Evaluating five forest models using multi-decadal inventory data from mountain forests

Florian Irauschek, Ivan Barka, Harald Bugmann, Benoit Courbaud, Che
Elkin, Tomáš Hlásny, Matija Klopcic, Marco Mina, Werner Rammer, Manfred
Lexer

HAL Id: hal-04042993
https://hal.inrae.fr/hal-04042993

Submitted on 23 Mar 2023

Evaluating five forest models using multi-decadal inventory data from mountain forests

Florian Irauschek [a,b,*], Ivan Barka [c], Harald Bugmann [d], Benoit Courbaud [e], Che Elkin [f], Tomáš Hlásny [g], Matija Klopcic [h], Marco Mina [d,i], Werner Rammer [a,j], Manfred J Lexer [a]

[a] *Institute of Silviculture, University of Natural Resources and Life Sciences (BOKU), Peter Jordan-Straße 82, 1190 Vienna, Austria*
[b] *Department of Forest Growth, Silviculture and Genetics, Austrian Research Centre for Forests (BFW), Seckendorff-Gudent-Weg 8, 1131 Vienna, Austria*
[c] *National Forest Centre – Forest Research Institute in Zvolen, T. G. Masaryka 22, 960 01 Zvolen, Slovakia*
[d] *Forest Ecology, Institute of Terrestrial Ecosystems, ETH Zürich, Universitätstraße 16, 8092 Zürich, Switzerland*
[e] *University Grenoble Alpes, INRAE, LESSEM, 2 rue de la Papeterie-BP 76, F-38402 St-Martin-d'Hères, France*
[f] *University of Northern British Columbia, 3333 University Way, Prince George, Canada*
[g] *Czech University of Life Sciences Prague, Faculty of Forestry and Wood Sciences, Kamýcká 129, 165 21 Prague 6, Czech Republic*
[h] *University of Ljubljana, Biotechnical Faculty, Department of Forestry and Renewable Forest Resources, Vecna pot 83, 1000 Ljubljana, Slovenia*
[i] *Centre for Forest Research, Université du Québec à Montréal UQAM, Montréal, Canada*
[j] *Ecosystem Dynamics and Forest Management Group, School of Life Sciences, Technical University of Munich, Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising, Germany*

ARTICLE INFO

ABSTRACT

Forest ecosystem models, being widespread science tools and used for forest management decision support are usually evaluated individually against field data sets, while model intercomparison and joint evaluation studies are rare. We tested five forest models according to a harmonized protocol against data from nine forest compartments in the Sněžnik region, in Slovenia. The suite of models included stand- and landscape-scale, empirical- and process-based models used across Europe. The test dataset originated from inventory data covering 50 years (tree measurements 1963, 1983 and 2013) and included annual harvesting records at tree level. Uncertainties in data and forest conditions were considered by defining 12 scenarios varying initial regeneration, browsing pressure and harvest modalities. We evaluated the models` ability to initialize forest conditions accurately, whether management interventions could be implemented based on harvest records, and how well basal area and diameter structure could be predicted. Simulation results for basal area development showed good to satisfactory performance for all models, at which SAMSARA2, SIBYLA and PICUS showed the best agreement. Comparison of simulated and observed diameter distributions showed good performance of ForClim, PICUS, SAMSARA2 and SIBYLA. Model output variability was between 6% and 24%, indicating the relevance to consider uncertainties that can be attributed to specific sources. There was no clear hierarchy between more empirical or more process-based models regarding accuracy of stand development projections. The cohort-based landscape model LandClim showed the lowest stand-level accuracy and scenario sensitivity, but results nevertheless qualified it for complementary application at landscape scale. Within individual-based models, spatially explicit models seemed to be more suitable for heterogeneous mixed mountain forests. The findings demonstrated the usefulness of inventory datasets for model testing and intercomparison.

1. Introduction

Forest simulation models are powerful tools for testing and evaluating the mid- to long-term implications of different management strategies on future forest development and related ecosystem service provisioning (Söderbergh and Ledermann 2003; Schelhaas et al., 2014). Changing environmental conditions and intensifying disturbance regimes (e.g., Seidl et al., 2011; Temperli et al., 2013) have increased the complexity in forest resource planning and management, and consequently the role of model-based decision support has drawn a lot of

attention recently (Muys et al., 2011; Linkevičius et al., 2019).

One response to this growing challenge has been an increased reliance on forest simulation models, which are potentially climate sensitive and allow for alternative and novel forest management strategies, to be evaluated. Concomitantly, there has also been an increase in testing the benefits, limits and credibility of forest models (Courbaud et al., 2015; Foster et al., 2017). Model evaluation studies are required to build trust and confidence in model outputs and are thus a prerequisite for any model application in practical decision support. In view of growing interest in complex forest structures, multi-species mixtures, the provisioning of various ecosystem services beyond timber production and the need to consider the effects of a changing climate, the demands being placed on forest models have grown considerably over the recent years. Silvicultural regimes that have been proposed to adapt forests to climate change often focus on small-scale silvicultural measures and the creation of heterogeneous stand conditions to foster forest resilience (Christensen 1997; Puettmann et al., 2009). Adequately representing these forest structures, and the underlying ecosystem processes that generate them, requires individual-based, climate-sensitive forest modeling approaches that allow the simulation of complex silvicultural tree selection and cutting patterns (Söderbergh and Ledermann 2003; Grimm et al., 2005; Mina et al., 2017).

From the perspective of a potential model user, it can be difficult to decide what the appropriate model for a specific task and location should be based on its original theoretical concept. The reason is that many models, being scientific tools under permanent development rather than ready products, are continuously refined, extended or hybridized with other models (Larocque et al., 2011). Therefore, to identify strengths and weaknesses of different models, intercomparison studies are recommended (Huber et al., 2013), where multiple models are tested within a harmonized framework against independent observational data. While model comparison studies are frequently published for carbon and water flux models (e.g., Ryan et al., 1996; Hanson et al., 2004; Jin et al., 2016; Thurner et al., 2017) multi-model evaluation studies for forest ecosystem models are rare. For instance, Badeck et al. (2001) tested six gap models against observed structure and species composition of a virgin forest. Other studies comparing models originating from the same region have been published by Härkönen et al. (2010), Huber et al. (2013) and McCullagh et al. (2017). Recently, Bugmann et al. (2019) compared the behavior of mortality algorithms implemented in several forest models. However, studies comparing forest models that originate from different countries and different ecological and management contexts against long-term observational data sets from managed forests are rare (but see e.g., Mäkelä et al., 2000; Lindner et al., 2005).

Consistent data sets from managed multi-species forests that extend over several decades are rare, particularly when tree-level data are required. Usually, data from silvicultural experiments are utilized for this purpose (Mäkelä et al., 2000; Yaussy 2000; Lindner et al., 2005; Seidl et al., 2005). When long-term observational data are used for a model evaluation study, the issue of information quality arises (Gadow 2000). Historical data from decades ago may be subject to uncertainty with regard to accuracy of measurements, and there may be gaps with regard to tree species-specific information, and unknown calipering thresholds. Moreover, usually no information about tree positions, forest structure or spatial species mixture types is available. The timing of harvests as well as the composition of harvested volume (species, dead and alive trees) may also not be known exactly. Given the relevance of legacies for future forest development, erroneous initial forest conditions may propagate over time and increase uncertainty, which limits the power of model evaluation studies.

In the FP7 project ARANGE (http://www.arange-project.eu), several forest models, originally developed for different European forest types and representing different conceptual modeling approaches, were employed to explore management alternatives for mountain forests in major European mountain ranges (Bugmann et al., 2017). In addition,

multi-decade forest inventory data from the Dinaric Mountains in Slovenia were available within the project consortium. This setting provided the opportunity to compare five established models in a model intercomparison study.

Specific questions of the study were:

(1) How well can forest models be initialized with historical inventory data?
(2) How well can forest models implement historical management schemes derived from harvest records?
(3) How well do observed stand trajectories and model simulations match with regard to volume, basal area and diameter structure?

2. Material and methods

2.1. Study area

The observational time series data comes from an area near the Snĕžnik Mountain (1796 m a.s.l.), in the northern part of the Dinaric Mountains, Slovenia, Europe. The Snĕžnik area is a karst limestone plateau, transformed in the last glacial period. The soils are mainly chromic Cambisols and rendzic Leptosols. The climate in the northern Dinaric Mountains has Mediterranean influences, with warm summer temperatures (long-term mean from July to August is 18.3 °C at 800 m a. s.l., and 14.9 °C at 1300 m a.s.l.) and low winter temperatures (mean January temperatures −0.6 °C to −4.1 °C). Mean annual temperature ranges from 6.8 °C at 800 m to 3.1 °C at 1300 m, with annual precipitation between 1670 mm and 1930 mm, respectively. Mean summer precipitation (May to September) ranges from 650 to 740 mm. The upper timberline is located at approximately 1550 m.

