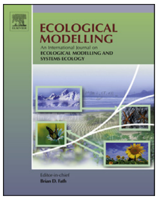
Ecological Modelling 314 (2015) 1–14

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Applying ecological model evaludation: Lessons learned with the forest dynamics model Samsara2

B. Courbaud[a,*], V. Lafond[a], G. Lagarrigues[a], G. Vieilledent[b], T. Cordonnier[a], F. Jabot[c], F. de Coligny[d]

[a] Irstea, UR EMGR, Centre de Grenoble, F-38402 St-Martin-d'Hères, France
[b] CIRAD, UPR BSEF, F-34398 Montpellier, France
[c] Irstea, UR LISC, F-63172 Aubière, France
[d] INRA, UMR AMAP, F-34000 Montpellier, France

ARTICLE INFO

Article history:
Received 26 February 2015
Received in revised form 25 June 2015
Accepted 26 June 2015

Keywords:
Validation
Evaluation
Sensitivity analysis
Calibration
ABC
Individual-based model

ABSTRACT

Ecological models are increasingly used as decision-making tools and their reliability is becoming a key issue. At the same time, the sophistication of techniques for model development and analysis has given rise to a relative compartmentalization of model building and evaluation tasks. Several guidelines invite ecological modelers to follow an organized sequence of development and analysis steps and have coined the term "evaludation" for this process. The objective of this paper is to assess the feasibility and the value of a structured evaludation process, based on the working example of the Samsara2 model, a spatially explicit individual-based forest dynamics model. We implemented the six steps of model design, process level calibration, qualitative evaluation, quantitative evaluation, global sensitivity analysis, and partial recalibration using approximate Bayesian computing. We then evaluated how the evaludation process revealed model strengths and weaknesses, specified the model's conditions of use, clarified how the model works, and provided insights into forest ecosystem functioning. Finally, the efficiency/cost ratio of the process and future improvements are discussed.

© 2015 Published by Elsevier B.V.

1. Introduction

Confidence in ecological model predictions is becoming a key issue considering the rapid development of their use to forecast the impact of environmental and social changes on ecosystems (Bennett et al., 2013; Clark et al., 2001; Schmolke et al., 2010; Purves and Pacala, 2008; Mouquet et al., 2015).

To increase the confidence in their models, modelers compare model predictions to independent knowledge or observations. For example, in forest dynamics models, it is common to compare the predictions of tree community composition to potential natural vegetation types (Bugmann, 1996; Lexer and Honninger, 2001; Botkin, 1993) or old growth forest plots (Pacala et al., 1996; Ruger et al., 2007), or to use historical records of forest inventories over several decades to compare the evolution of predicted and observed capital (e.g., basal area) (Wehrli et al., 2005; Wehrli et al., 2007) and/or distributions of trees in diameter classes (Seidl et al., 2005; Didion et al., 2009; Wehrli et al., 2005; Purves et al., 2008).

However, comparisons to observations can prove model reliability only within a limited range of cases already observed in the past, whereas we are most often interested in predictions either for new environmental conditions or for new management contexts. Basing a decision on the predictions of a model therefore requires a more detailed assessment of its reliability: the level of confidence of each of the model outputs, the ranges of conditions and time scales within which the predictions are considered reliable, and how the biases and uncertainties detected can be handled need to be determined.

Several guidelines invite ecological modelers to follow an organized sequence of development and analysis steps to check potential development pitfalls and characterize model performance comprehensively (Schmolke et al., 2010; Grimm et al., 2010; Vanclay and Skovsgaard, 1997; Prisley and Mortimer, 2004; Soares et al., 1995; Grimm and Railsback, 2005; Shaeffer, 1980; Rykiel, 1996; Augusiak et al., 2014; Grimm et al., 2014). This process has recently been called "evaludation" (Augusiak et al., 2014) and comprises steps that vary slightly from one author to another but that usually cover at least model design, calibration, prediction evaluation, and sensitivity analysis. This process is intended to increase model credibility by showing that the model has passed several

---

* Corresponding author. Tel.: +33 4 76 76 27 62.
E-mail address: benoit.courbaud@irstea.fr (B. Courbaud).

http://dx.doi.org/10.1016/j.ecolmodel.2015.06.039
0304-3800/© 2015 Published by Elsevier B.V.

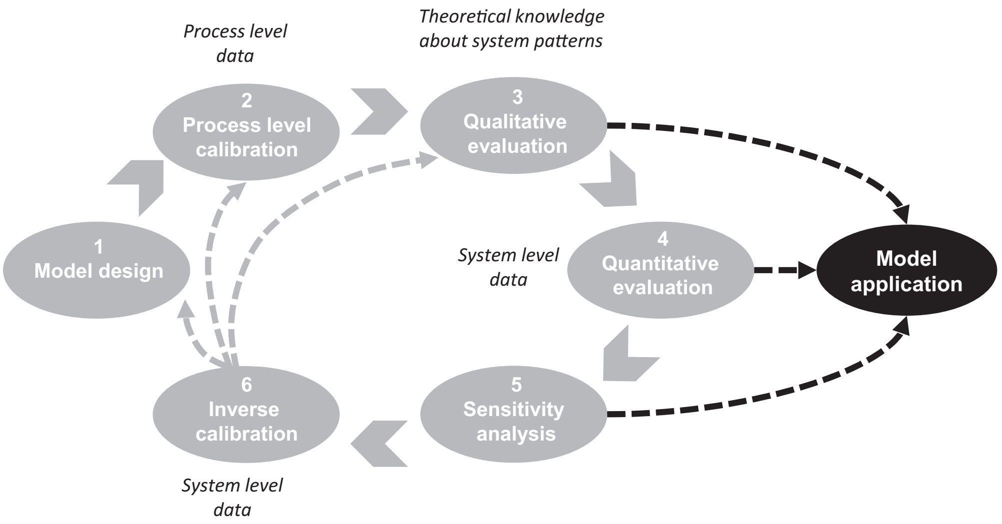


Fig. 1. *Flowchart of the evaludation process followed in this study.* Ellipses linked by plain arrows represent the sequence of evaludation steps followed in this study. Information about the system required at each step is added in italics. The evaludation process is cyclic. Dotted arrows show that it is flexible: some steps can be bypassed, especially if several evaludation cycles of are performed, and steps can be repeated at any moment with different methods.

verification procedures (Vanclay and Skovsgaard, 1997). It should demonstrate whether the model is acceptable for its intended use and assess the level confidence to place in inferences concerning the real system (Rykiel, 1996). Moreover, it should progressively increase our understanding of the process chain that led to the emergence of a model output (Augusiak et al., 2014; Grimm and Railsback, 2005). This last point is essential, because the evaluation process is unable to prove that a model accurately represents a natural system, simply because absolute proofs do not exist in science (Varenne and Silberstein, 2013; Godfrey-Smith, 2003; Stengers and Bensaude-Vincent, 2003). The only thing a model can do is to show what system dynamics can be expected, given the elementary processes that are simulated. Confidence in a model comes not only from the comparison of its predictions to empirical data, but also from the confidence we can have in its structure and in the elementary processes it assembles (Dutreuil, 2013). However, due to the variety of processes and the high number of input variables, complex models tend to rapidly become black boxes for the user. Heterogeneity among basic entities and nonlinearity in different elementary processes often lead to global behaviors that are difficult to predict qualitatively before running the simulations, resulting in the user losing a sense of “understanding” the model’s behavior (Braillard, 2013). A key role of the evaludation process should therefore be to clarify this relation between elementary processes and emergent patterns. Moreover, there is usually a tradeoff between model complexity and usability. If a complex model is used, it is important to show that it results in better predictions or improved generality, or that it can respond to new questions compared to simpler approaches.

These guidelines are nevertheless rather theoretical and rarely provide examples of complete evaludation processes concretely showing the costs and benefits of the approach, despite recent efforts in this direction (Grimm et al., 2014). Indeed, the whole approach is still rarely used (Schmolke et al., 2010). Sensitivity analysis steps are especially often lacking in model presentations (Pappas et al., 2013), limited to a few parameters only (Kunstler et al., 2013), or insufficiently exploited to better understand the models (Ruger et al., 2007). The long process of model improvement leads to the scattering of evaluation results, which makes building an overall vision of the strengths and weaknesses of a model difficult. Indeed, performing all these steps requires data, technical skills, time, and a supportive scientific environment. Ecologists need more working examples of the evaludation process to appraise its costs and benefits, and to apprehend the kind of practical organization required.

The objective of this paper is to demonstrate the feasibility and discuss the value of a structured evaludation process based on the working example of the Samsara2 forest dynamics model. As an individual-based, spatially explicit model, Samsara2 is intended to highlight the links between processes at the individual level and patterns at the community level. The long-term scales of forest dynamics make data for empirical evaluation difficult to collect. Moreover, the long computing time required to calculate competition for light among trees limits the techniques that can be used for model analysis.

After a quick presentation of the model, the six successive steps of the evaludation process followed are detailed: (1) model design, (2) process level calibration, (3) qualitative evaluation, (4) quantitative evaluation, (5) global sensitivity analysis, and (6) recalibration of the most influential parameters (Fig. 1). We then present the results obtained for each of these steps and discuss the advantages and disadvantages of the evaludation process. More specifically, the following questions are addressed: Did the evaludation process increase the credibility of model predictions by highlighting model strengths and weaknesses (Augusiak et al., 2014)? Did it clarify how the model works (Grimm and Railsback, 2005)? Did it specify conditions for using the model safely? Beyond the model itself, did the evaludation process increase our understanding of how individual-level processes drive ecosystem dynamics (Grimm and Railsback, 2005)? Finally, was the evaludation process cost-effective?

2. Method

2.1. Model design

The Samsara2 model is presented in detail in Annex A. Here, we give an overview of its structure and main processes.

The spatially explicit, individual-based Samsara2 model was designed to determine the relationships between stand structure and dynamics in uneven-aged mixed temperate forests and to predict the impact of management strategies (i.e., variations in the distribution of cuttings over time and space and among trees) at the population scale (i.e., a forest stand). This makes it possible to analyze the development of individual trees within a stand and the resulting collective dynamics, summarized by synthetic variables such as density, basal area (sum of tree cross-section areas at 1.30 m in height, a standard forest measure that combines density and tree size), distribution of trees among size classes, indices summarizing the spatial distribution of trees, cumulated harvests, and so on. The specific focus on the dynamics and management of uneven-aged stands composed of trees at different development stages requires the simulation of demographic processes (growth, mortality, and recruitment) simultaneously, as well as interactions among trees of different sizes (e.g., competition). We assumed that light interception by tree crowns is the key driver in uneven-aged stand dynamics because these stands present a strong vertical heterogeneity favoring asymmetric competition between trees and between the canopy and seedlings (Schütz, 1997). In the model, light distribution among trees, irradiance on the ground, and seed dispersion are spatially explicit and their spatial heterogeneity drives the changes in forest structure. Differences in environmental conditions among stands can enter the model indirectly through changes in model parameters. The development of Samsara2 was based on the experience and tools gained from previous model versions improved over several years (Dufour-Kowalski et al., 2012; Gauquelin et al., 2008; Cordonnier et al., 2008; Goreaud et al., 2005, 2006).

