# Bayesian calibration of a growth-dependent tree mortality model to simulate the dynamics of European temperate forests

Maxime Cailleret [iD] [1,2,3,7] Nicolas Bircher,[1] Florian Hartig [iD],[4,5] Lisa Hülsmann [iD],[5,6] and Harald Bugmann [iD][1]

[1]*Forest Ecology, Institute of Terrestrial Ecosystems, Department of Environmental Sciences, Swiss Federal Institute of Technology ETH, Universitätsstrasse 22, 8092 Zürich, Switzerland*
[2]*Forest Dynamics, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland*
[3]*UMR RECOVER, Aix Marseille University, IRSTEA, 3275 Route de Cézanne, 13182 Aix-en-Provence, France*
[4]*Department of Biometry and Environmental System Analysis, Albert-Ludwigs-University Freiburg, TennenbacherStraße 4, 79106 Freiburg, Germany*
[5]*Theoretical Ecology, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany*
[6]*Forest Resources and Management, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland*

*Abstract.* Dynamic vegetation models (DVMs) are important tools to understand and predict the functioning and dynamics of terrestrial ecosystems under changing environmental conditions. In these models, uncertainty in the description of demographic processes, in particular tree mortality, is a persistent problem. Current mortality formulations lack realism and are insufficiently constrained by empirical evidence. It has been suggested that empirically estimated mortality submodels would enhance DVM performance, but due to the many processes and interactions within a DVM, the claim has rarely been tested. Here, we compare the performance of three alternative growth-dependent tree mortality submodels in the DVM ForClim: (1) a mortality function with theoretical foundation (ForClim v3.3); (2) a mortality function with parameters directly estimated based on forest inventory data; and (3) the same function, but with parameters estimated using an inverse approach through Bayesian calibration (BC). Time series of inventory data from 30 ecologically distinct Swiss natural forest reserves collected over 35+ yr, including the main tree species of Central Europe, were used for the calibration and subsequent validation of the mortality functions and the DVM. The recalibration resulted in mortality parameters that differed from the direct empirical estimates, particularly for the relationship between tree size and mortality. The calibrated parameters outperformed the direct estimates, and to a lesser extent the original mortality function, for predicting decadal-scale forest dynamics at both calibration and validation sites. The same pattern was observed regarding the plausibility of their long-term projections under contrasting environmental conditions. Our results demonstrate that inverse calibration may be useful even when direct empirical estimates of DVM parameters are available, as structural model deficiencies or data problems can result in discrepancies between direct and inverse estimates. Thus, we interpret the good performance of the inversely calibrated model for long-term projections (which were not a calibration target) as evidence that the calibration did not compensate for model errors. Rather, we surmise that the discrepancy was mainly caused by a lack of representativeness of the mortality data. Our results underline the potential for learning more about elusive processes, such as tree mortality or recruitment, through data integration in DVMs.

*Key words: Bayesian calibration; data integration; dynamic vegetation model; forest inventory; growth; stand dynamics; tree mortality.*

## Introduction

Understanding and predicting how terrestrial ecosystems respond to environmental change are key issues in ecology, with wide-ranging consequences for stakeholders such as forest managers and policy makers. To achieve these goals, many scientific studies rely on process-based dynamic vegetation models (DVMs), which simulate demographic processes within plant communities over time, based on abiotic conditions and biotic interactions (e.g., Bonan et al. 2003, Sitch et al. 2008, Hartig et al. 2012). A challenge in using these models is that there is considerable uncertainty in parameters, drivers and processes (Beven 2016), and as a result in the

DVM projections of ecosystem properties in response to an altered climate (Moorcroft 2006, Friend et al. 2014). To improve the robustness of model projections, many authors have suggested reconsidering the design of DVMs, with a specific focus on adding more details to key demographic processes (e.g., Fisher et al. 2018). This applies particularly to tree mortality, which so far has been treated in a rather simplistic manner in most DVMs (e.g., Keane et al. 2001, McDowell et al. 2013), despite its strong impact on model behavior (Friend et al. 2014, Bugmann et al. 2019).

Improving process formulations in DVMs can be achieved by a better integration of known physiological or ecological mechanisms, and by a stronger link with empirical data. To achieve the latter, model parameters were traditionally determined by field or experimental observations, independently of the model. The outcome of this direct parameterization in terms of model behavior was then tested against reference data (benchmarking; cf. Pacala et al. 1996). However, such a direct estimate of all parameters in a DVM requires a large set of specialized observations (e.g., Le Roux et al. 2001) that are typically hard to achieve or not available at all (Hartig et al. 2012). Mortality parameters are a prime example of this. Tree death occurs rarely in forest ecosystems and is highly stochastic; thus, it is not yet well captured by forest surveys and long-term monitoring programs (Hartmann et al. 2018, McMahon et al. 2019). This is especially true for large trees, as the inventory data are usually dominated by small stems (Needham et al. 2018). Consequently, many mortality parameters in current DVMs are weakly constrained or originate from ecological assumptions about the underlying mechanisms (typically called theoretical approaches; see Bugmann et al. 2019) that are not empirically verified (Mäkelä et al. 2000, Manusch et al. 2012). For instance, some physiological-based DVMs predict mortality of a given tree to occur when its level of non-structural carbohydrates (NSC) falls below a given limit, while there is no experimental- nor field-based evidence of such a simplistic NSC threshold (Adams et al. 2017; McDowell et al. 2013).

A solution to better link models with empirical data that has become increasingly popular in recent years is to estimate the parameters inversely. By calibrating parameter values based on the match between empirical data and model outputs, inverse modeling approaches, such as Bayesian methods (Van Oijen et al. 2005) allow for harnessing complex data sources that would not be suited for direct parameter estimation (e.g., species distribution data, which depends on multiple and interacting processes, and therefore on many parameters at the same time), and therefore open up novel opportunities to constrain parameter-rich DVMs (Hartig et al. 2012). This approach can be based on multiple data types for calibration, which often improves final model performance and generality (cf. Grimm and Railsback 2012). Moreover, in the framework of Bayesian statistics, it is possible to combine direct parameter estimates from various sources (via the "prior distribution," e.g., Wang et al. 2013) and estimates that are generated inversely (via the "likelihood"). The result of a Bayesian calibration (BC) is a probability distribution (the "posterior") that represents the combination of both direct and inverse information for the respective parameter. The posterior can then be used for ecological interpretation and prediction while considering the uncertainty in parameter estimates (Wang et al. 2013, Reyer et al. 2016, Augustynczik et al. 2017).

Current applications of BC for statistical and process-based forest models have mainly been used to calibrate processes for which ecological knowledge was scarce and parameter uncertainty was large (e.g., O'Hara et al. 2002, Van Oijen et al. 2005, Larssen et al. 2006, Reinds et al. 2008, Hartig et al. 2014, Minunno et al. 2016), for estimating and propagating parametric model uncertainty (e.g., Augustynczik et al. 2017, Van Oijen 2017), or for model intercomparison exercises (e.g., Van Oijen et al. 2011, 2013). However, as pointed out by Hartig et al. (2012) and Van Oijen (2017), an interesting additional possibility of the approach is to compare direct (prior) and inverse (posterior) parameter estimates. A mismatch between direct and inverse parameter estimates may point at either (1) a structural problem in the model, (2) a systematic bias in the data, or (3) a discrepancy between the nature of the parameter in the model and the parameter that is measured in the field. Because a model is a simplification of reality, every parameter plays a different role in the model than its namesake in a true forest, and its "true" value cannot be measured exactly (Van Oijen 2017). For instance, species shade tolerance cannot be derived based on a single measurement or observation, but rather depends on different suites of functional traits (e.g., leaf and crown traits; Valladares and Niinemets 2008). Hence, calibration with direct and inverse information can lead to an improved understanding of ecological process interactions and their representation in a model.

An example of a DVM for which direct and inverse calibration data are available is ForClim, a forest gap model that predicts the dynamics of temperate forests (Bugmann 1996) by simulating establishment, growth and mortality of individual trees based on site- and species-specific environmental constraints and biotic interactions. Three recent studies have examined whether more empirically derived submodels of tree mortality reduce biases and uncertainties in ForClim predictions (Bircher et al. 2015, Hülsmann et al. 2018, Vanoni et al. 2019), which had been suggested by various authors (cf. Keane et al. 2001, Adams et al. 2013, Friend et al. 2014). The results of these studies highlighted that simulation results at multidecadal time scales were very sensitive to the choice of the mortality function; but, contrary to general expectations, the use of direct parameter estimates did not substantially improve model reliability. A better understanding of the mortality

function and its interactions with other ecological processes in the model (particularly tree growth) is therefore a key priority.

In this study, we evaluate whether the issues uncovered in these recent studies can be solved by inversely recalibrating the parameters of a growth-based mortality function incorporated in the DVM ForClim. We use inventory data from 9 and 21 unmanaged Swiss forest reserves for calibration and validation, respectively, which cover a wide variety of forest types and include all major tree species of central Europe. Our goal was to (1) determine if inversely calibrated mortality parameters match with empirical direct estimates; (2) identify the causes of potential mismatches; and (3) assess if the inverse calibration leads to improved model performance regarding historical short-term change in key forest characteristics as well as plausible simulations of Potential Natural Vegetation (PNV) along an extended environmental gradient in Switzerland (Bugmann and Solomon 2000, Rasche et al. 2012).

MATERIALS AND METHODS

*The ForClim model*

ForClim is an individual-based forest gap model developed for simulating the dynamics of managed and unmanaged temperate forests. It has been applied over Europe (Bugmann 1996, Mina et al. 2017, Huber et al. 2018) and on other continents (e.g., Bugmann and Solomon 2000, Gutiérrez et al. 2016). Tree growth, establishment, and mortality are simulated on independent patches ($\approx$800 m$^2$) in annual time steps, using parsimonious ecological assumptions regarding the influence of climate and ecological processes on tree demography. Averaging the results across all simulated patches allows for obtaining mean successional dynamics at the stand scale (Bugmann 2001). Annual growth is calculated using the carbon budget model by Moore (1989), in which a species-specific maximum growth potential is reduced if light availability, degree-day sum, soil moisture during the growing season, nutrient availability, and crown length are below an optimum (Bugmann 1996, Didion et al. 2009, Rasche et al. 2012). Tree recruitment, modeled as the species-specific rate of establishment of saplings with a diameter at breast height (DBH) of 1.27 cm, is also reduced when environmental conditions are unfavorable.

In ForClim v3.3 (Mina et al. 2017), tree mortality is modeled as a combination of (1) a constant "background" mortality, which mimics tree mortality induced by random local disturbances and assumes that 1% of a tree population will survive to the species-specific maximum age, and (2) a stress-induced mortality that is activated if the annual diameter increment of a tree falls below an absolute or relative threshold (0.3 mm and 10% of species-specific maximum growth, respectively) for more than two consecutive years (Solomon 1986). In

contrast to establishment and growth, which are operating on the level of tree cohorts (i.e., individuals of the same age and size), mortality is applied to each tree of a cohort individually using a stochastic approach, which assumes that a tree dies if a uniform random number [0…1] is below the prescribed cohort-specific mortality probability. A more detailed description of the mortality function is provided in Appendix S1, and in Bircher et al. (2015).

*Data for calibration and validation*

The inventory data used to calibrate and validate the growth-dependent mortality functions, and to validate ForClim's predictions, were recorded in multiple Swiss forest reserves (Brang et al. 2011, Hülsmann et al. 2018, Appendix S2) where management had been excluded at least since their foundation (cf. inventory period in Table 1). All reserves contain one or more permanent plots, where each tree with a DBH $\geq$ 4 cm is individually tagged and repeatedly measured in intervals of approximately 10 yr.

