A framework for evaluating forest landscape model predictions using empirical data and knowledge

Wen J. Wang, Hong S. He, Martin A. Spetich, Stephen R. Shifley, Frank R. Thompson III, William D. Dijak, Qia Wang

School of Natural Resource, University of Missouri, 203 ABNR Bldg, Columbia, MO 65201, USA
Arkansas Forestry Sciences Laboratory, USDA Forest Service, Southern Research Station, P.O. Box 1270, Hot Spring, AR 71902, USA
USDA Forest Service, Northern Research Station, 202 ABNR Bldg, Columbia, MO 65201, USA

Article history:
Received 26 December 2013
Received in revised form 21 August 2014
Accepted 1 September 2014
Available online

Keywords:
LANDIS PRO
Validation
U.S. Forest Service Inventory and Analysis (FIA) data
Stand density management diagrams (SDMDs)
Oak forests
Prediction

Evaluation of forest landscape model (FLM) predictions is indispensable to establish the credibility of predictions. We present a framework that evaluates short- and long-term FLM predictions at site and landscape scales. Site-scale evaluation is conducted through comparing raster cell-level predictions with inventory plot data whereas landscape-scale evaluation is conducted through comparing predictions stratified by extraneous drivers with aggregated values in inventory plots. Long-term predictions are evaluated using empirical data and knowledge. We demonstrate the applicability of the framework using LANDIS PRO FLM. We showed how inventory data were used to initialize the landscape and calibrate model parameters. Evaluation of the short-term LANDIS PRO predictions based on multiple metrics showed good overall performance at site and landscape scales. The predicted long-term stand development patterns were consistent with the established theories of stand dynamics. The predicted long-term forest composition and successional trajectories conformed well to empirical old-growth studies in the region.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Forest landscape models (FLMs) predict forest change that results from the complex interactions of endogenous dynamics (e.g., growth, competition, mortality) and exogenous drivers (e.g., climate, anthropogenic forces) (Mladenoff, 2004; Perry and Enright, 2006; Lischke et al., 2006; He, 2008). They have increasingly become useful tools to explore the effects of management (Syphard et al., 2011; Wang et al., 2013a), disturbance (Schumacher and Bugmann, 2006; Sturtevant et al., 2009), and climate change (He et al., 2005; Keane et al., 2008; Thompson et al., 2011; Liang et al., 2014) on forest composition and structure at landscape scales. However, effective applications of FLMs to inform stakeholders and policy makers largely depend on the credibility of predictions, thus making the evaluation of FLM predictions indispensable (Rykiel, 1996; Gardner and Urban, 2003; Shifley et al., 2009; Alexandrov et al., 2011; Bennett et al., 2013). In part, the success in mitigating and adapting to changes in disturbance and climate is dependent on our capacity to predict the consequences of these changes (Coreau et al., 2009; McMahon et al., 2011; Dawson et al., 2011; Cheaib et al., 2012).

A framework for evaluating FLMs predictions is currently lacking. In fact, FLMs share common features that enable a framework for model evaluation. FLM predictions emerge from interacting processes at site (raster cell) and landscape scales (Lischke et al., 2006; He, 2008; Seidl et al., 2012). For individual-based FLMs such as iLand (Seidl et al., 2012) and to some extent ED (Mountfort et al., 2001), site-scale dynamics are simulated as establishment, growth, competition for light, and mortality for each tree. Landscape dynamics are simulated as outcomes of landscape processes (e.g., seed dispersal and disturbances) and exogenous drivers (e.g., terrain, soil, land use change, and climate).
Most FLMs predictions are evaluated by comparing with published results, predictions from other models, and empirical data at a given site or time (Gardner and Urban, 2003; Busing et al., 2007; Blanco et al., 2007; Shifley et al., 2009). For example, the simulated aboveground net primary productivity and biomass at the landscape scale in LANDIS II were compared with reported values in published literature (Schellner et al., 2007). The simulated tree growth and mortality at site scale in iLand were evaluated against the observe values in FIA data and experimental data of old-growth stands (Seidl et al., 2012). In TreeMig, the simulated spatial pattern of species biomass for a theoretical landscape were evaluated using empirical knowledge and the simulated tree species spread in the region of the Valais, Switzerland were compared with the current species composition (Lischke et al., 2006). The simulated biomass, soil carbon, and NPP at the landscape scale in ED were compared with various databases and other model predictions (Moorecroft et al., 2001).

