree, but with rare (~0.1%) dispersal events >1,000 m (Fig. 2). The distance is equally highly correlated with the release height for both wind directions (*rho*=0.63, p<0.0001; Fig. 2).

All simulated values were highly correlated to the directly estimated distances (p>0.05, Table S2; Fig. 2). The mean log-transformed simulated distances in a north or south direction were not different when comparing releasing trees of the same height (p>0.05, Table S3).

The pollination events were mainly coming from the direction of the forcing winds: however, ~18% deviated from the forcing wind direction (Table 2). This variance is introduced by the formulae used for calculating the pollination probability for each tree and is further increased by the random (uniform) selection of a father out of a subset of all possible mature trees. The median distance along forcing winds of ~38 m is, in general, shorter by ~3-5 m than in other directions (Table 2).

The distances, simulated and expected from direct calculation, are more similar for those pollination directions with more samples, and only 4-6% of tests show significant differences. The higher values may be based on the uniform sampling of fathers and the associated smaller sample sizes in these directions. In all comparisons, we find no evidence that the differences are significant (Table S4).

3.2. Evaluation of dispersal processes

In north-central Siberia, the main wind directions observed during the vegetation period are a combination of both west and east (Fig. 3, upper row). In some years, one of these directions predominates, and is also characterised by stronger wind events. Accordingly, simulated seeds are dispersed into the general direction of the forcing wind data (Fig. 3, middle row). Dispersal distances can reach up to a maximum of several thousand kilometres, yet the majority of seeds fall within a few hundreds of metres, and these are dispersed over distances depicting the wind speeds as well.

The median pollen flight distances are generally larger than the seed‘s, with a technically fixed maximum of about the distance from the central plot to the borders (Fig. 3, lower row). Similar to seed dispersal, pollination follows the wind directions and fathers are positioned in the upwind direction of the main occurring winds.

3.3. Model performance

3.3.1. Memory consumption

The dynamic arrays need 120 bytes for each tree and 98 bytes for each seed. A further 54 bytes are needed for each of the environmental map tiles and another 117 bytes for the storage of output variables for each simulated year (Table S1). The constant containers use 390 bytes for the weather list and the parameter structures contain 642 bytes. On the basis of a simulated typical dense forest with ~92,000 seeds and ~25,000 tree individuals stored in the structures for each hectare, a simulation will need roughly ~15 MB of RAM in a setup of a 1,000 year initialising phase and a subsequent 80 year simulation phase.

3.3.2. Computation time

The simulation time increased with the number of trees in the simulation and for the contrasting simulation setups – either only wind-dependent seed dispersal SEED or also with the calculation of pollination +POLL (Fig. 4). The generalised additive model, including the number of seeds and a combination of the number of trees and number of seeds, explained the increase in computation time best and had the lowest AIC value among all simulation types (Table S5). All incorporated variables, namely number of trees and number of seeds, significantly explained the computation time. The number of trees is the most important explanatory variable at ~79.3%, followed by an interaction term of the number of trees and seeds at ~14.4%, and number of seeds at ~4.5% and a residual of ~1.7% unexplained variation.

Without inferring the pollination events, the computation takes ~0.6 s to calculate a year of a simulated plot on which 30,000 to 40,000 tree individuals are present (Fig. 4). In contrast, this increases to ~118 s yr^{-1} for a similar stand when estimating the pollen donor for each produced seed. The first implemented parallelisation of the pollination process (+POLL_PAR A) shortened the computation time by roughly half to ~60 s yr^{-1} when using eight cores. The second variant (+POLL_PAR B) outruns the first, and reaches a similar computation time using only four cores.

4. Discussion

The assumption of unlimited seedbeds – allowing species in models to grow as soon as climate space allows them – causes high uncertainty in future predictions with dynamic global vegetation models (e.g. Midgley et al., 2007; Neilson et al., 2005; Sato and Ise, 2012). Implementing time-lagged responses in such models highlighted the need for a proper understanding and implementation of processes that limit species’ migrations (Snell, 2014; Snell and Cowling, 2015). To reveal and understand the underlying processes that cause time lags, we designed the model LAVESI that represents all life-cycle stages of larches in high detail from seeds to mature trees, producing seeds themselves, which are then distributed in the environment (Kruse et al., 2016). Here we describe the model enhancements to achieve, for the first time, a coupled implementation of wind-driven seed dispersal and pollination in the larch forest simulator LAVESI. It can therefore be used for a very detailed evaluation of intra-stand processes determining migration speeds and help to improve abstract dynamic global vegetation models (e.g. Sato et al., 2007; Sitch et al., 2003) or regional forest gap models (e.g. Brazhnik and Shugart, 2015, 2016). Such a detailed representation of forest stands, as in the model presented here, is unlikely to be able to simulate forest dynamics on a continental to global scale (cf. Neilson et al., 2005). Nonetheless, the model can be used to parameterise dispersal kernels constraining inter-grid cell migration in DGVMs (Snell, 2014; Snell and Cowling, 2015) to achieve a better representation of processes constraining or enhancing the spread of a plant species.

We first tested the newly implemented functions for seed dispersal (Section 4.1) and pollination (4.2) by forcing simulations with one directional wind of constant speed. Then, the model was rerun with wind velocity profiles from the ERA-Interim data set (Balsamo et al., 2015) to test its behaviour under quasi-real conditions. In addition, we evaluated the model performance with the new processes (Section 4.3).