A subset of the permanent plots was selected for model calibration (both for the direct estimation of the mortality function and the inverse calibration of the same mortality parameters in ForClim) according to the following criteria: (1) plots were excluded that showed evidence of recent natural disturbances (e.g., wind storms), as such disturbances are not considered in the model; (2) plots had to have a minimum size of 0.2 ha to guarantee that forest structure and composition are representative for the site; (3) plots had to have a minimum record period of 35 yr to allow for an adequate consideration of forest dynamics; (4) only one plot was chosen per forest reserve to avoid pseudoreplication within the calibration data; (5) an adequate representation of all main tree species of central Europe and a variety of forest types was sought in the calibration data set. According to these criteria, nine permanent plots were selected (Table 1). This number was a compromise between the wish to have many plots for improving the robustness of the mortality function, and the constraints imposed by computational costs. The plot with the largest number of species was selected first (Girstel_04; Table 1). In a next step, we included those plots with the highest proportion of one main tree species in Switzerland such as European beech *Fagus sylvatica* (Adenberg_03), silver fir *Abies alba* (Leihubelwald_02), Norway spruce *Picea abies* (Nationalpark_07)*,* Norway maple *Acer pseudoplatanus* (St. Jean_01), Scots pine *Pinus sylvestris* (VormStein_02), Swiss stone pine *Pinus cembra* (Nationalpark_07), European larch *Larix decidua* (Nationalpark_07), and oak species *Quercus robur/petraea* (Bois de Chênes_02)*.* This set of calibration plots was complemented by two permanent plots with comparatively fast stand dynamics, i.e., with a high interdecadal rate of change in stem numbers: Fürstenhalde_01 and Tariche Haute Côte_04.

TABLE 1. Description of the permanent plots used for calibration (shown in boldface type) and validation, including information on location, plot size, elevation, climate (mean annual temperature and precipitation sum), and dominant tree species.

| Site | Climate | | | | | Stand diversity, main species | Survival rate (% 11-yr) | Stand dynamic | Slope/ Aspect | Inventory period (n) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Location (° N, ° E) | Area (ha) | Elevation (m a.s.l.) | Temperature (°C) | Rainfall (mm) | | | | | |
| Adenberg_01 | 47.6, 8.6 | 0.45 | 520 | 9.0 | 1,017 | 2, *F. sylvatica, Q. robur* | 77.4 | 0.023 | +1 | 1970–2012 (5) |
| Adenberg_02 | 47.6, 8.6 | 0.45 | 500 | 9.0 | 1,017 | 3, *F. sylvatica, Q .robur* | 80.0 | 0.020 | 0 | 1970–2012 (5) |
| **Adenberg_03** | **47.6, 8.6** | **0.45** | **505** | **9.0** | **1,020** | **2, *F. sylvatica, Q. robur*** | **74.7** | **0.026** | **0** | **1970–2012 (5)** |
| Adenberg_04 | 47.6, 8.6 | 0.45 | 520 | 9.1 | 1,006 | 2, *F. sylvatica, Q. robur* | 75.7 | 0.025 | +1 | 1970–2012 (5) |
| Bois de Chênes_01 | 46.4, 6.2 | 0.49 | 550 | 9.6 | 1,075 | 1, *F. sylvatica* | 90.2 | 0.004 | 0 | 1970–2007 (4) |
| **Bois de Chênes_02** | **46.4, 6.2** | **0.49** | **570** | **9.5** | **1,094** | **3, *Q. robur, F. sylvatica*** | **83.8** | **0.013** | **+1** | **1970–2007 (4)** |
| Bonfol_03 | 47.5, 7.2 | 0.53 | 440 | 9.5 | 1,003 | 2, *Q. robur* | 85.0 | 0.009 | 0 | 1962–2001 (5) |
| **Fuerstenhalde_01** | **47.6, 8.5** | **0.53** | **460** | **9.2** | **1,065** | **2, *F. sylvatica*** | **81.4** | **0.019** | **0** | **1971–2012 (4)** |
| Fuerstenhalde_02 | 47.6, 8.5 | 0.53 | 470 | 9.2 | 1,076 | 1, *F. sylvatica* | 78.5 | 0.018 | 0 | 1971–2012 (4) |
| **Girstel_04** | **47.3, 8.5** | **0.22** | **675** | **7.9** | **1,297** | **5, *P. sylvestris, F. sylvatica, P. abies*** | **85.1** | **0.016** | **−1** | **1964–2006 (5)** |
| Girstel_11 | 47.3, 8.5 | 0.14 | 720 | 8.1 | 1,270 | 3, *A. pseudoplatanus, F. sylvatica* | 80.8 | 0.018 | +1 | 1972–2007 (4) |
| **Leihubelwald_02** | **46.9, 8.1** | **0.25** | **1240** | **6.1** | **1,770** | **3, *A. alba, F. sylvatica, P. abies*** | **88.7** | **0.007** | **−1** | **1973–2011 (4)** |
| Leihubelwald_03 | 46.9, 8.1 | 0.24 | 1140 | 6.6 | 1,690 | 2, *A. alba, F. sylvatica* | 77.2 | 0.023 | +1 | 1973–2011 (4) |
| Leihubelwald_04 | 46.9, 8.1 | 0.25 | 1100 | 6.7 | 1,668 | 2, *P. abies, A. alba* | 87.9 | 0.008 | 0 | 1973–2011 (4) |
| Nationalpark_05 | 46.7, 10.2 | 0.61 | 1985 | 1.2 | 981 | 3, *L. decidua, P. cembra* | 83.9 | 0.006 | 0 | 1978–2013 (3) |
| **Nationalpark_07** | **46.7, 10.2** | **0.56** | **1860** | **2.3** | **915** | **2, *P. abies, P. cembra*** | **92.3** | **0.006** | **0** | **1977–2012 (3)** |
| Pfynwald_01 | 46.3, 7.6 | 0.19 | 575 | 10 | 670 | 2, *P. sylvestris* | 84.8 | 0.010 | +1 | 1956–2003 (6) |
| Scatlè_01 | 46.8, 9.0 | 3.47 | 1650 | 3.7 | 1,582 | 1, *P. abies* | 85.6 | 0.014 | +1 | 1965–2006 (4) |
| **St.Jean_01** | **47.1, 7.0** | **0.28** | **1375** | **4.7** | **1,520** | **2, *A. pseudoplatanus, P. abies*** | **80.2** | **0.011** | **0** | **1961–2006 (5)** |
| St.Jean_02 | 47.1, 7.0 | 0.44 | 1370 | 4.8 | 1,510 | 1, *P. abies* | 92.1 | 0.007 | 0 | 1961–2006 (5) |
| Tariche Haute Côte_03 | 47.3, 7.2 | 0.91 | 735 | 8.1 | 1,228 | 2, *F. sylvatica, A. alba* | 87.2 | 0.009 | −1 | 1974–2012 (4) |
| **Tariche Haute Côte_04** | **47.3, 7.2** | **0.56** | **740** | **7.9** | **1,250** | **3, *A. alba, F. sylvatica*** | **81.6** | **0.016** | **+1** | **1974–2012 (4)** |
| Tariche Haute Côte_06 | 47.3, 7.2 | 0.54 | 720 | 8.1 | 1,228 | 2, *F. sylvatica, A. alba* | 85.6 | 0.011 | 0 | 1976–2012 (4) |
| Tutschgenhalden_13 | 47.5, 8.8 | 0.25 | 600 | 9.1 | 1,151 | 2, *F. sylvatica, A. alba* | 87.0 | 0.012 | 0 | 1971–2013 (4) |
| Tutschgenhalden_14 | 47.5, 8.8 | 0.58 | 580 | 9.1 | 1,151 | 2, *F. sylvatica, P. abies* | 87.7 | 0.008 | 0 | 1971–2013 (4) |
| Vorm Stein_01 | 47.5, 8.5 | 0.25 | 545 | 8.9 | 1,114 | 3, *F. sylvatica P. abies, Q. robur* | 82.0 | 0.018 | +2 | 1972–2012 (4) |
| **Vorm Stein_02** | **47.6, 8.5** | **0.24** | **540** | **9.2** | **1,067** | **4, *Q. robur, F. sylvatica, P. sylvestris, P. abies*** | **86.7** | **0.011** | **+2** | **1972–2012 (4)** |
| Weidwald_02 | 47.4, 8.0 | 0.76 | 635 | 8.7 | 1,163 | 2, *F. sylvatica* | 88.3 | 0.010 | 0 | 1976–2011 (4) |
| Weidwald_03 | 47.4, 8.0 | 0.25 | 660 | 8.6 | 1,180 | 1, *F. sylvatica* | 82.7 | 0.013 | 0 | 1976–2011 (4) |
| Weidwald_04 | 47.4, 8.0 | 0.53 | 640 | 8.6 | 1,180 | 4, *F. sylvatica, Q. robur* | 86.0 | 0.014 | +1 | 1976–2011 (4) |

*Notes:* Elevation is reported as m above sea level (a.s.l.) Stand diversity refers to the number of species with a 10% proportion of stand basal area. Mean survival rates are calculated for an 11-yr period. Stand dynamic refers to the inter-annual rate of change in stem numbers across the entire inventory period. The slope/aspect parameter is an input of ForClim, defined as 0 = flat terrain, 1 = steep slope (10–30°), 2 = very steep slope (>30°); the sign denotes south- (+) and north-facing (−) slopes. The inventory period is the time between the first and the last inventory, with *n* indicating the number of inventories. Genera are spelled out in *Materials and Methods: Data for calibration and validation.*

From the remaining permanent plots, 21 were selected for validating ForClim. They encompass at least two inventories, a minimum record period of 35 yr, and no apparent effects of recent disturbances. The minimum area required was reduced to 0.1 ha to allow for sites with "extreme" climatic conditions (e.g., warm-dry in Pfynwald; Table 1).

