Luberon2 user manual - A forest demo-genetic simulation software on Capsis

Claire Godineau, Victor Fririon, Nicolas Beudez, François de Coligny, F. Courbet, Gauthier Ligot, Sylvie Oddou-Muratorio, Leopoldo Sanchez, Francois Lefèvre

HAL Id: hal-04234752

https://hal.inrae.fr/hal-04234752v1

Submitted on 10 Oct 2023 (v1), last revised 16 Jan 2025 (v2)

Luberon2 user manual

a forest demo-genetic simulation software on Capsis

Version : Capsis 4.2.7-19850 [20 July 2023]

Luberon2 was developed by Claire Godineau[*1], Victor Fririon[*1], Nicolas Beudez[2], François de Coligny[2], François Courbet[1], Gauthier Ligot[3], Sylvie Oddou-Muratorio[1,4], Leopoldo Sanchez[5], François Lefèvre[1§]
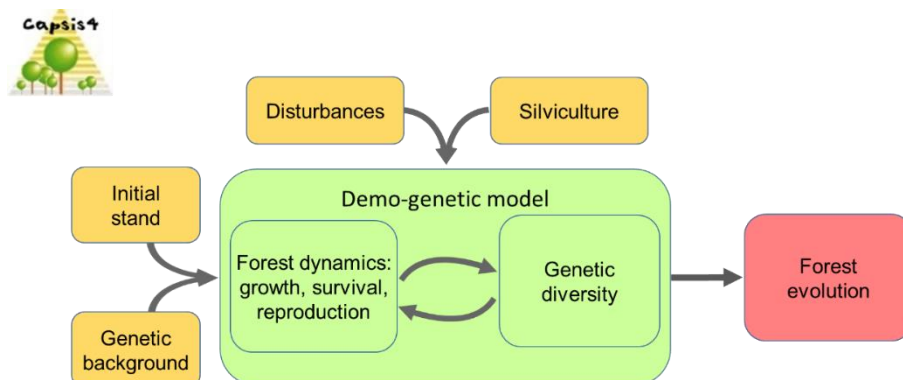
[*]co-first authors; [1]INRAE, URFM, France; [2]INRAE, AMAP, France; [3]KU Leuven, Gembloux Agro-Bio Tech, Belgium; [4]INRAE, ECOBIOP, France; [5]INRAE, BIOFORA, France; [§]contact: francois.lefevre.2@inrae.fr

Summary

Luberon2 is a forest dynamics model with heritable trait variation to simulate the joint effects of natural evolutionary processes including stochastic disturbance, thinning and cutting, and genetic diversity in monospecific stands: it is a demo-genetic agent-based model (Oddou-Muratorio et al, 2020; Lamarins et al, 2022). The current version runs either for cedar (*Cedrus atlantica*), Douglas fir (*Pseudotsuga menziesii*), Norway spruce (*Picea abies*), and larch (*Larix sp.*). Fir species (*Abies sp.*) are planned. Using a graphical interactive environment or a script mode, it can be used with various levels of expertise for communication, teaching, training or research purposes. The model is developed on the Capsis modeling platform[1], available upon request. This user manual contains a synthetic presentation of the model and more detailed documentation on the model. Some sections can be skipped depending on the level of use of the model: (1) simple simulations using pre-existing input files; (2) development of ad-hoc input files for personalized simulations; (3) advanced use of all parameters' flexibility; (4) extension of the model or re-use of parts of the code. Luberon2 is continuously evolving, the latest version of the user-manual and a quick tutorial to the graphical interface are available on the Capsis website[2].

---

[1] https://capsis.cirad.fr/

[2] https://capsis.cirad.fr/capsis/help_en/luberon2

Table of contents

I – Synthetic description of the model

About Capsis

Capsis (standing for "Computer-Aided Projections of Strategies In Silviculture") is a software platform to integrate forest growth and dynamics models for research, forest managers and training. It is a co-development oriented organization where modelers develop their own models and share common modeling tools, e.g. libraries and extensions (Figure 1). More than 80 projects are currently active on the platform, i.e. finalized models or under development, representing an international community of modelers. Models are basically written in Java, with possible links with other languages. The Capsis charter stipulates that all the source codes are freely accessible by all members in the Capsis community.
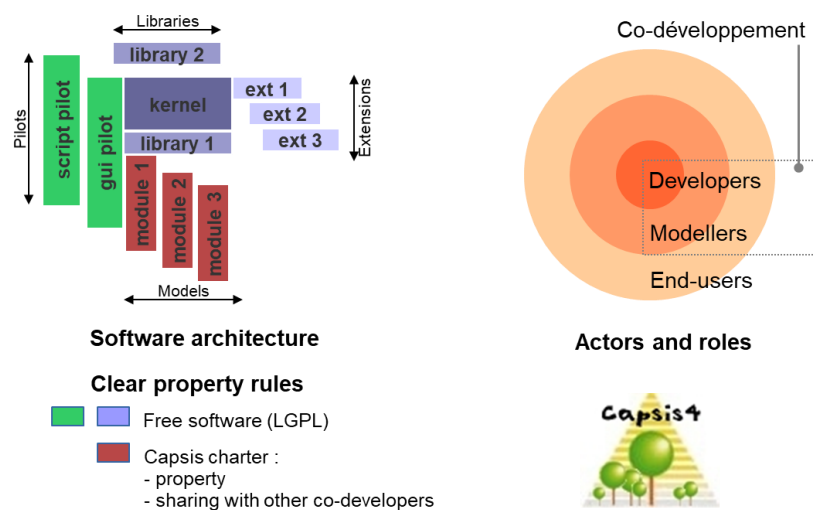
Figure 1: Co-development oriented organisation of the Capsis platform.

About Luberon2

Luberon2 is an individual-based forest dynamics model for even-aged (or even-aged per patch) monospecific stands with heritable trait variation and stochastic disturbance. Luberon2 integrates tree growth, reproduction and mortality processes, with two types of possible disturbance events affecting growth and mortality. Forest dynamics processes account for, and dynamically drive the quantitative genetic variation of three fitness-related traits: vigor and sensitivity to each type of disturbance[3]. The model includes stochasticity in all processes except tree growth. Diverse thinning interventions can be simulated. Graphical tools for visual analysis of simulation outputs are proposed.

The user has full control on the initial stand, on the succession of events occurring during an 'evolution' and on thinning interventions. Forest dynamics processes and evolutionary processes are driven by the model. The initial stand is loaded from an input file containing parcels characterized by their geometry and site index (top height at given age), trees with their individual genotypes, and genetic

---

[3] Beyond these three traits, the evolutionary consequences on other traits of interest can be approached through their genetic correlations with the model-driven fitness components, applying a classical quantitative genetics model of the response to indirect selection after Luberon2 simulations.

background information. The user defines each simulation scenario as a succession of temporal phases during which stochastic disturbance and regeneration are allowed or not. At each step, different types of thinning can be operated by the user. A large set of graphical tools are proposed to analyze the outputs and compare among different simulated scenarios, with regard to dendrometric and genetic variables. For more specific investigation, individual trees with their genetic and phenotypic attributes, or stand dynamics reports, can be exported at any step of each evolution.

Luberon2 integrates multiple processes: individual tree growth, sexual reproduction and natural regeneration, mortality through self-thinning as well as any types of random or selective thinning, diffuse mortality through stochastic disturbance regime which frequency and intensity can be configured, growth reduction through a second independent disturbance regime which can also reach mortality (e.g. drought events). Genetic diversity consist of individual genotypes at neutral genes and QTLs that control growth vigor and sensitivity to each disturbance factor following an additive inheritance model. Individual genotypes in the initial stand can be simulated using an heuristic algorithm to fit target values of populations and quantitative genetic parameters (heritability, differentiation, fixation index, etc.). The individual genotypes of new seedlings that appear during simulations directly result from sexual reproduction mechanisms that account for individual male and female fecundities as well as spatial distance between mates in the parent generation.

Used in a graphical interactive environment, the model provides graphical representations and tables of the evolution of dendrometric, demographic and genetic characteristics of the forest stand (Figure 2). Each graphical display has multiple customization options. The model can also be used in a script mode, with export of simulated data (i.e. individual tree characteristics or stand-level variables) for external analysis. A Capsis working 'session' may include several 'projects' to be compared: each project is determined by an initial scene to which several evolutions can be applied. Session and projects can be saved for later re-use. The software can be run in English or French.
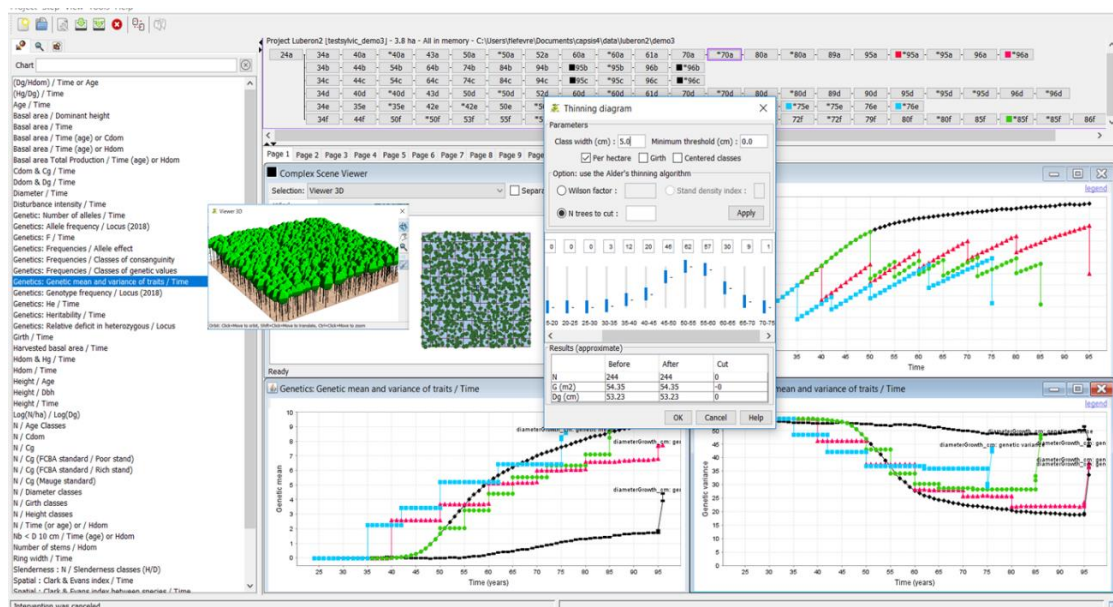
Figure 2*: screen-shot of Luberon2. Chrono-sequences at the top show the different scenarios of events (interventions, regeneration, with or without disturbances); left panel shows the choice of graphic representations; first front pop-up window is an interactive tool to proceed selective thinning based on the distribution of individual diameters; panels behind show 2D and 3D representations of the stand, graphical comparison of different thinning scenarios in terms of the evolutions of total basal area, genetic mean and variance of vigor during the simulation.*

Potential uses

Luberon2 can be used with different objectives:

- for basic research, to analyze the effects of interactions between silviculture, disturbance and initial genetic composition on stand dynamics, genetic quality and evolutionary potential;
- for applied research, to compare different thinning scenarios considering jointly classical forestry objectives, e.g. growth and production, with genetic diversity objectives, which not only include genetic improvement of vigor and resistance to disturbances but also preservation of the variance of adaptive traits and overall neutral genetic diversity;
- for training and teaching, to explain the pathways towards genetic impacts of silviculture practices;
- for communication, to raise awareness on the drivers of genetic evolutions in the forests.