A forest stand simulated in Samsara2 is based on a list of trees and a list of saplings that have explicit 3D coordinates on a plot. This plot is attributed a slope and an exposure value, and is divided into ground cells. Trees are characterized by their species, trunk diameter at breast height (DBH), crown dimensions, and location. Seedlings are simply characterized by their species, height, and location. Individual tree crown dimensions are calculated using allometric relationships relating total height, crown base height, crown base radius, and DBH (Vieilledent et al., 2010). The irradiance of each cell under canopy and the amount of radiation intercepted by each adult tree during a growing season are calculated together, in an integrated approach based on light ray interception by crowns in 3D (Courbaud et al., 2003). The annual basal area increment of a tree depends on the amount of radiation intercepted during a growing season. This relationship integrates both an ontogenetic effect (interception depends on tree size) and a competition effect (incident radiations depend on neighbors) on growth. Forest tree mortality is usually broken down into background mortality, linked to competition and senescence, and catastrophic mortality, caused by events such as windstorms, fires, or insect outbreaks. In Samsara2, the mortality submodel simulates only background mortality. The death of a tree is the result of a Bernoulli trial, the probability of mortality depending on DBH and local competition. Because of the small size and complex sampling scheme of our mortality data plots (see Section 2.2), mortality does not depend on light interception but on a simple competition index: the basal area of trees larger than the subject in a 15 m radius. Seed production depends on tree basal area and seeds are dispersed randomly in a circular area of fixed radius around seed producers. Seedling survival is the result of a binomial trial and depends on irradiance under canopy on the corresponding ground cell, with a bell-shaped relationship. This relationship is intended to represent an increase in survival with irradiance when it is low because of a positive effect on photosynthesis and a decrease in survival when irradiance is high because of indirect effects such as increased competition with ground vegetation and increased browsing in open areas. The relative height increment of seedlings depends on irradiance under canopy. When saplings reach an arbitrary height defined by the user, they are recruited as adult trees in the model.

Samsara2 is implemented in the Capsis simulation platform (de Coligny et al., 2003; Dufour-Kowalski et al., 2012) (http://capsis.cirad.fr/), which enables both interactive or automatic simulations and the visualization of simulation results. Harvests can be simulated using specific algorithms (Lafond et al., 2012, 2014). Typically, the area of a simulated stand is between 1 and 4 ha and ground cell area is 25 m^2 (5 m × 5 m). Radiation interception, which is the process requiring most of the computing time, is usually updated only every 5 years, whereas demographic processes are calculated on an annual basis. For a 1-ha area and average stand densities, a few minutes are typically required to calculate a 50-year simulation. Simulations are usually performed with fixed parameter values, but it is advised to run a simulation several times to take model stochasticity into account.

2.2. Initial independent calibration of the different processes

A common problem when calibrating complex ecological models is to find consistent data sets that inform all the processes included in the model. In the present case, the three growth, mortality, and regeneration processes represented different sampling constraints. The calibration of the growth submodel required large plots to compute the radiation intercepted by each tree with limited border effects. Mortality is a rare event that required very large tree samples available only in large-scale inventories. Seed production is a very fluctuating process in time and space that required at least a decade of annual observations. Seedling survival required specific small observation plots spanning a wide range of irradiance conditions.

Since these data were not available in a single setting, we decided to use data and results from previous studies to calibrate each process of the model separately (Annex 1, Table A.1). All the data sets came from uneven-aged forests belonging to comparable vegetation types: mixed forests dominated by Norway spruce (*Picea abies*) and European fir (*Abies alba*). We used equations and parameters obtained previously for light interception (Courbaud et al., 2003), allometries (Vieilledent et al., 2010), and growth (Vieilledent, 2009) in the northern French Alps. Tree crowns were measured two different years and increment cores provided 25 measurements of annual increments per tree. In addition, the submodel of mortality was estimated with data from the French National Forest Inventory, the seed production and seedling survival submodels with data from a large observation experiment conducted in Bavaria, Germany (Burschel et al., 1992) from 1976 to 1986, and the sapling growth submodel with observations made in forest gaps in the Carpathian mountains (Stancioiu and O'Hara, 2006). Management of these forests ranged from single tree selection to shelter-wood and provided a wide range of light conditions at the tree scale.

Allometries and growth were calibrated in a hierarchical Bayesian framework using the Winbug software (Gilks et al., 1994), and the mortality and regeneration processes were calibrated using the generalized linear model (glm) function, package {stats}, in R (R Core Team, 2014). Each demographic process was represented with rather simple equations (mostly power functions or logistic functions). Still, this led to a total of 21 calibrated parameters per species.

2.3. Qualitative evaluation versus natural population patterns

To assess the model's ability to produce qualitatively consistent population dynamic patterns emerging from tree interactions, the natural development of a forest stand following a major disturbance was simulated. The objective of this step was to analyze whether a consistent collective behavior would emerge at the 4-ha scale despite the stochastic components of the model (random regeneration and mortality events, random individual differences among trees in their allometries, and growth response to light). We considered an initial 4-ha stand, corresponding to a dense mixture of 3000 small trees per hectare (DBH between 2.5 and

7.5 cm, 50% spruce and 50% fir). We then ran simulations over 1000 years, first with only background individual mortality and second with additional minor disturbances creating gaps of about 3000 m^2 every 150 years, using an algorithm designed to simulate a variety of uneven-aged management strategies (Lafond et al., 2014). These intermediate-size disturbances are assumed to be representative of wind-induced disturbances that are frequent in European mountain forests and are thought to play a key role in the maintenance of the spruce–fir mix balance in these forests (Motta et al., 2011). The changes in different stand-level variables such as canopy cover, basal area, density, DBH distribution, spatial distribution, and species composition were then estimated, using a Monte Carlo approach to extract consistent patterns of variable changes despite the variability between simulation runs due to stochasticity. We made 100 realizations of each simulation using mean parameter estimates. We then estimated the predictive interval of each variable as the 2.5% and 97.5% quantiles of the different realizations. We then compared our simulation results to the patterns of natural forest development described in the literature based on the observation of forest stands of different ages, especially in Bormann and Likens (1979) and Oliver and Larson (1996). In addition, the levels of basal area obtained in the simulations were compared to field measures made in spruce–fir natural forests (Assmann, 1970; Korpel, 1982; Leibundgut, 1982).

2.4. *Quantitative evaluation versus stand level data in a managed forest*

To assess and quantify potential biases in model predictions in the applied context of uneven-aged forest management, we compared the simulated and observed dynamics of 22 spruce–fir forest stands on retrospective simulations over 50 years. These stands belong to the Queige forest, a spruce–fir forest managed with a single tree selection system at the montane level (670–1930 m) in Savoy, France. Data were composed of the distributions of spruce and fir trees per DBH class in 1931, 1949, and 1980, and of annual harvest reports indicating the total number of trees and total volume harvested. We reconstructed initial stands in 1931 and simulated their dynamics up to 1980, using an algorithm specifically designed to reproduce past management interventions (Lafond et al., 2012). We performed 100 realizations of each simulation using mean parameter estimates to check prediction consistency again despite model stochasticity. This time, model stochasticity covered not only demographic stochasticity, but also stochasticity in the selections of cut trees by the harvesting algorithm. Greater detail on the simulations and data sets is provided in Appendix B.

We assessed the capacity of the model to predict several stand characteristics: basal area (G, m^2/ha), density (N, t/ha), quadratic mean diameter (Dg, cm), tree species composition (MixR, % of spruce over total G), and size diversity (GiniG, Gini index on individual basal areas), all these variables being calculated for trees with DBH greater than 17.5 cm. In addition, the recruitment process was evaluated through pole tree density (PoleN: density of trees with DBH between 12.5 and 17.5) and pole composition (PoleMixR, % of spruce in poles). For each of these variables, we compared the predicted and observed values for 1980 and calculated the mean difference (on all repetitions), as well as bias, precision, and accuracy (see definitions and result tables in appendix B). We also compared predicted and observed diameter distributions using graphical comparisons (see Fig. 2, see also Appendix B-III.1) and chi-square tests (see Appendix B.2, Tables 12 and 13) (Reynolds et al., 1988). Chi-squares were compared after grouping trees in only three size categories (small, medium, and large trees) to avoid biases due to classes with low frequency. We applied this test both for each simulated distribution and the distribution averaged over the 100 realizations of a simulation.

2.5. *Sensitivity analysis in the range of uncertainty of parameter estimates*

A sensitivity analysis was undertaken to detect a potential impact of parameter estimate uncertainty on model predictions, assess the relative influence of these parameters, and detect key parameters that should be either recalibrated or considered as factors in simulation experiments. We based this step on the same retrospective data as in the quantitative evaluation step, but varied parameter values within the 95% confidence intervals of their estimates. We used four initial stands representing the observed stand diversity in our historical data set. These stands were chosen along an axis of stand heterogeneity (spruce-dominated stands with a nearly even-aged structure vs stands with a balanced species mix and an uneven-aged structure) and a density axis (low or high), based on the results of a principal component analysis on stand characteristics (Appendix C).

Because of the long computing time required by the model and the large dimension of the factorial space, we used the Morris method (Morris, 1991; Campolongo et al., 2007) that efficiently samples the factorial space with a limited number of combinations of factor values. The Morris method selects combinations of parameters encountered along r random paths by making $p + 1$ one-at-a-time discrete jumps (p is the number of parameters and only one parameter is varied at each jump) within the factorial space. We assessed the sensitivity of seven output variables to the variation of the 42 demographic parameters (21 per species) and two initial state factors (stand heterogeneity and density). We used $r = 100$ trajectories to explore the factorial space defined by these $p = 44$ input factors, leading to $r \times (p + 1) = 4500$ unique combinations of parameters (Appendix C). The size of the "jumps" depended on the range of variation and the number of levels defined for each parameter: four levels for demographic parameters within their 95% calibration confidence interval; two levels for allometry or growth individual effects (with/without individual effect); and two levels (low/high) for initial state factors. We used the *morris()* function of the "sensitivity" R package (Pujol et al., 2014) to generate an efficient experimental design (Campolongo et al., 2007) and to compute the sensitivity indices. W integrated the two Morris sensitivity indices (μ^* and σ) into a single sensitivity index, as suggested by (Ciric et al., 2012). This index was then used to rank the factors according to their respective influence on each output variable and to establish an overall ranking based on all output variables (Table 1).