### Direct parameter estimates

Individual tree mortality risk is influenced by both individual tree characteristics, and species functional traits and life-history strategies. Among others, tree size, recent growth rates, shade tolerance, and maximum longevity are assumed to be good predictors of tree mortality probability (Bigler and Bugmann 2004, Wunder et al. 2008, Cailleret et al. 2017, Hülsmann et al. 2018) and were therefore used as explanatory variables in a new empirical mortality function. Stand- or tree-level competition indices such as the basal area of larger trees were not directly included as predictors as they can produce ecologically inconsistent results (Bugmann et al. 2019, Thrippleton et al. 2019). Direct parameter estimates (DPE) for this model were derived using single-tree inventory data from the calibration plots ($n = 4{,}663$). The fitted logistic regression expresses the probability of tree $i$ to survive between two inventories from $t_1$ to $t_2$ ($P(Y_{i,t2} = 1)$; over 11 yr on average), according to tree DBH (cm) and annual relative basal area increment (relBAI; Eq. 1), which is then scaled to an annual resolution using a negative exponential survivorship model. Implementing species-specific values for these parameters would be interesting to account for the full diversity of growth patterns prior to death, but this would have drastically increased the number of parameters for recalibration, and consequently simulation time of the Bayesian approach. Thus, we included the categorical variable functionalGroup, which expresses changes in survival probability according to species shade tolerance and longevity (divided into three classes $k$) based on the size–growth–mortality relationships by Hülsmann et al. (2018), with the reference level being the shade-intolerant group with a short life span (see Appendix S3). Together with the intercept of the model ($\alpha_0$), functionalGroup also expresses the "background" mortality, i.e., the mortality independent of tree size and growth:

$$P(Y_{i,t2} = 1) \sim \mathrm{Binomial}(\alpha_0 + \beta_1 \times \log(\mathrm{DBH})_{i,t2} \\ + \beta_2 \times \mathrm{DBH}^2_{i,t2} + \beta_3 \times \mathrm{relBAI}_{i,t2} \\ + \mathrm{functionalgroup}_{i,k}). \quad (1)$$

Both log-transformed and quadratic terms of tree DBH were considered to allow for a flexible size–mortality relationship in both low and high DBH classes (Ruiz-Benito et al. 2013, Hülsmann et al. 2016, 2018). The variable relBAI$_{i,t2}$ was calculated based on stem radial growth between two inventories and the length of the respective growth period ($\Delta t = t_2 - t_1$; yr), assuming a constant growth rate during this period (see Cao [2000] for more sophisticated approaches)

$$\mathrm{relBAI}_{i,t2} = \left(\frac{\mathrm{BA}_{i,t2}}{\mathrm{BA}_{i,t1}}\right)^{\frac{1}{\Delta t}} - 1. \quad (2)$$

The ability of this mortality model to correctly identify tree status was considered as acceptable (area under the receiving operating characteristic curve, AUC = 0.787; Hosmer et al. 2013). In the following, we refer to the ForClim version containing these "direct parameter estimates" obtained from logistic regression as "ForClim-DPE."

### Bayesian calibration

The aim of Bayesian calibration (BC) was to recalibrate the empirical mortality submodel of ForClim-DPE by comparing ForClim outputs of forest structure and composition to the same observations from inventory data without incorporating directly fitted parameter estimates. The BC included all parameters of the empirical mortality function and, additionally, a multiplier of the annual diameter increment of a cohort (kGRateD; %), which is included in the growth submodel of ForClim. The parameter kGRateD scales the simulated annual growth of every cohort irrespective of its size and functional group, this procedure being done before simulating individual tree mortality. The possibility to adjust growth simultaneously with the mortality parameters allows for considering that annual growth may not be accurately simulated by ForClim, subsequently leading to erroneous mortality predictions. Jointly recalibrating several model parameters also provides more flexibility to the calibration approach. For instance, when using stand demographic processes as the unique calibration criterion, trade-offs between growth and mortality are likely to occur because a joint increase of both processes may lead to similar stand productivity, but with a faster turnover (cf. Hartig et al. 2014). It would therefore be uncommonly rigid to only calibrate mortality parameters when fitting a model to stand data. As for ForClim-DPE, the multi-annual cohort-based survival probability calculated by the mortality function is then adjusted to the time step of ForClim to calculate annual survival probability.

BC requires expressing prior knowledge about the parameter values in the prior distribution $P(\theta)$, where $\theta$ denotes the parameter vector of the model (Table 2). Although the estimate and confidence interval of each parameter of the mortality function were determined by logistic regression (which could have been used as a prior), we deliberately set uniform (i.e., non-informative) prior distributions 6–15 times wider than the direct estimates to obtain a picture of the parameters that would

TABLE 2. Direct estimates (maximum likelihood estimates MLE and their standard error se), prior, and posterior distributions for the eight recalibrated parameters.

| Parameter vector $\Theta$ [$\Theta(1)$...$\Theta(8)$] | | | Direct parameter estimates | | Prior probability distribution | | Posterior distribution | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Name | $\Theta(i)$ | Unit | MLE | SE | Min | Max | 2.5% | 97.5% | MAP | Correlated [$\Theta(i)$] |
| kIntercept ($\alpha$) | 1 | | −3.016 | 0.258 | −10 | 0.1 | −2.1058 | −2.1052 | −2.1057 | [−6] |
| klogDBH ($\beta_1$) | 2 | cm | 0.9708 | 0.095 | 0.5 | 1.5 | 1.0058 | 1.0071 | 1.0070 | [−3] |
| kDBH$^2$ ($\beta_2$) | 3 | cm$^2$ | 0.00009 | 0.00009 | −0.001 | 0.001 | −0.000291 | −0.000289 | −0.000291 | [−2] |
| kRelBAI ($\beta_3$) | 4 | % | 57.29 | 4.06 | 0 | 150 | 50.969 | 51.022 | 50.989 | [−7] |
| kFunctionalGroup2 | 5 | | 0.6272 | 0.1501 | −2 | 5 | 1.7716 | 1.7749 | 1.7739 | |
| kFunctionalGroup3 | 6 | | 1.444 | 0.132 | −2 | 5 | 1.6230 | 1.6240 | 1.6234 | [−1] |
| kGRateD | 7 | % | | | 1 | 200 | 56.78 | 56.86 | 56.82 | [−4] |
| kSD | 8 | | | | 0 | 3 | 0.4506 | 0.4931 | 0.4646 | |

*Notes:* The distributions of the priors between a minimum and maximum value were uniform. Posterior distributions are calculated based on 1000 samples after excluding the first 45,000 iterations, and are characterized by the maximum posterior estimates (MAP) and the 2.5% and the 97.5% quantile, respectively (see details in Appendix S4: Fig. S1). For each parameter, the numbers of correlated parameters are listed if the correlation is >0.3 or <−0.3, leading minus signs (−) indicating negative correlations.

result from the inverse calibration alone for the initial posterior estimation (see Table 2). For comparison, we additionally calculated posteriors based on informative priors (details see Appendix S4).

To increase model performance and generality, we constructed a likelihood function consisting of two components: the first part specifies the likelihood for basal area increment (BAI) between two inventories for each species at the plot level, defined by the probability $P(D_{\text{BAI}} \vee \theta)$ of obtaining observed BAI given the model with parameters $\theta$ and a normally distributed error model (cf. Van Oijen et al. 2005). The relative standard deviation of this normal distribution was included in the BC (parameter kSD in Table 2). The second part specifies the likelihood for the species-specific stem number distribution, defined as the probability $P(D_{\text{Stem}} \vee \theta)$ of obtaining the observed stem numbers in 4 cm wide DBH classes for each inventory year, given the model parameters and a Poisson error model per DBH class. The lowest DBH class (i.e., (0, 4] cm) was not considered, as it was not available in the observed data. The partial likelihood for stem numbers is strongly influenced by the very abundant trees in the low DBH classes, whereas medium- and large-sized trees have a stronger influence on basal area increment. Mixing both data types in one joint likelihood function should therefore create a balanced description of forest structure and dynamics in terms of stand density, productivity, and species composition.

Likelihoods were expressed as logarithmic values (log-likelihood), which means that the joint (total) likelihood $P(D \vee \theta)$ can simply be written as their sum. In this sum, we divided the partial likelihood for stem numbers per DBH class by 20 (number of DBH classes in the distribution), which has the result of giving the basal area increment approximately the same weight as the sum of the stem numbers over the DBH classes:

$$P(D|\theta) = P_{\text{BAI}}(D|\theta) + \frac{P_{\text{Stem}}(D|\theta)}{20}. \quad (3)$$

Implicitly, this means that we treat the basal area increment and the distribution of stem numbers per DBH class as one data point each. We acknowledge that this choice is somewhat ad hoc and does not fully account for the correlation within measured stem DBH distributions (see Finley et al. 2014). However, it avoids an overrepresentation of stem numbers in the likelihood, which is unproblematic in theory, but in practice often leads to unbalanced calibrations that satisfy the more data-rich calibration target, while disregarding the other.

Prior and likelihood are combined by Bayes' theorem, which states that the support given to a certain parameter combination, i.e., the posterior probability $P(\theta \vee D)$, is proportional to the prior distribution $P(\theta)$ and the joint likelihood $P(D \vee \theta)$. To estimate the posterior distribution and provide maximum a posteriori (BC-MAP) estimates for the eight parameters, we used the differential evolution Markov chain Monte Carlo (MCMC) algorithm (DEzs; terBraak and Vrugt 2008) from the *BayesianTools* package (Hartig et al. 2017) of the open-source statistical software R v3.4.3 (R Core Team 2017). The convergence of the MCMC was examined by visual inspection of the trace plots and calculating the Gelman-Rubin potential scale-reduction factors (Gelman and Rubin 1992). To speed up convergence of the algorithm, we reinitialized the $z$ matrix (i.e., the pool of past parameter values that are used for creating proposals in the DEzs algorithm) after ~300,000 iterations. We considered parameters with a potential scale-reduction factor smaller than 1.05 to be converged, which occurs after an additional ~60,000 iterations. The entire BC procedure took ~80 d on a single computer node of seven cores (RAM: 64GB). Below, we refer to the ForClim version containing the mortality function with "inverse parameter estimates" as "ForClim-IPE." For estimating

the posterior with informative priors, we did not rerun the entire MCMC with the ForClim model, but instead used MCMC sampling on a target distribution consisting of the informative priors, multiplied with the previously calculated posterior (which, due to the flat priors, is identical to the likelihood up to a constant). The procedure is mathematically identical to calculating the posterior (apart from numerical errors), except that it is much faster, because the model does not have to be run for the likelihood evaluations (for details see Appendix S4).

### Simulation set-up

The initialization of the ForClim simulations follows the procedure described in Bircher (2015) and Hülsmann et al. (2018). For each permanent plot, ForClim was initialized with single-tree data (species, DBH) from the first inventory. Since there was no spatial information about tree position on the plots, trees were allocated randomly and evenly to an initial set of patches, each with a size of 800 $m^2$ (cf. Wehrli et al. 2005). Depending on the ratio of the permanent plot area and patch size (Table 1), this resulted in the direct initialization of 2–44 patches per plot. This initial set of patches was then replicated to average out stochasticity across simulation runs. The final patch number was set to 200 for the validation sites but was reduced to 100 for the calibration sites to reduce simulation time. Some rare tree species that are not parameterized in ForClim, but are present in some forest reserves, albeit with very minor abundance, were not considered in the simulations.

### Evaluation of the mortality submodels in ForClim

We compared three model versions: ForClim with the mortality function based on theoretical assumptions (with both "background" and stress-induced components; v3.3), and two modified ForClim versions where this function was replaced by the new mortality submodel with direct and inverse parameter estimates (DPE and IPE, respectively). After running the models at the calibration and validation sites, we calculated the likelihood values as described previously. Then, we calculated the difference between measured and predicted basal area and stem numbers at the end of the simulation (i.e., at the time of the last inventory) for all species and each functional group. Finally, as model validation is constrained by the short length of the empirical data series (average of ~40 yr), we also evaluated the models' ability to simulate potential natural vegetation (PNV), i.e., the species composition expected in a pseudo-equilibrium state in the absence of anthropogenic influences and large-scale disturbances (Ellenberg and Leuschner 2010), along a well-studied environmental gradient in Switzerland (cf. Bugmann and Solomon 2000). Starting from bare ground, forest dynamics were simulated for 1,500 yr, and forest structure and composition at the

end of the simulation were examined semi-quantitatively for their plausibility (Rasche et al. 2012).

The R scripts developed to calibrate ForClim with a Bayesian approach, and the files required to run ForClim-IPE are available in Data S1.