Mountainous silver fir (*Abies alba Mill.*) - European beech (*Fagus sylvatica* L.) - Norway spruce (*Picea abies Karst.*) forests are the prevailing natural forest type, with frequent occurrence of sycamore maple (*Acer pseudoplatanus* L.) and wych elm (*Ulmus glabra Huds.*), while small-leaved lime (*Tilia cordata Mill.*), rowan (*Sorbus aucuparia* L.), common whitebeam (*Sorbus aria (L.) Crantz*), yew (*Taxus baccata* L.) and some other species can also be found sporadically. The first major regular utilization of these forests started in the second half of the 19th century when silver fir was promoted, while in line with the economic principles of that time, beech was weeded out and used for charcoal and potash production and wood distillation (Perko 2002). At the beginning of the 20th century, an uneven-aged single stem selection system (i.e., plenter system) was introduced (Schollmayer 1906). Due to a noticeable decrease in fir vitality and its insufficient regeneration and recruitment (Klopcic et al., 2010), a combination of single stem and small-scale irregular shelterwood system was introduced in the 1960s. Afterwards it was adapted to a more flexible, site and stand specific continuous cover system labeled "free style silviculture", combining elements of single stem selection, irregular shelterwood and shelterwood approaches, which has been applied since then (Mlinsek 1968; Boncina 2011).

Within the Sneznik area a set of nine compartments with a total area of 60.0 ha was chosen for the model evaluation study. The sites are located apart from each other on elevations between 800 m to 1300 m a. s.l..

2.2. Forest inventory dataset

Along with the introduction of uneven-aged forest management at the beginning of the 20th century, a permanent division of forests into compartments was established and since 1912 eight forest inventories have been conducted. Before 1973 inventories were implemented by fully callipering the compartments. In 1973 and 1983, inventories were executed as full callipering of a sample of compartments, while in 1993 permanent sample plots were established and used since then. The full-callipering data were available from inventories in 1963 and a follow-up

measurement in 1973 (compartment 17A) or 1983 (all other compartments). Due to low sampling densities in the 1993 and 2003 inventories, this data could not be utilized to calculate reliable values for individual compartments. Thus, for the current study, an angle-count sampling inventory (Bitterlich 1952) was conducted in 2013 on a 50 × 50 m grid to gather data compatible with the historical surveys in 1963, 1973 and 1983. Thus, compartment polygons did not change throughout the observation period starting in 1963. From all inventories, stem numbers per hectare were available per tree species in 5 cm DBH (diameter at breast height)-classes, starting at 10 cm DBH. Only live trees were recorded in the inventories. For details about the nine study compartments see Table SM1.

Depending on the site index, two sets of height functions relating tree height to DBH for individual species were assigned to the compartments to calculate initial tree height in 1963 (Table SM2).

Starting in 1963, a historical register of annual harvests per compartment was available, documenting the harvests in 5 cm DBH-classes. Table SM1 presents selected information about the harvesting activities in the forest compartments.

2.3. Climate data

The forest models require daily or monthly climate data to drive the simulations (see Table SM3). For each of the nine forest compartments a daily time series of climate data covering the period 1963–2013 was prepared based on the nearest grid cell (Lat. 45.625, Long. 14.375) of the E-OBS data set (van den Besselaar et al. 2011). The MT-CLIM routines (Running et al., 1987; Thornton and Running 1999) were used to adjust the E-OBS climate record for elevation, slope and aspect of the nine sites and to estimate incoming global radiation of the daylight period and vapor pressure deficit (see Thornton et al., 2000).

2.4. Forest models

The five models were the gap model ForClim (Bugmann 1996), the landscape model LandClim (Schumacher et al., 2004), the hybrid 3D patch model PICUS (Hönninger and Lexer 2001; Seidl et al., 2005) and the spatially explicit empirical models SAMSARA2 (Courbaud et al., 2015) and SIBYLA (Fabrika 2005). The models are briefly introduced and their key features summarized in Table SM3. For detailed descriptions, we refer to the original sources.

2.4.1. ForClim

ForClim is a climate-sensitive forest succession (gap) model that has been developed to simulate forest dynamics over a wide range of environmental conditions (Bugmann 1994). The model simulates establishment, growth and mortality of individual trees on small independent patches, using a minimum of ecological assumptions to capture the influence of climate and ecological processes on forest dynamics (Bugmann 1996; Didion et al., 2009b). ForClim is structured into four sub-models: weather, water, plant, and management. The PLANT sub-model is the core of ForClim, where establishment and growth of tree cohorts (i.e., trees of the same species and age) are simulated based on light availability, soil nutrients, browsing intensity and bioclimatic indices calculated within the sub-models WEATHER and WATER. Tree mortality is modeled as a combination of constant "background" mortality and a stress-induced component. The MANAGEMENT sub-model enables the simulation of a wide range of silvicultural treatments such as clearcutting, shelterwood, thinning, planting, and others. In this study, we applied ForClim version 3.0 (Rasche et al., 2011), complemented by an empirical harvesting algorithm for simulating removals of an exact number of stems for every tree species by diameter class (single stem removal; see description in Mina et al. (2017)). ForClim is currently parameterized for 31 European tree species and has been tested for the representation of natural forest dynamics of temperate forests of the Northern Hemisphere (e.g., Didion et al., 2009a).

2.4.2. LandClim

LandClim is a process-based forest landscape model (Schumacher et al., 2004; Schumacher et al., 2006) designed to simulate forest dynamics and disturbances at large spatial scales (10^3 to 10^6 ha) over long periods of time (hundreds to thousands of years). In LandClim, landscapes are represented as a 25 × 25 m grid with specific topographic and climatic input variables for each cell. Within each cell, a simplified forest gap model (Bugmann 2001) simulates establishment, growth, competition and mortality of trees on an annual time step. Similar to ForClim, trees are simulated using a cohort approach (i.e., a computational simplification where one representative individual is simulated for all trees of the same species and age within a cell (Bugmann 1996)). Tree growth is simulated using a logistic growth equation, where species-specific maximum growth rate and size are reduced by light availability, degree-day sum and a drought index (Schumacher et al., 2004). Establishment and mortality are stochastic processes. Each year, the potential for tree establishment is determined as a function of environmental filters (i.e., available light at the forest floor, minimum winter temperature, growing degree-day sum, drought index, and browsing). Mortality probability is determined as a combination of stress, density-dependent and intrinsic mortality. LandClim can simulate management in 10-year intervals on defined management areas by selecting and removing a percentage of trees fulfilling specified DBH constraints. The model has been tested and adapted to the European Alps (Briner et al., 2013; Elkin et al., 2013; Temperli et al., 2013), North American Rocky Mountains (Schumacher et al., 2006; Schwörer et al., 2016), and Mediterranean forests (Henne et al., 2015). For this study, the individual compartments were simulated in LANDCLIM as independent entities without landscape level interactions among them to produce results comparable to the stand-level models.

2.4.3. PICUS

The forest model PICUS version 1.5 (Lexer and Hönninger 2001; Seidl et al., 2005; Irauschek et al., 2017a), henceforth referred to as PICUS, is a hybrid of classical gap model components and process-based stand-level NPP algorithms (Landsberg and Waring 1997). The spatial core structure of PICUS is an array of 10 × 10 m patches with vertical crown cells of 5 m in height. Interactions between patches are considered via a three-dimensional light model and spatially explicit seed dispersal. Stand-level NPP is estimated with a model of light use efficiency (Landsberg and Waring 1997), which depends on intercepted radiation, temperature, precipitation, vapor pressure deficit, soil water and nutrient supply. Distribution of assimilates to individual trees is based on the relative competitive success of the individual trees. Tree mortality depends on age and stress conditions. Natural tree regeneration considers seed production and distribution, germination and establishment. Up to the height of 130 cm seedlings are simulated with a height class approach. Beyond that threshold, they are considered as individuals in the tree population (Irauschek et al., 2017b). PICUS includes a flexible management module enabling the implementation of silvicultural treatments at tree level depending on tree attributes and patch location. The model includes 17 parameterized tree species and has been validated (Seidl et al., 2005; Didion et al., 2009a) and applied in numerous studies all over Europe (Lexer et al., 2002; Maroschek et al., 2015; Pardos et al., 2015; Zlatanov et al., 2017).

2.4.4. SAMSARA2

SAMSARA2 is an individual-based and spatially explicit model to simulate regeneration, growth and mortality of individual trees in mixed and uneven-aged mountain forest stands (Courbaud et al., 2015). The model builds on the theory that light interception by tree crowns is a key driver in uneven-aged stands, because they present a strong vertical heterogeneity favoring asymmetric competition between trees and between the canopy and seedlings. In SAMSARA2, competition for light within a stand is calculated based on light ray interception by tree crowns. SAMSARA2 has been calibrated empirically for silver fir and

Norway spruce stands within the montane elevation belt of the Alps in France. Ecological factors other than light, such as climate and site conditions are not directly taken into account. Fertility of the site is taken into account indirectly through the value of the different demographic parameters. According to Courbaud et al. (2015) the model should be recalibrated if applied under different site conditions. Annual diameter increment of individual trees depends on their size and the amount of light intercepted by their crown during the growing season. Natural mortality depends on tree diameter and a competition index defined as the basal area of larger trees within a radius of 15 m. Seeds are produced by adult trees and germination, growth and survival of seedlings depend on the light reaching the ground calculated in the center of 25 m^2 cells. Trees participate in light interception above DBH of 7.5 cm. Specific management algorithms allow the simulation of detailed silvicultural strategies, varying both the characteristics of harvested trees and their spatial arrangement within a stand (Lafond et al., 2012; Lafond et al., 2014).