We have used Luberon2 in practice sessions with French forest managers to explore the mechanisms through which thinning can drive genetic selection on vigor. Participants could realize that: (1) different thinning scenarios may result in different genetic evolutions; (2) early interventions at juvenile stage are important drivers of selection because the impact of tree selection in the final thinning is constrained by the previous effects of juvenile selection; and (3) genetic impacts of management practices may be modified by the regime of natural hazard (Lefèvre et al., 2019). We have also used Luberon2 in an international Training Workshop on methods in forest conservation genetics organized by the H2020 GENTREE project to which the partners of B4EST were also invited[4]. Most participants were geneticists rather than forest managers, and it was possible in a 2h time slot to introduce them to the model, install the software on their machines and perform their own simulations using the graphical interactive environment.

These experiences showed the possible use for teaching and training purposes, in both forest managers and geneticists' communities. The simulator runs on a PC, it only requires installation of the Capsis software and appropriate Java version. A stand-alone installer can be prepared allowing straightforward installation, e.g. for training sessions.

Targeted users, accessibility and use of the model

The model is mainly dedicated to R&D services of public and private forest enterprises familiar with the Capsis modeling platform: these users will have the capacity to use all thinning simulation tools offered by the platform and create new initial inventory files to design their own simulation studies without any specific support other than the user manual. For these users, the graphical mode is appropriate for quick simulations and direct visualization of simulation outputs with the graphical charts; it also allows exporting individual genetic data at each step in a standard format[5] for external genetic data analysis. The script mode is appropriate for more complex simulation plans, e.g. with replicates, and various types of simulation outputs are proposed. In addition to the direct use of the model, all Capsis modelers (i.e. the Capsis community) can freely re-use the source code, with respect to the Capsis charter.

---

[4] https://b4est.eu/training-workshop-on-methods-in-forest-conservation-genetics
[5] The format used by GenPop software

The graphical interactive environment is intuitive and allows simple uses by people not familiar with the software, based on default initial inventory files and default options, e.g. for training where participants perform their own simulations or just for a visual demo. Still, it is possible for these users to develop their own specific case studies: in this case, INRAE (*francois.lefevre.2@inrae.fr*) can provide support in particular for the creation of the appropriate initial inventory files.

The model is simple to use but attention must be given with the most expert user mode since many options can be changed: a smart choice might require information difficult to obtain, and irrelevant choices could lead to absurd simulation results. **In any case, the interpretation of model outputs is under the responsibility of the user.**

Simulations start with an initial stand inventory made of individual trees spatialized in one or several parcels characterized by their site index. Each tree has initial dendrometric characteristics and a genotype. The model makes these initial trees grow, reproduce and die, and it generates the genotypes of new seedlings. Thus, a first set of user choices concerns the initial stand: environmental conditions (parcels with their site indexes), individual genotypes that can be provided directly by the user or simulated using the automatic generator of genotypes given some synthetic parameters provided by the user (number of genes controlling vigor and resistance to disturbance, heritabilities of these traits, etc.). A second set of user choices concerns the parameters of stand evolution: with or without natural regeneration, with or without stochastic disturbances (and parameters of disturbance regime). Finally, the user has to define thinning interventions at each step (e.g. which trees to remove).

II – Documentation

Luberon2 is coded in Java. It is based on pre-existing growth models implemented in the Capsis platform, which were extended by the addition of: (i) individual spatialisation and regeneration processes, (ii) genetic diversity and quantitative inheritance of growth and mortality traits, (iii) stochastic drought-like growth disturbance events and diffuse mortality disturbance, (iv) new charts to visualize the simulation outputs. Three traits are currently under quantitative additive genetic control: vigor and sensitivity to each type of disturbance. Both disturbances are simulated through their "regime of impact": either an impact on growth, which can ultimately result in tree death if growth reduction is too severe, or a direct impact on mortality without affecting growth. Their regime is defined as the probability of occurrence of different intensities.

Figure 3 shows the succession of forest dynamic processes. Luberon2 builds on distance-independent tree model and simulates regeneration in even-aged (by patches) stand structure. The forest is split into independent growth units where demographic processes occur, except pollen and seed dispersal operating at global scale. Interventions can be designed globally over the whole forest or specifically for each parcel.
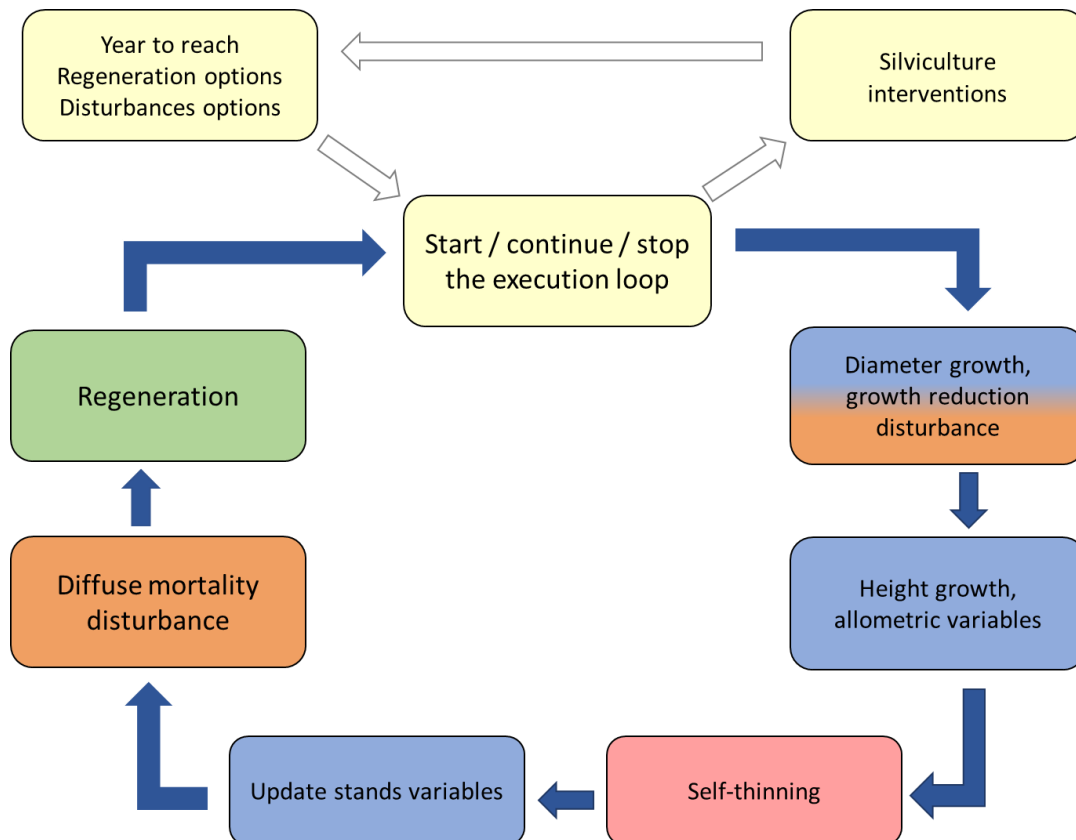
Figure 3: *Synthetic view of the execution loop of forest dynamic processes in Luberon2, yellow boxes represent the user choices, other colors are for dynamic processes (growth in blue, disturbance in orange, self-thinning in red and regeneration in green). Stochasticity occurs in all processes except in diameter growth for both species and in self-thinning for cedar.*

Main objects (Java classes)

In the Capsis platform, different types of common objects (Java classes) are defined that can be implemented more specifically by each model[6]. Here we briefly describe the main objects used in Luberon2, introducing the terminology illustrated in Figure 4.

Figure 4: Representation of a *C. atlantica* forest in SE France (Lacoste municipality, ~1km², on the Luberon mountain that gives its name to the model): _main parcels_ in the center of the _scene_ with their _pixel_ grid carrying _individual trees_, while _lateral parcels_ on each side are used to simulate external seed and pollen reservoirs through _multi-trees_ (underlined terms are described below).

Scene and plot

At each step of the simulation, the **scene** includes the lists of trees, fertile trees and individual female fecundities, while the **plot** includes the list of geographic elements, i.e. parcels and pixels (see below).

Main parcels and pixels

The simulated forest is composed of one or several **main parcels** representing spatial subunits with homogeneous environmental conditions and homogenous silviculture treatment where individual trees grow. Each main parcel is defined in the initial inventory by a spatial polygon and a site index (dominant height for a given age) and, during the evolution, thinning interventions can differ among parcels.

---

[6] For more information see the Capsis documentation https://capsis.cirad.fr/capsis/documentation

The even-aged growth and mortality models used in Luberon2 are not suitable for multi-cohort stand structure. Therefore, to achieve regeneration when seed trees and seedling co-exist in different parts of the forest, each main parcel is automatically split into a regular grid of square **pixels** with a pixel size given by the user: during the simulations, each pixel is modeled as an independent "growth unit" where dynamic processes occur with no impact of its neighbor pixels, except seed and pollen flow during the regeneration phase. In other terms, the main parcels are even-aged by pixel and growth dynamics variables (dominant height, etc.) are computed at pixel level.

Pixel size is a compromise defined by the user. On the one hand, partially irregular stand structure can be simulated with smaller pixels. On the other hand, splitting the parcel into independent pixels modifies the parcel-level predictions of growth and mortality all the more than pixel size is small. A default pixel size of 15x15m is proposed, which appears to be a good compromise with the *C. atlantica* growth model*: see Godineau, Fririon, et al. (2023) Appendix 2A.*

Individual trees, genotypes and genetic architecture of traits

Individual trees have fixed attributes like spatial coordinates, population of origin[7], genotype, parents' identity, consanguinity, etc. as well as dynamic attributes that evolve during simulation like age, dendrometric and fecundity variables, phenotypic value, etc. The initial individual trees are defined in the inventory given as input, then new individual trees are created by the model during the regeneration process.

The management of genetic information, i.e. effect of genotype on phenotype and hereditary transmission of nuclear and cytoplasmic genes, is piloted through the Genetics library of Capsis (Seynave & Pichot, 2004). Nuclear diploid individual genotypes are defined along a genetic map defined by the user, which contains multi-allelic neutral loci and diallelic quantitative trait loci (QTL) controlling three traits involved in stand dynamics processes: vigor, sensitivity to growth reduction disturbance and sensitivity to diffuse mortality disturbance (the following section explains how these traits are involved). Maternal and paternal cytoplasmic loci are multi-allelic.