4.1. Wind-dependent seed dispersal

The simulated seed dispersal strictly followed the wind forcing and seeds settled in a downwind direction as expected, and not, as in the original model, in a purely ballistic manner (Kruse et al., 2016). Seed dispersal is dependent on the release height of the seed (Matlack, 1987), which is low in the focus region (Wieczorek et al., 2017) and thus leads to low dispersal distances, compared to other taxa (González-Martínez et al., 2002, *Pinus*, 2006; *Picea*, Piotti et al., 2009). When winds are constantly blowing at 2.78 m s^{-1} (10 km h^{-1}), simulated distances seldom reach more than ~43 m and only on very rare occasions are they observed with distances exceeding 1,000 m. The dispersal kernel can thus be described as a combination of a Gaussian distribution, with its centre ~12 m metres from the releasing tree, and a long tail, best described by an exponential function. This aligns well with the implemented function in the model LAVESI (see details in Kruse et al., 2016). The model results are similar when driven with quasi-real wind data from the reanalysis data set ERA-Interim. The short seed dispersal distances depict well the generally observed values of other larch species. For example, Duncan 1954) found for *Larix laricina* in the northern USA that 94% of seeds fell within 18 m of the releasing trees. Furthermore, Pluess (2011) found in dense forests of *Larix decidua* in the Swiss Alps an effective seed dispersal distance of 2–48 m. Moreover, the directions are now more realistically represented and follow the predominant west and east winds as expected (Fig. 3).

The use of winds from only the vegetation period might have introduced a bias, but it is based on the observation that this is the time when seeds are primarily dispersed (Abaimov, 2010). However, secondary dispersal by winds, due to uplift in strong winds, or travel in winter on frozen surfaces over long distances (Nathan et al., 2011a; cf. Pluess, 2011), or due to wind-independent animal-mediated zoochory (Evstigneev et al., 2017), is currently not represented but could facilitate the migration into tundra further. Also the wind regimes could have shifted from the past to the current setting and might even change in the future (Lisitzin, 2012; Trenberth, 1990). This would further slowdown the recent potential migration rate. This could explain the slow response of the treeline in northern Siberia to global warming, in addition to the long life-cycle of larches, as well as prevailing seed limitation in the north (Kruse et al., 2016; Wieczorek et al., 2017).

4.2. Pollination coupled to prevailing wind conditions

The density-dependent probability function assigns pollen donors mostly in an upwind direction, but also has a small angular scattering, which was introduced by the use of the von Mises distribution to capture the stochasticity of this process (Gregory, 1961; Kuparinen et al., 2007). This uncertainty could lead to an overestimation of distances, but seems unlikely because simulated pollen are travelling distances of ~38 m, which is longer compared to seeds and which is in concordance with observations (e.g. Pluess, 2011). The pollination distance often reaches the maximal possible distance between two trees in our simulation setup, which was the diagonal of the simulated area. These hypothetical pollen grains could probably reach distances of several hundred metres to kilometres, which would be in the range of the general dispersal distance observed in larches (Dow and Ashley, 1996; Hall, 1986). However, the current version of LAVESI is not yet parameterised to field data, but the next important step would be to validate the modelled pollination process with field-based data, and finally, to apply this model to achieve realistic predictions of a future treeline.

4.3. Model performance

The individual-based approach of the model LAVESI-WIND, with the extension of wind-dependent seed dispersal and pollination, bears a high potential of knowledge gain, but this comes with some challenges: (1) repeated calculations for millions of individuals (seeds and trees) are computationally intense, and (2) they require a certain amount of memory during the simulation runs. Whereas the memory during each simulation run could be minimised to the needs of the simulation setup, the computational power was historically the limiting factor (Grimm and Railsback, 2005). But with the development of recent computer clusters with hundreds of CPUs, it seems very likely that one can overcome this, allowing us to use detailed and spatially explicit models at a regional scale (e.g. Paik et al., 2006; Zhao et al., 2013).

4.3.1. Memory consumption

We estimated the requirements for a hectare of a dense simulated forest as 15 MB of RAM. This means, on typical computer servers, even broad-scale simulation runs are easily feasible for 5,000 x 5,000 m, which would need ~38 GB RAM. The current LAVESI version was not fully optimised to lower the needs of memory and many variables that might not be needed for a specialised simulation experiment could be excluded. Although, the original simulation program was not intended to be run over continuous square kilometres of forests (Kruse et al., 2016), this is already possible with the current version. The programming language C++ and the process-based structure of the code support an easy and fast forward development of this model.

4.3.2. Computation time needed for millions of trees, seeds and pollination

The computational effort of pollination for each seed increases with the number of mature trees present on a simulated plot. Therefore, to allow simulations to be run on standard computers in manageable time, it was a major goal to minimise the time needed for each simulated year. To meet this requirement, we parallelised parts of the program code that are computationally intensive, namely the processes of pollination and seed dispersal. With our approach, we have been able to decrease the time so far by a factor of six when using 8 CPUs, in comparison to using only one. Still, overheads from using a standard template library (STL)-list container lead to a negative exponential progression of the computation time needed per year rather than linear improvements (Fig. 4). Additional gains for other not yet parallelised processes are much smaller than these, but there is further potential to reduce the computation time by using different implementations of the parallelisation.

4.4. Potential model applications

The new model version LAVESI-WIND allows for the evaluation of the importance of driving processes, which determine the response speed of tree stands growing at the treeline in Siberia. With this, we can approach novel research questions, such as "Do wind regime shifts explain faster or slower migration rates in past climate changes?" Furthermore, one could test how different treeline types determine the migration behaviour in changing environments. These can vary widely, based on the treeline type, being abrupt or with stand densities decreasing with the abiotic gradient. In addition, this may be influenced by single-tree stands growing ahead of the migration front (Holtmeier and Broll, 2005). Further interesting questions could address the role of refugia during past glacial periods and their influence on present-day tundra colonisation by trees (Wagner et al., 2015).

Before applying this new model version, however, a proper parameterisation is necessary. Molecular methods can help to improve the seed dispersal function, especially microsatellite markers, which can uncover connections among subpopulations and even kinships by parentage analyses at the stand level, which would make the effective seed dispersal distances directly inferable (Ashley, 2010; Dow and Ashley, 1996; Piotti et al., 2009; Pluess, 2011). Additionally, these methods can be used to estimate the fat tail of the dispersal function indirectly (Piotti et al., 2009).

The dispersal potential is influenced by species' traits, which are under selection and thus variable in time (Ackerly, 2003; Ronce, 2007). Trait variants that support a genetic lineage to disperse into new colonisable habitats can be superior when climate is warming and new habitats are available, and, *vice versa*, under colder climates, variants with a shorter dispersal distance potential that remain in refugia are superior. With this new model, the influence of this trait variability could be systematically assessed, because it is able to transmit maternal genetic information by seed dispersal and paternal genetic information via pollination.