2.4.5. SIBYLA

The SIBYLA model (Fabrika 2005) is based on the SILVA simulator (Pretzsch et al., 2002). It is an empirical, distance-dependent ecological niche-based model that simulates the growth of individual trees. The expected height increment is estimated from the potential height increment of the tree and a multiplier, which characterizes the effects of competition, soil and climatic conditions (see Pretzsch and Kahn 1998). SIBYLA was parameterized and validated using forest inventory data from Germany, Switzerland and Slovakia. The model is parameterized for five main European forest tree species – Norway spruce, silver fir, Scots pine, European beech and oak (*Quercus* sp.); other species can be simulated on the basis of their ecological and morphological similarity with the aforementioned species and using calibration functions. The model consists of sub-models for mortality, competition, growth, regeneration and thinning and a stand structure generator. Growth responses to environmental drivers (growing degree-days, annual temperature amplitude (°C), mean air temperature (°C) and precipitation in the growing season (mm), De Martonne (1925) index of aridity, soil moisture and site nutrient status) were formalized according to Kahn (1994). The model can simulate several cutting and thinning techniques typically applied in Central Europe (Fabrika and Ďurský 2005).

2.5. Simulation scenarios

2.5.1. Small tree initialization

For trees smaller than 10 cm DBH no information was available. However, most forest models consider trees of smaller sizes (in LandClim from DBH 5 cm, ForClim and SIBYLA from height 1.3 m and in PICUS from 10 cm seedling height). To allow a harmonized initialization of those models, two scenarios were defined for small tree initialization: (i) no small trees (A1, Table 1); (ii) small trees initialized using the threshold of 130 cm seedling height (A2). For the latter, the regeneration sub-model of SIBYLA (Fabrika 2005) was employed to estimate the initial number of trees in 1963 for diameter classes not covered by the historical inventory data (DBH 0–5 cm and 5–10 cm; see details of the

process in the Supplementary material) for use in all models.

2.5.2. Browsing by ungulates

Over the observation period (1963–2013), there was a considerable influence on regeneration by ungulate browsing (Klopcic et al., 2010). Because consistent browsing data was not available for each measurement period, browsing intensities per species over the entire simulation period were estimated based on detailed browsing inventories carried out between 1992 and 2000, expert knowledge and historical deer census data. To consider the uncertainties in browsing parameters, a total of three browsing scenarios were defined (B1, B2 and B3; see Table 1). Further details are given in Table SM4.

2.5.3. Harvesting

The historical register of harvests separately accounted for sanitary and regular logging. This means that harvests included, at least partly, dead or "near dead" trees (i.e. sanitary fellings). According to internal reports from the Sněžnik region, there is the tendency that in compartments with longer extraction distances less sanitary harvests are carried out due to economic reasons. This poses several problems for the specification of harvested trees in the simulations. To consider the potential effects of different modes of selecting trees for extraction, two harvesting scenarios were implemented: (i) only living trees harvested (C1); (ii) in a sanitary management scenario trees that had died in the simulations in the two years preceding a planned harvest were preferably selected for harvest (C2).

2.6. Analysis approach

We compared observed basal area in 1963 with the initial state of compartments in the model simulations (i.e., initial state), 1983 (except for compartment 17A, that was first remeasured in 1973; for simplicity we refer to the first remeasurement date in the results as "1983″ only) and 2013 with simulated basal area using the root-mean-square error (RMSE) and Nash-Sutcliffe Efficiency (NSE).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{9}(observed_i - simulated_i)^2}{9}} \qquad (1)$$

The root-mean-square error represents the square root of the quadratic mean of the differences between predicted (*simulated*) and observed basal areas for the 9 forest compartments (*i*) (Eq. (1)). The RMSE is always non-negative, where a value of 0 indicates a perfect fit to the data. RMSE can be related to the standard deviation of the measured values to provide a standardized index where values less than half the standard deviation may be considered low (Singh et al., 2004).

$$ME = \frac{\sum_{i=1}^{9}(observed_i - simulated_i)}{9} \qquad (2)$$

As an additional bias metric, we show the mean error (ME) (Eq. (2)).

Table 1

Simulation scenario overview. *N* = total number of factorial scenario combinations. Scenarios implemented by a model are marked by a tick.

Factor	Code	Scenario	ForClim N = 6	LandClim N = 6	PICUS N = 12	SAMSARA2 N = 12	SIBYLA N = 12
Initialization of small trees	A1	no small trees	✓	✓	✓	✓	✓
	A2	small trees initialized	✓	✓	✓	✓	✓
Browsing by ungulates	B1	no browsing	✓	✓	✓	✓	✓
	B2	current browsing	✓	✓	✓	✓	✓
	B3	1/3 browsing	✓	✓	✓	✓	✓
Harvesting	C1	regular logging (only living trees)	✗	✗	✓	✓	✓
	C2	sanitary and regular logging	✓	✓	✓	✓	✓

$$NSE = 1 - \frac{\sum_{i=1}^{9}\left(observed_i - simulated_i\right)^2}{\sum_{i=1}^{9}\left(observed_i - \overline{x}_{observed}\right)^2} \tag{3}$$

The Nash-Sutcliffe Efficiency is a normalized statistic that determines the relative magnitude of the residual variance compared to the variance in the measured data (Nash and Sutcliffe 1970) (Eq. (3)). NSE takes on values from -∞ to 1. Values close to 1 correspond to a perfect match of data and simulations. NSE of 0 indicates that the model predictions are as accur99ate as the mean of the observed data and values below 0 indicate that the observed mean is a better predictor than the model. The denominator in Eq. (3) determines to some extent the calculated NSE values. Because in the current study forest compartments were managed with an uneven-aged continuous cover regime, the variation of the observed basal area values was small and consequently to achieve high NSE values difficult. The R-package "hydroGOF" (Zambrano-Bigiarini 2017) was applied to calculate NSE.

$$DDE = \frac{1}{2}\sum_{j=1}^{14}\left|\frac{n\ observed_j}{N\ observed} - \frac{n\ simulated_j}{N\ simulated}\right| \tag{4}$$

To compare observed and simulated DBH distributions, we used the Diameter Distribution Error (DDE), described as "total variation distance index" in Levin et al. (2009) (Eq. (4)), where *j* are 14 diameter classes of 5 cm width (starting at 10 cm), n is the stem number per ha per diameter class, and N is the total sum of trees per ha (see also Saad et al., 2015). The sum of the differences of the relative frequency in the diameter classes is multiplied by ½ to scale the error between 0 and 100%, with optimum DDE at 0%. DDE was chosen because it is a distribution-free measure that calculates the diameter distribution error independent of the total number of trees. The diameter classes were the same as those in the inventory records.

For each compartment we calculated tree volume for the three measurement years, for all trees that had died during the simulations as well as for the harvested trees in a standardized approach by using local height curves (Table SM2) and individual tree stem volume functions (Pollanschütz 1974) to compare periodic volume growth and mortality, while avoiding possible biases due to biomass and volume estimation methods used by individual forest models only (Thurnher et al., 2013).

$$Total\ Growth_p = T\ Stock_{p(end)} - T\ Stock_{p(begin)} + \sum Harvest_p \\ + \sum Mortality_p \tag{5}$$

The total periodic production of timber (Total Growth) is calculated with (Eq. (5)). For a specific period (*p*), timber stock (*T Stock*) corresponds to the standing tree volume, *Harvest* is the standing tree volume of trees harvested in period (*p*) and *Mortality* is the standing tree volume of trees that had died during period (*p*).

2.7. Simulation protocol

The modeling groups jointly defined the scenario settings (cf. Table 1) and received the same data consisting of (i) tree data to initialize simulation stands representing the forest compartments in 1963, (ii) site and soil attributes, (iii) daily weather data for the period 1963–2013, and (iv) historical harvest data for the period 1963–2013. Each group performed the simulation runs (see Table 1) and delivered the model output data in a harmonized format to the lead author, who aggregated and analyzed the data. Tables SM1 and SM2 show site attributes available to the modeling groups and aggregated information on tree and harvesting data.

3. Results

3.1. Forest model initialization performance

Total basal area of trees with DBH above 10 cm in 1963, as initialized in PICUS, SAMSARA2 and SIBYLA, matched the observations very well (Fig. 1a). RMSE was 0.3 m² ha⁻¹ for PICUS, 0.2 m² ha⁻¹ for SAMSARA2, and 0.7 m² ha⁻¹ for SIBYLA (Table SM8). Initial basal area in LandClim showed the highest absolute deviations (RMSE 6.9 m² ha⁻¹), with lower basal area in all compartments compared with the inventory in 1963.

A comparison of initial DBH distributions in the models and the inventory in 1963 revealed a similar picture (Fig. 1b). The diameter distribution errors (DDE) for SAMSARA2 were below 1% and for PICUS and ForClim below 2%, with one department as exception for ForClim with DDE over 4%. DDE values for SIBYLA were between 1.5% and 3.5% indicating good agreement with the 1963 inventory. LandClim had moderate DDE values from 4 to 9%. Inspection of LandClim diameter distributions revealed that most of the mismatch in diameter structure occurred in the high DBH classes.

### 3.2. Simulation of compartment development

Starting with the initial compartment characteristics (see 3.1 above), the models simulated forest development according to the scenarios in Table 1. In Fig. 2 the simulated basal area development for each model is shown per compartment and compared to the inventory records. Most models slightly underestimated mean compartment basal area after 50 simulation years in 2013: ForClim at 27.0 m³ha⁻¹, SAMSARA2 26.3 m³ ha⁻¹, PICUS 26.1 m³ ha⁻¹ and SIBYLA at 25.5 m³ ha⁻¹ versus the mean inventory value of 28.4 m³ ha⁻¹. On average LandClim slightly overestimated basal area at the end of the simulation period in 2013 (29.3 m³ ha⁻¹). The mean error (ME) over all scenarios and compartments for the models was −3.0 m³ ha⁻¹ (ForClim), −3.2 m³ ha⁻¹ (PICUS), −1.6 m³ ha⁻¹ (SAMSARA2), −2.5 m³ ha⁻¹ (SIBYLA) and +3.2 m³ ha⁻¹ for LandClim. More details on simulated basal area are shown in Fig. 3.

