Environ Monit Assess (2024) 196:1034
https://doi.org/10.1007/s10661-024-13189-7

RESEARCH

Dynamic site index models sensitive to ecoregional variability for Scots pine stands in Western Black Sea Region of Türkiye

Fadime Sağlam · Oytun Emre Sakici

Received: 5 April 2024 / Accepted: 24 September 2024 / Published online: 8 October 2024

Abstract Site productivity, defined as the production amount of the stand at a specific age, has a significant impact on the growth of the stand and site index is used as an indicator of site productivity. The objective of this study is to develop ecoregion-based dynamic site index models for Scots pine (*Pinus sylvestris* L.) stands in the Kastamonu and Sinop regions of Türkiye. The mixed-effects modeling approach allowing for the inclusion of ecoregions in the models was used to develop dynamic site index models, and the models derived from seven base models were tested. The best model was selected based on statistical criteria. As a result of statistical analyses and graphical examinations, the King-Prodan model was found to yield the best predictive results in terms of growth patterns. The site index model based on the King-Prodan method produced a coefficient of determination (R^2) of 0.977. The statistical criteria for this model are as follows: Akaike information criterion (*AIC*) of 4931.052, Bayesian information criterion (*BIC*) of 4968.933, root mean square error (*RMSE*) of 1.218, and mean error (*ME*) of − 0.036. The *F*-test was used to test whether there was a statistically significant difference in dominant heights between ecoregions. The results demonstrated that the dominant heights exhibited statistically significant differences among the ecoregions. Consequently, it is of paramount importance to utilize ecoregion-based dynamic site index models in order to achieve reliable and accurate predictions.

Keywords Dominant height · Height growth · Ecoregions · Mixed-effects modeling · *Pinus sylvestris* L

F. Sağlam (✉) · O. E. Sakici
Department of Forest Engineering, Faculty of Forestry, Kastamonu University, 37150 Kastamonu, Türkiye
e-mail: fsaglam@kastamonu.edu.tr

O. E. Sakici
e-mail: oesakici@kastamonu.edu.tr

Introduction

Forest site productivity refers to the production capacity of a stand at a specific age and can also be defined as the combination of physical and biological factors that determine a site's capability to support stand growth (Günel, 1981; Skovsgaard & Vanclay, 2008). Site characteristics, which significantly affect growth, are influenced by factors such as climate, soil, topography, and the site's geographical location (Eler & Carus, 2006).

The site index, which is an indicator of stand productivity, is the dominant height value of a stand at a base age (Adame et al., 2006; Barrio-Anta & Diéguez-Aranda, 2005; Cieszewski & Strub, 2008; Kalıpsız, 1998; Mensah et al., 2022). Base age is a fundamental component of the site index models developed for the estimation of stand yield potential. Dynamic site index models, characterized by standard base-age invariance, ensure that height predictions remain unaffected by variations in base age, thus yielding successful predictions. They also have biologically realistic growth patterns. Two approaches are used to develop these models: the Algebraic Difference Approach (ADA) and the Generalized Algebraic Difference Approach (GADA) (Cieszewski, 2001; Cieszewski & Bailey, 2000; Diéguez-Aranda et al., 2006; Seki & Sakici, 2017).

The Algebraic Difference Approach, one of the approaches in developing site index models with base age-related invariance, was introduced to forestry literature by Bailey and Clutter (1974). In ADA, first, a base growth equation that can reveal the age-height relationship of the stands is selected. Then, it is algebraically rearranged according to a parameter of the growth model. Using this approach, site index models with either a multi-asymptote-anamorphic shape or a single asymptote-polymorphic shape can be developed. However, in terms of desirable biological presumptions, it is important to develop polymorphic and multi-asymptote site index models. To achieve this, algebraic adjustments in the base growth model should be made depending on at least two model parameters. Another approach in the development of site index models, which facilitates the creation of both polymorphic and multi-asymptote site index models while maintaining standard base-age invariance, is the Generalized Algebraic Difference Approach (GADA) developed by Cieszewski and Bailey (2000). In general, it is stated that the models developed using GADA yield reliable predictions and demonstrate compatibility with desirable biological presumptions (Cieszewski et al., 2007; Diéguez-Aranda et al., 2006; Ercanlı, 2010). The desirable presumptions are expressed by a number of researchers (Cieszewski & Bailey, 2000; Diéguez-Aranda et al., 2006; Ercanlı, 2010; Kahriman et al., 2018; Seki & Sakici, 2017) as follows: polymorphism, multiple asymptotes, base-age invariance, sigmoid growth pattern (curves form a wide s shape), and passing the origin ($h=0$ m for $t=0$).