2.6. *Partial recalibration using approximate Bayesian computation*

We finally added a recalibration step in the evaludation process in order to reduce the uncertainty on parameter estimates, calibrate jointly the growth and regeneration processes, which had been calibrated separately in step 1; and obtain local parameter estimations for different stands. We used the approximate Bayesian computation (ABC) technique (Beaumont et al., 2002) and once again based the simulations on our stand level historical Queige forest data set. A detailed presentation of this step can be found in (Lagarrigues et al., 2014). Because the structure of a population observed at a given time is the aggregated product of different demographic processes, there is a risk of factor confusion when trying to infer too many individual process parameters from these data. We therefore decided to focus the recalibration on a small number of influential parameters involved in complementary processes: one growth parameter (G_A) and one regeneration parameter (S_A) per species. The determination of the number of parameters that could be reliably recalibrated was obtained by a model selection procedure (Lagarrigues et al., 2014). We chose uniform prior distributions,

Table 1

Results of parameter calibration and sensitivity analysis Left: for each demographic and allometric parameter, the mean value and associated SE (standard error of the mean) estimates from the initial calibration. Right: for each output variables (*in columns*), the rank obtained by each input factor (*in lines*), based on the comparison of the sensitivity indices proposed by (Ciric et al., 2012). The last column indicates the overall order, based on the sum of the ranks obtained by each factor for all output variables. In addition to the factors being ranked, groups of factors have been identified depending on their relative degree of influence: very high (*dark gray*), high (*gray*), medium (*light gray*), and low or inexistent (*white*). Factors encompass demographic and allometric parameters for each species (0 for spruce and 1 for fir) as well as the two "initial state" factors (IS_F1 and IS_F2, respectively representing a heterogeneity and a density gradient. Demographic parameters were sampled along four levels in their uncertainty range (±1.96 SE around the mean values), except (*) for the individual variability parameters (in *gray*), which took only two levels: with or without individual variability. The output variables studied included stand density (*N*, t/ha) and basal area (*G*, m^2/ha), tree mean quadratic diameter (*Dg*, cm), stand diameter diversity (GiniG, no unit) and composition (MixR, % of spruce over total basal area), total pole density (PoleQty, t/ha), and the pole mixing ratio (PoleMixR, % of spruce poles over total pole density).

Samsara2 parameters initial calibration					Parameter rank regarding output sensitivity to variations within ±1.96 SE around the mean[*]								
	Factors	Definition	Mean	SE	*N*	*G*	*Dg*	GiniG	Pole Qty	MixR	Pole MixR	Rank Sum	Global rank
Growth	G_A_0	Spruce growth potential	−3.97E + 00	4.25E − 01	2	1	1	3	4	1	5	17	**1**
	G_A_1	Fir growth potential	−2.02E + 00	2.02E − 01	7	4	7	6	8	5	7	44	**6**
	G_b_0	Light effect on spruce	5.88E − 01	3.80E − 02	3	2	2	1	6	2	9	25	**2**
	G_b_1	Light effect on fir	4.16E − 01	1.90E − 02	8	5	5	5	11	4	27	65	**8**
	G_sB_0	SD of spruce log indiv. effects	6.00E − 01	4.92E − 02[*]	38	36	35	35	43	36	42	265	**37**
	G_sB_1	SD of fir log indiv. effects	6.02E − 01	5.39 E − 02[*]	39	44	36	39	39	40	40	277	**40**
	G_sL_0	SD of spruce temp. effects	1.05E − 01	1.91E − 02[*]	43	33	38	36	36	39	39	264	**36**
	G_sL_1	SD of fir temp. effects	1.84E − 01	2.98E − 02[*]	41	42	41	40	38	38	30	270	**38**
Crown B. Height	CBH_mK_0	Spruce log slope parameter	4.88E − 01	8.45E − 02	30	21	33	34	14	31	24	187	**32**
	CBH_mK_1	Fir log slope parameter	7.14E − 01	6.16E − 02	31	35	32	24	21	25	26	194	**34**
	CBH_sK_0	SD of spruce log indiv. effect	6.16E − 01	4.14E − 02[*]	37	40	37	43	40	37	37	271	**39**
	CBH_sK_1	SD of fir log indiv. effect	4.62E − 01	3.73E − 02[*]	35	43	43	42	44	44	36	287	**42**
Crown B. Radius	CBR_b_0	Spruce shape parameter	5.25E − 01	1.20E − 02	17	12	20	29	9	22	31	140	**20**
	CBR_b_1	Fir shape parameter	4.54E − 01	1.33E − 02	18	27	26	17	15	12	35	150	**21**
	CBR_mK_0	Spruce log scale parameter	−7.74E − 01	4.48E − 02	24	34	17	13	12	33	28	161	**23**
	CBR_mK_1	Fir mean log scale parameter	−3.54E − 01	4.30E − 02	22	28	15	9	33	17	15	139	**19**
	CBR_sK_0	SD of spruce log indiv. effect	1.58E − 01	1.34E − 02[*]	36	39	40	38	35	35	38	261	**35**
	CBR_sK_1	SD of fir log indiv. effect	1.22E − 01	1.14E − 02[*]	44	37	39	37	42	42	41	282	**41**
Height	H_mK_0	Spruce log max height	3.53E + 00	2.84E − 02	23	20	21	22	13	7	21	127	**13**
	H_mK_1	Fir log max height	3.28E + 00	3.59E − 02	26	29	22	15	27	26	23	168	**25**
	H_r_0	Spruce shape parameter	7.67E − 02	1.45E − 03	33	25	23	27	32	8	29	177	**29**
	H_r_1	Fir shape parameter	8.46E − 02	1.55E − 03	32	26	13	21	24	15	33	164	**24**
	H_sK_0	SD of spruce log indiv. Effects	1.04E − 01	1.12E − 02[*]	42	41	44	41	41	41	44	294	**44**
	H_sK_1	SD of fir log indiv. effects	1.77E − 01	2.10E − 02[*]	40	38	42	44	37	43	43	287	**42**
Regeneration (Saplings)	S_A_0	Spruce basal area effect	7.68E + 01	2.62E + 01	4	6	8	7	2	9	2	38	**5**
	S_A_1	Fir basal area effect	8.91E + 00	3.72E + 00	1	8	6	8	5	6	1	35	**4**
	S_B1_0	Spruce sapl. survival constant	−9.28E + 00	3.15E − 02	29	15	25	31	29	28	32	189	**33**
	S_B1_1	Fir sapl. survival constant	−6.11E + 00	2.67E − 02	21	32	18	19	16	11	34	151	**22**
	S_B2_0	Spruce light effect on sapl. survival	2.40E − 01	1.41E − 03	34	22	34	25	19	32	12	178	**31**
	S_B2_1	Fir light effect on sapl. survival	1.87E − 01	1.26E − 03	28	31	29	20	25	19	19	171	**27**
	S_B3_0	Spruce sq-light effect on sapl. survival	−1.89E − 03	1.23E − 05	25	16	30	28	30	27	18	174	**28**
	S_B3_1	Fir square light effect on sapl. survival	−1.50E − 03	1.15E − 05	20	19	16	18	20	16	22	131	**15**
	S_R1_0	Spruce sapling growth constant	4.65E + 00	1.51E − 02	11	10	12	32	7	20	4	96	**10**
	S_R1_1	Spruce sapling growth constant	4.55E + 00	1.32E − 02	27	24	31	33	17	23	14	169	**26**
	S_R2_0	Spruce light effect on sapl. growth	4.92E − 02	4.46E − 03	10	18	9	10	10	18	3	78	**9**
	S_R2_1	Fir light effect on sapl. growth	2.70E − 02	3.71E − 03	15	23	10	11	18	30	20	127	**13**
Mortality	M_P1_0	Spruce mortality constant	−3.60E + 00	1.35E − 01	9	13	24	23	31	24	11	135	**17**
	M_P1_1	Fir mortality constant	−3.21E + 00	2.54E − 01	12	14	19	14	22	13	17	111	**11**
	M_P2_0	Spruce dbh effect on adult mortality	−1.25E − 02	4.02E − 03	13	9	28	16	23	21	25	135	**17**
	M_P2_1	Fir dbh effect on adult mortality	−3.01E − 02	7.76E − 03	14	17	11	12	34	29	16	133	**16**
	M_P3_0	Spruce comp. effect on adult mortality	2.20E − 02	2.73E − 03	19	30	27	26	28	34	13	177	**29**
	M_P3_1	Fir comp. effect on adult mortality	2.28E − 02	5.36E − 03	16	11	14	30	26	10	10	117	**12**
Ini. State	IS_F1	Stand heterogeneity factor			5	3	3	2	1	3	8	25	**2**
	IS_F2	Stand density factor			6	7	4	4	3	14	6	44	**6**

[*] Parameters related to individual variability are not varied like other parameters but only take two levels: with/without individual variability.