Overall, SAMSARA2, SIBYLA and PICUS showed good model performance (NSE > 0.5; mean RMSE < 5 m² ha⁻¹) with PICUS being the only model performing well in both observation periods. LandClim had negative NSE values in both periods, meaning that the mean of the observations was a better predictor than the model. ForClim had the worst of all observed NSE values in 2013 (ranging from - 1.5 to −1.8), but NSE was highly influenced by poor performance in compartment 17A. Interestingly, this was the only compartment dominated by beech, while in the other compartments silver fir was dominating. Also other models (SIBYLA and SAMSARA2) showed the largest deviations from the observed values in compartment 17A for the observations in 2013.

It is interesting to look at the sensitivity of the models to the scenario assumptions. Overall, the variation in model output, as triggered by scenario A (initialization) and B (browsing), was smallest for LandClim (2% of the mean result) and largest for SAMSARA2 (8%); relative range in output for ForClim, PICUS and SYBILA was approx. 4 to 6% (Fig. 4). Adding scenarios C (harvesting mode) increased the variation to 24% (SAMSARA2), 11% (PICUS) and 6% (SYBILA) of the mean output (see also Figure SM1 and Table SM8).

Comparing simulated and observed diameter distributions at different time points is a rigorous test of the ability of the models to project forest structure over the 50-year period. Both in 1983 and 2013, LandClim showed the largest mean DDE value (23.6%, 69.8%) and ForClim the lowest (13.7%, 29.8%), with SYBILA, SAMSARA2 and PICUS having similar DDE values as ForClim. For all models DDE increased from 1983 to 2013 (Fig. 5).

What was the scenario that produced the best outcome with each of the five models? To answer this question the mean NSE and RMSE related to basal area over the entire simulation period were calculated for each simulation scenario and for each model the scenario with the maximum NSE was selected as "best" scenario" (Table SM6). In a similar
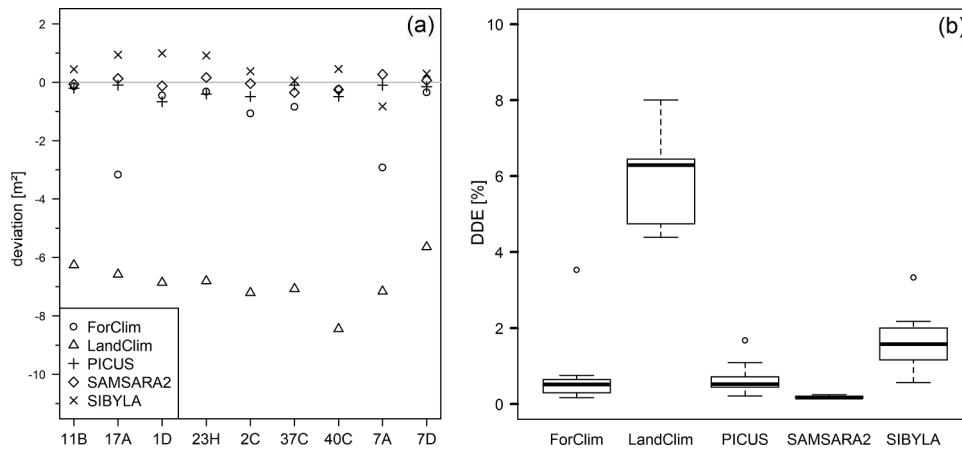
Fig. 1. Statistics for initialization of nine compartments (scenario A2; small trees initialized). Left (a): Deviation of observed and initialized basal area per hectare (trees > 10 cm DBH). Right (b): Diameter Distribution Error (DDE), maximum whiskers range is 1.5 times the interquartile range.
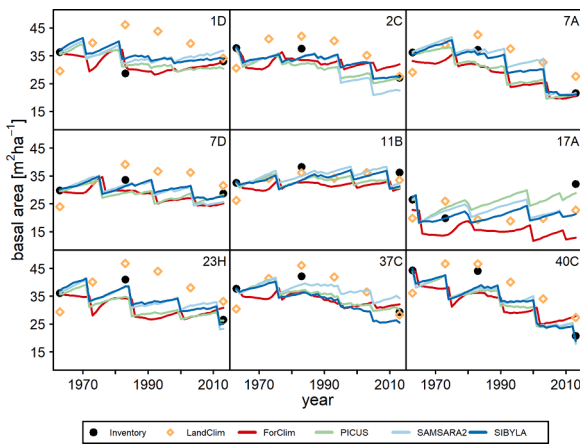
Fig. 2. Trajectories of simulated basal area development from 1963 to 2013 showing the scenario with the lowest error (NSE) in 2013 for each model. ForClim: A1 B1 C2, LandClim: A1 B1 C2, PICUS: A2 B3 C2, SAMSARA2: A2 B2 C1, SIBYLA: A1 B3 C1. Black dots denote inventory records.

procedure, mean DDE statistics were calculated for the entire simulation period and the scenario with the lowest DDE per model was selected (Table SM6). Regarding the simulated diameter distributions, scenario A2, which included small trees in the initialization, resulted in more

accurate simulated diameter distributions for all models. The scenario assumptions that produced the best overall match with regard to basal area differed among the five models. While for ForClim and LandClim "no browsing" (B1) combined with "sanitary management" (C2) produced better results, for PICUS, SAMSARA2 and SIBYLA the browsing pressure as observed in 2013 (B2) in combination with sanitary management (C2) resulted in more accurate results. For all models the two error statistics NSE and RMSE consistently ranked the same scenario at the top.

3.3. Realization of harvests

According to the harvest records, the compartments were treated regularly in an interval of 10 to 20 years. Moreover, in some compartments sanitation logging occurred almost every year (e.g., several operations in compartment 40C between 1985 and 1990 or in compartment 2C in the years 1980 to 1995) (see Fig. 2). When comparing the diameter structure of simulated harvests with the records, all models except LandClim performed well (not shown) with SYBILA showing the best match of simulated and observed diameter distribution of harvested trees. For all models, deviations occurred mainly in the smaller diameter classes below 20 cm DBH.

3.4. Total volume production

Utilizing the harvest records, the periodic harvest volumes *(Harvest)*
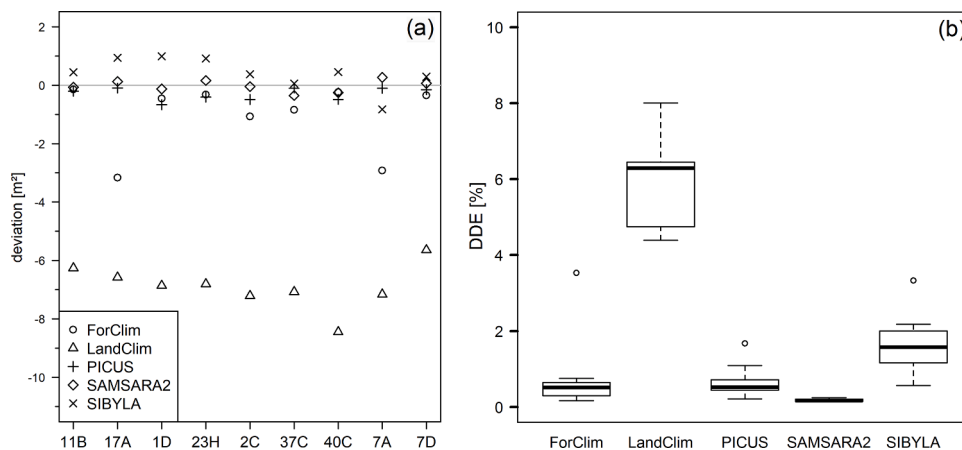
Fig. 3. Error indicators for basal area per model in 1983, 2013, and cumulated for both periods (1983 & 2013) over all compartments and scenarios. Left (a): Nash-Sutcliffe Efficiency (NSE). Right (b): Root Mean Squared Error (RMSE). *N* = 6 (scenarios).
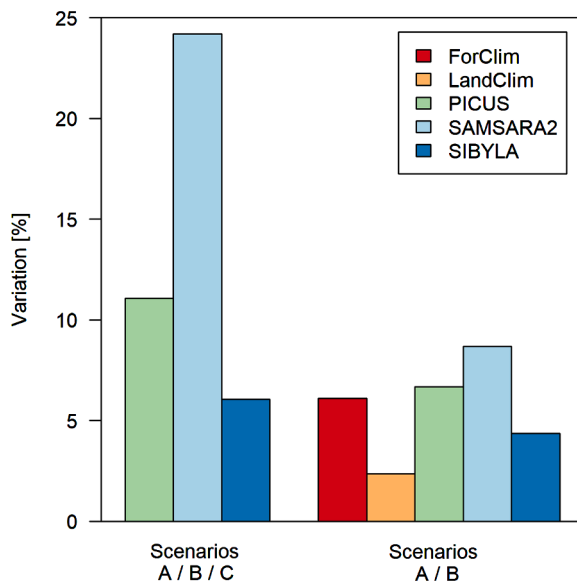
Fig. 4. Variation in basal area in the year 2013 as simulated in the scenarios from Table 1. Mean range of scenario results normalized by the mean scenario result for 2013. Scenario codes: *A* = Initialization of small trees, *B* = Browsing by ungulates, *C* = Harvesting.
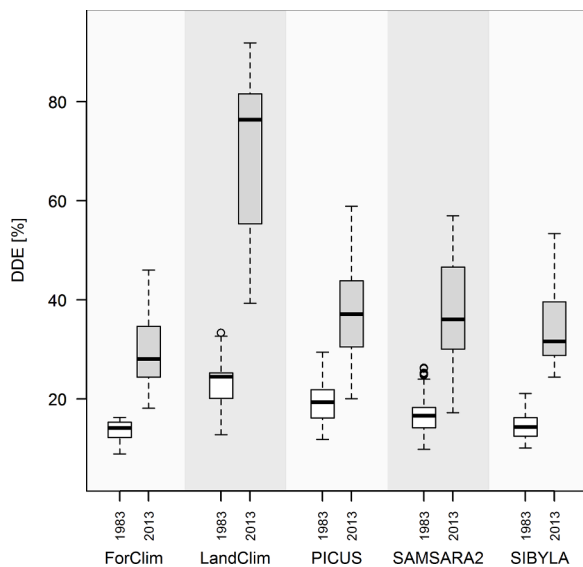