In general, site-scale processes (e.g., individual tree growth) utilize potential resources from the bottom up, whereas the total resources determined using physiological principles regulates growth of individual trees. On the other hand, landscape processes and exogenous drivers capture the landscape heterogeneity (Seidl et al., 2012). These models are the best available tools used by foresters, managers, and planners to evaluate the long-term predictive stand development trajectories (Larsen et al., 2010). SDMDs are therefore excellent exploratory tools in evaluating relationships among tree growth, self-thinning (competition-caused mortality), and yield over long periods of time (Jack and Long, 1996).

Our overall objective is to demonstrate a comprehensive framework of evaluating FLM predictions. Our framework involves evaluating short- and long-term predictions at both site-landscape-scales. Short-term model predictions are compared with extensive inventory data and long-term model predictions are evaluated using empirical data and knowledge. We use the LANDIS PRO FLM as an example to illustrate the framework that: (1) use historic FIA data to constrain the initial forest conditions and calibrate model parameters before predicting future changes, (2) evaluate the short-term predictions (30 years) against FIA data at site and landscape scales, (3) evaluate the long-term predictions (150 years) of stand development patterns using SDMDs and of successional trajectories against old-growth forest studies. This framework is not only relevant for forest landscape models but also for biogeochemical or ecophysiological models such as LPJ-DGVM (Smith et al., 2011), biome-BGC (Bond-Lamberty et al., 2005) and PnET II (Aber et al., 1997), and ecosystem demography models (e.g., and CAIN (Caspersen et al., 2011), which include both site-scale dynamics and exogenous drivers operating at broad scales.

2. Methods

2.1. Study area

The study area encompassed the entire Ozark Mountains and Boston Mountains of Northern Arkansas covering about 1012 ha. The boundaries correspond to FIA Survey Unit 5 in Arkansas (Fig. 1). The topography in the study area is deeply dissected and rugged, with elevations ranging from 213 m to 762 m. Soils in this region are mostly Ultisols. Average annual temperature and precipitation ranged from 14 to 17 °C and from 1150 to 1325 mm, respectively, and most rainfall occurred in spring and fall. Most of this area was covered by deciduous forest dominated by white oak (Quercus alba L.), post oak (Quercus stellata Wangenh.), chinkapin oak (Quercus muehlenbergii Engelm.), black oak (Quercus velutina Lam.), northern red oak (Quercus rubra L.), blackjack oak (Quercus marilandica Michx.), southern red oak (Quercus falcate Michxm.), pignut hickory (Carya glabra Sweet), and black hickory (Carya texana Buckl.). Shortleaf pine (Pinus echinata Mill) was abundant in the southern part of the study area. Majority of forest stands in this region regenerated following the extensive timber harvest during early 1900s. Those cut-over forests regenerated naturally, and with the aid of more than a half-century of effective forest management the stem density greatly increased to reach full stocking (Heitzman, 2003). The dominant and codominant oaks typically ranged in age from 60 to 90 years.