## RESULTS

### Comparison of parameter estimates from direct vs. inverse calibration

The Bayesian calibration resulted in substantially different parameter estimates than the direct calibration (Table 2). The inversely calibrated maximum a posteriori estimates for the coefficients of all parameters were outside the range of the mean ± SD of the direct estimates, except for klogDBH. Estimates for $kDBH^2$ moved from a positive ($9 \times 10^{-5}$) to a negative value ($-2.89 \times 10^{-4}$). In combination with the positive value of the klogDBH term, this results in an increase in survival probability with DBH until ~40 cm of DBH, after which the trend is reversed (Fig. 1b). With the direct estimates, survival probability gradually increased with DBH, such as reported in the inventory data (Fig. 1a). Thus, survival probabilities are lower for small trees (DBH < 50 cm), but higher for large ones compared to the inverse estimates (Fig. 1a). We also note that using log-transformed DBH as mortality predictor tends to overestimate mortality rates in low DBH classes whatever the parameter estimates: the survival probability of trees with DBH < 10 cm predicted by the mortality function with direct or inverse parameter estimates (Fig. 1b,c, respectively) was lower than the observed survival (Fig. 1a). Survival probabilities for the shade-intolerant species with a long life span and for the shade-tolerant species (functional groups 2 and 3, respectively) increased relatively to the shade-intolerant species with a short life span (reference class; functional group 1; Table 2; Fig. 1b; see also Appendix S3 for the species classification). This was particularly true for the shade-intolerant species with a long life span whose survival probability, everything else being equal, becomes higher than the shade-tolerant species (kFuntionalGroup2 from 0.63 to 1.77 and kFuntionalGroup3 from 1.44 to 1.62), which is not the case with the direct estimates (Fig. 1b). The new kGRateD estimate is quite low (0.57; Table 2), suggesting that species-specific maximum growth should be reduced by ~43% to fit the historical changes in stand BA and stem numbers with the new mortality function based on the inverse estimates. As indicated by the negative correlation (−0.77) between krelBAI and kGRateD, this decrease in individual BAI was compensated by a reduction in the effect of relBAI on survival probability (kRelBAI from 57.29 to 50.99), and by an increase in baseline survival probability (kIntercept from −3.02 to −2.11; Table 2).

Posterior distributions calculated with informative priors, corresponding to a mix between directly and the
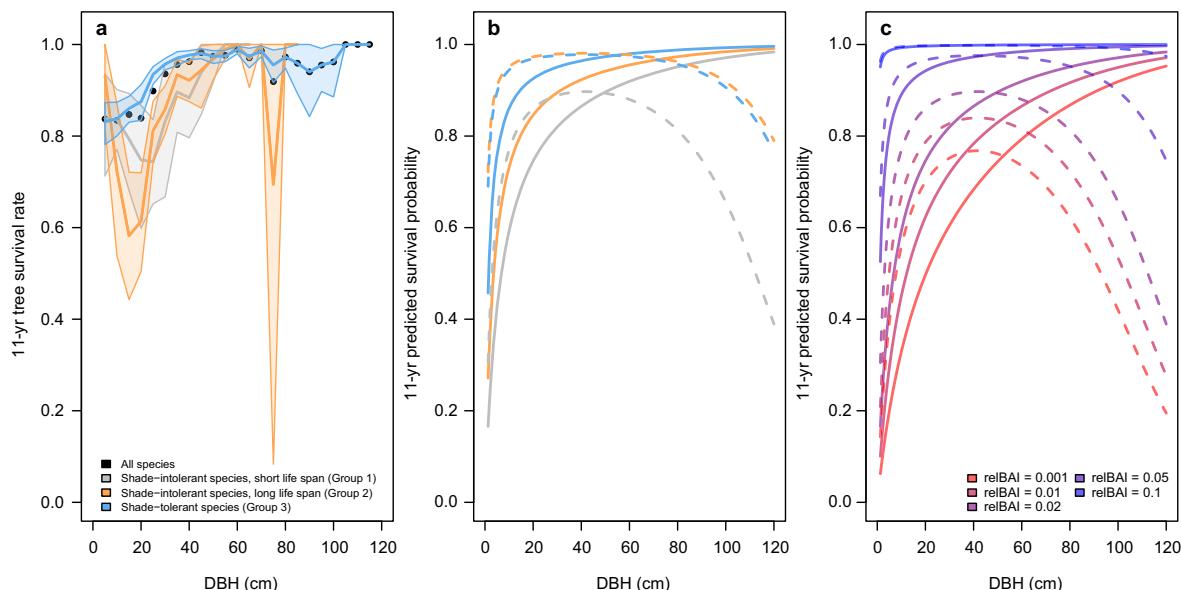
Fig. 1. Observed and simulated tree DBH–survival relationships. (a) Change in the mean survival rate (proportion) reported during an 11-yr period with tree DBH (in 5-cm bin) for each functional group (colored heavy lines) and considering all species (black dots). The 95% confidence intervals of the means (from bootstrap with 500 resamplings) were represented with the shaded areas. (b and c) Change in the 11-yr survival probability with tree DBH as predicted by the inventory-based mortality model with empirical direct parameter estimates (solid lines) and with the inverse parameter estimates obtained by Bayesian calibration (dashed lines) according to functional group (b; for annual relative basal area increment, relBAI = 0.02) and relBAI (c; for functional group 2). Measured DBH at the study sites range between 3.3 cm and 114.7 cm, while the minimum DBH simulated by ForClim is at 1.27 cm (DBH at which trees establish).

inverse estimates, did not lead to noteworthy changes compared to the uninformative inversion (Appendix S4). The reason is that the data were highly informative and the likelihood far narrower than the prior distributions (Table 2). Because all results for the uninformative inversion are therefore also representative for the informative inversion, we do not show this option separately when presenting the performance of different ForClim versions.

### Performance of the model versions on the historical inventory data

*Likelihood values.*—ForClim with the inverse parameter estimates (ForClim-IPE) showed the best average performance of all ForClim versions both at the calibration and independent validation sites (Table 3; Appendix S5), with the highest average total likelihood (−95.0 and −160.8 for the calibration and validation sites, respectively), followed by ForClim v3.3 (−107.7 and −183.6) and ForClim-DPE (−111.1 and −234.1). When looking at specific sites, ForClim-IPE showed the best performance of all models at three out of the nine calibration sites and 11 out of 21 validation sites. This was mainly due to the predictions of stem numbers, for which ForClim-IPE performed distinctly better (highest performance at 19 of the 30 sites) than ForClim-DPE (nine sites) and ForClim v3.3 (two sites). For BAI, all model versions had comparable averaged performance.

ForClim-IPE had the highest individual likelihood at the calibration and validation sites (−59.4 and −111.9, respectively) followed by ForClim v3.3 (−63.8 and −119.6) and ForClim-DPE (−69.7 and −182.9). Contrasting results were obtained when focusing on the number of sites with highest performance (Table 3).

*Simulated bias in basal area and stem numbers at the end of the simulation.*—At the end of the simulation period, ForClim v3.3 and ForClim-DPE generally overestimated total basal area at the validation sites (median in absolute difference: 1.7 and 6.2 m²/ha, respectively), while predictions by the ForClim-IPE version were very close to the measurements (median: −0.6 m²/ha, cf. Fig. 2a). These differences among models were mostly due to different predictions of the BA of shade-tolerant species (functional group 3), which were the main contributors to total stand BA (Fig. 3b). This was especially true for *Fagus sylvatica*, which was strongly overestimated by ForClim-DPE at most sites (Fig. 3a). For this group, all model versions predicted a higher stem number than the observations (+27.3, +32.2, and +71.4 stems/ha for ForClim v3.3, DPE, and IPE, respectively; Fig. 2c,d). This was mainly visible in DBH classes below 50 cm (Fig. 4) and was mainly caused by an underestimation of the mortality rates of small trees, especially by the IPE version (Appendix S6). A similar picture emerged for shade-intolerant species with a short lifespan (group 1). These are rare in the Swiss reserves compared to the

TABLE 3. Comparison of performance at calibration and validation sites for the three ForClim versions: the ForClim version 3.3, the model version featuring the empirical mortality function with the direct parameter estimates (ForClim-DPE), and the inverse parameter estimates obtained by Bayesian calibration (ForClim-IPE).

| Parameter | Stem numbers | | | Basal area increment | | | Total likelihood | | |
|---|---|---|---|---|---|---|---|---|---|
| | v3.3 | DPE | IPE | v3.3 | DPE | IPE | v3.3 | DPE | IPE |
| Average LL | | | | | | | | | |
| Calibration | −43.9 | −41.4 | −35.6 | −63.8 | −69.7 | −59.4 | −107.7 | −111.1 | −95.0 |
| Validation | −63.9 | −51.2 | −48.9 | −119.6 | −182.9 | −111.9 | −183.6 | −234.1 | −160.8 |
| Average rank | | | | | | | | | |
| Calibration | 2.8 | 2.0 | 1.2 | 2.1 | 1.9 | 2.0 | 2.0 | 2.0 | 2.0 |
| Validation | 2.5 | 1.7 | 1.8 | 1.8 | 2.5 | 1.7 | 2.0 | 2.5 | 1.5 |
| Number of sites with best performance | | | | | | | | | |
| Calibration | 0 | 1 | 8 | 2 | 4 | 3 | 3 | 3 | 3 |
| Validation | 2 | 8 | 11 | 9 | 4 | 8 | 6 | 4 | 11 |

*Notes:* The table shows the average values for the individual log-likelihood (LL) components (species stem numbers [divided by 20] and basal area increment) and the total log likelihood, the average rank of the model for each likelihood component, and the number of sites where a model version performed best. Detailed values for each site and standard errors of the average likelihoods are available in Appendix S5.

other species (Figs. 2a,c, 3a), but their correct representation in the model is still important, because they tend to dominate pioneer stages, e.g., after massive mortality events. For this group, BA predictions of ForClim-IPE were also better than those by the other two model versions (−40.4%, −46.8%, and +5.0% for ForClim v3.3, DPE, and IPE, respectively; Fig. 2b), although stem numbers were underestimated by all three versions
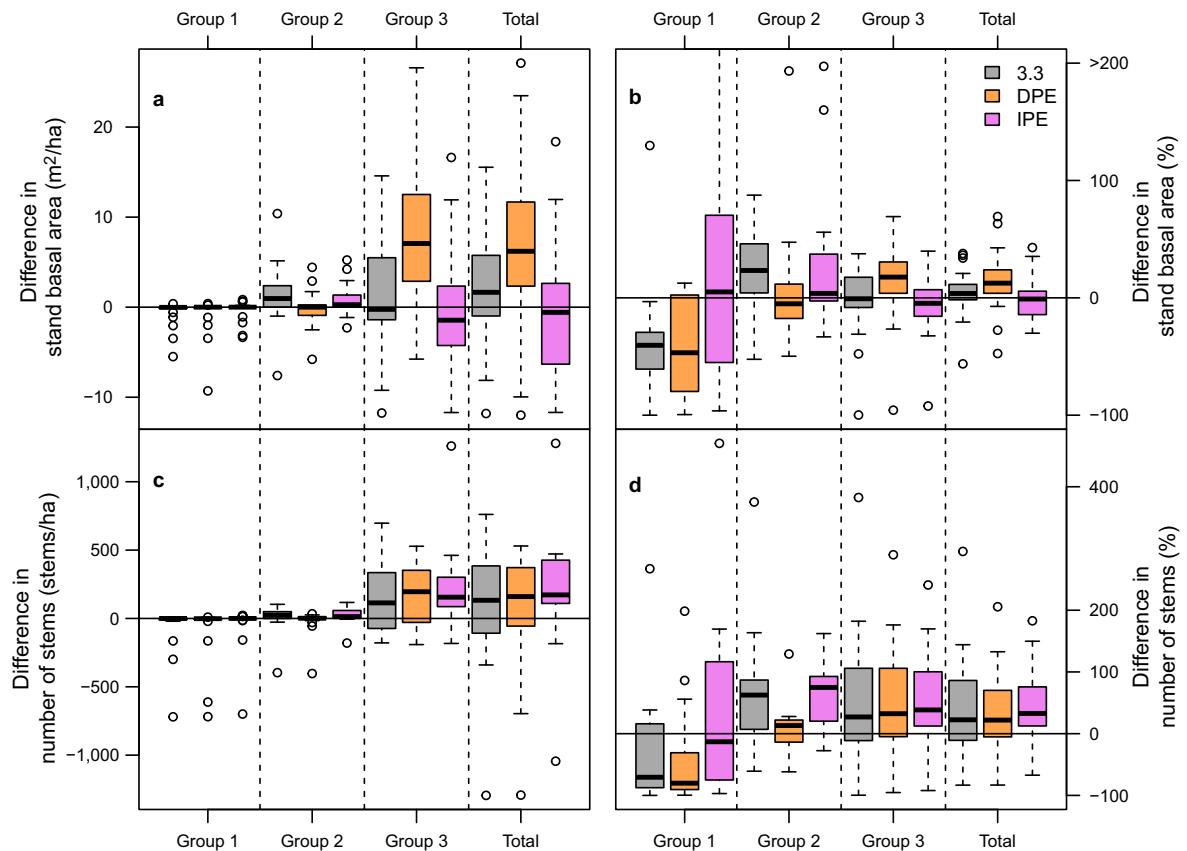