Fig. 5. Diameter distribution error (DDE) per model in 1983 and 2013 for all simulated scenarios and compartments.

can be calculated for the historical period, while tree *Mortality* which has not been extracted and thus is not included in *Harvest* is not known. Consequently, no observed *Total Growth* could be calculated for the nine study compartments. However, for comparison we could retrieve the local total mean annual volume growth from forest management plans (about 9 m³ ha⁻¹ yr⁻¹). This general estimate can be compared with *Total Growth* according to Eq. (4) for each model simulation. From Fig. 6 it is evident that there are differences among the models. Compared with the local estimate, ForClim, SAMSARA2 and SIBYLA simulated some-what lower *Total Growth* (5.6, 7.0, 7.1 m³ha⁻¹yr⁻¹), LandClim is clearly overestimating (15.7 m³ha⁻¹yr⁻¹), and PICUS is closest to the reference value of 9 m³ha⁻¹yr⁻¹ (8.4 m³ha⁻¹yr⁻¹). SAMSARA2 and SIBYLA reproduced the recorded harvest volumes accurately. PICUS and For-Clim show only minor underestimates, whereas LandClim clearly over-estimated *Harvests*. Large differences between the models are visible
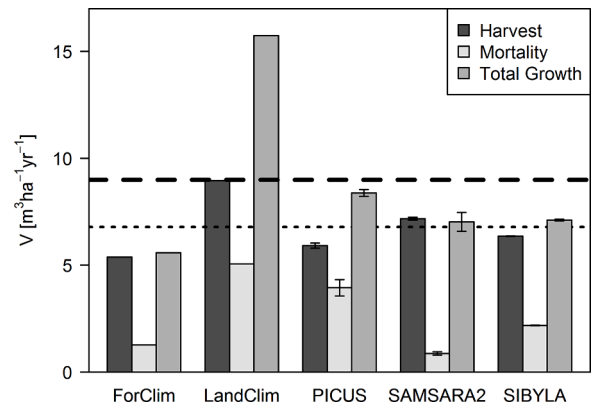

Fig. 6. Simulated yearly *Harvests, Mortality* and calculated *Total Growth* for the period 1963–2013 (mean over 9 compartments). *Dashed line* shows regional Total Growth estimate, *dotted line* shows historical harvests (mean over 9 compartments). For ForClim and LandClim bars show the results for scenario A2-B2-C2 (small tree initialized, current browsing, sanitary and regular log-ging). For PICUS, SAMSARA2 and SIBYLA bars show mean values for scenario A2-B2-C1 (regular logging) and A2-B2-C2 (sanitary and regular logging), the whiskers indicate results for C1 and C2.

regarding *Mortality*. Here, ForClim, SAMSARA2 and SIBYLA clearly simulated much lower tree mortality (ForClim 0.9 m³ha⁻¹yr⁻¹, SAMSARA2 1.3 m³ha⁻¹yr⁻¹ and SIBLYA 2.2 m³ha⁻¹yr⁻¹) than Land-Clim (5.1 m³ha⁻¹yr⁻¹) and PICUS (3.9 m³ha⁻¹yr⁻¹). The effect of the sanitary management scheme, causing a shift of volume from *Mortality* to *Harvest*, is small. Details for both inventory periods are shown in Table SM7.

4. Discussion

4.1. Assessment approach

One major challenge for increasing the acceptance of simulation tools for decision support is the objective validation of models, to build trust in the validity of model output (Muys et al., 2011). Multi-model evaluation exercises offer a great opportunity to compare detailed model outputs and point out differences, strengths and deficiencies. Recently studies have started to quantify parametric uncertainty (Hartig et al., 2012), and some results exist on the contributions of model sub-processes to overall parametric uncertainty (Augustynczik et al., 2017), but the quantification of structural uncertainty is less advanced.

A prerequisite for multi-model evaluation studies is that the participating model groups are provided with identical site, stand and climate information and follow a harmonized protocol that avoids spe-cific fitting of model settings to local conditions. In a comparative analysis of 15 forest models with regard to model sensitivity to different tree mortality algorithms, Bugmann et al. (2019) refrained from having all participating models run under identical conditions. However, this approach adds another source of uncertainty and makes it even harder to track those model sub-processes that contribute most to the variation in model output and to respond to the question of why a specific model does better or worse in a specific situation.

The historical long-term compartment-based dataset, in combination with management records as used in this study, offered great potential to test models in species-rich forest ecosystems in realistic management contexts. Inventories carried out by forest enterprises usually do not fulfill the criteria for long term model evaluation exercises. Here the focus usually lies on cost efficiency and less on continuous long-term datasets. Methodologies usually switched from full calipering of trees over wide forest areas to statistically more advanced and less cost intensive sampling methods starting from the 1950s onwards. As a result, continuous datasets consisting of tree data with sufficient

sampling densities over a long time period are very scarce. The tree and harvest data from the presented case study cover a remarkably long period of 50 years. They represent entire compartments and are thus better representatives of real forests compared to small plots of large-scale forest inventories (see Huber et al., 2013). However, to fully understand forest development in the nine compartments, information on tree mortality, stand structural information and tree regeneration would be essential. In an attempt to frame sources of uncertainty in the evaluation data, 12 scenarios were defined to specify explicit assumptions on the initial state of regeneration, browsing pressure by ungulates over the monitoring period and the inclusion of tree mortality in harvesting decisions (Table 1). What could be considered a weakness of the evaluation data was turned into a strength of the assessment approach because it added rigor to the comparison of model output and observations. It also facilitated analysis of the sensitivity of the models to assumptions that have to be made quite often, when data and model input from decades ago are used.

4.2. Model initialization

A first crucial step in practical model applications is to initialize stand structures from available information. To our knowledge, the ability of models to create realistic initial tree populations has not been considered before. All models involved in this study include special routines for initialization to avoid stand structures that are not compatible with internal processes dealing with competition for resources and space and may therefore result in undesired behavior of simulated tree populations, especially in the first simulation years (see Supplementary material). Generally, these routines are most important for complex forest stands with high tree density, vertical tree layers, diverse species mixtures, many size classes and patchy spatial structure. In the model SIBYLA, for example, the models internal consistency of the virtual tree population is evaluated by estimating the suitability of the tree positions according to the nearest and second nearest neighbor trees using a probabilistic approach (Pretzsch 1993). For rejected trees, new coordinates are generated until microstructural requirements are fulfilled. The PICUS model operates in a similar way within its 10×10 m resolution to avoid the initialization of non-viable tree neighborhoods. The gap model ForClim uses an algorithm to optimize the leaf area of virtual trees to avoid excessive shading. SAMSARA2 avoids initialization problems by using a uniform distribution for estimating coordinates, which may visually result in very regular stand structures. In LandClim, compartments are initialized spatially explicit at the level of the simulation cells (25 by 25 m). The effort that has been put into generating realistic initial forest states shows that, although so far not much attention has been paid to this issue, (i) creating realistic tree population structures may have crucial implications for simulated short- to midterm population behavior, and (ii) indicates the relevance of structural information and tree coordinates (see Table SM3). For instance, PICUS, SAMSARA2 and SIBYLA could have directly utilized tree coordinates for initializing forest stands. All models except LandClim were able to generate realistic forests regarding basal area density and diameter structure (compare Fig. 1 and Table SM8 for an overview). The landscape model LandClim, which had not been developed to simulate fine-grained tree populations at population level, initialized forests with substantially lower basal area compared to observations and also did not match well the initial diameter structure of the validation data set.

4.3. Forest simulation results

In simulating managed mixed mountain forests over five decades, beyond growth, regeneration and mortality of trees, also harvests and the response of tree populations to these human disturbances need to be mimicked realistically. The results showed that spatially explicit individual tree-based models performed better in simulating basal area development than the non-spatial gap model ForClim, and, not

surprisingly, the landscape model LandClim. Overall, the best performing models in our study yielded RMSE values for basal area over a simulation period of 50 years, which are about 10 to 12% of the compartment basal area. This simulation result is comparable to the range of basal area mean error values from large-scale forest inventories (e.g., Næsset 2007). The mean error (ME) values over all scenarios indicated that all models on average underestimated the observed basal area values by about 10% of the mean observed values, except the LandClim model which yielded a slightly larger positive bias.