Another interesting application would be to use this model to estimate the pollen influx in lakes (cf. Sugita, 2007). Pollen influx rates are widely used for vegetation reconstructions at the tundra-taiga transition zone (e.g. Klemm et al., 2016) and could now be used either to tune the dispersal parameters for a more precise population dynamics prediction, or inversely, to reconstruct ancient tree stands by simulations. Before the genetics or the influx rates are included in the model, however, a revision of boundary conditions for pollen in the model is necessary. This must include a relevant source area for the pollen (cf. Sugita, 2007) to determine to what extent genetic traits are delivered by pollen from beyond the borders of the simulated area. If this can be efficiently parameterised, the model could further be used to track genetic lineages in time.

Finally, by connecting the borders of a simulation plot along the meridional borders we already implemented boundary conditions that allow the simulation of south-to-north transects, which are representative of the treeline area where highest tree densities occur in the south and treeless areas in the north. Thus, with this model, past migration corridors and timings can be revealed by a landscape-scale simulation, potentially answering important questions of the past biogeography of larch species in Siberia.

5. Conclusions

We conclude that it is feasible to implement wind-driven seed dispersal and pollination in an individual-based model, which is then able to run across broader areas. However, the simulated area and duration of the simulation are constrained by available computer power and memory, and thus further effort is needed to minimise the computational load of this model in order to allow landscape-scale simulations on a standard computer. With the new model setup, further applications in combination with the genetics of the represented species are now feasible and can bring us detailed knowledge about the behaviour of the treeline and the biogeography of larch species through time.

6. Code availability

The source code of the host model is available at GitHub https://github.com/StefanKruse/LAVESI/releases/tag/v1.01, and stored in the zenodo database http://doi.org/10.5281/zenodo.1155486. The updated version presented here is named LAVESI-WIND and the first version 1.0 is accessible at GitHub at https://github.com/StefanKruse/LAVESI/tree/v1.0 and stored at http://doi.org/10.5281/zenodo.1165383.

7. Author contributions

S.K. and A.G. planned the study, N.K., A.G. and S.K. updated the model and implemented new functions. A.G. and S.K. performed the simulations and the statistical analysis. S.K. and A.G. wrote the manuscript. U.H. provided substantial advice in the process of data analysis and paper writing.

8. Competing interests

The authors declare no competing interests.

9. Acknowledgements

We acknowledge Sven Willner for valuable advice in the process of parallelising the program code and Cathy Jenks for proofreading and improving the manuscript. The position of Stefan Kruse is funded by the Helmholtz Initiative Fund.

10. References

Abaimov, A. P.: Geographical distribution and genetics of Siberian larch species, in Permafrost Ecosystems - Siberian Larch Forests, vol. 209, edited by A. Osawa, O. A. Zyryanova, Y. Matsuura, T. Kajimoto, and R. W. Wein, pp. 41–58, Springer, Netherlands, Dordrecht., 2010.

Abramowitz, M. and Stegun, I. A.: Handbook of mathematical functions: with formulas, graphs, and mathematical tables, Dover Books on Mathematics. Dover Publications., 2012.

Ackerly, D. D.: Community assembly, niche conservatism, and adaptive evolution in changing environments, Int. J. Plant Sci., 164(S3), S165–S184, doi:10.1086/368401, 2003.

Ashley, M. V.: Plant parentage, pollination, and pispersal: How DNA microsatellites have altered the landscape, CRC. Crit. Rev. Plant Sci., 29(3), 148–161, doi:10.1080/07352689.2010.481167, 2010.

Austerlitz, F., Jung-Muller, B., Godelle, B. and Gouyon, P.-H.: Evolution of coalescence times, genetic diversity and structure during colonization, Theor. Popul. Biol., 51(2), 148–164, doi:10.1006/tpbi.1997.1302, 1997.

Bader, J.: The origin of regional Arctic warming, Nature, 509(7501), 167–168, doi:10.1038/509167a, 2014.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Munõz-Sabater, J., Pappenberger, F., De Rosnay, P., Stockdale, T. and Vitart, F.: ERA-Interim/Land: A global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19(1), 389–407, doi:10.5194/hess-19-389-2015, 2015.

Ban, Y., Xu, H., Bergeron, Y. and Kneeshaw, D. D.: Gap regeneration of shade-intolerant Larix gmelini in old-growth boreal forests of northeastern China, J. Veg. Sci., 9(4), 529–536, doi:10.2307/3237268, 1998.

Bonan, G. B.: Forests and climate change: forcings, feedbacks, and the climate benefits of forests, Science (80-.)., 320(5882), 1444–1449, doi:10.1126/science.1155121, 2008.

Brazhnik, K. and Shugart, H. H.: 3D simulation of boreal forests: structure and dynamics in complex terrain and in a changing climate, Environ. Res. Lett., 10(10), 105006, doi:10.1088/1748-9326/10/10/105006, 2015.

Brazhnik, K. and Shugart, H. H.: SIBBORK: A new spatially-explicit gap model for boreal forest, Ecol. Model., 320, 182–196, doi:10.1016/j.ecolmodel.2015.09.016, 2016.

Burczyk, J., DiFazio, S. P. and Adams, W. T.: Gene flow in forest trees: How far do genes really travel?, For. Genet., 11(3–4), 179–192, doi:10.1016/j.foreco.2004.05.003, 2004.

Clark, J. S., Silman, M., Kern, R., Macklin, E., HilleRisLambers, J., Ecology, S. and Jul, N.: Seed dispersal near and far: Patterns across temperate and tropical forests, Ecology, 80(5), 1475–1494, doi:10.1890/0012-9658(1999)080[1475:SDNAFP]2.0.CO;2, 1999.

Clark, J. S. S.: Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord., Am. Nat., 152(2), 204–224, doi:10.1086/286162, 1998.

Dow, B. D. and Ashley, M. V.: Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*, Mol. Ecol., 5(5), 615–627, doi:10.1111/j.1365-294X.1996.tb00357.x, 1996.