Türkiye has quite a different climate, topography, and ecological characteristics, leading to notable regional variations. When this situation is evaluated regarding forestry, a tree species can spread across extensive areas characterized by markedly different climates and site conditions. Generally, increment and growth relationships differ between regions and ecoregions. Therefore, in situations requiring statistical differentiation, it is always preferable to develop models (ecoregion-based models or ecoregional models) specific to different ecoregions. Such regionalized models allow for individual ecoregions to be analyzed and can provide more reliable predictions on a regional basis (Huang, 1994; Huang et al., 2000; Klos et al., 2007).

When examining studies on ecoregion-based growth models, ecoregion-based site index models were developed by Calama et al. (2003), Álvarez-González et al. (2005), and Adame et al. (2006). The dummy variable method was used to test the ecoregional differences of the models developed in these studies, and it was revealed that the ecological differences had significant effects on the growth predictions. In employing this method, both full and reduced model structures are required to examine ecoregional differences, with comparisons between models conducted using the F-test. While the full model structures have different parameter values for each ecoregion, the reduced model has common parameter values. In general, full model structures are constructed with dummy variables; however, it is crucial to investigate this structure with mixed-effects models. Repeated measurements taken from the same tree or over time are often used in forest modeling. These measurements often exhibit autocorrelation, which means they violate the assumption that error terms are independent (Xu, 2012). Increasingly, mixed-effects models have been used as a means to overcome the autocorrelation problem (Fang & Bailey, 2001; Xu, 2012). Mixed-effects models can also provide a valuable perspective on incorporating differences among sampling units into models. Chenge (2021) stated that mixed-effects modeling is increasingly prominent due to its ability to account for variability in tree height predictions (the similar situation for site index models) across different subjects (e.g., forest type/sites, species, sample plots, clusters, as well as ecoregions for in this study). In the development of base site index models, though the mixed-effects modeling approach has been used in some studies (Lee et al., 2021; Nigh, 2015, 2017; Nothdurft et al., 2006; Senilliani et al., 2019; Sharma & Reid, 2018; Sharma et al., 2015; Wang et al., 2007, 2008, 2011, 2017; Westfall et al., 2017; Xu, 2012; Zhu et al., 2019), the use of this approach in the development of dynamic site index models has been relatively constrained (Fang & Bailey, 2001; García-Espinoza et al., 2019; Meng et al., 2009; Ni & Liu, 2008).

Scots pine (*Pinus sylvestris* L.) is one of the most important species ecologically and economically in Türkiye. Scots pine, the pine species with the widest distribution in the world (Dağdaş et al., 2019), also has an important distribution in Türkiye (1.41 million ha) (GDF, 2021). Scots pine, which has a very wide distribution both throughout Türkiye and in Kastamonu and Sinop regions, also grows under varied ecological conditions and plays a key role in protecting soil and water resources and reducing the negative effects of climate change in its distributed areas. For Scots pine, Alemdağ (1967) constructed site index tables. For the Kastamonu region, a dynamic site index model was developed by Şenyurt and Ercanlı (2013), based on 50 stem analysis data obtained by Şenyurt (2011). For Scots pine, which distributes under significantly different ecological conditions, it is important to develop ecoregional-based dynamic site index models based on new modeling techniques.

In this study, the aim is to develop ecoregion-based dynamic site index models for Scots pine stands distributed in Kastamonu and Sinop Regional Directorates of Forestry, where ecoregions formed by the integration of various ecological factors are also used as a base in modeling. In addition, a new approach using mixed-effects modeling has been tried in the development of ecoregion-based models.