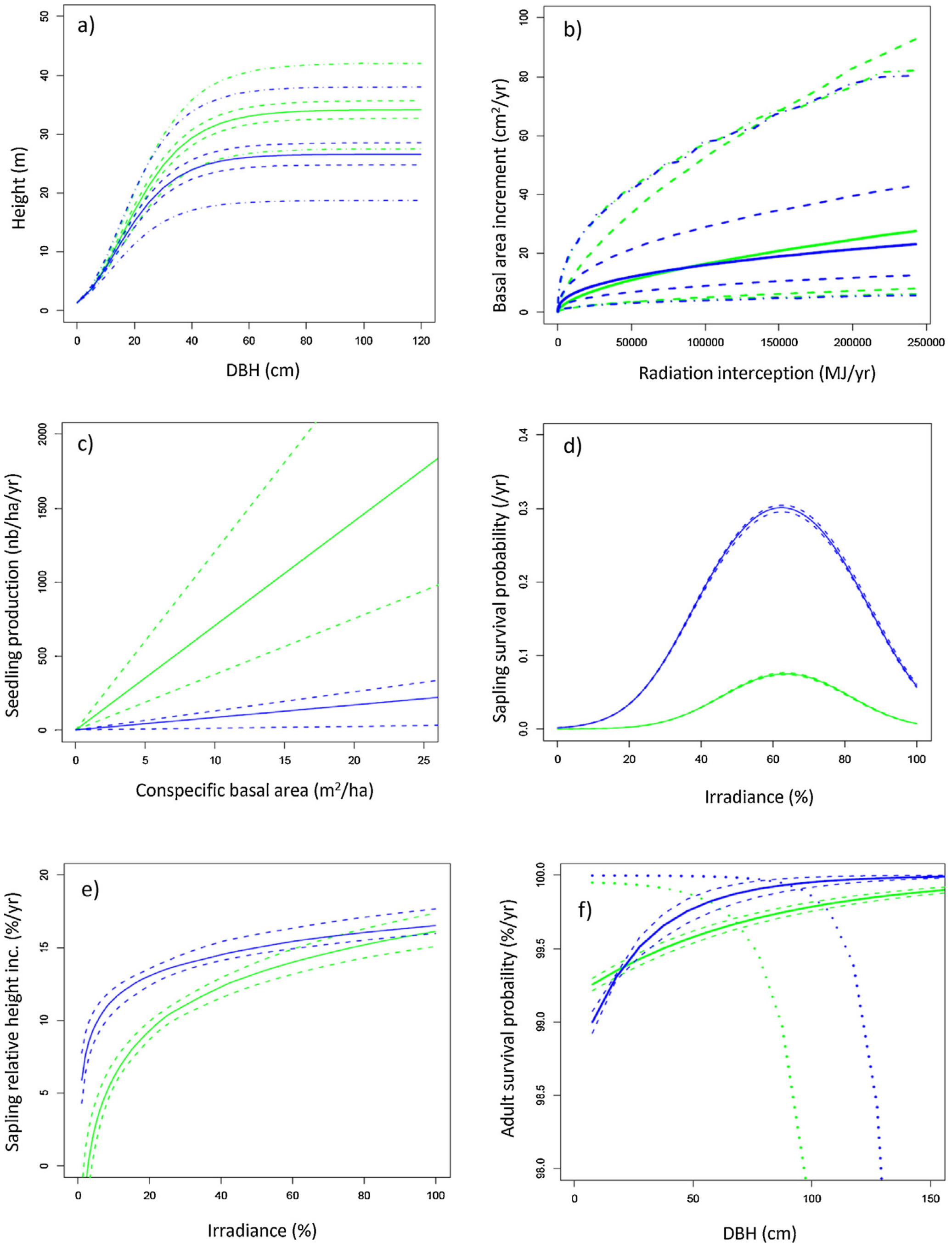


Fig. 2. *Demographic processes simulated in the Samsara2 model.* Green lines represent spruce and blue lines represent fir. Plain lines: mean relation, dashed lines: 95% confidence interval of the mean, dotted-dashed lines: 95% confidence intervals of individual effects. (a) Allometry between DBH and total tree height. (b) Relationship between the individual tree's annual basal increment and radiation interception during the vegetation season. (c) Relationship between seedling production and conspecific basal area, adapted to the alpine context. (d) Relationship between sapling survival probability and irradiance under canopy. (e) Relationship between sapling annual relative height increment and irradiance under canopy. (f) Relationship between adult survival probability and DBH. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

centered on the means of the estimates of the initial calibration and with a width ten times their standard deviations, so as to enable the ABC method to explore the parameter space widely around the initial calibration estimates. The ABC method consisted in iterating the three following processes: (1) values for the four parameters of interest were sampled randomly in their prior distributions, (2) a simulation was made using this parameter set, with the other parameters set at their default value, and (3) summary statistics were computed to compare simulation predictions to stand-level data in 1949 and 1980. We used the ten values of tree densities per 5-cm DBH class between 20 and 65 cm for each species, the total basal area of all trees with a DBH over 67.5 cm and the total volume of harvested trees as summary statistics and ran 5000 simulations for each stand. The 5% best simulations were selected and the corresponding parameter values extracted. In the basic ABC methodology, this set of accepted parameter vectors provides an empirical representation of the joint posterior parameter distribution. In addition, local regression techniques proposed in the "abc" R

Package were used (Csillery et al., 2012) to improve the accuracy of these distributions (Beaumont et al., 2002). We then evaluated the model with these new parameter estimates, following the quantitative evaluation process again (step 4, Fig. 1).

3. Results

3.1. Initial calibration using tree and patch level data

The detailed calibration results are given in Annex A and parameter estimates are summarized in Table 1.

Crown shapes were rather similar for the two species: large spruce trees were slightly higher than large fir trees (maximum height 95% CI 33–36 m for spruce and 25–29 m for fir; Fig. 2a), but mean crown base height and crown base radius for the same DBH did not significantly differ. In contrast, variability of crown size and shape was very high within each species, in relation to large individual effect variances; see also (Vieilledent et al., 2010). Variations among plots were taken into account by including a site effect. For the sake of homogeneity with other processes, we considered parameter estimates corresponding to the Queige plot (1358 m) as a reference.

Tree crown porosity was the only parameter to calibrate in the radiation interception model. We estimated leaf area density at 0.4 m^2/m^3 in the Les Arcs plot (Courbaud et al., 2003), corresponding to a crown porosity of about 20%. As discussed in (Ligot et al., 2014), this means that a ray is almost extinct after crossing a crown and that the distribution of radiation depends predominantly on tree spatial distribution and crown size.

Annual tree basal area increments increased as a convex power function of the amount of radiation intercepted during the vegetation season (Fig. 2b) (Vieilledent, 2009). Individual effects produced huge incremental differences among same-species individuals for the same amount of intercepted light. In contrast, the mean responses were not significantly different between spruce and fir, leading to a huge overlap among species: for example, for an average interception of 50,000 MJ/year, the 95% CI of individual tree basal area increments was 4–39 cm^2/year for fir and 3–35 cm^2/year for spruce.

The number of viable seeds per unit area was satisfactorily represented by a linear relationship to conspecific basal area in the surrounding stand (Fig. 2c). Spruce produced significantly more seeds than fir, despite the high uncertainty of the relationships. Seedling survival reached a maximum at around 60% full light for both species (Fig. 2d). A logistic model of seedling survival depending on light and squared light under canopy adjusted well to the data. Seedling survival was much greater for fir than for spruce at all light levels. Seedling density predicted with the parameters calibrated on Bavarian conditions appeared proportional to small tree density values observed on Alpine plots, but too high because of differences in stage (10-year-old seedlings in Bavaria, Trees 7.5-cm DBH in the Alps), and differences in environmental conditions. We therefore calculated a reduction factor that could adapt the model to the Alpine context. The relative height increment of seedlings increased as a convex response of light above canopy (Fig. 2e). Spruce and fir had about the same relative height growth at full light, around 15% per year. However, for light levels lower than 50%, the mean relative increment was significantly larger for fir than for spruce.

For small and medium-size trees, the probability of survival increased as DBH (Fig. 2f) and local competition increased. It was modeled using a logistic relationship of these variables. The probability of mortality appeared significantly higher for spruce than fir for most DBH levels and competition conditions. We considered that very large trees were insufficiently represented in the data set to calibrate their survival empirically, assuming instead that large tree survival decreased as a quadratic function of DBH (Fig. 2f) and used literature values of maximum spruce and fir DBH to estimate this function (Bugmann, 1996). The probability of mortality of a given tree was set at the maximum between the two functions (function shift around a DBH of 70 cm for spruce and 90 cm for fir).

3.2. Qualitative evaluation: Ability to simulate natural population dynamic patterns

The 1000-year simulation of stand dynamics without disturbance clearly showed the development stages described in (Oliver and Larson, 1996) (Fig. 3a). The starting stand represented an initial stage with a high density of small spruce and fir trees, randomly distributed in space. After 25 years, the stand entered the stem exclusion stage, characterized by an increase in basal area (Fig. 3a—stage 2), a phenomenon of self-thinning, and spatial regularization. During this stage, the log-density and the log-mean volume evolved approximately along a straight line, as expected from the self-thinning law (Pretzsch, 2006). After 200 years, an understory reinitiation stage started (Fig. 3a—stage 3) with a decrease in basal area and the creation of gaps due to the mortality of large trees. Radiation reached the ground in these gaps and clumps of seedlings started to develop. The spatial distribution of trees evolved towards more spatial aggregation and tree size distribution evolved towards bimodality. After 500 years, the stand had reached an old growth stage (Fig. 3a—stage 4), with relative stabilization of basal area, a reverse j-shaped distribution of sizes (Fig. 3b), and a spatial mosaic with randomly distributed large trees and small trees scattered in the understory (Fig. 3c). During these four stages, we observed a vegetation succession with the progressive elimination of spruce by the more shade-tolerant fir (Fig. 3a–c). Without disturbances, the stand basal area reached a peak at 90 m^2/ha and stabilized around 70 m^2/ha (Fig. 3a), which appeared consistent with the respective values of 71.4–110 m^2/ha and 47.1–62 m^2/ha reported by several authors in spruce–fir forests (Motta et al., 2011; Leibundgut, 1982; Korpel, 1982). However, the exclusion of spruce did not match with the balance between species usually observed in near-natural forests (Motta et al., 2011; Leibundgut, 1982; Korpel, 1982).

When disturbances were simulated in addition to baseline mortality, the dynamics switched to alternating stand collapse and regrowth phases (Fig. 3d), a pattern that recalls the fluctuations observed in European natural mountain forests subject to periodic wind disturbance (Korpel, 1982). Basal area varied within a 40- to 80-m^2/ha range. This time, spruce survived over the long term, even if at rather low densities (Fig. 3d–f). Both spruce and fir displayed j-shaped diameter distributions; however, at the 4-ha scale the relative proportion of small, medium, and large spruce trees varied between years within a collapse-regrowth cycle. Clumps of spruce regeneration appeared on cells receiving high irradiance because of canopy gaps (Fig. 3f). Fir regeneration also benefitted from canopy gaps but to a lesser degree than spruce, in agreement with field observations (Grassi et al., 2004; Stancioiu and O'Hara, 2006). This relative balance among species appears closer to observations in near-natural forests (Motta et al., 2011; Leibundgut, 1982; Korpel, 1982).

Both with and without disturbance, simulation stochasticity (due to random regeneration and mortality events, random individual differences among trees in their allometry values, and growth response to light) was relatively low on 4-ha plots, as shown by the relatively narrow predictive confidence intervals calculated over the different simulation runs on basal area dynamics (Fig. 3a and d) and DBH distributions (Fig. 3b and e).