Each diallelic QTL is characterized by an allelic effect, one allele providing the positive value of this effect and the other one the negative value. For each quantitative trait, the model assumes additive inheritance: the genotypic value of an individual is the sum of the effects of its alleles over the QTL controlling the trait and, therefore, a full heterozygous individual has a null genotypic value[8]. The additive genetic variance of each trait in the population is the sum of the contributions of its QTL (l):

$$VA = \sum_l (1 + F_l) 8 p_l q_l a_l^2$$

where F_l is the fixation index, p_l and q_l the allelic frequencies, and a_l the allelic effect of the QTL[9].

Individual phenotypic values are obtained by adding an individual environmental term made of two components: a fixed component and an "inter-step" component that randomly varies at each simulation step.

---

[7] The population of origin is only used to characterize the founder individuals in the initial inventory, the new individuals created during regeneration processes will have a "unspecified" population of origin

[8] The growth models were originally calibrated without considering genetic variation, which is retrieved in Luberon2 when all individual trees are clonal copies of a full heterozygous genotype at growth QTL.

[9] In this equation, the coefficient 8 is due to the fact that we use here the allelic effect a_l, and not the value of each homozygote $\alpha_l = 2a_l$ as classically used in Falconer's notation.

The user provides the genetic map and the initial genetic diversity (number of loci of each type and their allelic diversity). Individual trees in the initial inventory may belong to one population or several populations genetically differentiated. A first option is to provide initial individual genotypes and allelic effects in the inventory; in this case, initial quantitative genetic parameters (genetic mean, VA, h^2) are computed from these values. As an alternative option, the simulation tool Metatrom can be used to simulate individual genotypes fitting target values of population-level genetic parameters given by the user (see the preparation of input files, below). During simulations, the model monitors all population genetic changes through individual genotype information.

Lateral parcels and multi-trees

In addition to the main parcels and individual genotypes that are central to the simulation objectives, Luberon2 offers the possibility to use *lateral parcels* to represent surrounding gene pool, which may contribute to seed and pollen flow during regeneration. These lateral parcels are defined as polygons with homogeneous environmental conditions where the growth models are used without individual spatialisation and without regeneration, as originally developed. Contrary to the main parcels, each lateral parcel is an even-aged growth unit and growth dynamics variables are computed at this level. Thinning and diffuse mortality disturbance do not operate in lateral parcels but growth reduction disturbance does.

Lateral parcels harbor one or several *multi-trees*, each of them representing an even-aged cohort of trees belonging to a single population[10]. Multi-trees are similar to individual trees, except that they have a "number" attribute, their dendrometric variables are population-level predictions of the growth model, and their genotype is defined through allelic frequencies at each locus. Multi-trees in lateral parcels follow growth and self-thinning models and disseminate seeds and pollen, but there is no regeneration, no intervention and no disturbances in lateral parcels.

Individual tree growth

Original growth models

Luberon2 relies on original growth models calibrated for *Cedrus atlantica* (Courbet, 2002), *Pseudotsuga menziesii*, *Picea abies* and *Larix sp.* (Perin et al., 2013, 2017; Ligot et al. 2023) and implemented in the Capsis models CA1[11] and Gymnos[12], respectively. These growth models belong to the same family of distance-independent tree model for even-aged stands. In Luberon2, the simulated forest is split into independent even-aged growth units, i.e. pixels or lateral parcels if any, where individual tree growth and self-thinning processes occur.

Detailed descriptions of the original growth models are given in the references mentioned above. The basic principle is that current local stand characteristics determine the conditions of individual trees' growth, which in turn determines future local stand characteristics for the next step (those parts of the model are not detailed here). Individual tree growth follows a segmented model where annual increment in basal area is null for trees below a threshold circumference value, and increases linearly with the circumference above this threshold (Figure 5).

---

[10] In the graphical display, multi-trees are arbitrarily visualized at several locations randomly chosen by the software within each lateral parcel

[11] https://capsis.cirad.fr/capsis/help/ca1

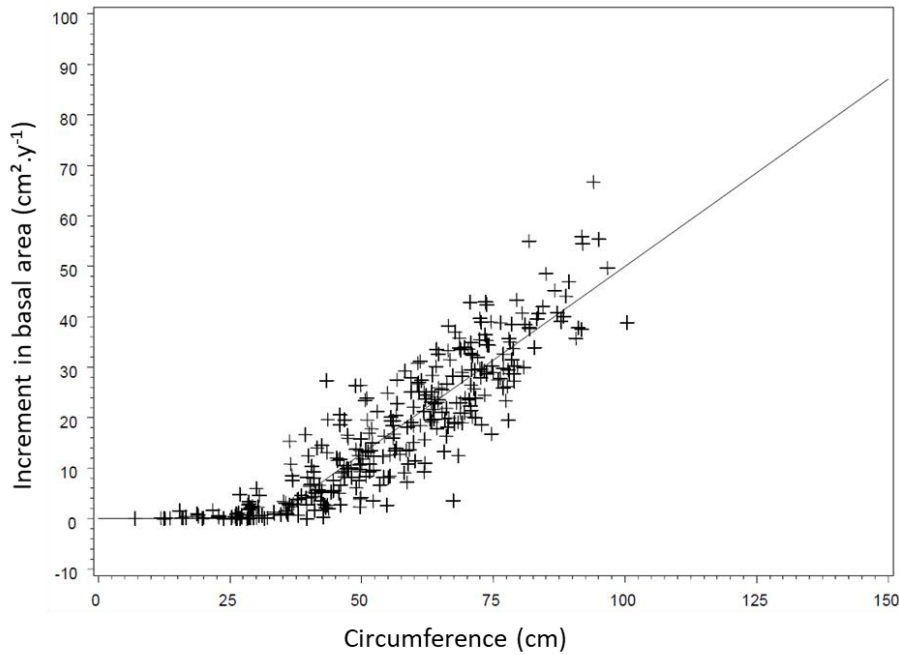[12] https://capsis.cirad.fr/capsis/help_en/gymnos

Figure 5: *Individual annual increment in basal area as a function of tree circumference in C. atlantica: empirical individual data and model prediction (taken from Courbet, 2002).*

In Luberon2, this segmented relationship is modeled using the nonlinear function proposed by Perin et al (2017) with three parameters: the threshold value, A, the slope of the relation between annual growth in basal area and circumference, P, and a flexibility parameter smoothing the break around the threshold value, $m(\geq 1)$, resulting in the following equation of growth in basal area (above the threshold):

$$Ig_i = 0.5P\left(C_i - mA + \sqrt{(mA + C_i)^2 - 4AC_i}\right)$$

where Ig_i is the increment in basal area of individual *i*, and C_i the initial circumference. These parameters evolve during stand development (see Perin et al, 2017), they are dynamically computed for each even-aged growth unit depending on age, stand density and site fertility, as well as species-specific constant values provided as input to the model.

Then, individual phenotypic variation in vigor is added to the diameter growth dD_i as follows:

$$dD_i = \frac{1}{\pi}\left(\sqrt{C_i^2 + 4\pi Ig_i} - C_i\right) + (Mean\_vigor + Phenotype\_vigor_i)$$

where $Mean\_vigor$ is a constant value common to all individuals and $Phenotype\_vigor_i$ is the individual phenotypic deviation to this value for individual *i*. Then, knowing the new diameter of each individual tree, allometric relationships are used to derive height growth, male and female fecundity variables (see below) as well as other values at individual level, such as crown height, or at growth unit level, such as leaf area index.

Note about proposed inventories: mean vigor is a scaling parameter set to 0 by default, which corresponds to the original growth model calibration; changing this value rescales the growth model out of its original calibration range, not recommended except in very specific situations. The default

within-stand phenotypic variance for vigor is set to an empirical estimate obtained for Douglas fir (Fririon et al, 2023), from which the within-stand additive genetic variance is deduced assuming 20 QTL involved with heritability h²=0.3 and 20% of the environmental variance between steps.

Warning: growth models were originally calibrated on empirical data without considering individual phenotypic variation in vigor. Adding this phenotypic variation changes the mean stand-level growth prediction when competitive self-thinning occurs: see further details in Godineau, Fririon, et al. (2023) Appendix 4.

Leaf area index (LAI)

This variable, required for the growth reduction disturbance process, was added to the original growth model. An allometric relationship calibrated for Douglas fir (Smith, 1993) has been implemented to estimate the LAI in each growth unit from tree DBH and growth unit density, as follows:

$$LAI_p = \frac{1}{Area_p} \sum LA_i$$

where the leaf area index of a growth unit *p* is the sum of the leaf areas of its individual trees *i* relatively to the growth unit area, with:

$$LA_i = 0.4781 \times D_i^{1.8659} \times e^{(-0.0829 \times RD_p)} \text{ , and } RD_p = \frac{Gha_p}{Dg_p^b}$$

where the leaf area of individual *i*, LA_i, is obtained by allometry from its DBH, D_i, and Curtis' relative density of the growth unit *p*, RD_p (Curtis, 1982). Curtis' relative density is obtained from the basal area, Gha_p, and the root mean square diameter, Dg_p, of the growth unit and a coefficient defined for each species, $b = 0.5$ for Douglas fir.

Note about LAI computing: The LAI equations are formalized for a continuous cover. In the case of isolated trees, the model applies a minimum LAI of 2. Since these equations tend to overestimate the individual leaf area for old trees, a maximum LAI is set at 16. In practice, this value is never reached in the domains of validity of the model. The LAI model calibrated for Douglas fir is also currently used for the other species, by default of other information.

Growth reduction disturbance

When drought-like growth disturbance is activated by the user, a stochastic disturbance regime occurs reducing the annual increment in diameter predicted by the previous equation, following a formalism proposed by Fririon et al. (2023 and in prep.). This disturbance simulates drought events, as follows. Within each growth unit, the distribution of the level of stress (*StressLevel*) is modeled as a Gamma distribution, Γ, defined with six parameters depending on the current LAI:

$$\Gamma(shape, scale) \text{ , with } shape = a_{shape}LAI^{b_{shape}} + c_{shape} \text{ , and } scale = a_{scale}LAI^{b_{scale}} + c_{scale}$$

where $a_{shape}, b_{shape}, c_{shape}, a_{scale}, b_{scale}, c_{scale}$ are six parameters defining the Gamma distribution.