2.2. The LANDIS PRO model

LANDIS PRO is a cellular automaton FLM that evolved over 15 years or research and development (Mladenoff and He, 1999; He and Mladenoff, 1999; He et al., 2002; Yang et al., 2011; Wang et al., 2013a,b, 2014). It simulates forest changes over large spatial (~106 ha) and temporal (~103 years) extents with flexible spatial (10–500 m) and temporal resolutions (1–10 years). Within each raster cell, tree species are recorded by number of trees by species age cohort; tree size (e.g., DBH) for a given species age cohort is derived from empirical age-DBH relationships (e.g., Loewenstein et al., 2000). LANDIS PRO simulates forest changes by incorporating species-, stand-, and landscape-scale processes (Wang et al., 2013b). Species-scale processes include tree growth, establishment, and mortality, which are simulated using species’ vital attributes (e.g., longevity, maximum DBH/SDI, and seedling establishment probability (SEP)) and species growth curves. Stand-scale processes simulate resource competition that regulates stand development patterns, seedling establishment, and self-thinning (Wang et al., 2013b). The intensity of competition among trees within each raster cell is quantified using tree density and size information to apply the Reineke stand density index (SDI) (Reineke, 1933) and compute the amount of growing space occupied (GSO) relative to the maximum growing space available (Maximum SDI) for each cell. This provides a metrics for the proportion of total growing space occupied as well as for the proportion currently utilized. Together GSO and tree size information for each cell govern progression through the stages of stand development described by (Oliver and Larson, 1996): stand initiation, stem exclusion, understory reiniitiation, and old-growth stages. Within LANDIS PRO, seedlings are established during the stand initiation stage of development depending on species shade tolerance and SEP that differ by ecological landtype and/or climate regime. When stands are modelled to exceed maximum growing capacity (MCSD), they enter the stem exclusion stage of development and the self-thinning process is modelled to mimic this period of intense competition (self-thinning) among trees (Oliver and Larson, 1996). Self-thinning is implemented using Yoda’s – 3/2 self-thinning theory (Yoda et al., 1963) as trees get larger the total number of trees declines. Within LANDIS PRO
this occurs as trees that are small, shade intolerant, or suppressed are predicted to
die due to completion from larger trees. LANDIS PRO does not use climate param-
eters directly as drivers of tree growth and survival. Instead, anticipated climate
effects are incorporated by altering the model parameters for SEP and MGSO. SEP
and MGSO can vary by landtype and change temporally. Values for those parameters
under a changing climate can be modeled outside LANDIS PRO platform using
ecosystem 160 process model (e.g. LINKAGES II, Wullschleger et al., 2003), which
uses climate and soil variables as drivers reflecting environment (or climate) change
resulting from nitrogen, CO2 fertilization, temperature and precipitation changes
(He et al., 2005). When those modified parameters are applied in LANDIS PRO they
model differences in species regeneration and maximum resource availability for
alternative climate regimes.

![Fig. 1.](image1.jpg)

Fig. 1. The 10^2 ha study area located in Northern Arkansas within FIA survey unit 5 (left panel). The area was dominated by oak forests (deciduous forest, right-top panel) and
topography was highly dissected with 5 landtypes (right-bottom panel).

![Fig. 2.](image2.jpg)

Fig. 2. Approach for initialization, calibration, and evaluation processes.
Landscape-scale processes modelled within LANDIS PRO include seed dispersal (including exotic species invasion), fire, wind, insect and disease spread, forest harvesting, and fuel treatments, each being an independent module (Wang et al., 2013a). LANDIS PRO simulates spatially explicit seed dispersal accounting for dispersal distance limitations and seed availability based on characteristics of trees at surrounding sites. Dispersal is simulated using a dispersal kernel determined by species’ maximum dispersal distances, where the seed dispersal probabilities follow a negative exponential decay function (He and Mladenoff, 1999). Seed availability for each species is accumulated from all available mature trees within the dispersal kernel. Number of potential germination seeds (NPGS) refers to the number of seeds with potential to germinate that are produced by one sexually mature tree per year. NPGS for each species is a user-defined parameter and is a state variable over the duration of the simulation. NPGS values which influence simulated tree density and basal area can be derived from (Burns and Honkala, 1990) and iteratively adjusted to ensure the predicted trees density, basal area, and species composition match observed forest inventory data. Exogenous disturbances (e.g., harvest or fire) were not simulated in this study, although it is possible to do so within LANDIS PRO. Further information of each disturbance module is reported elsewhere (e.g. Fraser et al., 2013a).