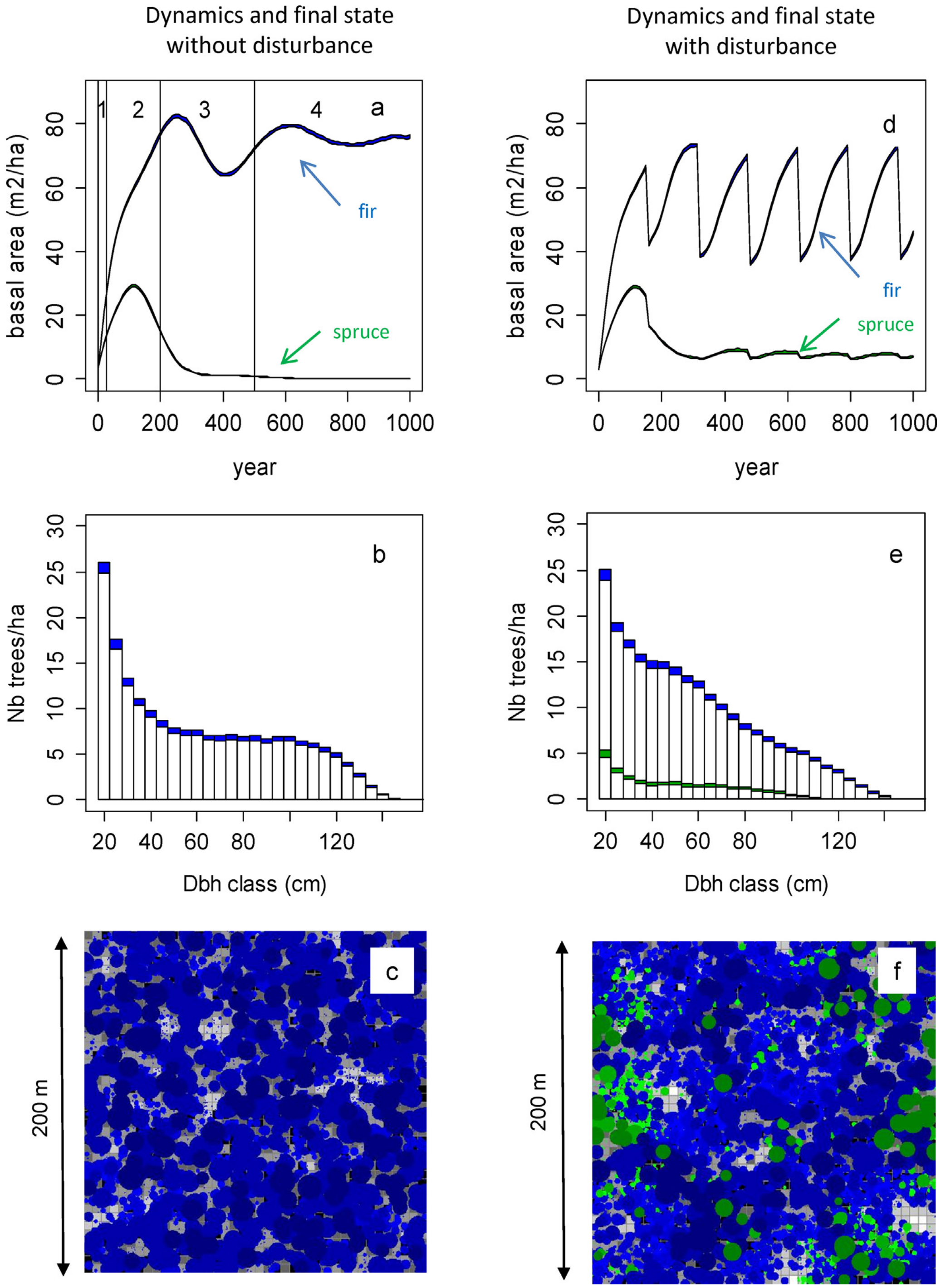


Fig. 3. *Simulated natural dynamics of a mixed spruce–fir stand over 1000 years.* (a) Dynamics of basal area on 4-ha plots, without disturbance. Blue and green lines represent fir and spruce respectively. Line thickness represents 95% confidence interval of the mean of 100 realizations of the simulations. Vertical lines indicate approximate limits between the four development stages: 1: initiation, 2: stem exclusion, 3: understory re-initiation, and 4: old growth. (b) Distribution of tree diameters at year 1000, without disturbance. Boxes indicate the 95% confidence interval of the mean of 100 realizations, in blue for fir and green for spruce. (c) Stand map at year 1000, without disturbance (one realization). Circles represent crown projections on the ground. Blue and green intensity represents categories of height, respectively, for fir and spruce. Grey squares represent levels of irradiance on the ground. (d) Dynamics of basal area on 4-ha plots, with disturbances. (e) Distribution of tree diameters at year 1000, with disturbances. (f) Stand map at year 1000, with disturbances. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. *Quantitative evaluation: Detection and quantification of prediction bias*

The analysis of the retrospective simulations based on historical records and the comparison of model predictions to field observations provided quantitative assessments of model performance. The predictive ability of the model depended on both the forest stand and the output variable considered (Fig. 4). Stand basal area (G, Fig. 4a) was relatively well predicted after 50 years, with an average prediction bias of +6.9% (mean difference, +0.96 m^2/ha) and nonsignificant differences between predicted and observed values for 19 of the 22 stands (*t*-test with 5% confidence; see Appendix B). More discrepancies between predictions and observations appeared when considering mean tree diameter (*Dg*, Fig. 4b).

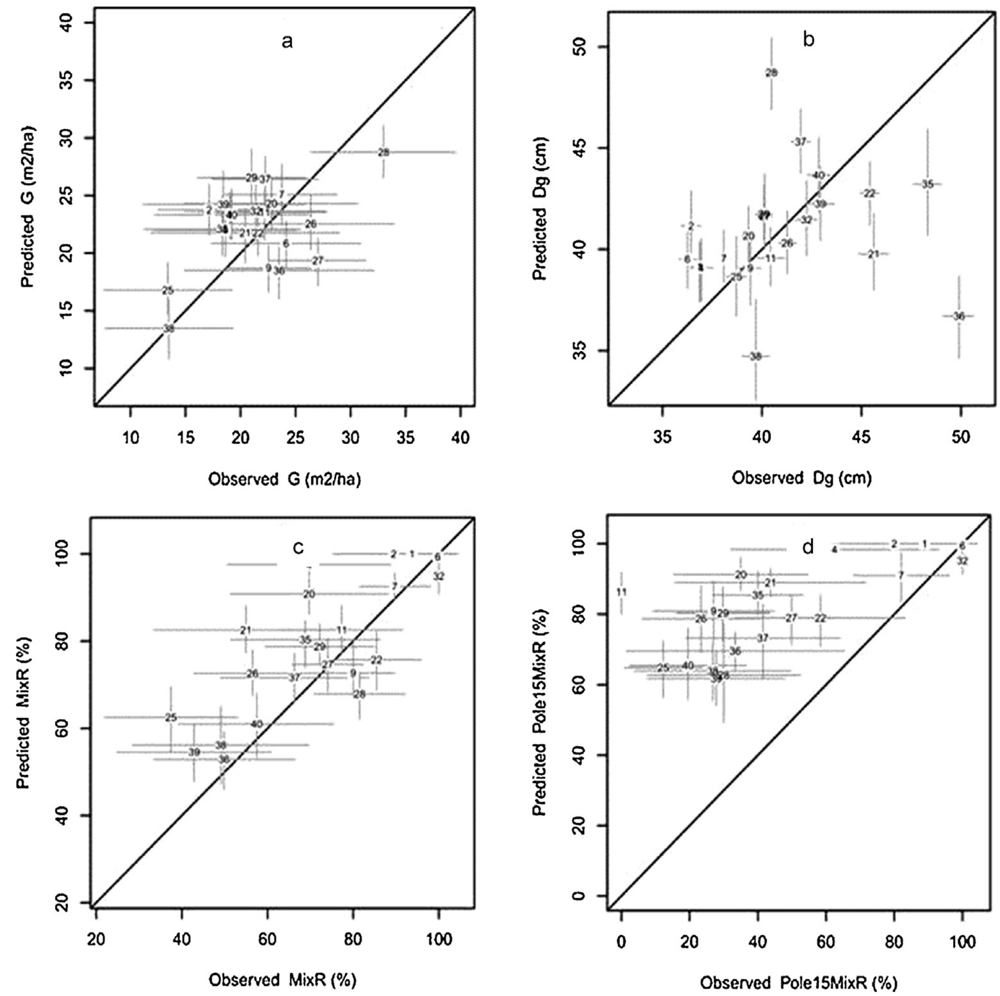


Fig. 4. *Comparison of predicted vs observed stand structure after 50 years.* Predicted vs observed values for basal area (*G*, m^2/ha), quadratic mean diameter (*Dg*, cm), spruce proportion over total basal area (MixR, %), Spruce proportion over total high pole numbers (i.e., trees in 15-cm DBH class) (Pole15MixR, %). Except for Pole15MixR, all stand characteristics were calculated on adult trees (diameter ≥ 17.5 cm). Values are indicated for each stand, with *stand number* positioned at coordinates corresponding to the mean observed (*x*) and mean predicted (*y*) values. Grey bars indicate 95% confidence intervals of predictions (*vertical bars*) and observations (*horizontal bars*).

The average bias was only 0.14% (mean difference, −0.20 cm), but the bias was significant for half of the stands, either positive or negative depending on whether the model tended to overestimate or underestimate the density of small, medium, or large trees. This result thus indicated that discrepancies in stand structure could hide behind good predictions of stocks.

The graphical analysis of the tree size (DHB) distributions (Fig. 5 and Appendix B-III-1) helped explain these biases, because it detected discrepancies between predicted and observed densities in each size class. A strong overestimation of the number of trees in the smallest size classes occurred in most stands (e.g., P11 in Fig. 5), pointing out an overestimation of the regeneration flow. The other size classes were quite well represented for most stands, with a few isolated discrepancies in some DBH classes (e.g., stand P11, Fig. 5). However, some stands showed larger or systematic discrepancies (the worst case being stand P28, Fig. 5), without any emerging common pattern. Chi-square tests were used to quantify this bias and revealed that the mean predicted distribution (average over the 100 runs) was identical to the observed distribution for 7 of the 22 stands (see Table 13 in Appendix B-III-2). For the 15 other stands, the predicted DBH distributions were systematically different from the observed distribution for only five stands (with 0% identical distributions over the 100 repetitions), while three stands had 19%, 28%, and 45% repetitions with good prediction of the DBH distribution (see Table 13 in Appendix B-III-2).