FIG. 2. Box plots of (a, c) absolute and (b, d) relative differences between simulated and observed (a, b) basal area and (c, d) stem numbers at the last inventory, considering all species (Total) and for each functional group. The three ForClim versions (ForClim v3.3, ForClim-DPE, and ForClim-IPE) are represented by different colors; they were run at all validation sites (*n* = 21). Box plot components are mid line, the median values; box edges, the upper and lower quartiles; whiskers, the highest and lowest observed values that lie within 1.5 times the interquartile distance of the upper and lower quartiles, respectively; and points, the outliers.
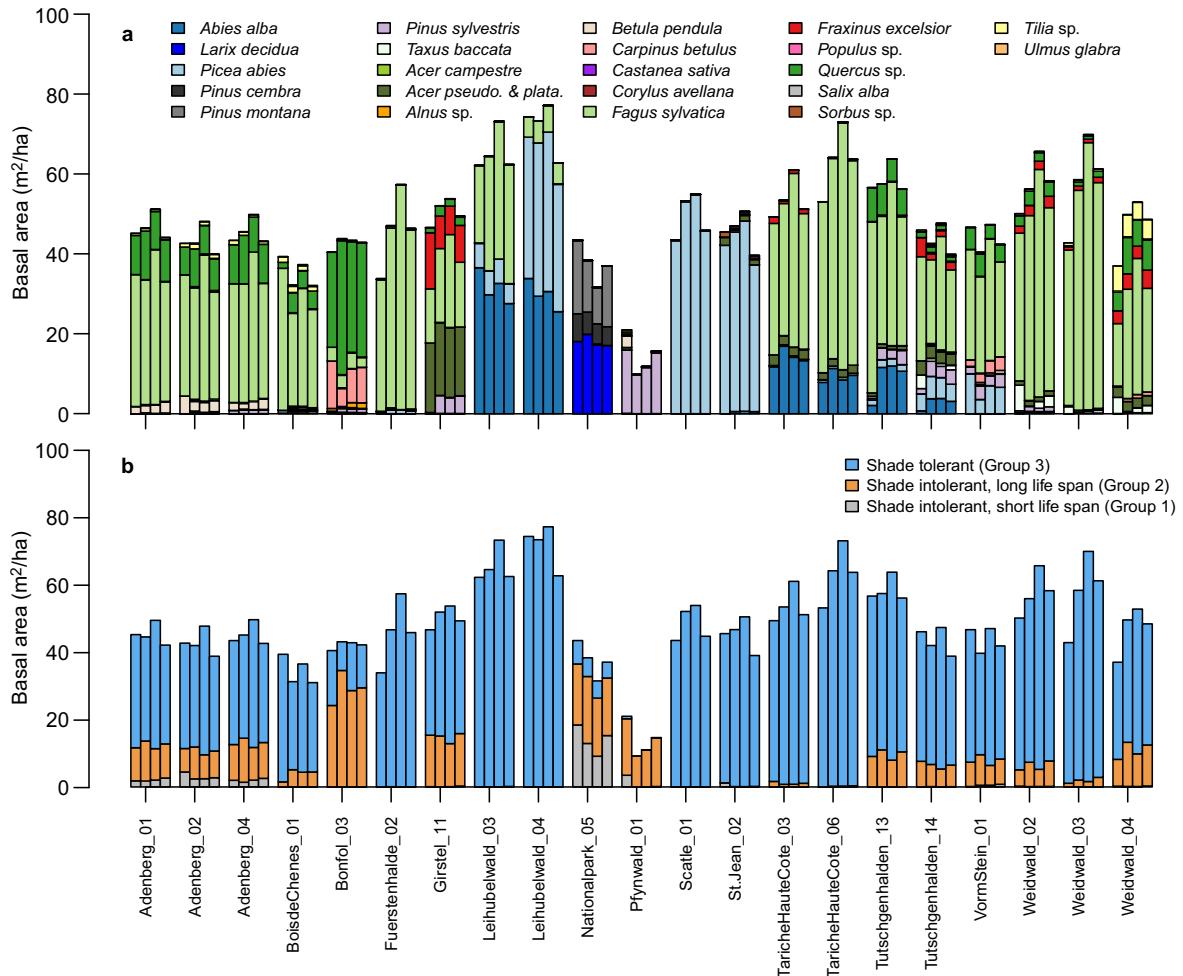
Fig. 3.  Stand composition and basal area measured (leftmost bar) and simulated by ForClim v3.3 (center-left bar), orClim-DPE (center-right bar), and ForClim-IPE (rightmost bar) at the validation sites at the last inventory. Stand composition is represented by (a) species and by (b) functional group and some species from the same genus being gathered in the former figure to improve its graphical quality. Abbreviations in key are pseudo., pseudoplatanus and plata., platanoides.

(−70.6%, −80.2%, and −13.1%; Fig. 2d). This was especially true for trees with a DBH below 24 cm (Fig. 4), caused by a slight overestimation of mortality rates (Appendix S6). For the shade-intolerant species with a long life span (group 2), ForClim-DPE tends to be more accurate than ForClim v3.3 and ForClim-IPE regarding stem numbers (+13.1% vs. +62.5% and +74.9%; Fig. 2d), as it underestimates mortality rates to a lesser extent (Appendix S6). For simulating BA of the functional group 2, ForClim-IPE was as accurate as ForClim-DPE, with relative biases of +3.8% and −5.1%, respectively, while ForClim v3.3 overestimated BA by +23.3% (Fig. 2b).

*Plausibility of long-term projections*

Species composition and BA predicted after 1,500 yr in the center of the Swiss environmental gradient differed considerably between ForClim v3.3, ForClim-DPE, and ForClim-IPE (i.e., at Adelboden, Huttwil, Bern, Schaffhausen, and Basel; Fig. 5; details in Appendix S7). Under these temperate conditions, vegetation is expected to be dominated by *Fagus sylvatica* in mixture with other hardwood species (e.g., *Acer* spp.) and some *Picea abies* and *Abies alba* (Bugmann and Solomon 2000, Rasche et al. 2012). ForClim v3.3 and ForClim-IPE accurately predict such mixed forests; however, they both tend to predict a surprisingly strong presence of either *Castanea sativa* (v3.3 in Schaffhausen and Basel) or *Populus* spp. and *Carpinusbetulus* (ForClim-IPE); mainly to the detriment of *Picea abies*. In contrast, PNV simulated by ForClim-DPE was overly dominated by *Fagus sylvatica* with unrealistic BA values above 100 m²/ha due to the presence of excessively large trees (e.g., DBH > 290 cm in Bern). The high BA of *Populus* spp. in the PNV predicted by ForClim-DPE and -IPE was also related to the occurrence of
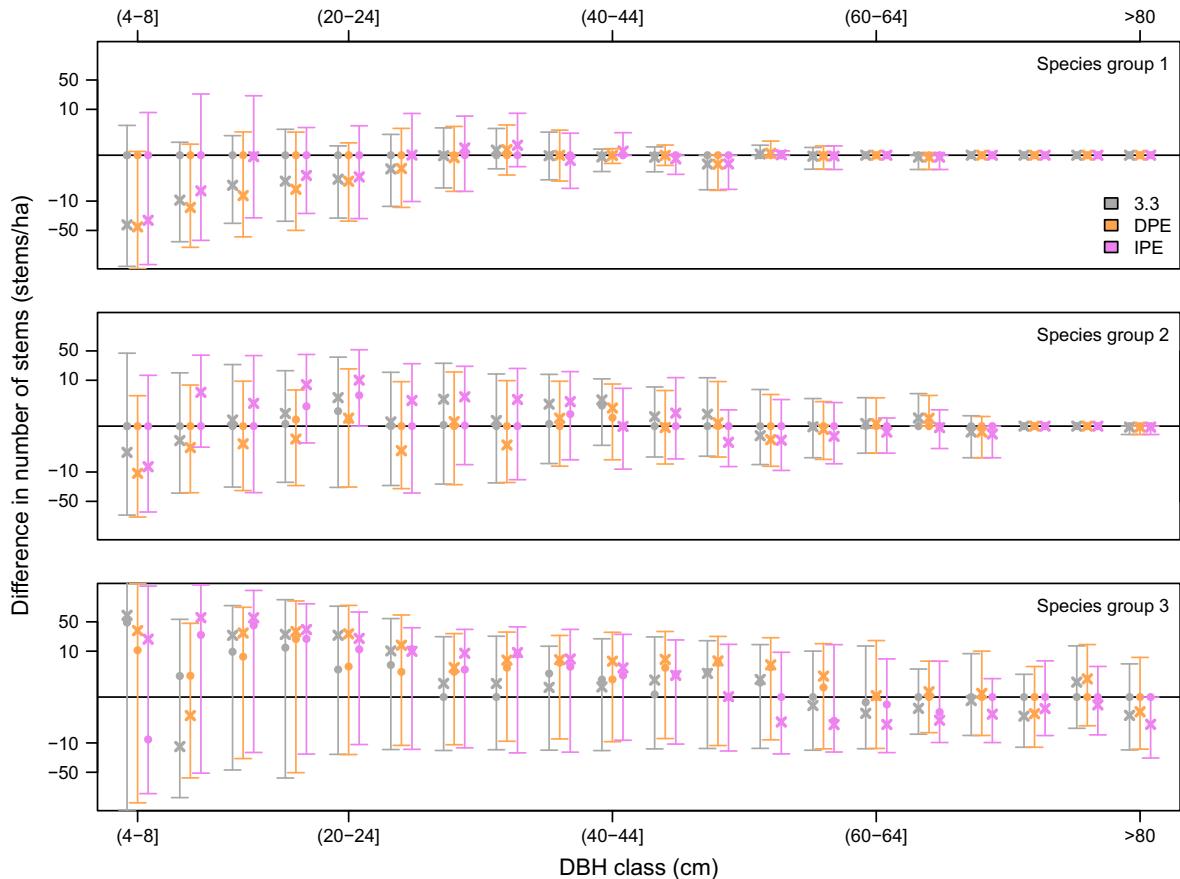
FIG. 4. Absolute differences between simulated and observed stem numbers for different model versions (ForClim v3.3, For-Clim-DPE, and ForClim-IPE): means (crosses), medians (dots), and 95% intervals (whiskers) across all validation sites were calculated for each DBH class and the three functional groups (different panels). Note the log-scale of stem numbers.

few large trees with DBH > 100 cm, which contributed strongly to species BA.