Duncan, D. P.: A study of some of the factors affecting the natural regeneration of tamarack (*Larix laricina*) in Minnesota, Ecology, 35(4), 498–521, doi:10.2307/1931040, 1954.

Eisenhut, G.: Untersuchungen über die Morphologie und Ökologie der Pollenkörner heimischer und fremdländischer Waldbäume, P. Parey., 1961.

Evstigneev, O. I., Korotkov, V. N., Murashev, I. A. and Voevodin, P. V.: Zoochory and peculiarities of forest community formation: A review, Russ. J. Ecosyst. Ecol., 2(1), 1–16, doi:10.21685/2500-0578-2017-1-2, 2017.

Fayard, J., Klein, E. K. and Lefèvre, F.: Long distance dispersal and the fate of a gene from the colonization front, J. Evol. Biol., 22(11), 2171–2182, doi:10.1111/j.1420-9101.2009.01832.x, 2009.

Frost, G. V, Epstein, H. E. and Walker, D. A.: Regional and landscape-scale variability of Landsat-observed vegetation dynamics in northwest Siberian tundra, Environ. Res. Lett., 9(2), 1–11, doi:10.1088/1748-9326/9/2/025004, 2014.

Frost, G. V. and Epstein, H. E.: Tall shrub and tree expansion in Siberian tundra ecotones since the 1960s, Glob. Chang. Biol., 20(4), 1264–1277, doi:10.1111/gcb.12406, 2014.

Fyllas, N. M., Politi, P. I., Galanidis, A., Dimitrakopoulos, P. G. and Arianoutsou, M.: Simulating regeneration and vegetation dynamics in Mediterranean coniferous forests, Ecol. Model., 221(11), 1494–1504, doi:10.1016/j.ecolmodel.2010.03.003, 2010.

González-Martínez, S. C., Gerber, S., Cervera, M. T., Martínez-Zapater, J. M., Gil, L. and Alía, R.: Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers, Theor. Appl. Genet., 104(8), 1290–1297, doi:10.1007/s00122-002-0894-4, 2002.

González-Martínez, S. C., Burczyk, J., Nathan, R., Nanos, N., Gil, L. and Alía, R.: Effective gene dispersal and female reproductive success in Mediterranean maritime pine (*Pinus pinaster* Aiton), Mol. Ecol., 15(14), 4577–4588, doi:10.1111/j.1365-294X.2006.03118.x, 2006.

Gregory, P. H.: The microbiology of the atmosphere., London: Leonard Hill. New York: Interscience Publishers., 1961.

Grimm, V. and Railsback, S. F.: Individual-based Modeling and Ecology, Princeton University Press, Princeton, NJ, USA., 2005.

Hall, J. P.: Growth and development of larch in Newfoundland, in Proceedings of the Sixth International Workshop on Forest Regeneration, pp. 21–26, U.S. Department of Agriculture, Forest Service, General Technical Reports., 1986.

Harsch, M. A., Hulme, P. E., McGlone, M. S. and Duncan, R. P.: Are treelines advancing? A global meta-analysis of treeline response to climate warming, Ecol. Lett., 12(10), 1040–1049, doi:10.1111/j.1461-0248.2009.01355.x, 2009.

Hastie, T.: gam: Generalized Additive Models, R Package, version 1.14-3, R Found. Stat. Comput. Vienna, Austria [online] Available from: http://cran.r-project.org/package=gam, 2017.

Holtmeier, F.-K. and Broll, G.: Sensitivity and response of Northern Hemisphere altitudinal and polar treelines to environmental change at landscape and local scales, Glob. Ecol. Biogeogr., 14(5), 395–410, doi:10.1111/j.1466-822x.2005.00168.x, 2005.

Jackson, S. T. and Lyford, M. E.: Pollen dispersal models in Quaternary plant ecology: Assumptions, parameters, and prescriptions, Bot. Rev., 65(1), 39–75, doi:10.1007/BF02856557, 1999.

Kaplan, J. O. and New, M.: Arctic climate change with a 2 °C global warming: Timing, climate patterns and vegetation change, Clim. Change, 79(3–4), 213–241, doi:10.1007/s10584-006-9113-7, 2006.

Kharuk, V. I., Ranson, K. J., Im, S. T. and Naurzbaev, M. M.: Forest-tundra larch forests and climatic trends, Russ. J. Ecol., 37(5), 291–298, doi:10.1134/S1067413606050018, 2006.

Klemm, J., Herzschuh, U. and Pestryakova, L. A.: Vegetation, climate and lake changes over the last 7000 years at the boreal treeline in north-central Siberia, Quat. Sci. Rev., 147(1–13), 422–434, doi:10.1016/j.quascirev.2015.08.015, 2016.

Kruse, S., Wieczorek, M., Jeltsch, F. and Herzschuh, U.: Treeline dynamics in Siberia under changing climates as inferred from an individual-based model for *Larix*, Ecol. Model., 338, 101–121, doi:10.1016/j.ecolmodel.2016.08.003, 2016.

Kuparinen, A., Markkanen, T., Riikonen, H. and Vesala, T.: Modeling air-mediated dispersal of spores, pollen and seeds in forested areas, Ecol. Model., 208(2–4), 177–188, doi:10.1016/j.ecolmodel.2007.05.023, 2007.

Levin, S. A., Muller-Landau, H. C., Nathan, R. and Chave, J.: The ecology and evolution of seed dispersal: A theoretical perspective, Annu. Rev. Ecol. Evol. Syst., 34(1), 575–604, doi:10.1146/annurev.ecolsys.34.011802.132428, 2003.

Lisitzin, A. P.: Sea-ice and iceberg sedimentation in the ocean: recent and past, Springer Science & Business Media., 2012.

Loarie, S. R., Duffy, P. B., Hamilton, H., Asner, G. P., Field, C. B. and Ackerly, D. D.: The velocity of climate change, Nature, 462(7276), 1052–1055, doi:10.1038/nature08649, 2009.

MacDonald, G. M., Kremenetski, K. V. and Beilman, D. W.: Climate change and the northern Russian treeline zone, Philos. Trans. R. Soc. B Biol. Sci., 363(1501), 2283–2299, doi:10.1098/rstb.2007.2200, 2008.