Each year, a stochastic value, representing climatic hazard intensity for the whole scene, is drawn between 0 and 1 in a uniform distribution. Within each growth unit, the model associates this stochastic value to the integral of the Gamma distribution between 0 and a given *stressLevel* value to provide the annual level of stress experienced by the growth unit considered in the current year (Figure

6). Therefore, through the six parameters, the user can simulate different drought regimes corresponding to different climates and soil water contents.
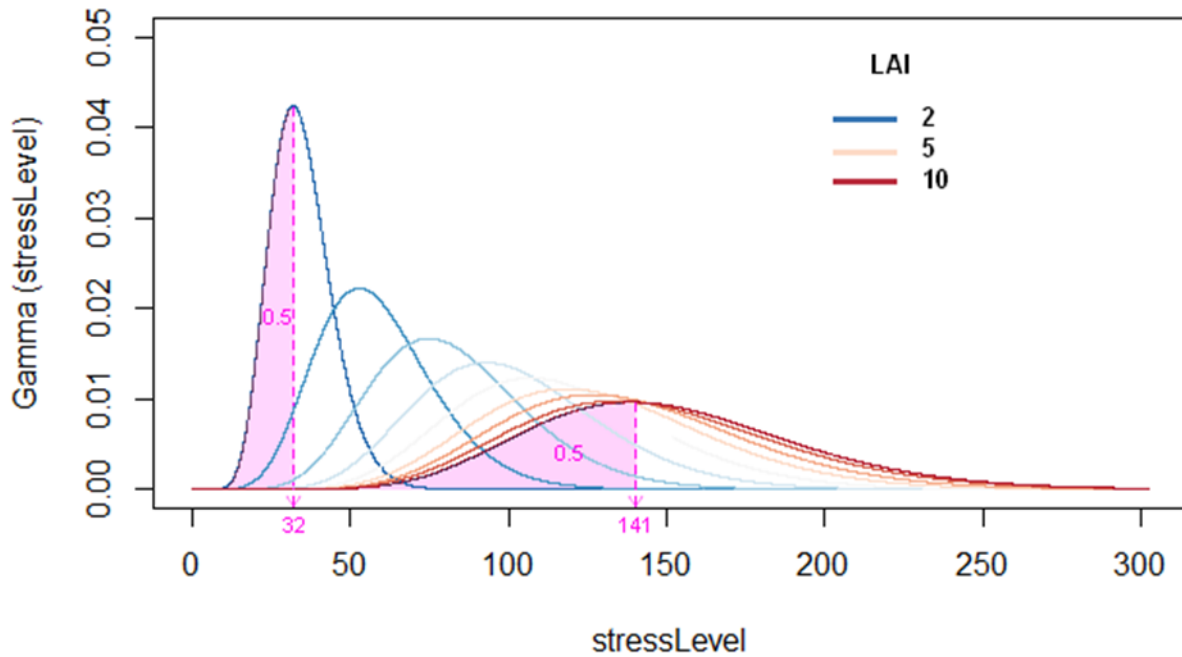

Figure 6*: Gamma distribution of the level of stress (StressLevel) according to LAI in a Douglas fir growth unit. The pink areas represent a hazard intensity of 0.5. For a LAI of 2, this hazard intensity results in a StressLevel of 32, whereas for a LAI of 10 it results in a StressLevel of 141.*

Finally, a reduction is applied to individual diameter growth accounting for phenotypic variation in sensitivity to growth disturbance, as follows:

$$dDimpacted_i = dD_i(1 - StressLevel \times (Mean\_sensitivity + Phenotype\_sensitivity\_growth_i))$$

where $dDimpacted_i$ is the diameter growth of individual tree i after reduction, dD_i the predicted diameter growth without stress, $Mean\_sensitivity$ a constant value common to all individuals used to scale the phenotypic plasticity model (by default identical for all species[13]), and $Phenotype\_sensitivity\_growth_i$ the individual phenotypic deviation to the constant value. To avoid negative values of effective sensitivity ($Mean\_sensitivity + Phenotype\_sensitivity\_growth_i$), which might occur for individual trees with highly negative phenotypic deviation to the constant value, the effective sensitivity is constrained to a minimum value of 10% of $Mean\_sensitivity$.

As a user choice, there is an option to allow poor growth, i.e. drastic relative growth reduction by disturbance below absolute and relative threshold values to result in tree death, following Bugmann (1996), with a given probability (the relative thresholds and probability are provided by the user).

Note about proposed inventories and default parameter values: StressLevel parameters, mean sensitivity and within-stand phenotypic variance of sensitivity to growth disturbance are set to empirical estimates obtained for Douglas fir (Fririon et al, 2023), from which the within-stand additive genetic variance is deduced assuming 20 QTL involved with heritability h²=0.3 and 20% of

---

[13] Analysing phenotypic variation in empirical data, Fririon et al. (2023) found no significant differences for this variable among five tree species, including Douglas fir but not the other species of Luberon2.

environmental variance between steps. Default values of relative growth reduction thresholds causing death and mortality rate are taken from Bugmann (1996) and M. Cailleret (pers. comm.). By default, the absolute growth reduction threshold causing death is set to 0: this can be changed but a strictly positive value would lead all poor growth individuals to be counted as "dead by drought stress" even if their poor growth is due to competition, or if their sensitivity is nul,l and when there is no stress.

Self-thinning

Within each growth unit, after annual tree growth, self-thinning occurs if the number of trees exceeds the maximum number of trees that depends on the root mean square of circumferences following a relation parameterized for cedar by Courbet (2002), unpublished for the other species. When self-thinning occurs, the smallest trees are progressively eliminated one by one, either systematically in the case of cedar or with some stochasticity giving higher removal probability to smaller trees in the other cases, until reaching a density value below the maximum. After self-thinning, the global characteristics of each growth unit are updated.

Diffuse mortality disturbance

The user has an option to add another disturbance event that directly kills trees, irrespectively of individual tree growth within a given growth unit, only depending on individual phenotypic value for sensitivity to this mortality event and a proxy of foliage density in the growth unit. In this case, a stochastic value is first drawn from a distribution combining frequent low values and rare high values (bimodal Weibull distributions where low values are five times more frequent than high values) as a potential probability of death for the whole scene, i.e. a measure of disturbance intensity. Then, this potential intensity of mortality is reduced in growth unit with lower foliage density as measured by the sum of crown heights (i.e. probability of death is reduced with fewer and/or smaller trees[14]). Finally, the probability of death is further adjusted for each individual tree based on its phenotypic value as shown on Figure 6.
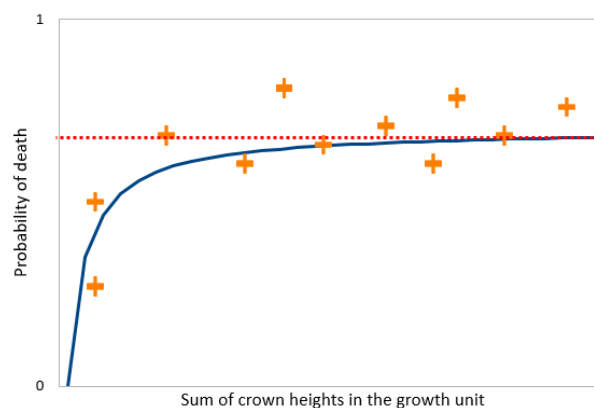
Figure 6: Schematic representation of diffuse mortality. The red dotted line represents the potential probability of death for the whole scene, i.e. disturbance intensity, the blue line represents the adjusted probability with regard to growth unit foliage density, and the orange crosses show the actual probability of mortality for individual trees in different growth units.

---

[14] Such trend is observed with abiotic stresses or biotic interactions under the resource concentration hypothesis.

The individual probability of death of a tree i $P(death)_i$ is computed as follows:

$$P(death)_i = \frac{intensity \times \sum crown\_heights}{(1 + \sum crown\_heights)} + Phenotype\_sensitivity\_mortality_i$$

Where $intensity$ is the potential probability of death for the whole scene, $\sum crown\_heights$ is computed for each growth unit and $Phenotype\_sensitivity\_mortality_i$ is the individual phenotypic value.

Note about proposed inventories and default parameter values: in the proposed inventories, parameters of diffuse mortality disturbance regime are set to reach ~2% of annual potential mortality rate (in addition to self-thinning mortality) following the review made by Cannon et al. (2022); within-stand phenotypic variance for sensitivity to diffuse mortality is an arbitrary values (3.33 10⁻⁴), the within-stand additive genetic variance is deduced assuming 20 QTL involved with heritability h²=0.3 and 20% of the environmental variance between steps.

Regeneration

Seedling establishment and parentage

The user decides on the temporal intervals during which natural regeneration occurs. By default, the regeneration may only occur in pixels with less than 3 pre-existing trees (this parameter can be changed). In those pixels, individual seedlings (not seeds) are randomly installed, as explained below, during a maximum period of time defined by the user (3 years by default). New seedlings are initialized with the dimension of the youngest individuals that can be recruited by the original growth models (i.e. 25 years-old for cedar, 20 years-old for spruce, 15 years-old for Douglas fir and 10 years-old for larch[15]), and don't grow until reaching this recruitment age. Until trees have reached recruitment age, the age of the pixel corresponds to the age of trees initially present or to 0 if the pixel was empty. When recruitment starts, i.e. as soon as a first seedling reaches recruitment age in the pixel, since the growth models are calibrated for even-aged stands, eventually pre-existing trees in the pixel are automatically removed to let recruited seedlings grow, self-thinning regulates the density of recruited seedlings, and the age of all cumulated seedlings over years is re-set to recruitment age.

During the regeneration period, in each pixel where regeneration occurs, all fertile mother-trees of the scene send a number of seedlings given by the product of their individual female fecundity (see below) and a dispersal probability that depends on their distance to the target pixel following a 2D-exponential seed dispersal kernel function that can be parametrized by the user:

$$Prob\_dispersal(dist) = \frac{b}{\left(2\pi\alpha^2\Gamma(^2/_b)\right)} e^{-\left|\frac{dist}{\alpha}\right|^b} \quad \text{with} \quad \alpha = \Delta\frac{\Gamma(^2/_b)}{\Gamma(^3/_b)}$$

where $Prob\_dispersal(dist)$ is the dispersal probability density, $dist$ the distance of the mother-tree to the center of the pixel (set to 5m if the mother-tree is exactly on the center), Δ the mean dispersal distance (240m by default for seeds, as calibrated for cedar[16]) and b the shape parameter of the dispersal kernel (0.22 for seeds, as calibrated for cedar), Γ the gamma distribution function.

---

[15] Initial seedling DBH is drawn at random from a Lognormal distribution for Douglas fir, spruce and larch, taking into account the seedling density and the site index of the given pixel, and from a Weibull distribution for cedar, taking into account seedling density.

[16] E. Klein, pers. comm.

Once a seedling is established, its father-tree is determined among fertile trees with a probability that depends on male fecundity (see below) and possibly distance to the mother tree. Three options are proposed for pollen dispersal: (i) panmixia (all fertile trees of the scene have equal probability to be the father), (ii) neighborhood (only fertile trees in the neighboring pixels of the mother tree can pollinate, depending only on male fecundity) or (iii) pollen dispersal kernel (all fertile trees in the scene are potential pollen donors, depending on male fecundity and distance to the mother tree through a pollen dispersal kernel similar as above with default values calibrated for cedar as 100m for the mean pollen dispersal and 0.3 for the shape parameter). As an option, selfing may be allowed with a parameter *selfingRate* defining how self pollen is treated in the different options for pollination: in the panmixia and neighborhood options, this parameter gives the relative efficiency, or rate of success, of self pollen compared to non-self; in the pollen dispersal kernel option, this parameter represents the proportion of the pollen of the mother tree that is not dispersed far away and remains available for selfing (default value set to 0.05).