Material and method

Study area

This study was conducted on pure Scots pine stands in the Kastamonu and Sinop regions of northwest Türkiye (Fig. 1). Forest areas cover approximately 65% of the study area, and Scots pine stands spread over an area of approximately 50,772 ha, which exhibit varying ecological characteristics, density levels, sites, and ages (Fig. 1). The study area is located in the Black Sea Region, and according to the data of the General

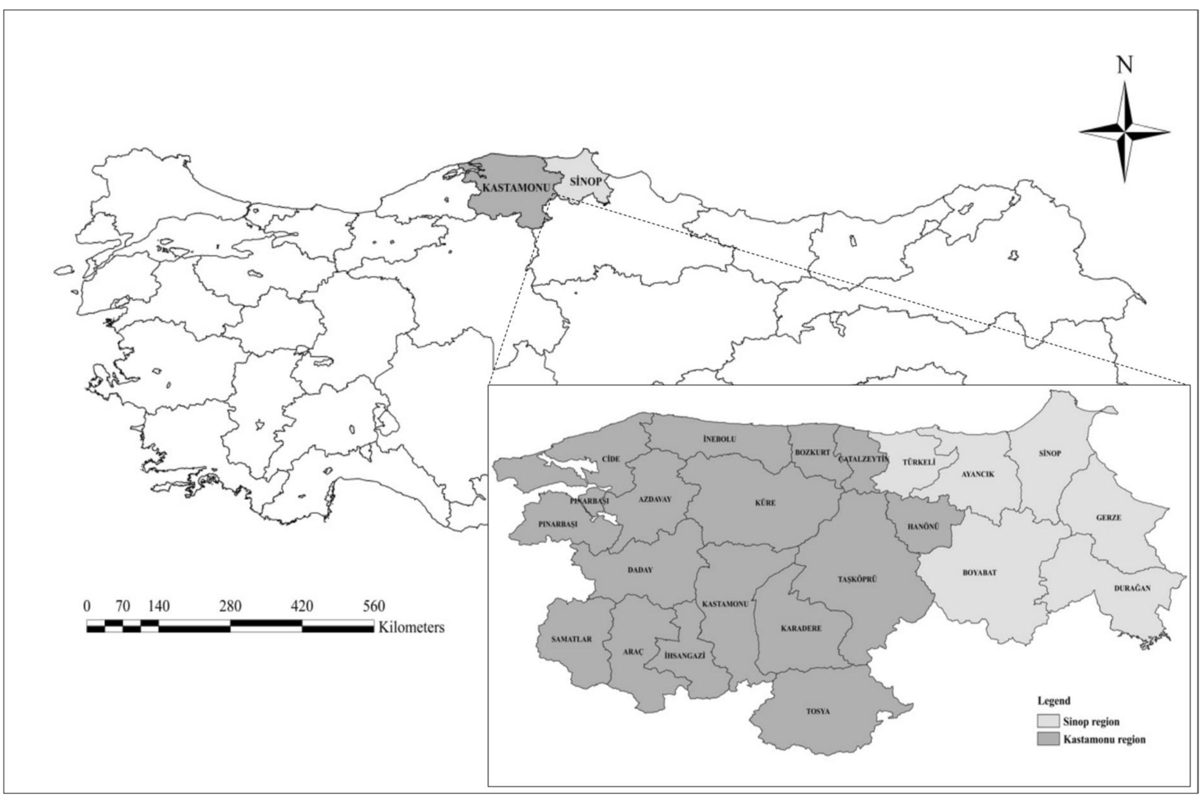

Fig. 1 Study area

Directorate of Meteorology, the average annual precipitation is 500.3 mm for Kastamonu Province and 705.3 mm for Sinop Province (URL-1, 2021).

Ecoregions

According to Atalay (2014), the Kastamonu and Sinop regions are located in the "Black Sea Climate Region," which is divided into five subregions (ecoregions, given below). In this ecological classification, the climate, soil, bedrock, topography, and biotic factors were considered.