2.3. FIA data for forest landscape initialization and calibration

In our study, FIA were available only for inventory years 1978, 1988, 1993, 2003, and 2008 (Fig. 2). Our study area was largely comprised of national forests, where fires were effectively suppressed and timber harvest was limited (less than 5% of FIA plots experienced disturbances over the past 30 years). Thus, we selected the plots that did not experience disturbance to construct historical landscape, calibrate model parameters, and evaluate model predictions. Only FIA inventory plots meeting the following conditions were included in the samples: (1) classified as forest, and (2) no evidence of disturbance including logging, insects, disease and fire since the prior measurement. We grouped the eleven most common species in our study area into six functional species groups, which accounted for 90% of total basal area: white oak (white oak and post oak), red oak (northern red oak and southern red oak), black oak, hickory, pine (shortleaf pine and loblolly pine (P. taeda L)), and maple (red maple and sugar maple) (Table 1). We initialized the forest species composition map for the study area containing number of trees by species age cohort in each cell directly from 1978 FIA data using Landscape Builder software (Dijak, 2013), which was developed specifically for LANDIS PRO (Fig. 2). This program stochastically selected and assigned a representative FIA plot to each cell according to their frequency in forest type, forest size class, and landform. We compiled the species’ vital attributes (Table 1), landscape map, and SEPs by landtype from existing data sets for the study area (Wang et al., 2013a,b). Digital input maps were gridded to 90 m resolution.

We iteratively adjusted parameters of the Landscape Builder software to ensure the initial basal area and density for the modelled landscape matched the summarized FIA data for 1978 (Fig. 2) (Wang et al., 2013a). We then used the initial landscape for 1978 as the starting point and simulated forest succession without disturbance until 2008 (30 years) to calibrate the model parameter (NPGS) for each species. Because FIA data were available for only a 30-year time period (1978–2008), we used a data-splitting approach for model calibration and evaluation in this study (Thuiller, 2004). We used 50% of the FIA plots for 1988, 1993, 2003 and 2008 for model calibration (calibration subset) and reserved the other 50% of FIA plots for those respective years for short-term model evaluation (evaluation subset) (Fig. 2). Specifically, we iteratively adjusted NPGS for each species until the predicted density and basal area by species group at 1988, 1996, 2003 and 2008 closely matched the observed changes (no differences based on a chi-square test (p > 0.05)) in the calibration subset for same time period at landscape scales (Wang et al., 2013a). The adjustment process was analogous to sensitivity analysis because the adjustments were incremental.

We then applied the calibrated model to simulate forest changes from 1978 to 2128 (150 years) without including any exogenous disturbances. We evaluated the short-term predictions of basal area and density by species group at 1988, 1993, 2003 and 2008 against the observed values from the FIA evaluation subset at site and landscape scales (Fig. 2). We also evaluated the long-term (150 year) LANDIS PRO predictions of forest composition, structure, stand development patterns, and successional trajectories using SDMDS and data from studies of old-growth oak forests (Fig. 2) (Richards et al., 1995; Shifley et al., 1995).

2.4. Evaluation of the short-term model predictions

2.4.1. Sampling design at site and landscape scales

The short-term model predictions at site scales were evaluated for two forest types: oak-hickory and loblolly-shortleaf pine. We first classified the LANDIS PRO predictions (raster cells) into two forest types and then randomly selected 5000 raster cells from each forest type for each evaluation year. We computed the basal area and density by species group for each subsample of sites from modelled landscape and compared the results with the FIA data for 1988, 1993, 2003, and 2008. Likewise, the FIA plots were stratified into two forest types. To ensure observed values from FIA evaluation subset were comparable to LANDIS PRO predictions, the density and basal area by species group for each FIA plot were extrapolated to a 90 m cell size using the FIA tree area expansion factors.

The short-term landscape-scale evaluations were conducted by stratifying predictions (raster cells) into landtypes, because landtypes were used to reflect exogenous forces in LANDIS PRO. Resources availability and species assemblages were assumed to be similar within a landtype and vary among landtypes. Five landtypes were included: southwest landtype, northeast landtype, ridgetop, upland drainage, and bottomland (Fig. 1). We only evaluated the model predictions at the southwest and northeast landtypes, because they were most abundant and comprised about 70 percentage of the total area. We aggregated the predicted density and basal area of individual cells into landtype polygons. In the FIA evaluation subset, we stratified FIA plots into southwest and northeast landtypes, and then used FIA area expansion factors to scale the FIA plot estimates of density and basal area to the area of the landtype polygons. We then randomly selected 5000 southwest and northeast landtype polygons from both simulated and FIA evaluation subset for each evaluation year to conduct significant tests.

2.4.2. Evaluating the short-term model predictions using forest inventory data

Due to the stochastic components included in the models (Mladenoff and He, 1999), FLMs are not designed to predict the occurrence of a given event or structure at a specific location. Thus only aggregated statistical properties can be estimated meaningfully across broad spatial and temporal scales (Levin et al., 1997). In this study, we used the mean of samples for the statistical comparisons.