Seedling genotypes are automatically derived from parents' genotypes, accounting for stochastic gamete sorting and recombination. Since this model is aimed to simulate short to mean term evolutions, i.e. few generations of trees, mutations are neglected.

Initial DBH of new seedlings are drawn at random, as previously explained, but genotypic values of vigor are derived from the parents: initial DBH and genotypic values are independent. Thus, the genotype starts to influence tree size only after the first step of growth. In other words, inter-individual variation in DBH resulting from all kinds of pre-recruitment processes is considered neutral with regard to the quantitative traits considered in the model, and genetic variation only acts on post-recruitment processes.

Note about default parameter values: seed and pollen dispersal kernels were calibrated for cedar in the multigeneration forest of Mont Ventoux by E. Klein (pers. comm.).

Warning: the parameter selfingRate has different meanings depending on the pollination option. Default value is appropriate for the pollen dispersal kernel option and default dispersal kernel parameters, but it may not be appropriate for other pollination options or other dispersal kernel parameters.

Warning: the number of seedlings received in a pixel is computed as the overall seed rain in 1m² at the center of the pixel from all mother-trees, given female fecundity and distance to the center of the pixel. Thus, relative contributions of all seed trees are determined by fecundity and dispersal kernel functions, which have been calibrated on empirical values, but total number of seedlings is arbitrarily scaled (1m²). This arbitrary scale is convenient with the default pixel size (15mx15m) but it may be an issue with other pixel sizes: see Godineau, Fririon, et al. (2023) Appendix 2B for further explanation and solutions.

Female and male fecundity

Trees are considered as fertile from recruitment age and above a threshold DBH value parameterized for each species, i.e. 15cm for cedar and 10cm for the other species. Female fecundity is determined as an ordinal variable related to tree DBH following the model originally calibrated for cedar by Bertrand (2004). Six levels of cone production are defined (in a log-scale) and, each year, the probability to reach each level depends on DBH, as shown on Figure 7. Once the level of cone production is drawn, the actual number of cones produced by this tree in this year is randomly drawn from a uniform distribution within the range of its fecundity level. Thus, female fecundity has a high inter-annual (for each tree) and inter-individual (among equal-size trees) stochasticity but, currently,

there is no synchrony implemented (i.e. masting). The potential number of seedlings generated by each cone produced is set to an estimate of the mean number of viable seeds per cone known for each species.

The probability of a tree i to fall into fecundity level j is given by the following equation:

$$P_i^j = \frac{e^{(m^j - 0.0114 DBH_i - 0.3914 crown\_height_i)}}{1 + e^{(m^j - 0.0114 DBH_i - 0.3914 crown\_height_i)}}$$

where m^j is a constant given for each fecundity level, DBH_i the diameter of the tree in mm and $crown\_height_i$ the crown height in m.
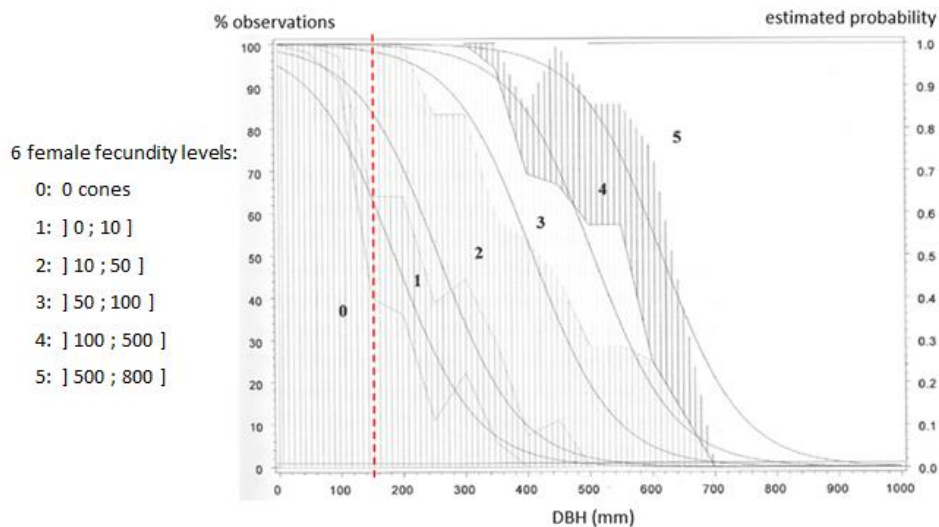
Figure 7: *female fecundity model in C. atlantica: observed data (hatching) and estimated probabilities (smooth lines) (Bertrand 2004). The red dotted line represents the sexual maturity threshold used in Luberon2, above this threshold fecundity level is randomly drawn each year depending on tree DBH (e.g. a tree with DBH=150mm will have higher probability to fall into level 0, 1 or 2).*

Individual "male fecundity" of each fertile tree, which is used to compute the probability of being father of a recruited seedling (see above), is a relative value simply computed as the square of DBH (in cm), it is not an absolute value.

Note about default parameter values: the female fecundity model was calibrated for cedar in the forest of Mont Ventoux by Bertrand (2004) and we use a mean number of 80 viable seeds per cone for this species (Gidoin, 2014). The female fecundity model accounts for high annual fluctuations of individual seed crops but does not account for masting (i.e. synchronized fluctuations of annual seed crop among individuals).

Warning: in the current version, in the absence of similar datasets to calibrate the female fecundity model for Douglas fir, spruce and larch, the same parameters as for cedar are used to predict the number of cones and the number of viable seeds per cone is arbitrarily set to 8 for these species[17]. Male fecundity estimates are still poorly documented for all these species.

---

[17] This value is deduced from the literature review in https://read.oecd-ilibrary.org/science-and-technology/safety-assessment-of-transgenic-organisms/section-4_9789264095434-8-en

III – Preparing input files

It is recommended to keep all input files in their directory by default "…/capsis4/data/luberon2/" with sub-directories "species" and "script".

A summary table of all parameters controlled by the user and default values is given in Annex 1, examples of genetic values are given in Annex 2.

Species files

For each species, the species files (*cedrus.txt*, *douglas.txt, picea_abies.txt, larix.txt)*, provide general parameters used in all simulations:

- regeneration parameters: minimum DBH of fertile trees, age of seedling recruitment, maximum number of years during which seedlings cumulate in a regeneration step, maximum number of pre-existing trees per pixel allowing regeneration, seed fecundity parameters, seed and pollen dispersal parameters, selfing parameter;
- parameters of mortality due to growth reduction: minimum absolute growth (for two consecutive years) and minimum growth rate below which mortality risk occurs, and death probability in this case.

Species files are proposed with default parameter values that can be changed, e.g. cedrus.txt contains the following:

```
# Species file for Luberon2
speciesName = Cedrus
diameterFertility = 15
recruitmentAge = 25
seedlingDuration = 5
maxNumberOfAdultTrees = 3
seedNumberPerFruit = 80
maxPollinationDistance = 150
deltaSeed = 240
bSeed = 0.22
deltaPollen = 100
bPollen = 0.3
selfingRate = 0.05
weakGrowthMortalityProbability = 0.368
weakGrowthThreshold_Rate = 0.10
weakGrowthThreshold_cm = 0.0
```

Inventories

For each simulation, an inventory file is used to define general simulation parameters and characterize the initial stand, providing following information:

(a) general simulation parameters: initial simulation year, pixel size, management of seed flow falling out of the parcels (ignore or repeat dispersal);
(b) diffuse mortality disturbance regime: defined by the parameters of the Weibull distributions representing frequent "weak" and rare "strong" mortality events;
(c) growth disturbance regime: defined by six parameters of the function relating the stress level Gamma distribution to growth unit LAI;
(d) mean phenotypic values of vigor and sensitivity to growth disturbance for the whole initial stand (initial mean sensitivity to diffuse mortality disturbance is 0)
(e) list of main parcels (and lateral parcels if any), with their polygon and site index defined as dominant height at a given age;

(f) list of initial individual trees (and multi-trees if any) with their population of origin (if mixed populations), spatial coordinates, initial DBH and height (optional), age (must be greater than the age of recruitment and homogeneous within the same growth unit), nuclear (required) and cytoplasmic (optional) genotype;

(g) linkage map, designed as a linear sequence of recombination rates, and allelic diversity for all neutral loci and QTL together;

(h) kinship relationship;

(i) quantitative genetic parameters for each trait in the initial stand: allelic substitution effect for each QTL, initial heritability or environmental variance (only one of these is needed because they are related through the additive genetic variance derived from allelic effects and frequencies), and ratio of environmental variance that varies through time.

Examples of inventory files are proposed, which can be used as templates to develop new ones with exactly the same structure and format (lines starting with # are comments). Files start with general simulation parameters and disturbance regime information (a) to (d):

```
# Initialisation file
speciesFileName = species/cedrus.txt

# General parameters
year = 1900
pixelSize = 15
ignoreSeedlingDispersedOutsideParcel = true

# Mortality disturbance parameters
mortalityDisturbance_alphaWeak = 2.9447201
mortalityDisturbance_betaWeak = 0.0221433
mortalityDisturbance_alphaStrong = 0.0000000001
mortalityDisturbance_betaStrong = 0.0000000001

# Stress level parameters
stressLevel_aShape = 55
stressLevel_bShape = -5
stressLevel_cShape = 7


stressLevel_aScale = -12.06
stressLevel_bScale = -1
stressLevel_cScale = 9.45

# Mean plasticity parameter, scaling the model of sensitivity to growth disturbance
growthDisturbanceMeanPlasticity = 0.001

# Mean vigor parameter, scaling the growth model
growthVigorMean_cm = 0
```

Then, all parcels are described, one single row for each parcel, with their name, type (main or lateral, M or L), site index (top height, HdomSI, at a given age, ageSI) and spatial coordinates of the vertices of polygon (the number of vertices may vary between parcels, the first and last vertices in the list must be the same). Additional information is required by the model: id is a unique number (generally in sequence) and pixelMapId can be set to 1 for all. Here is an example for a main parcel P12 and a lateral parcel Est (lines are split below for presentation only, should be one line per parcel):

```
# Parcels
# id   name typeM/L pixelMapId  HdomSI ageSI   polygon vertices
1      P12  M       1           12.06  50      {(835746,1871520);(835773,1871510);
                (835772,1871214);(835589,1871210);(835274,1871193);(835746,1871520)}

29     Est  L       1           12.06  50      {(836419,1871520);(836650,1871520);
             (836650,1870720);(836313,1870720);(836313,1870720); (836341,1870964);
                                (836318,1871205);(836419,1871438);(836419,1871520)}
```

Then all individual trees (coded IT) and multi-trees (coded MT) are described with their specific attributes, in one single row for each tree. Tree description contains mandatory and optional information as described in (f) above: missing information is coded -1 for a single value or { } for a list of values.