- Ecoregion 1. Humid mild broad-leaved forest subregion
- Ecoregion 2. The black sea coastal mountains humid cold coniferous forest subregion
- Ecoregion 3. Subhumid cold coniferous forest subregion of backward black sea plateau and mountains
- Ecoregion 4. Dry forest shrub subregion of backward black sea region
- Ecoregion 5. Mountain grass subregion

The distribution of pure and productive Scots pine stands across ecoregions is shown in Fig. 2. It is evident that these stands are predominantly concentrated within the three ecoregions spatially. In terms of distribution areas, the regions could be ranked from Ecoregion 3 encompassing 36,286.1 ha, to Ecoregion 1 encompassing 6645.6 ha, and to Ecoregion 2 encompassing 6322.6 ha. Conversely, the areas with pure and productive Scots pine stands in Ecoregion 4 and Ecoregion 5 are relatively small. For this reason, it was deemed appropriate to conduct this study on Scots pine stands in Ecoregion 1,

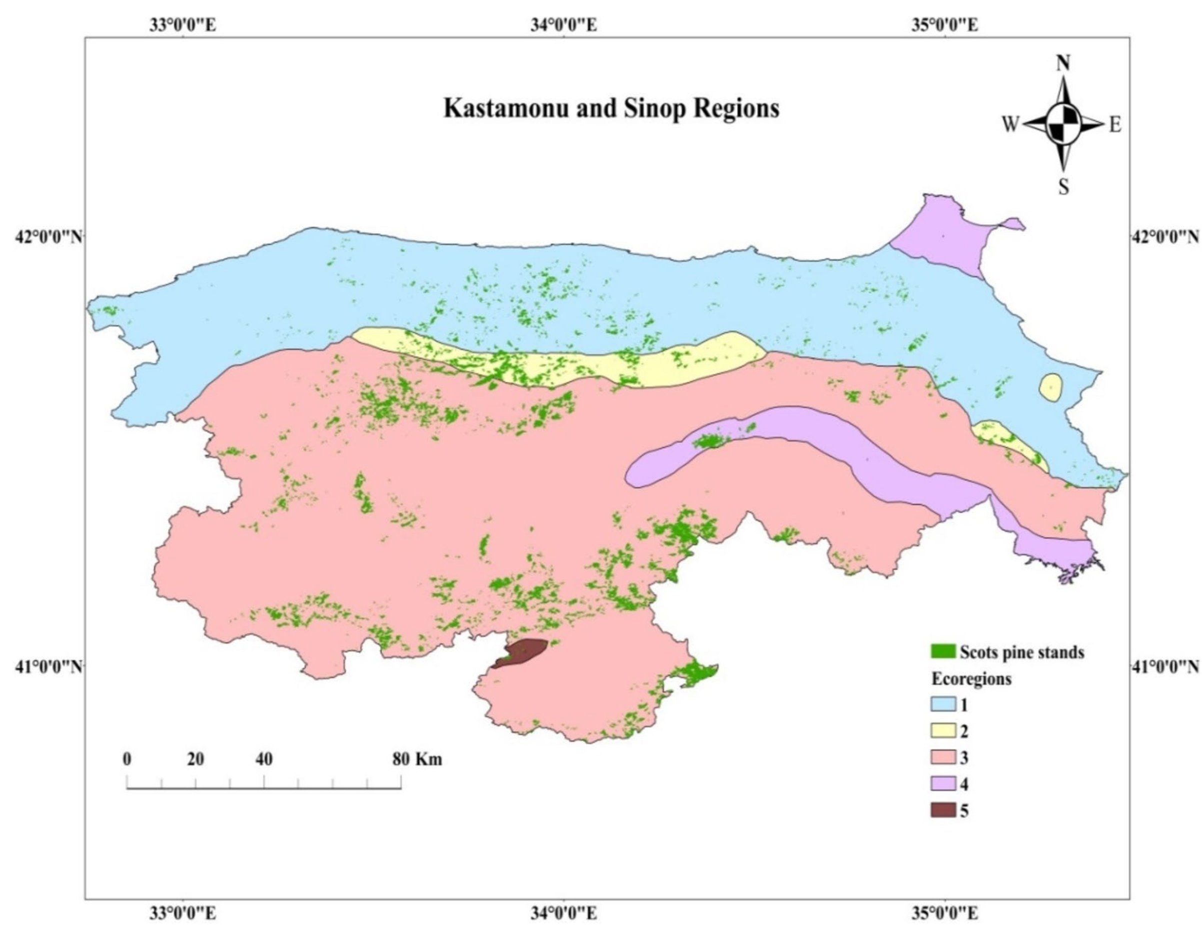

Fig. 2 Distribution of Scots pine stands in the study area to ecoregions

Ecoregion 2, and Ecoregion 3, which collectively represent approximately 97% of the distribution.