Species composition was relatively well predicted after 50 years when considering adult trees only (MixR; Fig. 4c), with a mean bias of the proportion of spruce basal area of +13.2% (mean difference, +7.1%). The overestimation of the spruce proportion was much higher for poles in almost all stands (Pole15MixR; Fig. 4d), with a mean bias of +114.1% (mean difference, +35.6%).

3.4. Sensitivity analysis: Identifying key parameters that drive prediction variability

The initial state (variations in initial stand density and heterogeneity) had a very strong and persistent effect on adult tree density (*N*), pole density (PoleQty), pole mixing ratio (PoleMixR), and diameter diversity (Gini) after 50 years of simulation (Table 1, bottom). The influence was more limited on basal area (*G*).

Three demographic parameters per species (six in total) had a high or very high influence on several stand structure variables (Table 1): the two parameters controlling growth (the scale parameter G_A and the shape parameter G_b), and the parameter driving seed production (S_A). Variations in the growth parameters G_A and G_b strongly affected all output variables, whereas variations of the regeneration parameters S_A had a strong effect only on stand density (*N*), pole density (PoleQty), and the pole mixing ratio (PoleMixR, % of spruce poles) but little influence on basal area (*G*), stand diameter diversity (Gini), and composition (MixR).

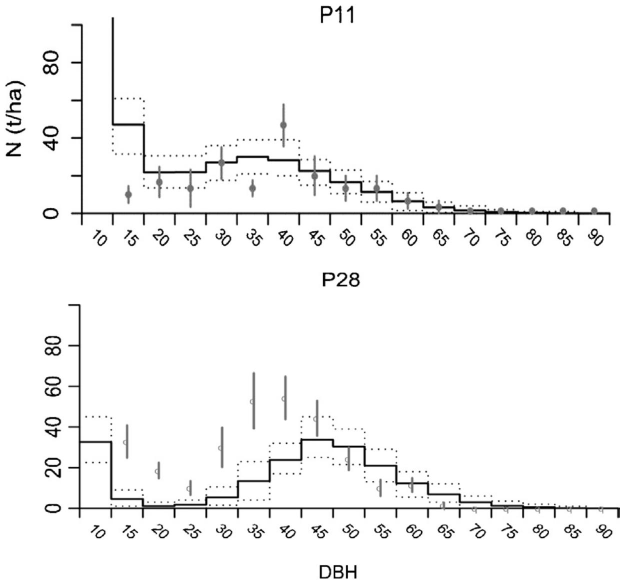


Fig. 5. *Comparison of predicted and observed diameter distributions after 50 years in stands P11 and P28.* The predicted number of trees/ha in 1980 (simulated since 1931) per DBH class are represented by black solid lines (mean) and black dotted lines (95% confidence interval). Observations from statistical field inventories are represented by gray dots (mean) and vertical gray bars (95% confidence intervals). The model correctly predicted the final structure of P11 except for an overestimation of poles (15-cm class) and slight smoothing between the 35-cm and 40-cm classes. In contrast, the model largely underestimated the number of small and intermediate trees in stand P28.

The effect of regeneration parameters was smaller than the effect of growth parameters, probably because of the short simulation time, which was not long enough for new recruited trees to impact the population of adult trees. The influence of these six demographic parameters was strong enough to affect forest dynamics substantially: extreme parameter values led to simulated basal areas ranging from 2 to 80 m^2/ha and total tree density ranging from 50 to 250–300 t/ha, whatever the initial stand.

We also detected parameters with a lower but non-negligible influence for at least one variable. Among these parameters, we identified two mortality parameters for both species (M_P1 and M_P2) that affected the total number of trees. Four other parameters were only influential for spruce: two allometric parameters affecting crown base radius (CBR_b) and height (H_mK), modulating light interception by crowns; and two parameters controlling spruce sapling height increment (S_R1 and S_R2) and therefore the delay before tree recruitment. Finally, we noted the systematically low influence of the parameters related to individual variability in allometries and growth (G_sB, CBH_sK, CBR_sK, H_sK).

3.5. Partial recalibration using ABC: Improving local model predictions

The new estimates of the growth parameter G_A and the regeneration parameter S_A recalibrated for the different stands were quite well distributed around the estimates of the initial calibration. This highlighted the consistency of the direct and indirect calibration approaches, despite the difference in the data sets in terms of organization levels and time scales. Differences in parameter values among stands showed substantial heterogeneity within the forest in terms of both regeneration conditions (recruitment proportional to S_A varying from 4 to 300 for spruce) and growth conditions (increment proportional to the exponential of G_A that varied from −5 to −3 for spruce). However, the recalibration produced unrealistic posterior distributions in four cases, with median values outside the prior distribution range, such as negative fecundity values. In these cases, we considered that the recalibration process had failed, which can be either due to the model's inability to simulate the dynamics of the stand, to incorrect historical data, or to unmodeled processes such as tending in the seedling stages.

We made new predictions of stand development between 1949 and 1980 using the recalibrated parameter estimates and these predictions appeared more consistent with the observations made at these two dates. Since the recalibration process took into account both the 1949 and 1980 inventories, the new trajectories expressed a balance of the influence of the two dates. Prediction bias in 1980 decreased in most cases. This was especially true for the pole mixing ratio, which improved greatly in 11 out of the 22 stands with a prediction bias divided by two.

4. Discussion

4.1. The evaludation process revealed model strengths and weaknesses

The evaludation process succeeded in progressively building a consistent understanding of the model's strengths and weaknesses. First, it showed that the Samsara2 model was able to predict the dynamics of uneven-aged spruce–fir mountain forests correctly over both the short and long term. For time scales of a few centuries, the model predicted natural forest dynamic patterns that were qualitatively consistent with the literature (Bormann and Likens, 1979; Oliver and Larson, 1996; Leibundgut, 1982) and a response to disturbance consistent with field descriptions of spruce and fir strategies (Grassi et al., 2004; Stancioiu and O'Hara, 2006). For time scales of a few decades, the model predicted stock dynamics with a quantitative bias less than 7% after 50 years in a forest managed with a selection system. This bias is within the lower range of the order of magnitude reported for comparable models (Wehrli et al., 2007; Pacala et al., 1996; Seidl et al., 2005; Didion et al., 2009). However, the number of small trees was overestimated in most stands. The difficulty of correctly predicting stand structure details has been reported for all forest dynamic models of uneven-aged stands (Didion et al., 2009). This is disappointing, however, in that it remains a major objective for individual-based models.

Second, the evaludation process showed that at the scale of forest stands (4 ha in our simulations), predictions were consistent despite model stochasticity. The quantitative evaluation step and the recalibration step also showed that the model was able to adapt correctly to different forest stands: with standard parameters, stand predictions were within the range of observation uncertainty in most cases. Unfortunately, a reason for this result might be that observation uncertainty was relatively high in the historical data, because of the size and heterogeneity of forest stands that make accurate field inventories difficult.

Third, the local parameter recalibration showed that parameter estimates could vary substantially among stands. This variation seems reasonable for the growth parameter (±25% around the mean). It is more problematic for the regeneration parameter, because an estimate close to zero in some stands meant that regeneration was impossible in these stands, whatever the stand density. This could reveal either a problem with certain historical data or confusion between the seedling production and seedling survival processes, limiting the possibility of simulating the effect of silviculture changes in these stands. By integrating explicitly environmental variables, several forest dynamics models attempt to address these local variations more completely (Bugmann, 2001), which is an attractive direction for improvement, because it would help constrain parameters with larger data sets and because it would improve prediction capacity in new stands.

4.2. The evaludation process clarified how the model works

An important point in building confidence in the model is that the evaludation process clarified how the model produces changes in forest structure over time. One of the main objectives of individual-based models is to show how interactions among individuals generate emergent patterns at the population level (Grimm and Railsback, 2005). The qualitative evaluation step was especially informative here because it showed, step by step, how tree-level processes generated changes in stand structure and in return how stand structure regulated individual processes. The exclusion stage was thus characterized by a homogeneous canopy where the dominant tree level processes were growth and competition, whereas the re-initiation stage was characterized by a heterogeneous canopy where old tree mortality and regeneration dominated. With disturbances, stand spatial heterogeneity became the key regulator of dynamic processes with a concentration of regeneration within gaps and a regulation of species balance by gap size. Moreover, the simulations showed how the model produced shifts between stages. For example, at the end of the exclusion stage, the incapacity of crown growth to close the space between large trees produced a shift to the expression of a new process, regeneration.

The sensitivity analysis step also increased our understanding of the model by pointing out that the uncertainties on the estimates of two growth parameters were the leading factors responsible for prediction uncertainties. This was not necessarily expected, given that growth is not the most difficult process to calibrate. Moreover, analyses of other models have highlighted either an influence of very specific parameters such as crown width (Ruger et al., 2007) or a comparable influence of every process (Pacala et al., 1996). However, the dominant influence of growth in our model is consistent with the fact that all the demographic processes in the model are related to tree size either directly or indirectly via radiation availability and interception. Canopy tree growth has a key regulating role in the model by determining changes in competition levels that propagate to mortality and regeneration processes.

4.3. The evaludation process specified the conditions for using the model safely

Beyond giving a sense of overall model performance, the evaludation process clarified the kind of questions that can be studied with the model and some precautions for its use. It indeed indicated how to make the most of this model, while reducing its limitations. The qualitative and quantitative evaluations showed that the model responded quite well to a range of disturbance and management regimes. The model therefore was a useful tool to study the effect of disturbance or management on the spatiotemporal dynamics of uneven-aged mixed spruce–fir stands. However, the biases observed on regeneration and stand structure predictions over 50 years calls for the limitation of quantitative predictions to a few decades in order to avoid the propagation of the regeneration bias to the adult population.

Plots covering a few hectares appeared sufficiently large to have little model stochasticity. However, a complete uncertainty analysis was lacking in this study. The sensitivity analysis stressed the strong impact of the uncertainty of the estimates of some demographic parameters, especially two related to growth. This result calls for a complete quantification of the uncertainties in model predictions caused not only by model stochasticity, but also by the propagation of uncertainties in parameter estimates. This would produce more conservative predictive intervals than when taking only model stochasticity into account and more accurately represent the model's true ability to predict forest dynamics. Moreover, (Fortin et al., 2009; Fortin and Langevin, 2012) have shown that nonlinearities in a model can interact with the propagation of uncertainties, leading to small differences between the average predictions of a stochastic model and the predictions of a deterministic model. However, this approach would require many more simulation runs to sample the high dimensional parameter space correctly. Because of the consequential computing time required by the model, it is unlikely that we can use this approach in standard model applications. Moreover, covariance among parameters should be taken into account in this type of analysis, which would be relatively simple for parameters in the same demographic process but difficult between processes calibrated independently. We then focused on the model's standard application conditions, with a limited number of realizations focused on averaging model stochasticity. A more complete uncertainty analysis could take place, however, in the evaludation process after the recalibration step, restricting the analysis to the propagation of uncertainty of the most influential parameters and using the covariance estimates between regeneration and growth processes produced by the joint recalibration. Unfortunately, in the ABC approach, local regressions that improve parameter estimations make it impossible to estimate parameters and prediction uncertainty at the same time as in a classical MCM context. The strong propagation of growth and regeneration uncertainty in model predictions also calls for consideration of their impact on simulation experiment results. When comparing management alternatives, growth and regeneration parameters can be treated as explicit experimental factors. This could reveal useful variations in forest response to management along growth and regeneration gradients.