The specific format for an individual tree is illustrated below (lines are split below for presentation only, should be one line per individual). This tree, named 334, belongs to population 1. Spatial coordinates x and y are mandatory (z is not used in Luberon2 but it must be there with any value). From these coordinates, the individual tree will be automatically assigned to a main parcel (and will be discarded with a message if its coordinates do not fit into any main parcel). In this example, tree DBH and age are provided but not height, which will be automatically computed by the model[18]. The diploid nuclear genotype is given as a list of alleles for each locus in their order along the genome map. QTL alleles are coded 1 and 2, neutral alleles coding starts at 11. In this example, the nuclear genotype has five QTL and three neutral loci, the individual is heterozygous at the second QTL and all neutral loci. Cytoplasmic DNA is coded with one maternal and one paternal loci, each haplotype is given as a list between { } with a single value: here, there is no genotype provided for maternal cDNA but it has the allele 33 of the paternal cDNA locus. The last three attributes are generally not given for the initial trees but must be present, these attributes are used for new individuals created by the model during regeneration: identity of the mother and father and date of creation.

```
# Trees   (individual)
# They belong to main parcels
# Height is recomputed if set to -1
# Code  Id    PopId  x(m)     y(m)     z(m)    Dbh(cm)      Height(m)    Age(year)
IT     334   1      836052   1870483  0       27.25        -1           45

              nucDNA
              {2;2;1;2;1;1;2;2;2;2;15;13;11;13;15;13}

              motherCytDNA         fatherCytDNA
              {}                   {33}

              motherId    fatherId    creationDate
              -1          -1          -1
```

The specific format for a multi-tree is slightly different. A multi-tree has no precise spatial coordinates but the identity of the lateral parcel where it is located is mandatory (cannot be a main parcel). A multi-tree is an even-aged cohort, its age must be given. Its dendrometric characteristics are tree density (Nha) and basal area per ha (Gha). Genetic characterization of a multi-tree is given through the number of copies of each allele of each locus, and the model uses these numbers to compute the number of trees in the multi-tree: if a multi-tree represents n trees, it has 2n alleles in total for each nuclear locus and n alleles for each cytoplasmic locus. The genotype is provided has a list between { } of arrays between [ ], one array per locus that includes the number of copies of each of its alleles. Thus, the sum of all allele copies per locus is the same for all nuclear loci, which is exactly twice the sum for cytoplasmic loci (there is an error message if numbers are not consistent across loci). For multi-allelic neutral and cytoplasmic loci, the number of copies are given in the order of allele coding, e.g. if an allele is absent in the multi-tree it must appear with a 0. Below is an example matching with the individual tree above: multi-tree 1 represents 1000 individuals, there are five diallelic QTL and three

---

[18] For Douglas fir only, missing DBH (set to -1) will be automatically estimated by the model given the individual tree age and the site index of the given parcel provided in the inventory. For cedar, missing DBH will systematically be replaced by a value estimated for a fixed age of 25 years, similarly as for a young seedling (which would not be appropriate if another age is indicated in the inventory).

neutral nuclear loci with five alleles each, the complete list of the 33 alleles for the paternal cDNA is truncated here for clarity.

```
# MultiTrees (cohorts)
# Code  Id      PopId           ParcelId        Age     Nha     Gha
MT      1       1               29              45      9.0     0.68

                nucDNA
                {[161,839];[85,915];[613,387];[401,599];[535,465];
                    [199,208,210,197,186];[196,202,203,191,208];[200,202,195,206,197]}

                motherCytDNA            fatherCytDNA
                {}                      {[2,91,49,24,2,95, …/… ,1,18,47,38]}

                motherId        fatherId        creationDate
                -1              -1              -1
```

Then, as mentioned in (g) above, a single line provides the genetic map and allelic diversity in the initial population (this line is split below for presentation only). The genetic map is a list of recombination probability between nuclear loci, in the same order as used for individual tree genotypes and identifies the list of allele names for each nuclear and cytoplasmic locus: names of QTL alleles must be 1 and 2, names must start at 11 for neutral nuclear loci. Note that names of nuclear alleles must be followed by ".0", not for cytoplasmic alleles. The model checks that all the alleles declared in the individual tree genotypes are also present on this line, or an error message is thrown.

```
# Species
# speciesName         recombinationProbabilities
Cedrus                {0.5;0.5;0.5;0.5;0.5;0.5;0.5;0.5}

                allelesNuclear
                {[1.0,2.0];[1.0,2.0];[1.0,2.0];[1.0,2.0];[1.0,2.0];
                                [11.0,12.0,13.0,14.0,15.0];[11.0,12.0,13.0,14.0,15.0];
                                                [11.0,12.0,13.0,14.0,15.0]}

                allelesMCytoplasmic
                {}

                allelesFCytoplasmic
                {[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,
                                                26,27,28,29,30,31,32,33]}
```

The kinship relation among all individuals is generally not provided (but it might be given if needed in very particular situations), and a general consanguinity value is provided, set to 0 by default.

```
# Phi (consanguinity)
# initialPhiArray    defaultPhi
{}                   0
```

Finally, quantitative genetic parameters are given for traits controlled by the QTL, one line for each. For each variable trait, a list of arrays provides, for each of its QTL, the QTL location on the genome map and allelic effects of the two alleles: by convention allele 1 has a negative effect and allele 2 has a positive effect with equal absolute value. Allele effects given in the array are constrained to be integer values for computational reasons, and appropriate scaling of actual effects is achieved through an *effectCoefficient*. Maternal and paternal cytoplasmic loci can also have an effect on traits. Either heritability or environmental variance should be provided as preferred, and the last parameter is the proportion of environmental variance that varies between steps (e.g. due to climatic fluctuation in time). In the example below (the single line is split for presentation only), the five QTL declared in the genotypes above all control vigor, and there is no genetic variation for the two other traits related to disturbance. QTL 3 and 4 have a weaker effect than QTL 1, 2 and 5. The heritability is set to 0.3 and 20% of the environmental variance is inter-step.

```
# Parameter effects
# Parameter          effectCoefficient   nuclearEffect
growthVigor_cm       1.0E-5               {[1.0,-1391.0,1391.0];[2.0,-1585.0,1585.0];
                     [3.0,-971.0,971.0];[4.0,-339.0,339.0];[5.0,-1192.0,1192.0]}

          mCytoplasmicEffect  pCytoplasmicEffect
          {}                  {}

          heritability   environmentalVariance   interEnvironmentalVariance
          0.3            -1.0                     0.2
```

Initialization of quantitative genetic data with Metatrom

Genetic architecture of quantitative traits and initial individual tree genotypes can be arbitrarily determined by the user. But a more common situation is to have empirical or theoretical estimates of synthetic quantitative genetics variables such as additive genetic variance, heritability or distribution of QTL contributions. These synthetic quantitative genetic variables can be estimated or inferred from empirical data, and it is important to remember that these empirical estimates are instantaneous values of dynamic variables. However, fixed parameters required by the model, i.e. exhaustive list of QTL controlling a quantitative trait and their allelic effects as well as individual genotypes for each QTL are not empirically accessible. It is therefore necessary to translate the quantitative genetics estimates into fixed parameters and individual genotypes used by Luberon2.

Luberon2 proposes an option to create an inventory file using Metatrom at initialisation of a project. Metatrom is an optimization algorithm to simulate neutral and QTL nuclear genotypes for individual trees and multi-trees defined in the inventory. For neutral loci, the input is the number of loci with their allelic diversity. For the three quantitative traits, Metatrom uses the number of diallelic QTL involved[19] provided by the user, and target values of synthetic quantitative genetics variables for each trait, to simulate allelic effects and allelic frequencies fitting these target values, assuming L-shaped distribution of QTL contributions following Bost et al (1999, 2001), and to simulate individual genotypes respecting these allelic frequencies. Metatrom gets its inputs from an inventory file as described above with complementary information at the end that provide target values of desired variables and some optimization parameters. The output of Metatrom is an inventory file where individual nuclear genotypes and allelic effects from the input file are replaced by the results of the optimization, with a report on optimization at the end of the file. Metatrom can simulate individual genotypes sampled from several differentiated subpopulations identified by the "population attribute" of each individual in the inventory.

In case of a single population, for each trait, Metatrom first draws the contribution of each QTL (l) to the target additive genetic variance in a Gamma distribution (VA_l). Then, for each QTL, the frequencies of the two alleles (p_l, q_l) are drawn at random in a Gamma distribution, and allelic effect a_l is defined as: $a_l = \sqrt{\frac{VA_l}{8p_l q_l}}$. These allelic effects replace previous values in the input inventory. For each QTL, alleles are drawn at random following the allelic frequencies and distributed among individual genotypes with two conditions: (i) the achieved QTL contribution to the additive genetic standard deviation does not depart from the expected value by more than 1%, (ii) the difference between achieved genetic mean of the population and 0 does not exceed a threshold value provided by the user[20]. If necessary, alleles are resampled and reshuffled among individuals until these conditions are met. For each neutral nuclear locus, allelic frequencies and drawn at random in a uniform distribution

---

[19] This assumption is under the responsibility of the user: this number is generally unknown and it has evolutionary consequences (Cubry et al 2022), in the proposed inventories 20 QTL are assumed for each trait.
[20] This is a security to avoid far away local optima.

and distributed among individual genotypes with the condition that Hardy-Weinberg equilibrium chi-square test remains not significant within each subpopulation at a p-value provided by the user. If necessary, alleles are reshuffled among individuals until this condition is met. Metatrom can also simulate multi-trees genotypes, i.e. allelic frequencies for each QTL and neutral locus with no genotypic distribution.

In the case of multiple populations, target values of F_{ST}, for neutral loci, and Q_{ST}, for each trait, must be provided by the user. These values are approached by reshuffling alleles among individuals, while preserving the previous conditions, with a fixed limited number of iterations in the algorithm.

In the example below, there is just one population (expected F_{ST} is set to 0 in this case), and p-value threshold for Hardy-Weinberg test is given to Metatrom. Then, for each trait, a single line (split here for presentation only) provides the tolerance for population mean around 0, the target additive variance (warning: the variance target given to Metatrom is not scaled by the effectCoefficient) and the expected Q_{ST} (set to 0.0 in this case). The last comment lines are the report in the output inventory after optimization.

```
# Metatrom configuration, global
# Genetic structuration
# Fst global
# expectedNeutralFst       Hardy Weinberg Equilibrium threshold
  0                        0.01

# Metatrom configuration, traits level
# Trait              threshold_value_for_additive_mean
growthVigor_cm       0.05

                          expectedAdditiveVariance  expectedQst
                          35.0                      0.0


# Metatrom diagnostic report
# These lines are written by Metatrom for user control
#
# Metatrom growthVigor_cm realizedAdditiveVariance: 35.4474983215332
```

Warning: The current version of Metatrom allows simulating genetic data for a single trait: it must be run for each trait one after the other, before merging outputs to reconstruct full genetic information and complete individual genotypes. An improved version allowing multi-trait simulations is in preparation. Furthermore, it produces an output inventory file with inappropriate heading lines (components (a) to (e) above, currently corresponding to a previous version of inventory files). Therefore, after running Matatrom, these lines (a) to (e) must be retrieved from a proper inventory template.