An even greater challenge for forest models is to capture the development of the size structure of the tree population. With the exception of LandClim, all models performed reasonably well. Furthermore, the study did not confirm two general beliefs: First, and interestingly, empirical models (SYBILA, SAMSARA2) did not perform significantly better regarding the accurate prediction of DBH structure than the process-based models ForClim and PICUS (compare Guisan and Zimmermann 2000; Fontes et al., 2011). A similar finding for the PICUS model had already been reported in Seidl et al. (2005). Second, none of the models had been developed for Dinaric mountain forests. As the model evaluation protocol avoided the calibration of the tested models to local tree growth data, the expectations towards accurate model results in simulating 50 years of tree growth were low. However, the stand-level models performed reasonable (compare Fig. 7). The spatially explicit models PICUS, SAMSARA2 and SYBILA performed overall better than the non-spatial gap model ForClim and showed very concordant results. On the other hand, ForClim showed the best performance regarding the diameter distributions.

The current study clearly showed that a model should perform well in a series of aspects beyond tree growth that drive forest development. Over the 50-year simulation period, the interacting processes tree growth, regeneration, mortality and tree selection and harvesting determine the standing stock as well as the DBH distribution at a specific point in time. Hülsmann et al. (2018) point out that mortality subroutines in forest models are typically rather fundamental and lack empirical basis. In a recent study Bugmann et al. (2019) concluded that the sensitivity of forest models to changes in mortality algorithms is
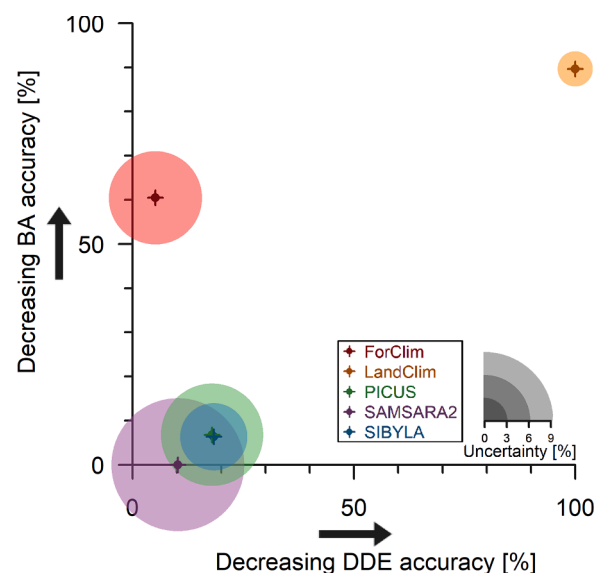

Fig. 7. Relative model performance scaled to the respective maximum and minimum values. Maximum BA accuracy and DDE accuracy in origin (0/0). BA accuracy = sum of relative RMSE for initialization and Nash-Sutcliffe Efficiency (NSE) for simulation results (pooled 1983 and 2013). DDE accuracy = sum of relative diameter distribution error (DDE) for initialization and simulation results (mean 1983 and 2013). The circle diameters correspond to the relative variation resulting from scenario assumptions (results 2013, scenarios A (Initialization of small trees) and B (Browsing by ungulates). Data in Table SM8.

larger than the sensitivity to climate change signals. Unfortunately, the testing data used for our study did not include tree mortality. The extent to which harvests included a share of natural tree mortality was also not known. However, the scenarios C1 (harvests include only living trees) and C2 (sanitary management, harvests also include dead trees) allowed to estimate the effect of the underlying assumptions. Our results showed that a seemingly simple process such as harvesting trees might produce quite different results among the models.

All models except LandClim could handle harvest prescriptions on an annual basis. However, harvest algorithms within the models differ in the way they handle tree removals in specific DBH-classes. Some algorithms search for specific trees (size and species) in the simulated forest. As differences between observed and simulated DBH-structures increase with increasing simulation length, more flexible algorithms (e.g., trees may be taken from neighboring DBH-classes) may be beneficial, especially for very detailed harvest operations such as single-tree or sanitary management records (Mina et al., 2017). Our results indicated that harvest algorithms incorporating multiple criteria, such as basal area, tree diameter and diameter distributions, can improve overall modeling accuracy (Lafond et al., 2012). The importance of consistently considering natural tree mortality in harvest algorithms will even increase with increasing relevance of disturbances in future climates (Seidl and Rammer 2017).

Framing the uncertainty in the evaluation data (initialization of small trees, browsing by ungulates, tree selection for harvesting) by means of 12 scenarios proved to be an adequate approach to (i) consider data uncertainty, and (ii) test the sensitivity of the models towards the scenario assumptions (see Figure SM1 and Tables SM5 and SM6). It was interesting to note that the browsing scenarios (B1, B2, B3) had a strong influence on the predictions, and were predominantly affecting diameter structures, whereas basal area was less affected. SAMSARA2 seemed to be particularly sensitive to the harvesting assumptions. What is a useful threshold for model sensitivity? This question cannot be answered easily. As a general rule, the sensitivity of a model to a driving gradient must allow to reproduce observable data along the gradient (e.g., Lexer and Hönninger 2004). Interestingly, the average variation, caused by the full set of scenarios (as simulated by SAMSARA2, PICUS and SIBYLA), is approximately in the same order of magnitude as the error in basal area or volume estimates of large-scale forest inventories.

LandClim, the only landscape model included in the comparison set, delivered satisfactory estimates of basal area development per compartment, though it operates on coarser resolutions regarding initialization, management and growth simulation, and was designed to simulate landscape level disturbances and processes that were not the focus of this study. The advantage of landscape models is that they can be used to complement stand-level models to explore forest development at multiple spatial scales from stand to landscape scale, which is gaining importance with growing confidence that disturbance regimes will intensify under a warming climate. Interacting disturbance agents operate at multiple spatial scales, which is beyond the reach of stand-level models (e.g., Elkin et al., 2013; Hlásny et al., 2019).

5. Conclusions

Several conclusions can be drawn from this study:

First, a strong effort should be made by research institutes and forest owners, to access and store historical inventory datasets and harmonize contemporary inventories with historical data to be able to capitalize on the benefits of a long consistent time series. In particular, detailed harvest records are crucial to understand time series data of stocks, which are the usual monitoring focus. Knowing harvested trees by tree species and by status (live, dead) would be a huge leap forward for further development and testing of forest ecosystem models.

Second, there is no clear hierarchy between more empirical or more process-based models regarding the accuracy of stand development projections.

Third, individual-based models are more precise for stand-level predictions than the tested landscape model LandClim, but the latter performed reasonably well at the stand scale, qualifying it for a complementary application at the landscape scale. This serves as a great example for a finding from a model intercomparison study, where the ultimate goal is not that every contestant shows excellent performance regarding standard outputs. More important is a critical comparison between models to differentiate why certain models are better under specific conditions observed in the case study for specific outputs. Following a detailed description of underlying model assumptions and objective evaluation of outputs, recommendations can then be given regarding the specific applicability of models and their limitations. However, model performance demonstrated in a specific ecological and management setting should not be generalized carelessly to diverging conditions. Repeated successful model evaluation experiments in different settings will build trust in model performance.

Fourth, in model evaluation studies a comprehensive set of model output variables should be assessed. There is a fast-growing demand by decision makers for reliable prediction of stand structural features beyond volume or basal area, because for forest management planning ecosystem service provisioning is based on indicators on species composition and structure of the tree population (e.g., Maroschek et al., 2015). The current study has shown that there are forest models available that qualify for forest management decision support.

Fifth, as a very relevant mission joint model intercomparison and evaluation studies compare currently available ecosystem models by challenging them with applications in novel case study settings and foster scientific collaboration among the modeling groups. This prepares the ground for multi-model applications in decision support, a prerequisite to quantify and explain uncertainty from model assumptions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2021.109493.

References

Augustynczik, A.L.D., Hartig, F., Minunno, F., Kahle, H.P., Diaconu, D., Hanewinkel, M., Yousefpour, R., 2017. Productivity of Fagus sylvatica under climate change – A Bayesian analysis of risk and uncertainty using the model 3-PG. For Ecol Manage 401, 192–206. https://doi.org/10.1016/j.foreco.2017.06.061.

Badeck, F.W., Lischke, H., Bugmann, H., Hickler, T., Hönninger, K., Lasch, P., Lexer, M. J., Mouillot, F., Schaber, J., Smith, B., 2001. Tree species composition in european pristine forests: comparison of stand data to model predictions. Clim Chang Chang 51, 307–347.

Bitterlich, W., 1952. Die Winkelzählprobe. Forstwissenschaftliches Cent 71, 215–225.

Boncina, A., 2011. History, current status and future prospects of uneven-aged forest management in the Dinaric region: an overview. Forestry 84, 467–478. https://doi.org/10.1093/forestry/cpr023.

Briner, S., Huber, R., Bebi, P., Elkin, C., Schmatz, D.R., Grêt-Regamey, A., 2013. Trade-offs between ecosystem services in a mountain region. Ecol Soc. https://doi.org/10.5751/ES-05576-180335.

Bugmann, H., 1994. On the Ecology of Mountainous Forests in a Changing climate: a Simulation Study. Eidg. Technische Hochschule, Zürich.

Bugmann, H., 1996. A simplified forest model to study species composition along climate gradients. Ecology 77, 2055–2074. https://doi.org/10.2307/2265700.

Bugmann, H., 2001. A review of forest gap models. Clim Change 51, 259–305. https://doi.org/10.1023/A:1012525626267.