At both ends of the gradient, simulation results are more similar among the three model versions and close to expectations. Simulated PNV at the cold high-elevation sites was dominated by *Pinus cembra* (Bever) or *Picea abies* (Grande Dixence and Davos). At the latter site, the presence of *Populus* spp. predicted by ForClim-IPE was unexpected and had similar causes as mentioned above. Under a dry and warm bioclimate (Sion), all models simulate a forest dominated by *Pinus sylvestris* alone (ForClim v3.3) or mixed with *Quercus* spp. and *Castanea sativa*; both combinations being typical of the study area, and thus realistic (Rigling et al. 2013).

## DISCUSSION

### Comparison of direct and inverse mortality parameter estimates

The parameter estimates of the growth-dependent mortality model obtained directly based on inventory data did not consistently match the ones estimated

inversely with the target of reproducing observed forest structural dynamics by the DVM ForClim. Combining both estimates into a single posterior (direct estimates as informative prior) did not qualitatively change the results compared to the uninformative inversion, as the inverse signal (likelihood) was substantially stronger than the prior (i.e., the posterior parameter estimates were strongly informed by the data). Direct and inverse parameter estimates generally show good agreement for the intercept, logDBH, and relBAI; the latter two having a positive impact on tree survival. However, the BC strongly modified the parameter estimates for $DBH^2$, GRateD, and FunctionalGroups. The fact that the inverse estimates differ quite strongly from the direct estimates either suggests discrepancies between the data used for direct and inverse estimates (e.g., measurement error, non-representativeness of the mortality information used for calibration), or structural problems in the empirical-based mortality function itself or in ForClim (cf. Hartig et al. 2012), as discussed further below.

The BC estimated the effect of $DBH^2$ on tree survival to be negative, which is in line with reports of increased mortality rates of large trees when analyzed in
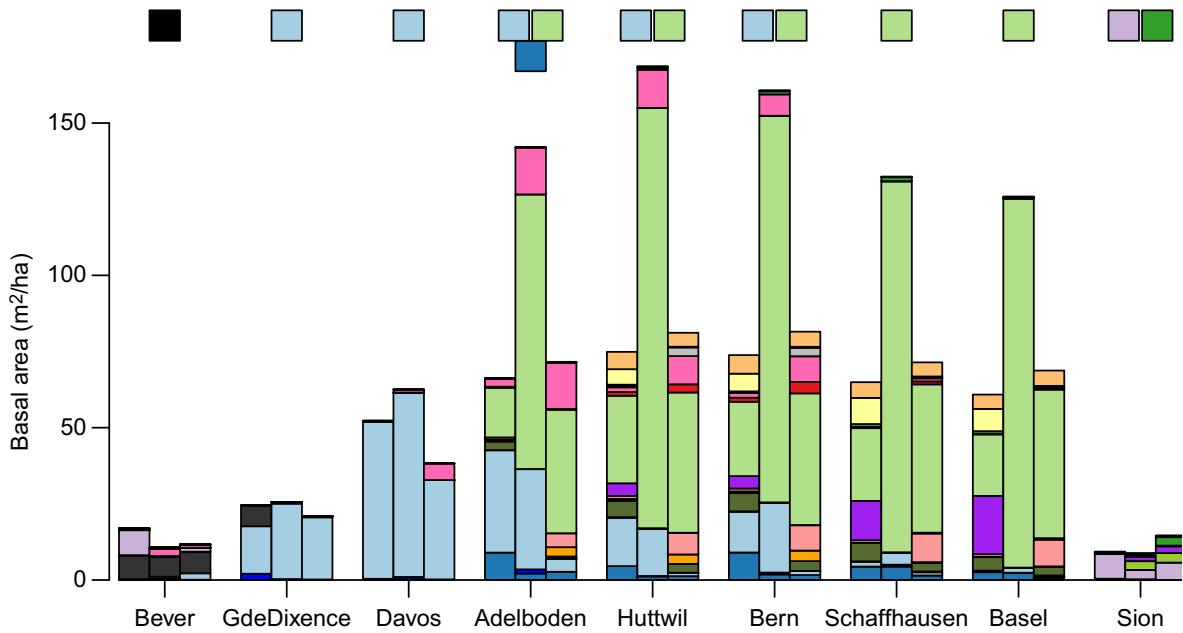
Fig. 5. Potential natural vegetation (PNV, basal area per species, m²/ha) simulated along the Swiss environmental gradient at the end of a 1,500-yr run using ForClim 3.3 (left bar), ForClim-DPE (center bar), and ForClim-IPE (right bar). Expectations of PNV are indicated at the top of each group of bars based on Rasche et al. (2012). Sites were ordered from cold high-elevation (Bever) to dry-warm bioclimates (Sion). For color legend, cf. Fig. 3. Details on the temporal change in species composition and basal area over the 1,500 yr are available for each site and model version in Appendix S7.

combination with the positive linear effect of DBH (e.g., Lorimer and Frelich 1984, Monserud and Sterba 1999, Holzwarth et al. 2013). In contrast, the growth and mortality information from the inventory data at the forest reserves (direct parameter estimates) revealed a weak positive effect of DBH$^2$ on survival, and consequently a typical J-shaped mortality curve with decreasing mortality probability along tree size (see also Hülsmann et al. 2018). The reason for this divergence may be that the forest reserves in Switzerland are still in younger development stages (except Scatlè; Brang et al. 2011), and large trees that would show the right tail of the U-shape curve are largely lacking in the calibration data set (see Appendix S8). Thus, an increased mortality risk due to higher sensitivity to drought-induced xylem embolism, mechanical instability, or insect attacks (Franklin et al. 1987, Bennett et al. 2015, Das et al. 2016) cannot be captured, in contrast to data from true old-growth stands (e.g., U-shaped curves were obtained in a "primeval forest" in Ukraine; Hülsmann et al. 2016). In low DBH classes, the empirical-based mortality function tends to underestimate survival rates, mainly due to the log-transformation of the DBH effect, which seems too strong (Fig. 1). This weakness is less pronounced after the BC thanks to the increase in the model intercept.

The positive effect of the relative basal area increment (relBAI) on survival obtained by both estimation methods reflects that trees with low productivity are usually more prone to die than those that are highly productive (Monserud 1976, Wyckoff and Clark 2002, Cailleret

et al. 2017). Inverse estimates of kRelBAI are lower than direct estimates, mainly due to the strong negative correlation with kGrateD. BC attempts to find an adequate balance between growth and mortality in ForClim to match the calibration data in terms of basal area increment and stem numbers. That is, a decrease in the growth parameter was compensated by an increase in the survival probability at low growth rates.

In the direct empirical parameterization, shade-tolerant species (group 3) had a higher survival probability than shade-intolerant species (groups 1 and 2), in accordance with the theory on life-history strategies of pioneer vs. late-successional species (Grime 1977). The inverse estimates suggest that this difference is not straightforward but further depend on species' maximum longevity, as shade-intolerant species with a long life span (group 2) showed comparable survival probability as shade-tolerant species, all other variables being equal (Fig. 1). Although senescence effects on physiological mechanisms are limited in trees (Mencuccini and Munné-Bosch 2017) due to the low number of somatic mutations (Schmid-Siegert et al. 2017), there is clear evidence that some species live longer than others (Di Filippo et al. 2015). This pattern can be explained by differences in hydraulics and mechanic constraints, and in investment to defenses against pathogens; they are most likely more related to size rather than to age effects (see above; Mencuccini and Munné-Bosch 2017).

The uncertainty in the parameter estimates resulting from the BC, and thus in the predictions of ForClim-

IPE, was relatively low, as reflected by their narrow posterior distributions (Table 2; Appendix S9) compared to other DVM calibrations (e.g., Lagarrigues et al. 2015, Reyer et al. 2016, Augustynczik et al. 2017). We suggest that the main reasons for these low uncertainties are (1) that few parameters were calibrated, (2) the relatively large calibration data set, and (3) the low mortality and regeneration rates of these "young" forest reserves (see above), which was not fully compensated by the selection of highly "dynamic" sites for the calibration, i.e., sites with the with high inter-annual change in stem numbers. Hence, we see no particularly noteworthy pattern in this result as such. However, the uncertainty estimates would have been wider and probably more realistic if we had (1) calibrated all model parameters; (2) considered more complex functional relationships for the mortality; or (3) considered in particular spatial or hierarchical variation in the model parameters or the data through spatial or random effects. For example, there is good evidence for variation in growth–mortality relationships between individuals, sites, and species (e.g., Kane and Kolb 2014, Cailleret et al. 2016), and it would have been possible to include such structures in the likelihood. Yet, those parameters would also have to be estimated, which would considerably increase computation time and is currently not feasible for the ForClim model. In summary, the true predictive uncertainties are likely wider than those we reported here, due to the issues discussed, but we do not see this as a limitation for the conclusions drawn in this study, as these are mainly affected by mean parameter estimates, and not their uncertainty.

### Prediction of short- and long-term forest dynamics

The ForClim-IPE version showed the highest likelihood values at the calibration and validation sites in terms of species-specific BAI and especially stem size distributions. The fact that the BC procedure succeeded in improving the fit also for the independent validation data shows that it did not simply "tune" the parameters to better fit the calibration data (over-fitting), but rather found generally acceptable parameter values for predicting forest dynamics at the decadal time scale.

While good performance of model outputs that were used for calibration is a basic requirement for a good calibration method, a more demanding test is if the calibration also improves the performance of alternative model outputs that were not calibrated (extrapolation). One would expect particularly poor performance if the model has structural problems or biases, which would allow for tuning one model output without improving others. In our case, we did not find any signals of such structural problems. The better performance of ForClim-IPE over ForClim-DPE regarding decadal forest dynamics translated directly into a better performance for centennial-scale forest dynamics (PNV). We see these results as encouraging evidence that, given adequate model

structures, short-term data can improve long-term predictions of DVMs.

When looking at the details of the performance differences, we can make further interesting observations. Focusing on stand BA at the end of the short-term simulations, ForClim-IPE performed best, while ForClim v3.3 was the version with the lowest overestimation of stem numbers (Fig. 2). The overestimation of the BA of shade-tolerant species (functional group 3, especially *Fagus sylvatica*) that was observed for ForClim-DPE was most likely due to the J-shaped mortality curve that is typical of the Swiss reserve data set (Hülsmann et al. 2018). Such an overestimation was not observed any more after BC, due to the strong reduction in the growth rate parameter GRateD and the fact that $kDBH^2$ became negative. Using a U-shaped instead of a J-shaped mortality curve led to an increase in mortality rates of large trees, which contribute the most to stand BA, without strongly decreasing stem numbers, which mainly depend on the small size classes. This modification also improved long-term PNV simulations. The strong prevalence of *Fagus sylvatica* simulated by ForClim-DPE in the center of the Swiss environmental gradient due to the presence of few very big trees with very low mortality probability (see also Hülsmann et al. 2018) was eliminated by the BC, thus leading to more realistic PNV predictions. *Fagus sylvatica* did not outcompete the other species, allowing for the establishment of mixed stands. Yet, some problems remained such as the probable overestimation of *Populus* spp. and *Carpinus betulus*, but these were most likely due to an inappropriate parameterization of shade tolerance in ForClim v3.3 (e.g., according to Landolt et al. 2010, *Populus nigra* should be less shade tolerant than *Picea abies* while their tolerance is identical in ForClim v3.3; see Appendix S3).