Note about default parameter values and recommendations: The optimization process may be very long or even fail to converge in certain conditions. The threshold p-value for the HW test has a strong impact on convergence, it is recommended to stick to the default value. Similarly, the threshold tolerance on the population genetic mean around 0 influences convergence, it should not be too stringent (the stringency of this threshold depends on the additive genetic variance of the population). Convergence is also more difficult with fewer QTL and fewer individuals: we recommend using no less than 10 QTL and 50 individuals per population. The names of alleles defined by the user in the inventory as "allelic diversity" must follow a standard rule (the "allelic diversity" of the inventory is not corrected by Metatrom): QTL alleles named 1 and 2, multiple alleles at neutral loci named 11 to n. In the current version, it is recommended to use the same number of alleles for all neutral loci.

IV Script mode and exports

Luberon2 can be used in a script mode, where a command file gives instructions for: (i) the inventory to use, (ii) the sequence of steps to simulate with their options and (iii) the types of outputs.

There is currently one script mode proposed by default, named "ScriptVictor2022". Advanced users can develop other types. Script mode "ScriptVictor2022" requires a command file structured as follows. Each command file has a set of unique values for some general parameters that are applied to each command line, i.e. each text line is interpreted to run one simulation. The general parameters are: the number of repetitions for each simulation, the duration of the regeneration phase between two rotations, the option chosen for pollen dispersal, and the outputs. The names of the output files are automatically determined by the model. Within a command file, each command line is determined by initial inventory (required), silvicultural scenario within each rotation (sequence of thinning interventions, optional), number of successive rotations to simulate with the same sequence of interventions (required), and option chosen for the growth reduction disturbance regime (required, there is no diffuse mortality in this script). To launch the proposed script mode with a command file named "testCommandFile.txt", type (depending on windows or linux versions):

capsis -p script luberon2.myscripts.ScriptVictor2022 data\luberon2\script\testCommandFile.txt
sh capsis.sh -p script luberon2.myscripts.ScriptVictor2022 data\luberon2\script\testCommandFile.txt

Warning: By default, the memory allocated to Capsis is 1024 because higher values might not work with low-performance computers using Windows (crash at start). However, this may not be enough for intensive simulations (crash after a few annual steps). Running the setmem command to redefine the memory allocated to Capsis, e.g. 'setmem 8000' (under Windows) or 'sh setmem.sh 8000' (under Linux and Mac)[21].

A command file is illustrated below. Note that the choice of an option for pollen dispersal is made by commenting the other options. Additional silvicultural scenarios similar to "s1" can be defined in successive lines. The definition of a silvicultural scenario is explained in the commented lines: each intervention can be applied to a single parcel, or to all parcels independently one by one, or to the whole scene. Multiple simulations can be launched via the same command file with additional command lines (last line below).

```
# A command file for luberon2
repetitionNumber = 1
regenerationDuration = 5

#fatherSelectionMethodName = PANMIXIA
#fatherSelectionMethodName = NEIGHBORHOOD
fatherSelectionMethodName = POLLEN_DISPERSION_KERNEL

# saveProjects should be false in normal mode
saveProjects = false
saveIndividualDynamicsExport = false
savePixelDynamicsExport = true
saveSceneDynamicsExport = true

# Sylvicultural scenarios
# name encodedScenario
# encodedScenario is a list between {} of (age,level,targetDensity,type)
                                                          separated by ;
```

---

[21]

https://capsis.cirad.fr/capsis/documentation/faq#i_need_more_memory_for_my_simulations_how_can_i_do

```
# age: triggers the thinning
# level: parcelName, or PARCEL_ONE_BY_ONE, or SCENE
# if targetDensity is an int, it is a targetNha, else it is a thinning intensity
                                                            between 0 and 1
# type is between 0 (from bottom) and 1 (from top)
# name {(age,level,targetNha,type);...}
# name {(age,level,intensity_01,type);...}
s1      {(20,PARCEL_ONE_BY_ONE,600,0.5);(30,P2,0,0.5);(50,SCENE,50,0.5)}

# The inv files below are in the same dir than this file
# growthDisturbanceType: NONE, GROWTH_DISTURBANCE_ONLY or
                                                GROWTH_DISTURBANCE_WITH_MORTALITY
# inventoryFileName    scenario    rotationNumber    growthDisturbanceType
Inventory.txt          s1          2                 GROWTH_DISTURBANCE_ONLY
```

The contents of the three types of export files are described in the following box:

```
SceneDynamicsExport    [global stand-level values]
      date
      interventionResult    [true if thinning intervention, false otherwise]
      Nha    [tree density, including trees before recruitment]
      deadBySelfThinning    [Nb of dead trees by this cause until this date]
      deadByGrowthDisturbance    [ibid.]
      deadByDiffuseMortalityDisturbance    [0 with this script]
      deadByRecruitmentManagement    [Nb of seed trees automatically removed]
      deadTotal    [total of the previous four values]
      growthVigor_cmMean    [genetic mean]
      growthVigor_cmObservedVariance    [observed genetic variance²²]
      growthVigor_cmPanmicticVariance    [panmictic genetic variance]
      growthDisturbanceIndividualSusceptibilityMean    [genetic mean]
      growthDisturbanceIndividualSusceptibilityObservedVariance    [observed
      genetic variance]
      growthDisturbanceIndividualSusceptibilityPanmicticVariance    [panmictic
      genetic variance]
      mortalityDisturbanceIndividualSusceptibilityMean    [genetic mean]
      mortalityDisturbanceIndividualSusceptibilityObservedVariance    [observed
      genetic variance]
      mortalityDisturbanceIndividualSusceptibilityPanmicticVariance    [panmictic
      genetic variance]
      corG_growthVigor_growthSusceptibility    [genetic correlation]
      He_mean_neutralLoci    [neutral gene diversity]
      He_mean_qtl    [QTL gene diversity (all QTL for all traits together)]
      mortalityDisturbanceIntensity    [0 with this script]
      growthDisturbanceIntensity    [overall disturbance intensity, in [0,1]]
      fecundityMean    [mean female fecundity, 0 for years without regeneration]
      fecundityVariance    [variance of female fecundity, 0 when no regeneration]
      growthDisturbanceStressLevelMean    [mean stress level across pixels]
      growthDisturbanceStressLevelVariance    [variance of stress level per pixel]


PixelDynamicsExport    [values for each pixel]
      date
      interventionResult    [true if thinning intervention, false otherwise]
      parcelName    [parcel to which the pixel belongs]
      parcelType    [only main parcels are exported]
      parcelHdomSI    [site index (top height)]
      parcelAgeSI    [site index (age)]
      pixelId
      age
      Hdom    [top height]
      Ddom    [top diameter]
      CBHdom    [top crown-base height]
      dg    [root mean diameter]
      hg    [root mean height]
      vg    [root mean volume]
```

---

[22] The observed genetic variance is the variance of genotypic values, it accounts for both allelic frequencies and linkage disequilibrium, while the panmictic genetic variance is computed from allelic frequencies only.

```
      CBHg    [root mean crown-base height]
      diameterGrowthAlpha    [slope parameter of the growth model]
      diameterGrowthThreshold    [threshold parameter]
      diameterGrowthThresholdFlexibility    [flexibility parameter]
      Nha    [tree density, NOT including trees before recruitment]
      Gha    [basal area per ha]
      Vha    [volume per ha]
      prodGha    [cumulated Gha since the beginning of the simulation]
      prodVha    [cumulated Vha since the beginning of the simulation]
      lai    [Leaf Area Index]
      growthDisturbanceStressLevel    [stress level]


IndividualDynamicsExport    [individual tree-level values]
      date
      interventionResult    [true if thinning intervention, false otherwise]
      parcelName    [parcel in which the tree is located]
      pixelId    [pixel in which the tree is located]
      id
      popId
      age
      adult    [true if fertile tree]
      dbhIncrement_cm    [annual growth]
      dbh
      height
      fecundity    [female fecundity]
      paternityNumber    [0, unused here]
      ageOfDeath    [-1 for alive trees, only alive trees are exported]
      dateOfDeath    [-1 for alive trees, only alive trees are exported]
      causeOfDeath    [NOT DEAD for alive trees, only alive trees are exported]
      number    [1 for individual trees, only individual trees are exported]
      volume_m3
      leafArea
      weakGrowthInDiameter    [true if beyond growth reduction mortality threshold]
      mId    [mother id]
      pId    [father id]
      consanguinity
      globalConsanguinity    [-1, unused for individual trees]
      G_growthVigor_cm    [genotypic value]
      G_growthDisturbanceIndividualSusceptibility    [genotypic value]
      G_mortalityDisturbanceIndividualSusceptibility    [genotypic value]
      P_growthVigor_cm    [phenotypic value]
      P_growthDisturbanceIndividualSusceptibility    [phenotypic value]
      P_mortalityDisturbanceIndividualSusceptibility    [phenotypic value]
      Efixed_growthVigor_cm    [fixed environmental value]
      Efixed_growthDisturbanceIndividualSusceptibility    [fixed envir. value]
      Efixed_mortalityDisturbanceIndividualSusceptibility    [fixed envir. value]
      nuclearDNA    [nuclear genotype]
      mCytoplasmicDNA    [maternal cytoplasmic genotype]
      pCytoplasmicDNA    [paternal cytoplasmic genotype]
```

Warning: using this script, some export files are huge, only ask for the ones really needed. In particular, saving the whole project allows to re-open it with the graphical interface but it should be avoided in general: it can be useful for a preliminary check of the thinning scenario. Above all, the individual export file contains individual information for all trees at each step, it can be huge: if this export is needed, we strongly recommend a preliminary test on a simple scenario to estimate the total size of expected outputs.

Warning: in the pixel export, empty pixels have their dendrometric variables set to 0 not to missing data.