Martínez, I., Wiegand, T., Camarero, J. J., Batllori, E. and Gutiérrez, E.: Disentangling the formation of contrasting tree-line physiognomies combining model selection and Bayesian parameterization for simulation models., Am. Nat., 177(5), E136–E152, doi:10.1086/659623, 2011.

Matlack, G. R.: Diaspore size, shape, and fall behavior in wind-dispersed plant species, Am. J. Bot., 74(8), 1150–1160, doi:10.2307/2444151, 1987.

Midgley, G. F., Thuiller, W. and Higgins, S. I.: Plant Species Migration as a Key Uncertainty in Predicting Future Impacts of Climate Change on Ecosystems: Progress and Challenges, in Terrestrial Ecosystems in a Changing World, edited by J. G. Canadell, D. E. Pataki, and L. F. Pitelka, pp. 129–137, Springer Berlin Heidelberg, Berlin, Heidelberg., 2007.

Mitchell, T. D., Carter, T. R., Jones, P. D., Hulme, M. and New, M.: A comprehensive set of high-resolution grids of monthly climate for Europe and the globe : the observed record (1901-2000) and 16 scenarios (2001-2100)., , (July), 2004.

Montesano, P. M., Sun, G., Dubayah, R. O. and Ranson, K. J.: Spaceborne potential for examining taiga-tundra ecotone form and vulnerability, Biogeosciences, 13(13), 3847–3861, doi:10.5194/bg-2015-575, 2016.

Moran, E. V. and Clark, J. S.: Between-site differences in the scale of dispersal and gene flow in red oak, PLoS One, 7(5), e36492, doi:10.1371/journal.pone.0036492, 2012.

Nathan, R. and Muller-Landau, H. C.: Spatial patterns of seed dispersal, their determinants and consequences for recruitment, Trends Ecol. Evol., 15(7), 278–285, doi:10.1016/S0169-5347(00)01874-7, 2000.

Nathan, R., Katul, G. G., Bohrer, G., Kuparinen, A., Soons, M. B., Thompson, S. E., Trakhtenbrot, A. and Horn, H. S.: Mechanistic models of seed dispersal by wind, Theor. Ecol., 4(2), 113–132, doi:10.1007/s12080-011-0115-3, 2011a.

Nathan, R., Horvitz, N., He, Y., Kuparinen, A., Schurr, F. M. and Katul, G. G.: Spread of North American wind-dispersed trees in future environments., Ecol. Lett., 14(3), 211–9, doi:10.1111/j.1461-0248.2010.01573.x, 2011b.

Navascués, M., Depaulis, F. and Emerson, B. C.: Combining contemporary and ancient DNA in population genetic and phylogeographical studies., Mol. Ecol. Resour., 10(5), 760–72, doi:10.1111/j.1755-0998.2010.02895.x, 2010.

Neilson, R. P., Pitelka, L. F., Solomon, A. M., Nathan, R., Midgley, G. F., Fragoso, J. M. V., Lischke, H. and Thompson, K.: Forecasting Regional to Global Plant Migration in Response to Climate Change, Bioscience, 55(9), 749, doi:10.1641/0006-3568(2005)055[0749:FRTGPM]2.0.CO;2, 2005.

Nishimura, M. and Setoguchi, H.: Homogeneous genetic structure and variation in tree architecture of *Larix kaempferi* along altitudinal gradients on Mt. Fuji, J. Plant Res., 124(2), 253–263, doi:10.1007/s10265-010-0370-1, 2011.

Pacala, S. W. and Deutschman, D. H.: Details that matter: The spatial distribution of individual trees maintains forest ecosystem function, Oikos, 74(3), 357, doi:10.2307/3545980, 1995.

Pacala, S. W., Canham, C. D., Saponara, J., John, A. S. J., Kobe, R. K. and Ribbens, E.: Forest Models Defined by Field Measurements : Estimation, Error Analysis and Dynamics, Ecol. Monogr., 66(1), 1–43, 1996.

Paik, S. H., Moon, J. J., Kim, S. J. and Lee, M.: Parallel performance of large scale impact simulations on Linux cluster super computer, Comput. Struct., 84(10–11), 732–741, doi:10.1016/j.compstruc.2005.11.013, 2006.

Petrovskii, S. and Morozov, A.: Dispersal in a statistically structured population: Fat tails revisited, Am. Nat., 173(2), 278–289, doi:10.1086/595755, 2009.

Piao, S., Friedlingstein, P., Ciais, P., Viovy, N. and Demarty, J.: Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades, Global Biogeochem. Cycles, 21(3), doi:10.1029/2006GB002888, 2007.

Piotti, A., Leonardi, S., Piovani, P., Scalfi, M. and Menozzi, P.: Spruce colonization at treeline: where do those seeds come from?, Heredity (Edinb)., 103(2), 136–45, doi:10.1038/hdy.2009.42, 2009.

Pluess, A. R.: Pursuing glacier retreat: genetic structure of a rapidly expanding *Larix decidua* population, Mol. Ecol., 20(3), 473–485, doi:10.1111/j.1365-294X.2010.04972.x, 2011.

Polezhaeva, M. A., Lascoux, M. and Semerikov, V. L.: Cytoplasmic DNA variation and biogeography of *Larix* Mill. in northeast Asia, Mol. Ecol., 19(6), 1239–1252, doi:10.1111/j.1365-294X.2010.04552.x, 2010.

Prentice, I. C.: Pollen representation, source area, and basin size: Toward a unified theory of pollen analysis, Quat. Res., 23(1), 76–86, doi:10.1016/0033-5894(85)90073-0, 1985.

Ray, N. and Excoffier, L.: A first step towards inferring levels of long-distance dispersal during past expansions, Mol. Ecol. Resour., 10(5), 902–914, doi:10.1111/j.1755-0998.2010.02881.x, 2010.

Roberts, D. R. and Hamann, A.: Climate refugia and migration requirements in complex landscapes, Ecography (Cop.)., 39(12), 1238–1246, doi:10.1111/ecog.01998, 2016.