Bugmann, H., Cordonnier, T., Truhetz, H., Lexer, M.J., 2017. Impacts of business-as-usual management on ecosystem services in European mountain ranges under climate change. Reg Environ Chang 17, 3–16. https://doi.org/10.1007/s10113-016-1074-4.

Bugmann, H., Seidl, R., Hartig, F., et al., 2019. Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale. Ecosphere 10, e02616. https://doi.org/10.1002/ecs2.2616.

Christensen, N.L., 1997. Managing for heterogeneity and complexity on dynamic landscapes. In: Pickett, STA, Ostfeld, RS, Shachak, M, Likens, GE (Eds.), The Ecological Basis of Conservation. Springer, Boston, MA, pp. 167–186.

Courbaud, B., Lafond, V., Lagarrigues, G., Vieilledent, G., Cordonnier, T., Jabot, F., de Coligny, F., 2015. Applying ecological model evaludation: lessons learned with the forest dynamics model Samsara2. Ecol Modell 314, 1–14. https://doi.org/10.1016/j.ecolmodel.2015.06.039.

De Martonne, E., 1925. Traité de géographie physique, Vol. I: notions generales, climat, hydrographie. Geogra Rev 15 (2), 336–337.

Didion, M., Kupferschmid, A.D., Lexer, M.J., Rammer, W., Seidl, R., Bugmann, H., 2009 a. Potentials and limitations of using large-scale forest inventory data for evaluating forest succession models. Ecol Modell. 220, 133–147. https://doi.org/10.1016/j.ecolmodel.2008.09.021.

Didion, M., Kupferschmid, A.D., Zingg, A., Fahse, L., Bugmann, H., 2009 b. Gaining local accuracy while not losing generality — Extending the range of gap model applications. Can J For Res 39, 1092–1107. https://doi.org/10.1139/X09-041.

Elkin, C., Gutiérrez, A.G., Leuzinger, S., Manusch, C., Temperli, C., Rasche, L., Bugmann, H., 2013. A 2 °C warmer world is not safe for ecosystem services in the European Alps. Glob Chang Biol 19, 1827–1840. https://doi.org/10.1111/gcb.12156.

Fabrika, M., 2005. Simulátor Biodynamiky Lesa SIBYLA, koncepcia, Konštrukcia a Programové Riešenie. Technical University in Zvolen.

Fabrika, M., Ďurský, J., 2005. Algorithms and software solution of thinning models for SIBYLA growth simulator. J For Sci 51, 431–445.

Fontes, L., Bontemps, J.-D., Bugmann, H., Van Oijen, M., Gracia, C., Kramer, K., Lindner, M., Rötzer, T., Skovsgaard, J.P., 2011. Models for supporting forest management in a changing environment. For Syst 3, 8. https://doi.org/10.5424/fs/201019S-9315.

Foster, A.C., Shuman, J.K., Shugart, H.H., Dwire, K.A., Fornwalt, P.J., Sibold, J., Negron, J., 2017. Validation and application of a forest gap model to the southern Rocky Mountains. Ecol. Modelling 351, 109–128.

Gadow, K v., 2000. Evaluating risk in forest planning models. Silva Fenn 34, 181–191.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S., Thulke, H., Weiner, J., Wiegand, T., DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310, 987–991. https://doi.org/10.1126/science.1116681.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol Modell 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9.

Hanson, P.J., Amthor, J.S., Wullschleger, S.D., et al., 2004. Oak forest carbon and water simulations: model intercomparisons against independent data. Ecol Monogr 74, 443–489. https://doi.org/10.1890/03-4049.

Härkönen, S., Mäkinen, A., Tokola, T., Rasinmäki, J., Kalliovirta, J., 2010. Evaluation of forest growth simulators with NFI permanent sample plot data from Finland. For Ecol Manage 209, 573–582. https://doi.org/10.1016/j.foreco.2009.11.015.

Hartig, F., Dyke, J., Hickler, T., Higgins, S.I., O'Hara, R.B., Scheiter, S., Huth, A., 2012. Connecting dynamic vegetation models to data - an inverse perspective. J Biogeogr 39, 2240–2252. https://doi.org/10.1111/j.1365-2699.2012.02745.x.

Henne, P.D., Elkin, C., Franke, J., Colombaroli, D., Caló, C., La Mantia, T., Pasta, S., Conedera, M., Dermody, O., Tinner, W., 2015. Reviving extinct Mediterranean forest communities may improve ecosystem potential in a warmer future. Front Ecol Environ 13, 356–362. https://doi.org/10.1890/150027.

Hlásny, T., Krokene, P., Liebhold, A.M., Montagné-Huck, C., Jörg, M., Qin, H., Raffa, K., Schelhaas, M.-J., Seidl, R., Svoboda, M., Viiri, H., 2019. Living With Bark beetles: impacts, Outlook and Management Options. From Science to Policy 8. European Forest Institute.

Huber, M.O., Eastaugh, C.S., Gschwantner, T., Hasenauer, H., Kindermann, G., Ledermann, T., Lexer, M.J., Rammer, W., Schörghuber, S., Sterba, H., 2013. Comparing simulations of three conceptually different forest models with National Forest Inventory data. Environ Model Softw 40, 88–97. https://doi.org/10.1016/j.envsoft.2012.08.003.

Hülsmann, L., Bugmann, H., Cailleret, M., Brang, P., 2018. How to kill a tree: empirical mortality models for 18 species and their performance in a dynamic forest model. Ecol Appl 28, 522–540. https://doi.org/10.1002/eap.1668.

Irauschek F., Rammer W., Lexer M.J. (2017a) Can current management maintain forest landscape multifunctionality in the Eastern Alps in Austria under climate change? Reg. Environ. Chang 17:33–48. doi: 10.1007/s10113-015-0908-9.

Irauschek, F., Rammer, W., Lexer, M.J., 2017b. Evaluating multifunctionality and adaptive capacity of mountain forest management alternatives under climate change in the Eastern Alps. Eur J For Res 136, 1051–1069. https://doi.org/10.1007/s10342-017-1051-6.

Jin W., He H.S., Thompson F.R. (2016) Are more complex physiological models of forest ecosystems better choices for plot and regional predictions? Environ. Model Softw.75:1–14. doi: 10.1016/j.envsoft.2015.10.004.

Kahn, M., 1994. Modellierung der Hohenentwiklung ausgewählter Baumarten in Abhängigkeit vom Standort. Forstliche Forschungsberichte München, Band 141.

Klopcic, M., Jerina, K., Boncina, A., 2010. Long-term changes of structure and tree species composition in Dinaric uneven-aged forests: are red deer an important factor? Eur J For Res 129, 277–288. https://doi.org/10.1007/s10342-009-0325-z.

Lafond, V., Cordonnier, T., De Coligny, F., Courbaud, B., 2012. Reconstructing harvesting diameter distribution from aggregate data. Ann For Sci 69, 235–243. https://doi.org/10.1007/s13595-011-0155-2.

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, 173–186. https://doi.org/10.1007/s13595-013-0291-y.

Landsberg, J.J., Waring, R.H., 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. For. Ecol. Manage. 95, 209–228.

Larocque, G.R., Bhatti, J.S., Ascough, J.C., Liu, J., Luckai, N., Mailly, D., Archambault, L., Gordon, A.M., 2011. An analytical framework to assist decision makers in the use of forest ecosystem model predictions. Environ Model Softw 26, 280–288. https://doi.org/10.1016/j.envsoft.2010.03.009.

Levin, D.A., Peres, Y., Wilmer, E.L., 2009. Introduction to Markov Chain Mixing. Markov Chains and Mixing Times. American Mathematical Society, Providence, USA, pp. 47–62.

Lexer, M.J., Hönninger, K., 2001. A modified 3D-patch model for spatially explicit simulation of vegetation composition in heterogeneous landscapes. For Ecol Manage 144, 43–65. https://doi.org/10.1016/S0378-1127(00)00386-8.

Lexer M.J., Hönninger K. (2004) Effects of error in model input : experiments with a forest patch model. 173:159–176. doi: 10.1016/j.ecolmodel.2003.08.018.

Lexer, M.J., Hönninger, K., Scheifinger, H., Matulla, C., Groll, N., Kromp-Kolb, H., Schadauer, K., Starlinger, F., Englisch, M., 2002. The sensitivity of Austrian forests to scenarios of climatic change: a large-scale risk assessment based on a modified gap model and forest inventory data. For Ecol Manage 162, 53–72. S.

Lindner, M., Lasch, P., Badeck, F.W., Beguiristain, P.P., Junge, S., Kellomäki, S., Peltola, H., Gracia, C., Sabate, S., Jäger, D., Lexer, M.J., Freeman, M., 2005. SilviStrat model evaluation exercise. In: Kellomäki, S, Leinonen, S (Eds.), Management of European Forests Under Changing Climatic Conditions. University of Joensuu, Faculty of Forestry, Research Notes 163, pp. 117–157.

Linkevičius, E., Borges, J.G., Doyle, M., Pülzl, H., Nordström, E., Vacik, H., Brukas, V., Biber, P., Teder, M., Kaimre, P., Synek, M., Garcia-Gonzalo, J., 2019. Linking forest policy issues and decision support tools in Europe. For Policy Econ 103, 4–16. https://doi.org/10.1016/j.forpol.2018.05.014.

Mäkelä, A., Sievänen, R., Lindner, M., Lasch, P., 2000. Application of volume growth and survival graphs in the evaluation of four process-based forest growth models. Tree Physiol 20, 347–355. https://doi.org/10.1093/treephys/20.5-6.347.