References

Bertrand A, 2004. Dynamique naturelle du Cèdre de l'Atlas (*Cedrus atlantica* Manetti) au Mont Ventoux. Approche spatio-temporelle et éléments de modélisation. DEA Aix-Marseille. https://capsis.cirad.fr/capsis/help_en/luberon2

Bost B.C., Dillmann C., De Vienne D., 1999. Fluxes and metabolic pools as model traits for quantitative genetics. I. L-shaped distribution of gene effects. Genetics, 153:2001-2012. https://doi.org/10.1093/genetics/153.4.2001

Bost B., De Vienne D., Hospital F., Moreau L., Dillmann C., 2001. Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. Genetics, 157:1773-1787. https://doi.org/10.1093%2Fgenetics%2F157.4.1773

Bugmann H.K.M., 1996. A Simplified forest model to study species composition along climate gradients. Ecology, 77:2055-2074. http://www.jstor.org/stable/2265700

Cannon C.H., Piovesan G., Munné-Bosch S., 2022. Old and ancient trees are life history lottery winners and vital evolutionary resources for long-term adaptive capacity. Nature Plants, 8:136-145. https://doi.org/10.1038/s41477-021-01088-5

Courbet F. 2002. Modélisation de la croissance et de la qualité du Cèdre : intégration des modèles dans Capsis. In : Rapport convention MAPA-DERF / INRA n° 61.45.47/01. https://hal.archives-ouvertes.fr/hal-03654591

Cubry P., Oddou-Muratorio S., Scotti I., Lefèvre F., 2022. Interactions between micro-environment, selection and genetic architecture drive multiscale adaptation in a simulation experiment. Journal of Evolutionary Biology, 35:451-466. https://doi.org/10.1111/jeb.13988

Curtis R.O., 1982. A simple index of stand density for Douglas-fir. Forest Science, 28:92-94. https://www.fs.usda.gov/treesearch/pubs/7202

Fririon V., Davi H., Oddou-Muratorio S., Lebourgeois F., Lefèvre F., 2023. Within-stand phenotypic variation in growth vigor and sensitivity to drought stress in five temperate tree species. Forest Ecology and Management, 531:120754. https://doi.org/10.1016/j.foreco.2022.120754

Gidoin C., 2014. Impact écologique d'une invasion d'insectes séminiphages dans un écosystème méditerranéen. PhD thesis, Univ. Montpellier II. http://www.theses.fr/2014NSAM0015

Godineau C., Fririon V., Beudez N., de Coligny F., Courbet F., Ligot G., Oddou-Muratorio S., Sanchez L., Lefèvre F. (2023). A demo-genetic model shows how silviculture reduces natural density-dependent selection in tree populations. Evolutionary Applications (in press).

Lamarins A., Fririon V., Folio D., Vernier C., Daupagne L., Labonne J., Buoro M., Lefèvre F., Piou C., Oddou-Muratorio S., 2022. Importance of interindividual interactions in eco-evolutionary dynamics: the rise of demo-genetic agent-based models. Evolutionary Applications, 15:1988–2001. https://doi.org/10.1111/eva.13508

Lefèvre F., Godineau C., Beudez N., de Coligny F., Courbet F., Oddou-Muratorio S., Sanchez L., Deleuze C., Pichot C., Chartier M., Musch B., François D., Girard S., Guyot J., Le-Legard Moreau L., Riou-Nivert P., Rousselle Y., Salvaudon A., Sédilot-Gasmi C., 2019. Comprendre comment la sylviculture modifie la qualité génétique et les capacités d'adaptation des peuplements. Forêt-entreprise, 249:49-51. https://www.foretpriveefrancaise.com/publications/voir/742

Ligot G., Gheysen T., Perin J., Candaele R., de Coligny F., Licoppe A., Lejeune P., 2023. From the simulation of forest plantation dynamics to the quantification of bark-stripping damage by ungulates. European Journal of Forest Research, 142:899-916. https://doi.org/10.1007/s10342-023-01565-w

Oddou-Muratorio S., Davi H., Lefèvre F., 2020. Integrating evolutionary, demographic and ecophysiological processes to predict the adaptive dynamics of forest tree populations under global change. Tree Genetics & Genomes 16:67. https://doi.org/10.1007/s11295-020-01451-1

Perin J., Hébert J., Brostaux Y., Lejeune P., Claessens H., 2013. Modelling the top-height growth and site index of Norway spruce in Southern Belgium. Forest Ecology and Management, 298:62-70. https://doi.org/10.1016/j.foreco.2013.03.009

Perin J., Claessens H., Lejeune P., Brostaux Y., Hébert J., 2017. Distance-independent tree basal area growth models for Norway spruce, Douglas-fir and Japanese larch in Southern Belgium. European Journal of Forest Research, 136:193-204. https://doi.org/10.1007/s10342-016-1019-y

Seynave I., Pichot C., 2004. Bibliothèque Genetics de Capsis, Documentation Internal report. https://capsis.cirad.fr/capsis/help_en/genetics#documentation

Smith N.J., 1993. Estimating leaf area index and light extinction coefficients in stands of Douglas-fir (*Pseudotsuga menziesii*). Canadian Journal of Forest Research, 23:317-321. https://cdnsciencepub.com/doi/abs/10.1139/x93-043?journalCode=cjfr

Annex 1 – Overview of input parameters controlled by the user

The table below synthesizes all input parameters controlled by the user, their location, and default values. We recommend not to change default values in grey except in advanced use of the model.

Default species files and some inventories are proposed in "capsis4/data/luberon2".

Parameters	Where is the control	Default values Cedar	Default values Douglas fir
Species characteristics			
Minimum diameter for fecundity (diameterFertility)	*Species file*	*15 cm*	*10 cm*
Age of recruited seedlings (recruitmentAge)	*Species file*	*25 years*	*15 years*
Duration of seedling recruitment (seedlingDuration)	*Species file*	*5 years*	*5 years*
Maximum nb of trees per pixel allowing seedling recruitment (maxNumberOfAdultTrees)	*Species file*	*3 trees*	*3 trees*
Nb of viable seeds per cone (seedNumberPerFruit)	*Species file*	*80 seeds*	*8 seeds*
Pollen dispersal kernel parameters	*Species file*	*deltaPollen = 100 bPollen = 0.3*	*deltaPollen = 100 bPollen = 0.3*
Maximum pollen distance (maxPollinationDistance)	*Species file*	*150 m*	*150 m*
Seed dispersal kernel parameters	*Species file*	*deltaSeed = 240 bSeed=0.22*	*deltaSeed = 240 bSeed=0.22*
Selfing rate (selfingRate)	*Species file*	*0.05*	*0.05*
Minimum relative growth rate compared to the potential, below which mortality occurs (weakGrowthThreshold_Rate)	*Species file*	*0.1*	*0.1*
Absolute growth threshold for 2 consecutive years, below which mortality occurs (weakGrowthThreshold_cm)	*Species file*	*0.0 cm*	*0.0 cm*
Probability of death when mortality due to growth disturbance occurs (weakGrowthMortalityProbability)	*Species file*	*0.368*	*0.368*
Simulation management options			
Pixel size (pixelSize)	*Inventory*	*15 m*	*15 m*
Management of seed flow falling out of the parcels (ignoreSeedlingDispersedOutsideParcel)	*Inventory*	*true*	*true*
Mean sensitivity to growth stress level (growthDisturbanceMeanPlasticity)	*Inventory*	*0.002 Mpa-1*	*0.002 Mpa-1*
Mean vigor (growthVigorMean_cm)	*Inventory*	*0 cm*	*0 cm*
Disturbances regimes			
Diffuse mortality disturbance regime (only frequent weak events, no rare strong events, ~2% annual mortality)	Inventory	alphaWeak = 2.945 betaWeak = 0.022 alphaStrong = 10^{-10} betaStrong = 10^{-10}	alphaWeak = 2.945 betaWeak = 0.022 alphaStrong = 10^{-10} betaStrong = 10^{-10}

Growth reduction disturbance regime (moderate stress level, mortality threshold reached exceptionally)[23]	Inventory	aShape = 55.9 bShape = -5 cShape = 7.41 aScale = -12.06 bScale = -1 cScale = 9.45	aShape = 55.9 bShape = -5 cShape = 7.41 aScale = -12.06 bScale = -1 cScale = 9.45
Initial stand characteristics			
Geometry and site index (top height at a given age) per parcel (including lateral parcels, if any)	Inventory	- No default values - Required	- No default values - Required
Individual tree coordinates, age and population of origin (multi-trees in lateral parcels, if any)	Inventory	- No default values - Required	- No default values - Required
Individual DBH and height	Inventory	- No default values - DBH required, or simulated only for the age 25 - height optional	- No default values - if not provided, DBH simulated for the indicated age - height optional
General kinship coefficient among individuals (defaultPhi), or a full kinship matrix (initialPhiArray)	*Inventory*	*0*	*0*
Initial genetic data			
Individual genotypes at nuclear and cytoplasmic loci, consanguinity	Inventory	- No default value - Required (can be created with Metatrom)	- No default value - Required (can be created with Metatrom)
Genetic map of nuclear loci, cytoplasmic loci, nb of alleles per locus	Inventory	- No default values - Required	- No default values - Required
Allelic effects at each QTL (diallelic), and possibly cytoplasmic loci, for each trait	Inventory	- No default value - Required (can be created with Metatrom)	- No default value - Required (can be created with Metatrom)
Heritability or environmental variance (choose one) and rate of the "inter-step" environmental variance that fluctuates between steps, for each trait	Inventory	- No default values - Required	- No default values - Required
Genetic parameters for Metatrom (if needed to create initial genetic data)			
Target additive variance and tolerance around 0 on the mean genetic value, for each trait	Inventory	- No default values - Required	- No default values - Required
Target heritability or environmental variance and ratio of the "inter-step" environmental variance that fluctuates between steps, for each trait	Inventory	- No default values - Required	- No default values - Required
Target neutral F_{ST} and target Q_{ST} for each trait (if multiple populations)	Inventory	- No default values - Required	- No default values - Required
Threshold p-values for HW test within-each population	*Inventory*	*0.01*	*0.01*

---

[23] High stress levels with : ashape=0.34; bshape=1.3; cshape=3.74; ascale=-189; bscale=-3.5; cscale=26.31

Annex 2 – Examples of genetic values for the three traits

Values of within-stand phenotypic variance and ratio of "inter-step" environmental variance given below for vigor and sensitivity to growth disturbance were inferred from the data analysed in Fririon et al. (2023). Values of sensitivity to diffuse mortality disturbance are arbitrary. The choice of heritability is arbitrary but in line with literature review.

Due to computational constraint, the model uses integer values of allelic effects. Appropriate scaling of QTL effects is obtained by using the *effectCoefficient* parameter. The scaling *effectCoefficient* must be provided by the user in the inventory file, if not provided a default value 1 is considered but this is not appropriate if Metatrom is used to generate genetic data. In its current version, Metatrom systematically multiplies the simulated allelic effects by 10^3 which should provide a floor integer value > 0 in many cases: in this case the *effectCoefficient* should be put to 10^{-3} to rescale allelic effects appropriately. But this might not be enough for traits with low values and, in a more general situation, the target value of additive variance (V_A) given to Metatrom should first be unscaled by a coefficient 10^{2k} in order to obtain a target variance higher than 10^{-2}, and then, re-scaling is obtained by using the *effectCoefficient*=$10^{-(k+3)}$. In the examples below k=2 for vigor, k=3 for sensitivity to growth disturbance, and k=1 for sensitivity to diffuse mortality disturbance.

	Vigor	Sensitivity to growth reduction disturbance	Sensitivity to diffuse mortality disturbance
Within-stand phenotypic variance V_P	0.14	$2\ 10^{-6}$	$3.33\ 10^{-4}$
Ratio of the "inter-step" environmental variance	0.13	0.2	0.2
Supposed heritability h^2	0.3	0.3	0.3
Within-stand additive variance $V_A=h^2V_P$	0.0042	$6\ 10^{-7}$	10^{-4}
Unscaled target V_A provided to Metatrom	42	0.6	0.01
Re-scaling allelic effects (effectCoefficient)	10^{-5}	10^{-6}	10^{-4}