Data collection

The material for this study comprised stem analysis data obtained from a total of 268 dominant trees selected from Scots pine stands and distributed across ecoregions with varying site conditions. Stem analysis was conducted on these 268 sample trees, which were chosen from dominant or co-dominant trees representing the respective sites, to develop the site index models. These sample trees needed to be either dominant or co-dominant free-growing trees, showing no clear signs of growth abnormalities or damage. The sample trees were felled as close to the ground as possible, and cross-sections were taken at different cutting heights from the bottom to the top of the trees. Descriptive statistics of sample trees according to ecoregions are given in Table 1.

Candidate models and statistical analysis

In order to develop generalized dynamic site index models, the models (Bertalanffy-Richards, Hossfeld, King-Prodan, Schumacher, Bailey-Clutter, and Lundqvist-Korf) arranged by Cieszewski (2002), Cieszewski (2004), Krumland and Eng (2005), and Trim et al. (2020) were selected (Table 2). The base growth model called Bertalanffy-Richards is also referred to as Richards in some sources and Chapman-Richards model in others (García, 2008; Vanclay, 1994).

The data used in developing the site index models were obtained from the stem analysis. Since these data consist of height values in consecutive periods, they have time series data features. An issue encountered in time series data is autocorrelation, wherein errors are correlated with each other, presenting an undesirable scenario. The fact that the errors are related to each other contradicts the assumption of regression analysis (Ercanlı et al., 2014; Grégoire et al., 1995). For this reason, a mixed-effects modeling approach is often recommended in the literature to eliminate the effect of the autocorrelation problem on the predictions. Autoregressive (AR(p)) and autoregressive moving average (ARMA(p,q)) structures were incorporated into the mixed-effects models, and depending on the improvement in statistical criteria, it was decided which of AR (1) or ARMA (1,1) structures would be used (Fang & Bailey, 2001; Pinheiro & Bates, 2000; Sharma & Reid, 2018; Wang et al., 2017; Xu, 2012).

In the development of the models, ecoregions were evaluated as “random effects”, that is, “subject” since

Table 1 Descriptive statistics for sample trees

	N	Minimum	Maximum	Mean	Standard deviation
Ecoregion 1					
Tree diameter ($d_{1,30}$, cm)	61	20.0	66.0	37.4	8.8
Tree height (h, m)	61	13.0	27.7	20.3	3.8
Tree age (t, year)	61	36	173	88	36.2
Ecoregion 2					
Tree diameter ($d_{1,30}$, cm)	60	18.2	90.2	34.0	14.5
Tree height (h, m)	60	11.1	35.3	20.6	7.5
Tree age (t, year)	60	30	186	90	54.9
Ecoregion 3					
Tree diameter ($d_{1,30}$, cm)	147	17.7	71.5	39.2	10.4
Tree height (h, m)	147	11.5	32.0	22.5	5.3
Tree age (t, year)	147	31	185	104	33.2
Total					
Tree diameter ($d_{1,30}$, cm)	268	17.7	90.2	37.6	11.4
Tree height (h, m)	268	11.1	35.3	21.6	5.7
Tree age (t, year)	268	30	186	97	41.2