4.4. The evaludation process highlighted important ecological processes

When simulations are run to explore the impact of new management strategies that have never been implemented in the field or to predict the effect of new environmental conditions, the knowledge they provide is clear. During the evaludation process, we attempted to reproduce patterns already observed in the field. We feel that it nevertheless provided valuable insights on forest dynamics.

The evaludation process showed that ontogeny and competition for light were regulators of demographic processes able to produce collective dynamics close to observed patterns. Similar results have been produced by other models of uneven-aged forest dynamics despite variations in model details (Pacala et al., 1993). This strengthens our hypothesis that ontogeny and competition for light are structuring factors of uneven-aged forest dynamics. It also gives confidence in the understanding gained about the model and how it represents the relationships between stand structure and how demographic fluxes can be transposed to some extent to real ecosystems even if it is known that the model remains a highly simplified representation of reality.

The dominant influence of growth revealed by this study is also striking, because this process is ignored in most theoretical models of population dynamics. Simple models such as the Lotka–Voltera competition model or the Levins patch model consider competition as independent of individual size (Begon et al., 1996). In contrast, our model indicates that growth is a strong regulator of demographic fluxes in forests. It shows that competition cannot be discussed in forests without considering variations in tree size and time. In addition, it suggests that gradients of growth conditions could potentially be key drivers of variations of forest dynamic patterns.

The model showed that despite slight differences in parameter values among species and a tradeoff between fecundity and seedling survival, these differences could be sufficient to produce the exclusion of spruce by fir. Moreover, their relative dynamics responded to disturbances, in agreement with field observations (Grassi et al., 2004; Antos and Parish, 2002). However, the balance among young trees of both species was highly sensitive to model parameters. This highlights the complexity of the relation among these species, changing drastically with altitude levels in mountain areas. Additional processes such as seed immigration, usually higher for spruce because of its lighter seeds and allelopathy, have also been described. How they modulate the patterns shown by the model is a promising research perspective.

Nevertheless, the evaludation process remains only a preliminary step in the use of a model that must be followed by more focused simulation experiments. To unravel principles of forest dynamics or progress in the understanding of the effect of silviculture on ecosystem services (Lafond et al., 2014), we have to analyze simulation experiment results and extract from them simple relations directly relating subsets of key parameters and response indicators. Understanding is indeed improved when one no longer needs the simulation model to predict qualitatively the response of the system to input variations (Braillard, 2013). Such simulation experiments are also a good preparatory step before field experiments, because they can help select modalities efficiently, by taking into account expected nonlinearity in stand dynamics responses. Models will be upgraded with new processes and indicators as new questions arise (Kimmins et al., 2008).

4.5. Improvements and perspectives

We expect that the widespread use of evaludation will gradually lead to much more efficient strategies of model development and evaludation. From the experience gained in this study, we believe that having the different steps of evaludation in mind from the beginning changes the hierarchy of priorities, organizes work more efficiently, and saves a great deal of time. Data should therefore be collected immediately with the evaluation step in mind, not only the calibration step. A preliminary sensitivity analysis step can be carried out very early, based on parameter ranges found in the literature, in order to focus calibration efforts on the most influential processes. During calibration, covariance matrices among parameters are often overlooked by ecologists (Clark et al., 2011), but it is essential to record them to include parameter uncertainty in simulations. Simulators must also be developed early in the process with functionalities facilitating simulations with different parameter values, since this is necessary in the sensitivity and uncertainty analysis steps. The automated simulation of management must also be considered early in the process if evaluation is to be included within a management context. Moreover, all these steps must be considered as a cycle of model development (Grimm and Railsback, 2005; Schmolke et al., 2010), with the idea that each step can be done several times, possibly using methods of increasing complexity. For example, the sensitivity analysis process can be conducted in two steps: first screening to identify a reduced set of influent parameters based on methods requiring few simulations such as factorial designs or the Morris method, and second a method based on the decomposition of variance, more precise but requiring more simulations (Saltelli et al., 2000, 2008).

The development of evaludation approaches also raises new challenges for the modeling community. Methods are currently multiplying in the fields of sensitivity analysis and experimental design for numerical experiments (Saltelli et al., 2008). Such methods are increasingly available in statistical software. However, determining which method is the most appropriate for a given simulation model and problem remains difficult. Guidelines and packages facilitating the sequential use of increasingly complex methods would be considerably helpful for modelers. Moreover, theoretical progress is required in difficult questions such as how to take into account covariance among factors in sensitivity analyses and when it is appropriate to use meta-models, i.e., approximations of the simulation models, in analyses and predictions. Incremental methods able to make use of simulations performed during a previous step of the evaludation process, possibly completed only by a limited set of additional simulations, would be highly valuable to reduce computing time. The development of simulation platforms shared by different modelers (Dufour-Kowalski et al., 2012) is also a key perspective because they provide environments facilitating the simulation replication work, sampling in parameter distributions, and multiple simulation synthesis for the modeler. Last, the development of environments facilitating the distribution of simulations on super-calculators or clusters of computers is essential for the use of time-expensive models.

5. Conclusion

We have shown that beyond revealing the strengths and weaknesses of a model, a comprehensive evaludation process considerably improves the understanding of the model and its credibility. By highlighting the potential effects of key ecological processes, the evaludation also increases ecological knowledge. We expect that the widespread use of evaludation will gradually lead to much more efficient model development strategies, mostly because modelers will think of the different steps from the beginning and organize their data, their model, and their research questions accordingly. The development of packages and simulation platforms favoring the cyclic development of models and the different evaludation steps will make this process easier. This cyclic process illustrates, within the modeling context, the dialog between observation, inference, theory, and prediction that characterizes the scientific process.

Acknowledgements

This work was funded by the GeForHet project, from the French “Biodiversity, Forest Management and Public Policy” research program (BGF, no. E23/2010) and by the FP7 European “Advanced multifunctional forest management in European mountain ranges” project (ARANGE, no. 289437).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2015.06.039

References

Antos, J.A., Parish, R., 2002. Dynamics of an old-growth, fire-initiated, subalpine forest in southern interior British Columbia: tree size, age, and spatial structure. Can. J. For. Res. 32 (11), 1935–1946 (Revue Canadienne De Recherche Forestiere).

Assmann, E., 1970. The Principles of Forest Yield Study. Pergamon Press, Oxford.

Augusiak, J., Van den Brink, P.J., Grimm, V., 2014. Merging validation and evaluation of ecological models to ‘evaludation’: a review of terminology and a practical approach. Ecol. Model. 280, 117–128.

Beaumont, M.A., Zhang, W.Y., Balding, D.J., 2002. Approximate Bayesian computation in population genetics. Genetics 162 (4), 2025–2035.

Begon, M., Harper, J.L., Townsend, C.R., 1996. Ecology. Blackwell Science, Oxford.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20.

Bormann, F.H., Likens, G.E., 1979. Pattern and Process in a Forested Ecosystem. Springer, New York, NY.

Botkin, D.B., 1993. Forest dynamics. In: An Ecological Model. Oxford University Press, New York, NY.

Braillard, P.-A., 2013. Que peut expliquer un modèle complexe et peut-on le comprendre? In: Varenne, F., Silberstein, M. (Eds.), Modéliser et simuler.

Epistémologies et pratiques de la modélisation et de la simulation, vol. 1. Matériologiques, Paris, France, pp. 275–295.

Bugmann, H., 2001. A review of forest gap models. Clim. Change 51, 259–305.

Bugmann, H.K.M., 1996. A simplified forest model to study species composition along climate gradients. Ecology 77 (7), 2055–2074.

Burschel, P., El Kateb, H., Mosandl, R., 1992. Experiments in mixed mountain forests in Bavaria. In: Kelty, M.J., Oliver, C.D., Larson, B.C (Eds.), The ecology and silviculture of mixed-species forests. Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 183–215.

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Model. Softw. 22 (10), 1509–1518.

Ciric, C., Ciffroy, P., Charles, S., 2012. Use of sensitivity analysis to identify influential and non-influential parameters within an aquatic ecosystem model. Ecol. Model. 246, 119–130.

Clark, J.S., Bell, D.M., Hersh, M.H., Kwit, M.C., Moran, E., Salk, C., Stine, A., Valle, D., Zhu, K., 2011. Individual-scale variation, species-scale differences: inference needed to understand diversity. Ecol. Lett. 14 (12), 1273–1287.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., Lodge, D.M., Pascual, M., Pielke, R., Pizer, W., Pringle, C., Reid, W.V., Rose, K.A., Sala, O., Schlesinger, W.H., Wall, D.H., Wear, D., 2001. Ecological forecasts: an emerging imperative. Science 293 (5530), 657–660.

Cordonnier, T., Courbaud, B., Berger, F., Franc, A., 2008. Permanence of resilience and protection efficiency in mountain Norway spruce forest stands: a simulation study. For. Ecol. Manage. 256, 347–354.

Courbaud, B., de Coligny, F., Cordonnier, T., 2003. Simulating radiation distribution in a heterogeneous Norway spruce forest on a slope. Agric. For. Meteorol. 116, 1–18.

Csillery, K., François, O., Blum, M.G.B., 2012. ABC: an R package for approximate Bayesian computation (ABC). Methods Ecol. Evol. 3, 475–479.

de Coligny, F., Ancelin, P., Cornu, G., Courbaud, B., Dreyfus, P., Goreaud, F., Gourlet-Fleury, S., Meredieu, C., Saint-André, L., 2003. CAPSIS: Computer-Aided Projection for Strategies in Silviculture: advantages of a shared forest modelling platform. In: Amaro, A., Reed, D., Soares, P. (Eds.), Modelling Forest Systems. CABI Publishing, Wallingford, UK, pp. 319–323.

Didion, M., Kupferschmid, A.D., Zingg, A., Fahse, L., Bugmann, H., 2009. Gaining local accuracy while not losing generality—extending the range of gap model applications. Can. J. For. Res. 39 (6), 1092–1107 (Revue Canadienne De Recherche Forestiere).