Ronce, O.: How does it feel to be like a rolling stone? Ten questions about dispersal evolution, Annu. Rev. Ecol. Evol. Syst., 38(1), 231–253, doi:10.1146/annurev.ecolsys.38.091206.095611, 2007.

Sato, H. and Ise, T.: Effect of plant dynamic processes on African vegetation responses to climate change: Analysis using the spatially explicit individual-based dynamic global vegetation model (SEIB-DGVM), J. Geophys. Res. Biogeosciences, 117(G3), n/a-n/a, doi:10.1029/2012JG002056, 2012.

Sato, H., Itoh, A. and Kohyama, T.: SEIB–DGVM: A new dynamic global vegetation model using a spatially explicit individual-based approach, Ecol. Model., 200(3–4), 279–307, doi:10.1016/j.ecolmodel.2006.09.006, 2007.

Scheiter, S., Langan, L. and Higgins, S. I.: Next-generation dynamic global vegetation models: learning from community ecology, New Phytol., 198(3), 957–69, doi:10.1111/nph.12210, 2013.

Semerikov, V. L., Iroshnikov, A. I. and Lascoux, M.: Mitochondrial DNA variation pattern and postglacial history of the Siberian larch (*Larix sibirica* Ledeb.), Russ. J. Ecol., 38(3), 147–154, doi:10.1134/S1067413607030010, 2007.

Semerikov, V. L., Semerikova, S. A., Polezhaeva, M. A., Kosintsev, P. A. and Lascoux, M.: Southern montane populations did not contribute to the recolonization of West Siberian Plain by Siberian larch (*Larix sibirica*): a range-wide analysis of cytoplasmic markers, Mol. Ecol., 22(19), 4958–4971, doi:10.1111/mec.12433, 2013.

Shugart, H. H., Leemans, R. and Bonan, G. B.: A Systems Analysis of the Global Boreal Forest, edited by H. H. Shugart, R. Leemans, and G. B. Bonan, Cambridge University Press, Cambridge, UK., 1992.

Shuman, J. K., Shugart, H. H. and O'Halloran, T. L.: Sensitivity of Siberian larch forests to climate change, Glob. Chang. Biol., 17(7), 2370–2384, doi:10.1111/j.1365-2486.2011.02417.x, 2011.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Chang. Biol., 9(2), 161–185, doi:10.1046/j.1365-2486.2003.00569.x, 2003.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C. and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five dynamic global vegetation models (DGVMs), Glob. Chang. Biol., 14(9), 2015–2039, doi:10.1111/j.1365-2486.2008.01626.x, 2008.

Sjögren, P., Edwards, M. E., Gielly, L., Langdon, C. T., Croudace, I. W., Merkel, M. K. F., Fonville, T. and Alsos, I. G.: Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland, New Phytol., 213(2), 929–941, doi:10.1111/nph.14199, 2017.

Snell, R. S.: Simulating long-distance seed dispersal in a dynamic vegetation model, Glob. Ecol. Biogeogr., 23(1), 89–98, doi:10.1111/geb.12106, 2014.

Snell, R. S. and Cowling, S. A.: Consideration of dispersal processes and northern refugia can improve our understanding of past plant migration rates in North America, J. Biogeogr., 42(9), 1677–1688, doi:10.1111/jbi.12544, 2015.

Sugita, S.: Theory of quantitative reconstruction of vegetation I: Pollen from large sites REVEALS regional vegetation composition, Holocene, 17(2), 229–241, doi:10.1177/0959683607075837, 2007.

Sugita, S., Parshall, T., Calcote, R. and Walker, K.: Testing the Landscape Reconstruction Algorithm for spatially explicit reconstruction of vegetation in northern Michigan and Wisconsin, Quat. Res., 74(2), 289–300, doi:10.1016/j.yqres.2010.07.008, 2010.

Svenning, J.-C., 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. M. S., Pagel, J. and Normand, S.: The influence of interspecific interactions on species range expansion rates, Ecography (Cop.)., 37(12), 1198–1209, doi:10.1111/j.1600-0587.2013.00574.x, 2014.

Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G. F., Paterson, J., Schurr, F. M., Sykes, M. T. and Zimmermann, N. E.: Predicting global change impacts on plant species' distributions: Future challenges, Perspect. Plant Ecol. Evol. Syst., 9(3–4), 137–152, doi:10.1016/j.ppees.2007.09.004, 2008.

Trenberth, K. E.: Recent observed interdecadal climate changes in the Northern Hemisphere, Bull. Am. Meteorol. Soc., 71(7), 988–993, doi:10.1175/1520-0477(1990)071<0988:ROICCI>2.0.CO;2, 1990.

Wagner, S., Litt, T., Sánchez-Goñi, M. F. and Petit, R. J. R. J.: History of Larix decidua Mill. (European larch) since 130 ka, Quat. Sci. Rev., 124, 224–247, doi:10.1016/j.quascirev.2015.07.002, 2015.

Wieczorek, M., Kruse, S., Epp, L. S., Kolmogorov, A., Nikolaev, A. N., Heinrich, I., Jeltsch, F., Pestryakova, L. A., Zibulski, R. and Herzschuh, U.: Dissimilar responses of larch stands in northern Siberia to increasing temperatures-a field and simulation based study, Ecology, 98(9), 2343–2355, doi:10.1002/ecy.1887, 2017.