All model versions underestimated the stem numbers of shade-intolerant species with a short life span (group 1), especially for trees with a DBH < 24 cm. This can be explained by the overestimation of mortality rates by ForClim v3.3 and ForClim-DPE (Appendix S6), combined with an underestimation of regeneration rates. This bias did not substantially affect the short- and long-term predictions of stand BA and species composition, as we focused on demographic processes that occur during forest succession without external large-scale disturbances such as wind-throws or bark-beetle outbreaks. After massive disturbance-induced mortality events, forest characteristics depend much more on pioneer species (Connell and Slatyer 1977), especially on their regeneration, growth, and mortality rates during the juvenile stages, and would thus be most likely inadequately simulated by all three model versions.

### Perspectives on the use of Bayesian procedures to improve mortality models in DVMs

Science and management require DVMs that provide reliable projections of forest dynamics (Dietze 2017). A necessary condition for having confidence in such

projections is that the models accurately predict past vegetation dynamics, but also that ecological processes, such as tree mortality, are simulated with a high level of mechanistic realism and/or with functions based on large and high-quality empirical data sets (e.g., Larocque et al. 2011, Adams et al. 2013, McDowell et al. 2013). When including empirically based submodels in DVMs, one needs to carefully test the behavior and accuracy of the overall model, especially when the parameters of the submodels are derived independently of the DVM (Hülsmann et al. 2018). Discrepancies between direct and inverse parameter estimates need to be examined and resolved to detect potential structural problems in the model or in the data (cf. Hartig et al. 2012).

Our study highlights that parameter estimates from BC may differ considerably from those deriving from an independent empirical fitting procedure. All model versions reasonably predict decadal-scale stand dynamics on independent validation data sets (median in predictive errors ≤6.2 m²/ha after ≥35 yr of simulation), but the calibration outperformed direct estimates from inventory data, especially for long-term projections.

The low performance of the model with independent field-based parameters may sound counterintuitive, as one would expect direct field estimates to always provide better information about ecological processes, without being subject to interactions of the, possibly flawed, structure of the DVM. However, this pattern could be expected for several reasons. First, in contrast to the IPE estimates, the mortality submodels of ForClim v3.3 and DPE were not calibrated to fit stand basal area and stem numbers, but were based on theoretical assumptions and qualitative knowledge, or calibrated on tree-level mortality rates, respectively. Secondly, ecological models are necessarily abstractions of reality (Bugmann 2001), and their formulations often average ecological variability over nonlinear processes (e.g., intraspecific variability; Chesson 1998), which gives rise to higher-level dynamics and patterns (idea of emergence; e.g., Levin 1992). Even for a structurally correct model, the "best" parameters may not always be those derived directly from field data. Thirdly, direct observations are not always as reliable as one may think. Apart from the obvious possibility of measurement error, field data may also lack representativeness, for example if they are not covering the full range of forest functions and dynamics necessary for the model (here, large tree data are strongly under-represented, leading to J-shaped mortality curves that are inappropriate for long-term simulations). Monitoring many individuals across large areas, with a focus on large trees is highly needed to better capture the spatiotemporal variability in mortality rates and in size-mortality relationships (McMahon et al. 2019). Similarly, the decadal-scale resolution of forest inventory data may mask important ecological processes (e.g., abrupt diebacks induced by short series of drought years). Improving the temporal resolution of the growth and mortality observations may be a solution to this,

e.g., by combining tree-ring and inventory data. However, such sophisticated mortality submodels were found to not necessarily perform better (cf. Vanoni et al. 2019).

Also, the inverse calibration procedure may lead to more accurate higher-level dynamics in such a situation without necessarily being based on an accurate model structure or identifying the correct corresponding parameter values (see Van Oijen 2017). At least partly, the new parameter values may compensate for suboptimal parameter choices elsewhere in the model (e.g., inaccurate shade tolerance parameter for *Populus nigra* as mentioned in the section Prediction of short- and long-term forest dynamics), or for flaws in the model structure (Hartig et al. 2012). For instance, the reduction of the species-specific maximum growth rate that is advocated by the BC (−43.2%; details in Appendix S3) may be exaggerated as ForClim-IPE tends to underestimate individual BAI of large trees for most species (Appendix S10). This highlights the need to better understand the dependencies between different processes (e.g., tree growth and mortality) through improved data acquisition procedures and more realistic model assumptions.

From an ecological perspective, considering the scarcity of information on the mortality process (Hartmann et al. 2018), we believe that fitting models to forest structure can be an interesting route to improve our understanding of mortality patterns across time and space and their underlying mechanisms. Inverse methods could be used to detect and implement DVM-specific growth–mortality relationships or physiological thresholds leading to tree death (Davi and Cailleret 2017), even without dedicated physiological information on the dead trees. Doing so, grouping species into PFTs reduces the number of parameters to be calibrated, but this requires that appropriate functional traits are available that can be linked to stress tolerance and life-history strategies. Ideally, these traits should be close to the physiological mechanisms, such as the species' hydraulic safety margin, which reflects drought vulnerability (O'Brien et al. 2017). This would compensate for unbalanced species coverage (e.g., with *Fagus sylvatica* in the present study), and would allow for the simulation of species for which we do not have mortality information.

Overall, to properly capture mortality responses, flexible mortality functions should be preferred and growth and mortality modules should be re-calibrated simultaneously to allow for nonlinear relationships between growth, size, and mortality. As an alternative to the functions used here, which most likely overestimate mortality rates in low DBH classes, fitting different equations to the low and high DBH classes (e.g., Needham et al. 2018) would provide more flexibility in the size-mortality relationship, but more parameters would have to be estimated. Similarly, even though not considered in the present paper, size or growth effects on mortality rates should differ among the functional groups (see Hülsmann et al. 2018). If growth and mortality rates are estimated, we recommend including variables with

different scales and levels of integration in the combined likelihood function, such as individual-level variables (e.g., tree-level BAI) and stand-scale characteristics (e.g., total basal area).

Finally, using a large variety of sites for the calibration and validation based on forest inventory data does not necessarily support the model's capability of providing accurate predictions beyond this time horizon. Long-term simulations under varying environmental conditions (e.g., under different climate change scenarios; Bircher et al. 2015, Bugmann et al. 2019) or along environmental gradients (e.g., PNV simulations) are needed and highly instructive. In addition, a specific focus on sites with high stress-induced mortality rates (see Hartmann et al. 2018) is recommended to improve model fit and the reliability of model projections.

## Conclusion

We compared two alternative approaches to calibrate a new mortality submodel in the ForClim DVM based on forest inventory data. Our results show that the inverse Bayesian calibration outperformed the calibration based on directly estimated parameters as benchmarked against independent short-term data as well as in long-term projections along an environmental gradient. Although it is difficult to generalize from a single case to DVMs in general, we interpret the higher performance of the mortality function with inversely estimated parameters as evidence for two important conclusions: first, tree mortality is a highly sensitive process in dynamic vegetation models that crucially determines predictions of forest development (Bugmann et al. 2019); and secondly, direct measurements are not always ideal to constrain parameters of complex system models in a way that the models produce satisfactory emergent patterns. Keeping in mind its high computational cost, our results show that Bayesian inference may substantially improve ecological forecasts by merging direct measurements with inverse constraints by comparing emerging outputs to data (see also Dietze 2017, Van Oijen 2017).

## Literature Cited

Adams, H. D., A. P. Williams, C. Xu, S. A. Rauscher, X. Jiang, and N. G. McDowell. 2013. Empirical and process-based approaches to climate-induced forest mortality models. Frontiers in Plant Science 4:438.

Adams, H. D., et al. 2017. A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. Nature Ecology & Evolution 1:1285.

Augustynczik, A. L., F. Hartig, F. Minunno, H. P. Kahle, D. Diaconu, M. Hanewinkel, and R. Yousefpour. 2017. Productivity of *Fagus sylvatica* under climate change—A Bayesian analysis of risk and uncertainty using the model 3-PG. Forest Ecology and Management 401:192–206.

Bennett, A. C., N. G. McDowell, C. D. Allen, and K. J. Anderson-Teixeira. 2015. Larger trees suffer most during drought in forests worldwide. Nature Plants 1:15139.

Beven, K. 2016. Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. Hydrological Sciences Journal 61:1652–1665.

Bigler, C., and H. Bugmann. 2004. Predicting the time of tree death using dendrochronological data. Ecological Applications 14:902–914.

Bircher, N. 2015. To die or not to die: forest dynamics in Switzerland under climate change. Thesis. ETH Zürich, Zurich, Switzerland. https://doi.org/10.3929/ethz-a-010596194

Bircher, N., M. Cailleret, and H. Bugmann. 2015. The agony of choice: different empirical mortality models lead to sharply different future forest dynamics. Ecological Applications 25:1303–1318.

Bonan, G. B., S. Levis, S. Sitch, M. Vertenstein, and K. W. Oleson. 2003. A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics. Global Change Biology 9:1543–1566.

Brang, P., C. Heiri, and H. Bugmann. 2011. Waldreservate. 50 Jahre natürliche Waldentwicklung in der Schweiz. Eidg. Forschungsanstalt WSL, Birmensdorf; ETH Zürich, Zurich, Swirtzerland.

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

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

Bugmann, H. K., and A. M. Solomon. 2000. Explaining forest composition and biomass across multiple biogeographical regions. Ecological Applications 10:95–114.

Bugmann, H., et al. 2019. Tree mortality submodels drive simulated long-term forest dynamics: an assessment across 15 models from the stand to the global scale. Ecosphere 10:e02616.

Cailleret, M., et al. 2016. Towards a common methodology for developing logistic tree mortality models based on ring-width data. Ecological Applications 26:1827–1841.

Cailleret, M., et al. 2017. A synthesis of radial growth patterns preceding tree mortality. Global Change Biology 23:1675–1690.

Cao, Q. V. 2000. Prediction of annual diameter growth and survival for individual trees from periodic measurements. Forest Science 46:127–131.

Chesson, P. 1998. Making sense of spatial models in ecology. Pages 151–166 in J. Bascompte and R. Sole, editors. Modeling spatiotemporal dynamics in ecology. Landes Bioscience, Austin, Texas, USA.

Connell, J. H., and R. O. Slatyer. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. American Naturalist 111:1119–1144.

Das, A. J., N. L. Stephenson, and K. P. Davis. 2016. Why do trees die? Characterizing the drivers of background tree mortality. Ecology 97:2616–2627.

Davi, H., and M. Cailleret. 2017. Assessing drought-driven mortality trees with physiological process-based models. Agricultural and Forest Meteorology 232:279–290.

Di Filippo, A., N. Pederson, M. Baliva, M. Brunetti, A. Dinella, K. Kitamura, H. D. Knapp, B. Schirone, and G. Piovesan. 2015. The longevity of broadleaf deciduous trees in Northern Hemisphere temperate forests: insights from tree-ring series. Frontiers in Ecology and Evolution 3:46.

Didion, M., A. D. Kupferschmid, M. J. Lexer, W. Rammer, R. Seidl, and H. Bugmann. 2009. Potentials and limitations of using large-scale forest inventory data for evaluating forest succession models. Ecological Modelling 220:133–147.