Dufour-Kowalski, S., Courbaud, B., Dreyfus, P., Meredieu, C., de Coligny, F., 2012. Capsis: an open software framework and community for forest growth modelling. Ann. For. Sci. 69 (2), 221–233.

Dutreuil, S., 2013. Comment le modèle Daisyworld peut-il contribuer à l'hypothèse Gaïa? In: Varenne, F., Silberstein, M. (Eds.), Modéliser et simuler. Epistémologies et pratiques de la modélisation et de la simulation, vol. 1, pp. 189–254.

Fortin, M., Bedard, S., DeBlois, J., Meunier, S., 2009. Assessing and testing prediction uncertainty for single tree-based models: a case study applied to northern hardwood stands in southern Quebec, Canada. Ecol. Model. 220 (20), 2770–2781.

Fortin, M., Langevin, L., 2012. Stochastic or deterministic single-tree models: is there any difference in growth predictions? Ann. For. Sci. 69 (2), 271–282.

Gauquelin, X., Courbaud, B., Fay, J., Berger, F., Mermin, E., 2008. Conduite de peuplements mélangés en forêt de montagne: exemple d'un transfert chercheurs-gestionnaires. Rev. For. Française LX (2), 207–214.

Gilks, W.R., Thomas, A., Spiegelhalter, D.J., 1994. A language and program for complex Bayesian modelling. Statistician 43, 169–177.

Godfrey-Smith, P., 2003. Theory and reality: An Introduction to the Philosophy of Science. The University of Chicago Press, Chicago, IL.

Goreaud, F., Alvarez, I., Courbaud, B., de Coligny, F., 2006. Long-term influence of the spatial structure of an initial state on the dynamics of a forest growth model: a simulation study using the Capsis platform. Simulation 82, 475–495.

Goreaud, F., de Coligny, F., Courbaud, B., Dhôte, J.F., Dreyfus, P., Pérot, T., 2005. La modélisation: un outil pour la gestion et l'aménagement en forêt. VertigO 6 (2), 1–12.

Grassi, G., Minotta, G., Tonon, G., Bagnaresi, U., 2004. Dynamics of Norway spruce and silver fir natural regeneration in a mixed stand under uneven-aged management. Can. J. For. Res. 34 (1), 141–149 (Revue Canadienne De Recherche Forestiere).

Grimm, V., Augusiak, J., Focks, A., Frank, B.M., Gabsi, F., Johnston, A.S.A., Liu, C., Martin, B.T., Meli, M., Radchuk, V., Thorbek, P., Railsback, S.F., 2014. Towards better modelling and decision support: documenting model development, testing, and analysis using TRACE. Ecol. Model. 280, 129–139.

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol a review and first update. Ecol. Model. 221 (23), 2760–2768.

Grimm, V., Railsback, S.F., 2005. Individual-Based Modeling and Ecology. Princeton U.P., Princeton.

Kimmins, J.P., Blanco, J.A., Seely, B., Welham, C., Scoullar, K., 2008. Complexity in modelling forest ecosystems: how much is enough? For. Ecol. Manage. 256 (10), 1646–1658.

Korpel, S., 1982. Degree of equilibrium and dynamical changes of the forest on example of natural forests of Slovakia. Acta Facult. For., Zvolen–Czechoslovakia XXIV, 9–31.

Kunstler, G., Allen, R.B., Coomes, D.A., Canham, C.D., Wright, E.F., 2013. Sustainable management, earthquake disturbances, and transient dynamics: modelling timber harvesting impacts in mixed-species forests. Ann. For. Sci. 70 (3), 287–298.

Lafond, V., Cordonnier, T., de Coligny, F., Courbaud, B., 2012. Reconstructing harvesting diameter distribution from aggregate data. Ann. For. Sci. 69, 235–243.

Lafond, V., Lagarrigues, G., Cordonnier, T., Courbaud, B., 2014. Uneven-aged management options to promote forest resilience for climate change adaptation: effects of group selection and harvesting intensity. Ann. For. Sci. 71 (2), 173–186.

Lagarrigues, G., Jabot, F., Courbaud, B., 2014. Approximate Bayesian computation to recalibrate individual-based models with population data: illustration with a forest simulation model. Ecol. Modell. 306, 278–286.

Leibundgut, 1982. Europäische Urwälder der Bergstufe. Haupt, Bern.

Lexer, M.J., Honninger, K., 2001. A modified 3D-patch model for spatially explicit simulation of vegetation composition in heterogeneous landscapes? For. Ecol. Manage. 144 (1-3), 43–65.

Ligot, G., Balandier, P., Courbaud, B., Claessens, H., 2014. Forest radiative transfer models: which approach for which application? Can. J. For. Res. 44 (5), 391–403.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.

Motta, R., Berretti, R., Castagneri, D., Dukic, V., Garbarino, M., Govedar, Z., Lingua, E., Maunaga, Z., Meloni, F., 2011. Toward a definition of the range of variability of central European mixed Fagus–Abies–Picea forests: the nearly steady-state forest of Lom (Bosnia and Herzegovina). Can. J. For. Res. 41 (9), 1871–1884 (Revue Canadienne De Recherche Forestiere).

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P., Jabot, F., Jarne, P., Joly, D., Julliard, R., Kéfi, S., Kergoat, G.J., Lavorel, S., Le Gall, L., Meslin, L., Morand, S., Morin, X., Morlon, H., Pinay, G., Pradel, R., Schurr, F.M., Thuiller, W., Loreau, M., 2015. Improving predictive ecology in a changing world. J. Appl. Ecol., http://dx.doi.org/10.1111/1365-2664.12482

Oliver, C.D., Larson, B.C., 1996. Forest Stand Dynamics. John Wiley & Sons, New York, NY.

Pacala, S.W., Canham, C.D., Saponara, J., Silander, J.A., Kobe, R.K., Ribbens, E., 1996. Forest models defined by field measurements: estimation, error analysis and dynamics. Ecol. Monogr. 66 (1), 1–43.

Pacala, S.W., Canham, C.D., Silander, J.A., 1993. Forest models defined by field measurements: I. The design of a northeastern forest simulator. Can. J. For. Res. 23, 1980–1988.

Pappas, C., Fatichi, S., Leuzinger, S., Wolf, A., Burlando, P., 2013. Sensitivity analysis of a process-based ecosystem model: Pinpointing parameterization and structural issues. J. Geophys. Res. Biogeosci. 118 (2), 505–528.

Pretzsch, H., 2006. Species-specific allometric scaling under self-thinning: evidence from long-term plots in forest stands. Oecologia 146 (4), 572–583.

Prisley, S.P., Mortimer, M.J., 2004. A synthesis of literature on evaluation of models for policy applications, with implications for forest carbon accounting. For. Ecol. Manage. 198 (1–3), 89–103.

Pujol, G., Iooss, B., Janon, A., 2014. Package Sensitivity., pp. 52, ⟨http://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf⟩.

Purves, D., Pacala, S., 2008. Predictive models of forest dynamics. Science 320, 1452–1453.

Purves, D.W., Lichstein, J.W., Strigul, N., Pacala, S.W., 2008. Predicting and understanding forest dynamics using a simple tractable model. Proc. Natl. Acad. Sci. U.S.A. 105 (44), 17018–17022.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reynolds, M.R.J., Burk, T.E., Huang, W.-C., 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. For. Sci. 34 (2), 373–399.

Ruger, N., Gutierrez, A.G., Kissling, W.D., Armesto, J.J., Huth, A., 2007. Ecological impacts of different harvesting scenarios for temperate evergreen rain forest in southern Chile—a simulation experiment? For. Ecol. Manage. 252 (1–3), 52–66.

Rykiel, E.J., 1996. Testing ecological models: the meaning of validation. Ecol. Model. 90 (3), 229–244.

Saltelli, A., Chan, K., Scott, E.M. (Eds.), 2000. Sensitivity Analysis. John Wiley & Sons, Hoboken, NJ.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis—the Primer. John Wiley & Sons, Hoboken, NJ.

Schmolke, A., Thorbek, P., DeAngelis, D.L., Grimm, V., 2010. Ecological models supporting environmental decision making: a strategy for the future. Trends Ecol. Evol. 25 (8), 479–486.

Schütz, J.P., 1997. Sylviculture 2. La gestion des forêts irrégulières et mélangées. Presses Polytechniques et Universitaires Romandes, Lauzane.

Seidl, R., Lexer, M.J., Jager, D., Honninger, K., 2005. Evaluating the accuracy and generality of a hybrid patch model. Tree Physiol. 25 (7), 939–951.

Shaeffer, D.L., 1980. A model evaluation methodology applicable to environmental assessment models. Ecol. Model. 8, 275–295.

Soares, P., Tomé, M., Skovsgaard, J.P., Vanclay, J.K., 1995. Evaluating a growth model for forest management using continuous forest inventory data. For. Ecol. Manage. 71, 251–265.

Stancioiu, P.T., O'Hara, K.L., 2006. Regeneration growth in different light environments of mixed species, multiaged, mountainous forests of Romania. Eur. J. For. Res. 125, 151–162.

Stengers, I., Bensaude-Vincent, B., 2003. 100 mots pour commencer à penser les sciences. Les empêcheurs de penser en rond, Paris, France.

Vanclay, J.K., Skovsgaard, J.P., 1997. Evaluating forest growth models. Ecol. Model. 98, 1–12.

Varenne, F., Silberstein, M. (Eds.), 2013. Modéliser et simuler. Epistémologies et pratiques de la modélisation et de la simulation. Matériologiques, Paris, France.

Vieilledent, G., 2009. Structurer l'incertitude et la variabilité dans les modèles de dynamique forestière. Application à la coexistence du Sapin et de l'Epicéa en forêt de montagne. AgroParisTech (Vol. Thèse de doctorat).

Vieilledent, G., Courbaud, B., Kunstler, G., Dhôte, J.F., Clark, J.S., 2010. Individual variability in tree allometries determines light resource allocation in forest ecosystems—a hierarchical Bayes approach. Oecologia 163, 759–773.

Wehrli, A., Weisberg, P.J., Schoenenberger, W., Brang, P., Bugmann, H., 2007. Improving the establishment submodel of a forest patch model to assess the long-term protective effect of mountain forests. Eur. J. For. Res. 126 (1), 131–145.

Wehrli, A., Zingg, A., Bugmann, H., Huth, A., 2005. Using a forest patch model to predict the dynamics of stand structure in Swiss mountain forests? For. Ecol. Manage. 205 (1–3), 149–167.