Goodness-of-fit measurements were used to quantify the accuracy of LANDIS PRO short-term predictions: the relative mean error (RMSE) (Eq. (1)), the relative absolute error (MAE%) (Eq. (2)), and the Nash-Sutcliffe index of model efficiency (ME) (Eq. (4)) (Walther and Moore, 2005; Miehle et al., 2006; Bennett et al., 2013). RMSE estimated the mean bias and the accuracy of model predictions, whereas MAE% and RMSE% measured the prediction accuracy using absolute prediction errors on an individual level. Since RMSE% was based on squared prediction errors, it was more sensitive to outliers than MAE% that was a linear function of the errors. The greater the difference between MAE% and RMSE% was, the greater was the likelihood of significant prediction errors (Walther and Moore, 2005). The ME index examined the agreement of individual predicted and observed values; the closer the computed value of ME to +1, the better was the predicted accuracy (Miehle et al., 2006).

Table 1

<table>
<thead>
<tr>
<th>Species group</th>
<th>Longevity (years)</th>
<th>Mean maturity (years)</th>
<th>Shade tolerance (class)</th>
<th>Fire tolerance (Class)</th>
<th>Maximum seeding distance (m)</th>
<th>Vegetative reproduction probability</th>
<th>Minimum sprouting age (years)</th>
<th>Maximum sprouting age (years)</th>
<th>Maximum DBH (cm)</th>
<th>Maximum SDI (trees/ha)</th>
<th>Number of potential germination seeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine</td>
<td>200</td>
<td>20</td>
<td>3</td>
<td>4</td>
<td>200</td>
<td>0.5</td>
<td>1</td>
<td>47</td>
<td>60</td>
<td>990</td>
<td>50</td>
</tr>
<tr>
<td>Black oak</td>
<td>120</td>
<td>30</td>
<td>3</td>
<td>3</td>
<td>200</td>
<td>0.4</td>
<td>1</td>
<td>10</td>
<td>70</td>
<td>570</td>
<td>90</td>
</tr>
<tr>
<td>Red oak</td>
<td>150</td>
<td>20</td>
<td>3</td>
<td>3</td>
<td>200</td>
<td>0.4</td>
<td>1</td>
<td>10</td>
<td>70</td>
<td>570</td>
<td>90</td>
</tr>
<tr>
<td>White oak</td>
<td>200</td>
<td>20</td>
<td>3</td>
<td>3</td>
<td>200</td>
<td>0.4</td>
<td>1</td>
<td>10</td>
<td>70</td>
<td>570</td>
<td>90</td>
</tr>
<tr>
<td>Hickory</td>
<td>250</td>
<td>20</td>
<td>3</td>
<td>3</td>
<td>200</td>
<td>0.5</td>
<td>1</td>
<td>10</td>
<td>70</td>
<td>570</td>
<td>90</td>
</tr>
<tr>
<td>Maple</td>
<td>200</td>
<td>20</td>
<td>5</td>
<td>1</td>
<td>200</td>
<td>0.3</td>
<td>1</td>
<td>10</td>
<td>70</td>
<td>570</td>
<td>90</td>
</tr>
</tbody>
</table>

Species life history parameters used in the forest landscape model LANDIS PRO in Northern Arkansas.
Prior to calibrating model parameters (NPGS), the predicted density and basal area from 1978 to 2008 were significantly different from observed values from the same period of FIA data. Following the calibration of the NPGS, there were no significant differences in species density and basal area at landscape scales between LANDIS PRO predictions and observed FIA estimates for 1988, 1992, 2003 and 2008 (Fig. 4a–d). For example, the Chi-Square tests results for species density at 2008 were $\chi^2 = 1.04$, $df = 5$, $P = 0.96$ at southwest landscape, and $\chi^2 = 2.68$, $df = 5$, $P = 0.75$ at northeast landscape; The Chi-Square results for basal area at 2008 were $\chi^2 = 3.70$, $df = 5$, $P = 0.59$ at southwest landscape, and $\chi^2 = 1.85$, $df = 5$, $P = 0.87$ at northeast landscape. Thus, the calibrated model parameters predicted reasonable outcomes.