Maroschek, M., Rammer, W., Lexer, M.J., 2015. Using a novel assessment framework to evaluate protective functions and timber production in Austrian mountain forests under climate change. Reg Environ Chang 15, 1543–1555. https://doi.org/10.1007/s10113-014-0691-z.

McCullagh, A., Black, K., Nieuwenhuis, M., 2017. Evaluation of tree and stand-level growth models using national forest inventory data. Eur J For Res 136, 251–258. https://doi.org/10.1007/s10342-017-1025-8.

Mina, M., Bugmann, H., Klopcic, M., Cailleret, M., 2017. Accurate modeling of harvesting is key for projecting future forest dynamics: a case study in the Slovenian mountains. Reg Environ Chang 17, 49–64. https://doi.org/10.1007/s10113-015-0902-2.

Mlinsek D. (1968) Sproscena tehnika gojenja gozdov. Serbia Poslovno zdruzenje gozdnogospodarskih organizacij. Jugoslavenski Poljoprivredno Sumarski Centar pg, Belgrade.

Muys, B., Hynynen, J., Palahi, M., Lexer, M.J., Fabrika, M., Pretzsch, H., Gillet, F., Briceño, E., Nabuurs, G.-J., Kint, V., 2011. Simulation tools for decision support to adaptive forest management in. Europe. For Syst 3, 86. https://doi.org/10.5424/fs/201019S-9310.

Næsset, E., 2007. Airborne laser scanning as a method in operational forest inventory: status of accuracy assessments accomplished in Scandinavia. Scand J For Res 22, 433–442. https://doi.org/10.1080/02827580701672147.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. J Hydrol 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Pardos, M., Calama, R., Maroschek, M., Rammer, W., Lexer, M.J., 2015. A model-based analysis of climate change vulnerability of Pinus pinea stands under multiobjective management in the Northern Plateau of Spain. Ann For Sci 72, 1009–1021. https://doi.org/10.1007/s13595-015-0520-7.

Perko, F., 2002. Zapisano v branikah: Gozdovi in Gozdarstvo Od Sneznika do Nanosa skozi Cas. Gozdarsko drustvo Postojna, Postojna.

Pollanschütz, J., 1974. Formzahlfunktionen der Hauptbaumarten Österreichs. Allg Forstzeitung, Wien 85, 341–343.

Pretzsch H. (1993) Analyse und Reproduktion räumlicher Bestandesstrukturen. Versuche mit dem Strukturgenerator STRUGEN. In: Schriften Aus Der Forstlichen Fakultät Der Universität Göttingen und Der Niedersächsischen Forstlichen Versuchsanstalt, Bd. 114. Sauerländer, Frankfurt am Main, p 87.

Pretzsch, H., Biber, P., Ďurský, J., 2002. The single tree-based stand simulator SILVA: construction, application and evaluation. For Ecol Manage 162, 3–21. https://doi.org/10.1016/S0378-1127(02)00047-6.

Pretzsch H., Kahn M. (1998) Konzeption und Konstruktion des Wuchsmodells SILVA 2.2 - Methodische Grundlagen. In: Abschlußbericht Projekt W28, Teil2. München, p 277.

Puettmann, B.K.J., Coates, K.D., Messier, C., Press, I., Puettmann, K., 2009. A Critique of Silviculture : Managing for Complexity. Island Press, Washington.

Rasche, L., Fahse, L., Zingg, A., Bugmann, H., 2011. Getting a virtual forester fit for the challenge of climatic change. J Appl Ecol 48, 1174–1186. https://doi.org/10.1111/j.1365-2664.2011.02014.x.

Running, S.W., Nemani, R.R., Hungerford, R.D., 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. Can J For Res 17, 472–483.

Ryan, M.G., Hunt, E.R., McMurtrie, R.E., Agren, G.I., Aber, J.D., Friend, A.D., Rastetter, E.B., Pulliam, W.M., Raison, R.J., Linder, S., 1996. Comparing models of ecosystem function for temperate conifer forests. I. Model description and validation. Glob Chang Eff Conifer Grasslands 13, 313–362.

Saad, R., Wallerman, J., Lämås, T., 2015. Estimating stem diameter distributions from airborne laser scanning data and their effects on long term forest management planning. Scand J For Res 30, 186–196. https://doi.org/10.1080/02827581.2014.978888.

Schelhaas M.J., Nabuurs G.J., Hengeveld G.M., Reyer C., Hanewinkel M., Zimmermann N.E., Cullmann D. (2014) Adaptive Forest Management to Account For Climate Change-Induced Productivity and Species Suitability Changes in Europe. Reg Environ Chang submitted. doi: 10.1007/s10113-015-0788-z.

Schollmayer, H., 1906. Direktiven Für Die Bestandesaufnahme und Die Betriebseinrichtung auf Der F.C. Herrschaft Schneeberg. Kleinmayr-Bamberg, Ljubljana.

Schumacher, S., Bugmann, H., Mladenoff, D.J., 2004. Improving the formulation of tree growth and succession in a spatially explicit landscape model. Ecol Modell 180, 175–194. https://doi.org/10.1016/j.ecolmodel.2003.12.055.

Schumacher, S., Reineking, B., Sibold, J., Bugmann, H., 2006. Modeling the impact of climate and vegetation on fire regimes in mountain landscapes. Landsc Ecol 21, 539–554. https://doi.org/10.1007/s10980-005-2165-7.

Schwörer, C., Fisher, D.M., Gavin, D.G., Temperli, C., Bartlein, P.J., 2016. Modeling postglacial vegetation dynamics of temperate forests on the Olympic Peninsula (WA, USA) with special regard to snowpack. Clim Change 137, 379–394. https://doi.org/10.1007/s10584-016-1696-z.

Seidl, R., Fernandes, P.M., Fonseca, T.F., Gillet, F., Jönsson, A.M., Merganičová, K., Netherer, S., Arpaci, A., Bontemps, J.D., Bugmann, H., González-Olabarria, J.R., Lasch, P., Meredieu, C., Moreira, F., Schelhaas, M.J., Mohren, F., 2011. Modelling natural disturbances in forest ecosystems: a review. Ecol Modell 222, 903–924. https://doi.org/10.1016/j.ecolmodel.2010.09.040.

Seidl, R., Lexer, M.J., Jager, D., Honninger, K., 2005. Evaluating the accuracy and generality of a hybrid patch model. Tree Physiol 25, 939–951. https://doi.org/10.1093/treephys/25.7.939.

Seidl, R., Rammer, W., 2017. Climate change amplifies the interactions between wind and bark beetle disturbances in forest landscapes. Landsc Ecol 32, 1485–1498. https://doi.org/10.1007/s10980-016-0396-4.

Singh J., Knapp H.V., Demissie M. (2004) Hydrologic modeling of the Iroquois river Watershed using HSPF and SWAT. In: ISWS Contract Rep. CR-2004-08. http://hdl.handle.net/2142/94220. Accessed 15.5.2019.

Söderbergh, I, Ledermann, T., 2003. Algorithms for simulating thinning and harvesting in five european individual-tree growth simulators: a review. Comput Electron Agric 39, 115–140. https://doi.org/10.1016/S0168-1699(03)00022-X.

Temperli, C., Bugmann, H., Elkin, C., 2013. Cross-scale interactions among bark beetles, climate change, and wind disturbances: a landscape modeling approach. Ecol Monogr 83, 383–402. https://doi.org/10.1890/12-1503.1.

Thornton, P.E., Hasenauer, H., White, M., 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Australia. Agric For Meteorol 104, 255–271.

Thornton, P.E., Running, S.W., 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agric For Meteorol 93, 211–228. https://doi.org/10.1016/S0168-1923(98)00126-9.

Thurner, M., Beer, C., Ciais, P., Friend, A.D., Ito, A., Kleidon, A., Lomas, M.R., Quegan, S., Rademacher, T.T., Schaphoff, S., Tum, M., Wiltshire, A., Carvalhais, N., 2017. Evaluation of climate-related carbon turnover processes in global vegetation models for boreal and temperate forests. Glob Chang Biol 23, 3076–3091. https://doi.org/10.1111/gcb.13660.

Thurnher, C., Gerritzen, T., Maroschek, M., Lexer, M.J., Hasenauer, H., 2013. Analysing different carbon estimation methods for Austrian forests Eine Analyse unterschiedlicher Kohlenstoffschätzmethoden für Wälder in Österreich. Austrian J For Sci 3, 141–166.

van den Besselaar, E.J.M., Haylock, M.R., van der Schrier, G., Klein Tank, A.M.G., 2011. A European daily high-resolution observational gridded data set of sea level pressure. J Geophys Res 116, D11110. https://doi.org/10.1029/2010JD015468.

Yaussy, D.A., 2000. Comparison of an empirical forest growth and yield simulator and a forest gap simulator using actual 30-year growth from two even-aged forests in Kentucky. For Ecol Manage 126, 385–398. https://doi.org/10.1016/S0378-1127(99)00111-5.

Zambrano-Bigiarini M. (2017) hydroGOF: goodness-of-fit functions for comparison of simulated and observed hydrological time series. R package version 0.3-10. URL https://github.com/hzambran/hydroGOF.

Zlatanov, T., Elkin, C., Irauschek, F., Lexer, M.J., 2017. Impact of climate change on vulnerability of forests and ecosystem service supply in Western Rhodopes Mountains. Reg Environ Chang 17, 79–91. https://doi.org/10.1007/s10113-015-0869-z.