Dietze, M. C. 2017. Ecological forecasting. Princeton University Press, Princeton, New Jersey, USA.

Ellenberg, H., and C. Leuschner 2010. Vegetation Mitteleuropas mit den Alpen. Eugen Ulmer Verlag, Stuttgart, Germany.

Finley, A. O., S. Banerjee, A. R. Weiskittel, C. Babcock, and B. D. Cook. 2014. Dynamic spatial regression models for space-varying forest stand tables. Environmetrics 25:96–609.

Fisher, R. A., et al. 2018. Vegetation demographics in Earth System Models: a review of progress and priorities. Global Change Biology 24:35–54.

Franklin, J. F., H. H. Shugart, and M. E. Harmon. 1987. Tree death as an ecological process. BioScience 37:550–556.

Friend, A. D., et al. 2014. Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric $CO_2$. Proceedings of the National Academy of Sciences USA 111:3280–3285.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–472.

Grime, J. P. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. American Naturalist 111:1169–1194.

Grimm, V., and S. F. Railsback. 2012. Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. Philosophical Transactions of the Royal Society B 367:298–310.

Gutiérrez, A. G., R. S. Snell, and H. Bugmann. 2016. Using a dynamic forest model to predict tree species distributions. Global Ecology and Biogeography 25:347–358.

Hartig, F., J. Dyke, T. Hickler, S. I. Higgins, R. B. O'Hara, S. Scheiter, and A. Huth. 2012. Connecting dynamic vegetation models to data—an inverse perspective. Journal of Biogeography 39:2240–2252.

Hartig, F., C. Dislich, T. Wiegand, and A. Huth. 2014. Approximate Bayesian parameterization of a process-based tropical forest model. Biogeosciences 11:1261–1272.

Hartig, F., F. Minunno, and S. Paul. 2017. BayesianTools: General-Purpose MCMC and SMC samplers and tools for Bayesian statistics. R package version 0.1.0. https://CRAN.R-project.org/package=BayesianTools

Hartmann, H., et al. 2018. Research frontiers for improving our understanding of drought-induced tree and forest mortality. New Phytologist 218:15–28.

Holzwarth, F., A. Kahl, J. Bauhus, and C. Wirth. 2013. Many ways to die–partitioning tree mortality dynamics in a near-natural mixed deciduous forest. Journal of Ecology 101:220–230.

Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant. 2013. Applied logistic regression. Wiley, Hoboken, New Jersey, USA.

Huber, N., H. Bugmann, and V. Lafond. 2018. Global sensitivity analysis of a dynamic vegetation model: model sensitivity depends on successional time, climate and competitive interactions. Ecological Modelling 368:377–390.

Hülsmann, L., H. K. Bugmann, B. Commarmot, P. Meyer, S. Zimmermann, and P. Brang. 2016. Does one model fit all? Patterns of beech mortality in natural forests of three European regions. Ecological Applications 26:2463–2477.

Hülsmann, L., H. Bugmann, M. Cailleret, and P. Brang. 2018. How to kill a tree–Empirical mortality models for eighteen species and their performance in a dynamic forest model. Ecological Applications 28:522–540.

Kane, J. M., and T. E. Kolb. 2014. Short-and long-term growth characteristics associated with tree mortality in southwestern mixed-conifer forests. Canadian Journal of Forest Research 44:1227–1235.

Keane, R. E., M. Austin, C. Field, A. Huth, M. J. Lexer, D. Peters, A. Solomon, and P. Wyckoff. 2001. Tree mortality in gap models: application to climate change. Climatic Change 51:509–540.

Lagarrigues, G., F. Jabot, V. Lafond, and B. Courbaud. 2015. Approximate Bayesian computation to recalibrate individual-based models with population data: Illustration with a forest simulation model. Ecological Modelling 306:278–286.

Landolt, E., B. Bäumler, A. Erhardt, O. Hegg, F. Klötzli, W. Lämmler, et al. 2010. Flora indicativa= Ecological indicator values and biological attributes of the flora of Switzerland and the Alps: ökologischeZeigerwerte und biologischeKennzeichenzur Flora der Schweiz und der Alpen. Haupt Verlag, Bern, Switzerland.

Larocque, G. R., L. Archambault, and C. Delisle. 2011. Development of the gap model ZELIG-CFS to predict the dynamics of North American mixed forest types with complex structures. Ecological Modelling 222:2570–2583.

Larssen, T., R. B. Huseby, B. J. Cosby, G. H st, T. H gÍsen, and M. Aldrin. 2006. Forecasting acidification effects using a Bayesian calibration and uncertainty propagation approach. Environmental Science and Technology 40:7841–7847.

Le Roux, X., A. Lacointe, A. Escobar-Gutiérrez, and S. Le Dizès. 2001. Carbon-based models of individual tree growth: a critical appraisal. Annals of Forest Science 58:469–506.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Lorimer, C. G., and L. E. Frelich. 1984. A simulation of equilibrium diameter distributions of sugar maple (*Acer saccharum*). Bulletin of the Torrey Botanical Club 193–199.

Mäkelä, A., J. Landsberg, A. R. Ek, T. E. Burk, M. Ter-Mikaelian, G. I. Ågren, C. Oliver, and P. Puttonen. 2000. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiology 20:289–298.

Manusch, C., H. Bugmann, C. Heiri, and A. Wolf. 2012. Tree mortality in dynamic vegetation models–a key feature for accurately simulating forest properties. Ecological Modelling 243:101–111.

McDowell, N. G., et al. 2013. Evaluating theories of drought-induced vegetation mortality using a multimodelexperiment framework. New Phytologist 200:304–321.

McMahon, S. M., G. Arellano, and S. J. Davies. 2019. The importance and challenges of detecting changes in forest mortality rates. Ecosphere 10:e02615.

Mencuccini, M., and S. Munné-Bosch. 2017. Physiological and biochemical processes related to ageing and senescence in plants. Pages 257–283 in R. P. Shefferson, O. R. Jones, and R. Salguero-Gómez, editors. The evolution of senescence in the tree of life. Cambridge University Press, Cambridge, UK.

Mina, M., H. Bugmann, M. Klopcic, and M. Cailleret. 2017. Accurate modeling of harvesting is key for projecting future

forest dynamics: a case study in the Slovenian mountains. Regional Environmental Change 17:49–64.

Minunno, F., M. Peltoniemi, S. Launiainen, M. Aurela, A. Lindroth, A. Lohila, I. Mammarella, K. Minkkinen, and A. Mäkelä. 2016. Calibration and validation of a semi-empirical flux ecosystem model for coniferous forests in the Boreal region. Ecological Modelling 341:37–52.

Monserud, R. A. 1976. Simulation of forest tree mortality. Forest Science 22:438–444.

Monserud, R. A., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. Forest Ecology and Management 113:109–123.

Moorcroft, P. R. 2006. How close are we to a predictive science of the biosphere? Trends in Ecology & Evolution 21:400–407.

Moore, A. D. 1989. On the maximum growth equation used in forest gap simulation models. Ecological Modelling 45:63–67.

Needham, J., C. Merow, C. H. Chang-Yang, H. Caswell, and S. M. McMahon. 2018. Inferring forest fate from demographic data: from vital rates to population dynamic models. Proceedings of the Royal Society B 285:20172050.

O'Brien, M. J., et al. 2017. A synthesis of tree functional traits related to drought-induced mortality in forests across climatic-zones. Journal of Applied Ecology 54:1669–1686.

O'Hara, R. B., E. Arjas, H. Toivonen, and I. Hanski. 2002. Bayesian analysis of metapopulation data. Ecology 83:2408–2415.

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

R Core Team 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org

Rasche, L., L. Fahse, A. Zingg, and H. Bugmann. 2012. Enhancing gap model accuracy by modeling dynamic height growth and dynamic maximum tree height. Ecological Modelling 232:133–143.

Reinds, G. J., M. Van Oijen, G. B. Heuvelink, and H. Kros. 2008. Bayesian calibration of the VSD soil acidification model using European forest monitoring data. Geoderma 146:475–488.

Reyer, C. P., M. Flechsig, P. Lasch-Born, and M. Van Oijen. 2016. Integrating parameter uncertainty of a process-based model in assessments of climate change effects on forest productivity. Climatic Change 137:395–409.

Rigling, A., et al. 2013. Driving factors of a vegetation shift from Scots pine to pubescent oak in dry Alpine forests. Global Change Biology 19:229–240.

Ruiz-Benito, P., E. R. Lines, L. Gómez-Aparicio, M. A. Zavala, and D. A. Coomes. 2013. Patterns and drivers of tree mortality in Iberian forests: climatic effects are modified by competition. PLoS ONE 8:e56843.

Schmid-Siegert, E., et al. 2017. Low number of fixed somatic mutations in a long-lived oak tree. Nature Plants 3:926–929.

Sitch, S., et al. 2008. Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). Global Change Biology 14:2015–2039.

Solomon, A. M. 1986. Transient response of forests to CO2-induced climate change: simulation modeling experiments in eastern North America. Oecologia 68:567–579.

terBraak, C. J., and J. A. Vrugt. 2008. Differential evolution Markov chain with snooker updater and fewer chains. Statistics and Computing 18:435–446.

Thrippleton, T., L. Hülsmann, M. Cailleret, and H. Bugmann. 2019. Projecting forest dynamics across Europe: potential and pitfalls of empirical mortality algorithms. Ecosystems. https://doi.org/10.1007/s10021-019-00397-3

Valladares, F., and Ü. Niinemets. 2008. Shade tolerance, a key plant feature of complex nature and consequences. Annual Review of Ecology, Evolution, and Systematics 39:237–257.

Van Oijen, M. 2017. Bayesian methods for quantifying and reducing uncertainty and error in forest models. Current Forestry Reports 3:269–280.

Van Oijen, M., J. Rougier, and R. Smith. 2005. Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiology 25:915–927.

Van Oijen, M., D. R. Cameron, K. Butterbach-Bahl, N. Farahbakhshazad, P. E. Jansson, R. Kiese, K.-H. Rahn, C. Wernerbe, and J. B. Yeluripatif. 2011. A Bayesian framework for model calibration, comparison and analysis: application to four models for the biogeochemistry of a Norway spruce forest. Agricultural and Forest Meteorology 151:1609–1621.

Van Oijen, M., et al. 2013. Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. Forest Ecology and Management 289:255–268.

Vanoni, M., M. Cailleret, L. Hülsmann, H. Bugmann, and C. Bigler. 2019. How do tree mortality models from combined tree-ring and inventory data affect projections of forest succession? Forest Ecology and Management 433:606–617.

Wang, D., D. LeBauer, and M. Dietze. 2013. Predicting yields of short-rotation hybrid poplar (*Populus* spp.) for the United States through model–data synthesis. Ecological Applications 23:944–958.

Wehrli, A., A. Zingg, H. Bugmann, and A. Huth. 2005. Using a forest patch model to predict the dynamics of stand structure in Swiss mountain forests. Forest Ecology and Management 205:149–167.

Wunder, J., B. Brzeziecki, H. Żybura, B. Reineking, C. Bigler, and H. Bugmann. 2008. Growth–mortality relationships as indicators of life-history strategies: a comparison of nine tree species in unmanaged European forests. Oikos 117:815–828.

Wyckoff, P. H., and J. S. Clark. 2002. The relationship between growth and mortality for seven co-occurring tree species in the southern Appalachian Mountains. Journal of Ecology 90:604–615.