3.2. Evaluation of the short-term model predictions

Differences between predicted and observed density and basal area at site scales in 1988, 1993, 2003, and 2008 were small for the 50% of data reserved for model evaluation (Fig. 5a,b). The minor difference between MAE% and RMSE% indicated there were no extreme prediction errors at 1988, 1993, 2003, and 2008. ME values close to 1 indicated a reasonable level of predicted accuracy. Specifically, there was smaller bias and better accuracy for loblolly-shortleaf pine sites (Fig. 5a) than for oak-hickory sites (Fig. 5b).

There were small differences between the predicted and the observed values at landscape scales that were within 10% of $\%$, MAE%, and RMSE% (Fig. 5c,d). The small differences between MAE% and RMSE% indicated there were no extreme prediction outliers. Values for ME were close to 1.0. Specifically, the predicted accuracy on the northeast landscape (Fig. 5c) was higher than that on the southwest landscape (Fig. 5d). Furthermore, there was also smaller bias and better accuracy of predicted species density than basal area (Fig. 5a–d), and better predicted accuracy at landscape scales (Fig. 5cd) than at site scales (Fig. 5a,b). The comparisons of predictions at 1988, 1993, 2003, and 2008 demonstrated that the bias increased and predicted accuracy decreased over time from 1993 to 2008 (Fig. 5a–d). Overall, the short-term predictions (1978–2008) showed a reasonable level of performance with accuracy better at landscape than the site scales, and there were greater discrepancies in predicted basal area than density.

3.3. Evaluation of the long-term model predictions

3.3.1. Evaluating the predicted stand development patterns at site scales

The predicted stand development patterns from 1978 to 2128 plotted on Gingrich stocking charts and Reineke density diagrams illustrated changes over time for three typical groups that encompassed a wide range of initial stand conditions (Fig. 6). Group I represented the development of stands initialized at the stand initiation stage (Fig. 6). The initial stands in group I were typically characterized by relatively fewer trees, lower basal area, and lower stocking percent. As succession proceeded, more seedlings became established resulting in an increase of tree density, a decrease of the mean diameter, and a slight increase in basal area. When those stands reached the stem exclusion stage, self-thinning resulted in a rapid decrease of trees density. The remaining live trees increased in diameter and thus basal area increased over time. Group II represented the development of stands initialized at the stem exclusion stage (Fig. 6). They had more trees and higher basal area than group I. The modelled self-thinning process decreased tree density while the basal area increased slightly. Continued tree growth resulted in a rapid increase in mean diameter and basal area later in the prediction period. Group III represented the development of stands initialized at the late-stand initiation stage with high
stocking but a small mean diameter (Fig. 6). These stands started at or above the maximum stocking line on Gingrich stocking charts and Reineke density diagrams, so the predicted mortality rates were high due to intense self-thinning within the stand. After growing space was released and stocking decreased below the maximum, the subsequent tree growth led to increases in basal area and mean diameter. The simulated rates of mortality were not significantly different compared to the theoretical models of self-thinning mortality by Reineke (1933) and Yoda et al. (1963) (Fig. 6b). Therefore, the comparisons between the SDMDs and the established theories of forest stand development suggested that LANDIS PRO predicted reasonable patterns of stand development for stands representing a wide range of initial conditions.

3.3.2. Evaluating the predicted forest composition, structure, and successional trajectories at landscape scales

The predicted density of white oak, red oak, black oak, hickory, and pine species groups decreased over the 150 simulation years as a result of self-thinning and forest maturation (Fig. 7a,b). However, the predicted density for maple species group gradually increased over the 150 years, because the lack of simulated disturbance favored the establishment of shade tolerant species. The predicted total basal area, biomass, and carbon increased from 1978 to a peak at 2098, followed subsequently by slight declines from 1998 to 2128 (Fig. 7c–h). The declines after 2098 resulted from the predicted mortality of a large proportion of trees in the red oak and black oak species groups that reached maximum longevity, died, and were replaced by young trees. The predicted basal area reached a maximum of 23 m²/ha on the southwest landtypes and 28 m²/ha on the northeast landtypes (Fig. 7c,d). These values were consistent with the basal area estimates of 23.5–28 m²/ha reported by Shifley et al. (1995) and Richards et al. (1995) for mature, undisturbed upland oak forests in the Ozark Highlands.