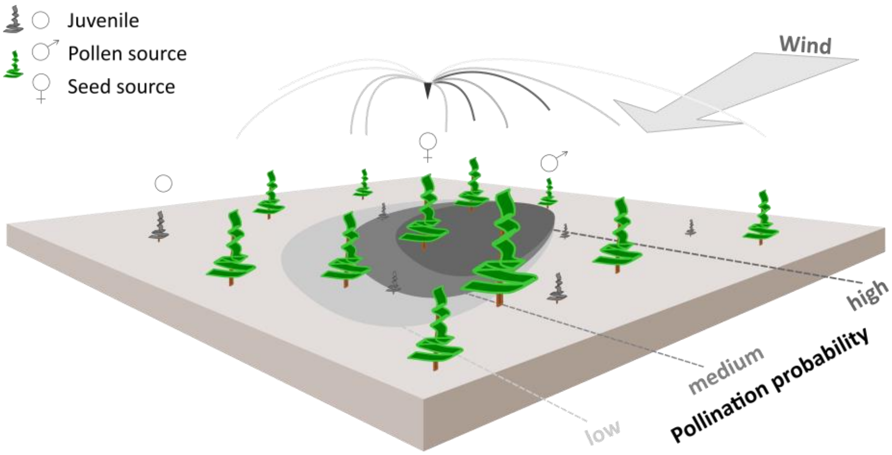
Yu, Q., Epstein, H. and Walker, D.: Simulating the effects of soil organic nitrogen and grazing on arctic tundra vegetation dynamics on the Yamal Peninsula, Russia, Environ. Res. Lett., 4(4), 45027, doi:10.1088/1748-9326/4/4/045027, 2009.

Zhang, J., Zhou, Y., Zhou, G. and Xiao, C.: Structure and composition of natural Gmelin larch (*Larix gmelinii* var. *gmelinii*) forests in response to spatial climatic changes., PLoS One, 8(6), e66668, doi:10.1371/journal.pone.0066668, 2013.

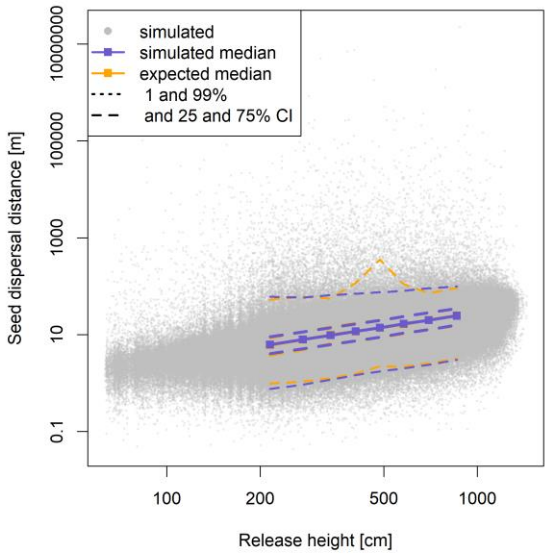
Zhang, N., Yasunari, T. and Ohta, T.: Dynamics of the larch taiga–permafrost coupled system in Siberia under climate change, Environ. Res. Lett., 6(2), 24003, doi:10.1088/1748-9326/6/2/024003, 2011.

Zhao, G., Bryan, B. A., King, D., Luo, Z., Wang, E., Bende-Michl, U., Song, X. and Yu, Q.: Large-scale, high-resolution agricultural systems modeling using a hybrid approach combining grid computing and parallel processing, Environ. Model. Softw., 41, 231–238, doi:10.1016/j.envsoft.2012.08.007, 2013.

11. Figures



620 **Figure 1. Schematic representation of wind pollination as newly implemented in the LAVESI model. Based on actual winds, a distance-dependent pollination probability is estimated for each adult tree (father) and for each seed source (mother) in the simulated area. The shaded areas on the ground represent the pollination probability for winds from the upper-right corner. These are generally higher for adult trees in upwind direction of the central seed source.**



625 **Figure 2. Dispersal distances of seeds are wind dependent and positively correlated with the height of the releasing tree. The simulated and hypothetically calculated dispersals were compared across evenly distributed height classes; the results are similar for north and south winds, and here the results with north winds are presented.**