Our model predictions also indicated that without disturbances the white oak species group would continually dominate the landscape from 1978 to 2128 (Fig. 7c–f). Trees in the red oak species...
and black oak group declined in basal area after 2098, because many trees of those species were established in the early to mid-1900’s and were predicted to experience increased mortality as they approached their maximum longevity (Fig. 7c,d). The basal area for maple species was predicted to gradually increase from 1978 to 2128. These predicted successional trajectories were consistent with previous studies in oak-dominated forest and empirical knowledge. In the absence of disturbance, oak-dominated forests in this region typically transition to a greater proportion of longer-lived white oak species and shade-tolerant species such as maple (Johnson et al., 2009).

4. Discussion

4.1. Result implications

Our study demonstrates a process for extensively evaluating FLM predictions at site and landscape scales using forest inventory data, SDMDs, and empirical studies. Evaluation results for the calibrated model demonstrated reasonable performance, and were encouraging for subsequent applications of the model. Overall, the prediction accuracy at landscape scales was higher than that at site scales. This was consistent with previous studies because the variance in model prediction decreased as the predicted attributes were aggregated into a higher spatial hierarchy (Guisan et al., 2007). Prediction accuracy was greater for northeast landtypes than for southwest landtypes, which may be because there was greater variation in FIA data for the lower-quality sites found on southwest landtypes (Gordon et al., 2004). The prediction accuracy decreased as the simulations continued over time because uncertainties (e.g., parameter uncertainty and model stochasticity) accumulated through time (Xu et al., 2009). Our results also showed that the prediction accuracy for species density was higher than basal area. This was because predicted density was largely determined by a single parameter (NPGS) in the model. Besides the NPGS, the predicted basal area was additionally affected by species

![Fig. 5. Measures of goodness-of-fit for evaluating LANDIS PRO short-term predictions (including species density and basal area) against reserved 50% FIA data (validation subset) at 1988, 1993, 2003 and 2008 at site (a,b) and landscape scales (c,d) in Northern Arkansas.](image)

![Fig. 6. The representative predicted stand development patterns from 1978 to 2128 plotted on Gingrich (1967) stocking chart (a) and Reineke (1933) density diagrams (b). They fell into three general groups from LANDIS PRO predictions for a landscape in Northern Arkansas.](image)
diameter growth curves that introduced additional uncertainties into basal area estimates.

Our results showed that the predicted patterns of forest succession conformed well to theories and empirical knowledge of old-growth forest conditions in this region. Likewise, the predicted long-term patterns of stand development when measured jointly by changes in mean size, tree density, and basal area and examined in the SDMD framework were consistent with both theoretical and empirical knowledge of forest stand development. We showed that FLM predictions can be directly linked with SDMDs that were commonly used by forest managers and planners. This helps establish the credibility and utility of model predictions for informing management and policy decisions. This type and detail of model evaluation represents a significant advance in forest landscape modeling.

We used the differences between predicted results and observed values from FIA data to quantify the prediction accuracy of LANDIS PRO at site and landscape scales. Quantifying such differences is essential for effective applications of FLMs for scenario analyses. The impracticality of conducting real landscape-scale forest ecosystem experiments has resulted in increasing use of FLMs for scenario modeling to analyze the effects of different management actions on forest landscapes (Mladenoff and He, 1999). Model scenarios are generally created by altering input parameters to reflect changes in climate, disturbance, and/or management while the other calibrated model relationships remain unchanged (He, 2008; Coreau et al., 2009; Schmolke et al., 2010). Thus, quantifying the differences between simulation results and the real world data provides a basis to separate whether a response is due to the different simulated scenarios or inherent uncertainty in the model. Only if we quantify and understand the uncertainties in the initial conditions, model internal algorithms, and stochastic modeling components can we legitimately analyze the effects of different model scenarios.

4.2. Approach implications

We proposed a framework for evaluating FLM predictions, which involved evaluating short- and long-term predictions at both site and landscape scales. Evaluating site-scale predictions is conducted through comparing predicted results within raster cells with observed values in inventory plots randomly sampled across
the landscape. Evaluating landscape-scale predictions is conducted through comparing predicted results stratified by extraneous drivers (e.g., weather, soil, or terrain) with observed values in inventory plots aggregated by extraneous drivers. The short-term FLM predictions are evaluated using forest inventory data whereas the long-term FLM predictions are evaluated using stand density management diagrams (SDMDs) and empirical studies. We have demonstrated the applicability of this framework by using LANDIS PRO. However, such framework is also applicable to other large-scale models such as landscape and regional models. The response variables for evaluating LANDIS PRO predictions are basal area and density by species group. These variables can be different for different models. For example, Seidl et al. (2012) evaluated short-term, site-level predictions of iLand by comparing site index at age 100 (growth) and mortality from the simulated and FIA data. Landscape evaluation was conducted by replicating the site-level evaluation method over the elevation transects in the Eastern Alps. Long-term, site-scale model predictions were evaluated by comparing the simulated and observed mortality rate of old-growth forests. Lischke et al. (2006) evaluated the long-term predictions of TreeMig at a small and a very large landscape scale, by comparing the simulated species distribution patterns with the empirical understanding of these patterns. They showed that the model was capable of producing different endogenously driven pattern as a result of seed production, dispersal, and regeneration, as well as species competition for resources and environmental change.

Spatial and temporal autocorrelations in FIA data may present some limitations of using FIA data since the short-term model predictions were initialized and evaluated using the same set of FIA plots that were re-measured over time. The data splitting method for initializing and evaluating model predictions increased the average distances between plots and consequently reduced the spatial autocorrelation. However, since the time span for FIA data is only 30 years apart, temporal autocorrelation may still be high, which may lead to optimistic estimates of predictive ability. While independent landscape-scale data sets with a longer time series of repeated measurements would have been desirable, these FIA data provided a rare record of observed changes over several decades and allowed direct comparisons between observed and simulated results for a real forest landscape. This type of data-intensive evaluation has not been previously attempted at this level of detail in forest landscape modeling. It helps improving the realism of model assumptions, algorithms, and parameters (Araújo et al., 2005).

Long-term predictions of forest change require accounting for climate change effects. We did not include climate change in this study, because the objective of this study was to demonstrate a framework of evaluating FLM predictions. Our premise is that the model must be able to produce acceptable results under current climate, before it will be plausible for projections under a changed climate. In addition, the field data from the old-growth studies used to evaluate the long-term predictions corresponded to the past and current climate; comparable evaluation datasets under a changing climate do not exist. Thus, we only simulated the forest growth and succession under the current climate. However, we acknowledge that climate change is an important factor, especially for long-term predictions.

Our study responds to an unprecedented demand in the current data-rich era to combine inventory or observational data from the long-term accumulation with ecological models to improve predictions of future change towards a predictive science (Clark and Gelfand, 2006; Moorcroft, 2006; Peng et al., 2011). Advanced ecological forecasting is critical for informing natural resource policy and management decisions concerning ecosystem management and climate change (Coreau et al., 2009; Schmolke et al., 2010; Cheaib et al., 2012). In our study, FIA data were integrated with FLMS to constrain the initial landscape and model parameters, and ultimately to improve model predictions. Establishing appropriated initial conditions is critical, because they can greatly affect the subsequent forest dynamics (Luo et al., 2011). The accurate representations of the initial landscape and model parameters that approximate reality as close as possible improve model predictions (Peng et al., 2011).

Finally, in this study we evaluated predicted results only for forest succession in the absence of disturbances. Quantitative evaluation of cumulative effects for landscape with exogenous disturbances is more difficult because most FLMS employ stochastic methods to simulate disturbances. Thus far, the effects of disturbance have been widely explored at stand scales using observational data (e.g., Johnstone et al., 2010; Fraser et al., 2013; Luo et al., 2014). However, few studies have actually used a data infusion approach to validate predicted disturbance effects or the interaction of disturbance and succession at landscape scales. New approaches are yet to be developed on this front.

Acknowledgments

This research was funded by the U.S Forest Service Southern Research Station and Northern Research Station, and the University of Missouri GIS Mission Enhancement Program.

References