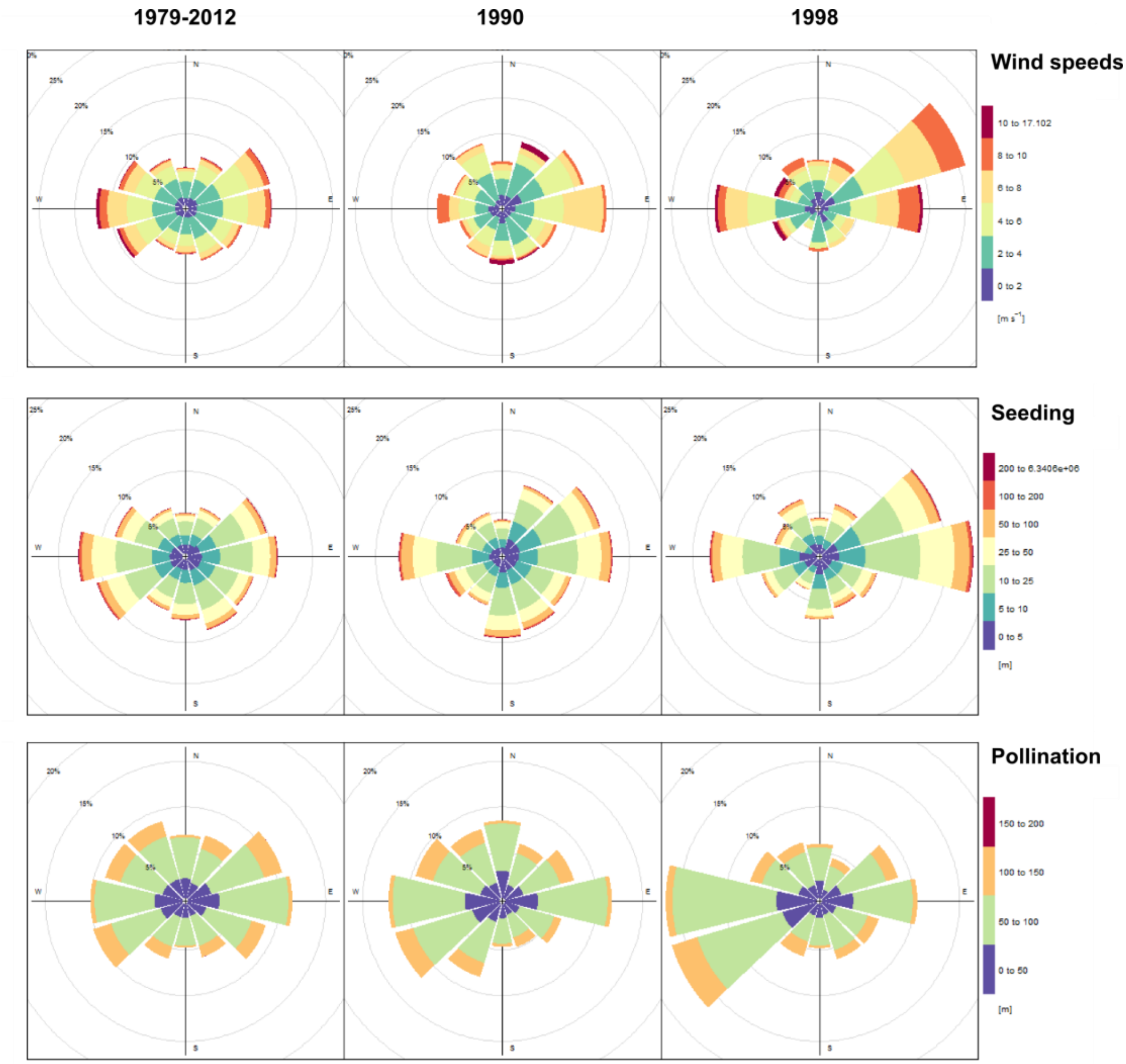


Figure 3. Wind forcing (upper row), simulated seed dispersal (middle row) and pollination distances (lowest row) by distance and cardinal direction.

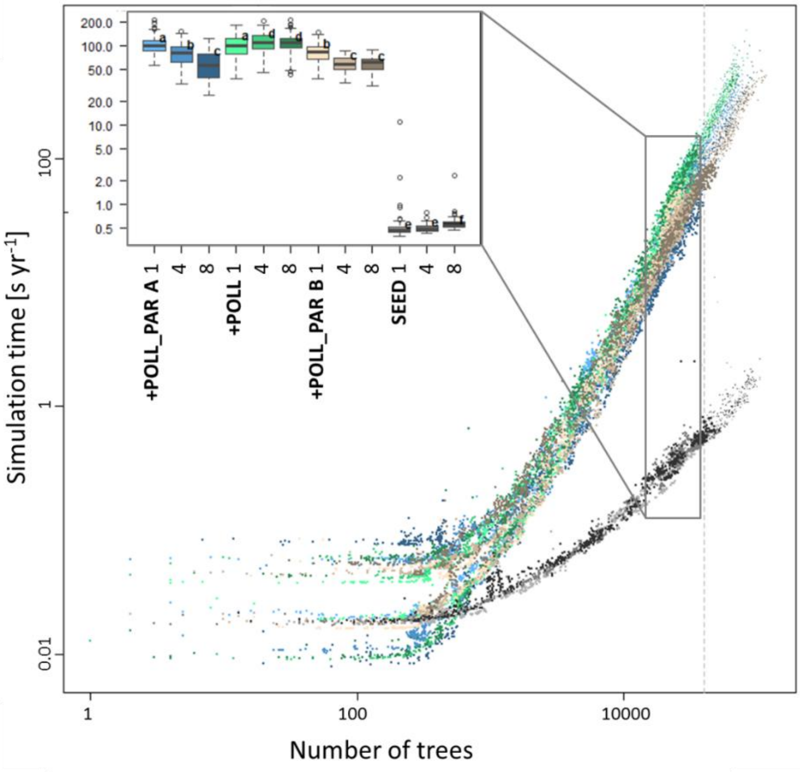


Figure 4. Simulation consumption time in relation to the number of trees present, the number of CPUs used and for different types of parallelisation of the code. The time increases exponentially with the number of trees and more quickly when simulating the additional pollination (+POLL) compared with just the explicit seed dispersal (SEED). The inset summarises the simulation time for simulated typical northern taiga stands, ranging between 30,000 and 40,000 trees. The letters next to the boxes indicate similar groups inferred with a Wilcoxon-test and Holm correction for multiple testing.

12. Tables

Table 1. Overview of model parameters and processes for *L. gmelinii* individuals that are different from the original version (Kruse et al., 2016).

Parameter	Value and dimension	References
Growth		
Quadratic term of the equation for diameter growth rate	-0.003 ln(cm)/cm²	data-based estimate similar to Fyllas *et al.* (2010)
Linear term of the growth function	0.030 ln(cm)/cm	
Constant term of the growth function	-1.98 ln(cm)	
Seed production, dispersal and establishment		
Factor of seed productivity	8	literature-based estimate (Kruklis & Milyutin, 1977, cited in Abaimov, 2010)
Background germination rate	0.01	estimated
Horizontal seed dispersal distance at wind speed of 10 km/h	60.1 m	estimated after Matlack (1987)
Seed descent rate	0.86 m/s	estimated descent rate based on Matlack (1987)
Mortality		
Background mortality rate	0.0001 yr^{-1}	data-based estimate
Current tree growth influence factor on tree mortality	0.0	estimated
Weather influence factor on tree mortality	0.1	estimated
Density influence factor on tree mortality	2.0	estimated
Seed fertility	2 years	Ban *et al.* (1998)
Mean temperature of the coldest month (January) at the border of the species‘ geographical range	-45 °C	Shugart *et al.* (1992)
Exponent scaling the height influence	0.2	estimated
Weather processing		
Exponent scaling the influence of surrounding density for a tree	0.1	estimated
Exponent scaling the density value	0.5	estimated

Table 2. Pollination event comparison between simulated and expected values. The statistics are based on 50 simulation repeats. Significant differences between expected and simulated distances tested with a Welch two-sample t-test.

Cardinal direction of winds	Direction of pollination [°]	Distance (simulated/expected) [m]	Fraction of observations (simulated/expected) [%]	Fraction of significant differences
North	135-165	44.32/43.79	9.00/8.39	20%
	165-195	39.03/38.90	81.53/82.68	4%
	195-225	40.03/40.19	9.47/8.93	14%
South	315-345	40.94/40.47	7.55/7.92	12%
	-15-15	36.98/37.90	82.92/82.8	6%
	15-45	42.76/40.84	9.51/9.28	22%
	45-75	*53.80/-*	0.01/0.00	*

* only one observation, thus excluded from further analyses.