Table 2 Base and dynamic site index models

Model	Base model	GADA model
M1	Bertalanffy-Richards $h = b_0\left(1-\exp\left(-b_1 t\right)\right)^{b_2}$	$h = \exp\left(X_0\right)\left[1-\exp(-at)\right]^{\left(b+1/X_0\right)}$ $X_0 = \frac{1}{2}\left(\left(\ln h_0 - bL_0\right) + \sqrt{\left(\ln h_0 - bL_0\right)^2 - 4L_0}\right)$ $L_0 = \ln\left(1-\exp\left(-at_0\right)\right)$
M2	Hossfeld $h = \frac{b_0}{1+b_1 t^{-b_2}}$	$h = \frac{a+X_0}{1+\left(b/X_0\right)t^{-c}}$ $X_0 = \frac{1}{2}\left[h_0 - a - \sqrt{\left(h_0-a\right)^2 + 4bh_0 t_0{}^{-c}}\right]$
M3		$h = \frac{a+X_0}{1+bX_0 t^{-c}}$ $X_0 = \frac{h_0-a}{1-bh_0 t_0{}^{-c}}$
M4	King-Prodan $h = \frac{t^{b_0}}{b_1+b_2 t^{b_0}}$	$h = \frac{t^a}{b+cX_0+X_0 t^a}$ $X_0 = \frac{\left(\frac{t_0{}^a}{h_0}\right)-b}{c+t_0{}^a}$
M5	Schumacher $h = \exp\left(b_0 + b_1 t^{-1}\right)$	$h = \exp\left(X_0 + \left(a+bX_0\right)t^{-1}\right)$ $X_0 = \frac{\ln h_0 - at_0{}^{-1}}{1+bt_0{}^{-1}}$
M6	Bailey-Clutter $h = \exp\left(b_0 + b_1 t^{b_2}\right)$	$h = \exp\left(a + X_0 + bX_0 t^c\right)$ $X_0 = \frac{lnh_0 - a}{1+bt_0{}^c}$
GADA7	Lundqvist-Korf $h = b_0\exp\left(-b_1 t^{-b_2}\right)$	$h = \exp\left(X_0\right)\exp\left[-\left(b_1 + \frac{b_2}{X_0}\right)t^{-b_3}\right]$ $X_0 = \frac{1}{2}\left(at_0{}^{-c} + \ln h_0 + L_0\right)$ $L_0 = \sqrt{\left(at_0{}^{-c} + \ln h_0\right)^2 + 4bt_0{}^{-c}}$

b_0, b_1, and b_2, parameters of base models; a, b, c, parameters of site index models obtained with GADA; h and h_0, height and dominant height; t and t_0, age and base age

it was investigated whether there were significant differences in terms of dominant heights between ecoregions. The "nlme" package in the R program was used in the development of the models (Pinheiro & Bates, 2000; R Core Team, 2021).

Model selection

Initially, all possible combinations of the random effects parameters were fitted and the best combination was determined using Akaike information criterion ($AIC = -2LL + 2p$), and Bayesian information criterion ($BIC = -2LL + plogn$, where n is the number of observations, LL is the log-likelihood of the fitted model, and p is the number of model parameters). After determining the best combination for the models, the coefficient of determination (R^2), AIC, root mean square error ($RMSE$), and mean error (ME) were used to select the best model using the relative ranking suggested by Poudel and Cao (2013). Then, estimations were made with the model with significant parameters and satisfactory in terms of statistical criteria, and the distributions of the residuals were examined.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(h_i - \hat{h}_i\right)^2}{\sum_{i=1}^{n}\left(h_i - \overline{h}_i\right)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(h_i - \hat{h}_i\right)^2} \quad (2)$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(h_i - \hat{h}_i\right) \quad (3)$$

where h_i, $\widehat{h}_i$, and $\overline{h}_i$ represent the observed, estimated, and average height values, respectively, and n the number of observations.

In order to check the compatibility of the mixed-effects dynamic site index models with the desirable biological attributes, the age-height relationships related to the estimations made by the models should be examined. For this purpose, relative error ($RE\%$) values for various base ages were calculated and assessed to determine the most appropriate base age. $RE\%$ values were calculated using Eq. 4.

$$RE\% = 100\frac{\sqrt{\sum_{i=1}^{n}\left(h_i - \widehat{h}_i\right)^2/(n-p)}}{\overline{h}} \tag{4}$$

where h_i, $\widehat{h}_i$, and $\overline{h}$ represent the observed, estimated, and average height values, respectively; n is the number of data; and p is the number of parameters in the model.

In addition, one of the factors that should be considered when deciding on the base age is that the base age value should not be greater than the rotation age of the tree species (Goelz & Burk, 1992).

Testing ecoregion differences

The "Non-Linear Sum of Extra Squares Test," also known as the F-test, was used to determine whether there were differences between the ecoregions in terms of dominant heights. This test requires both "full" and "reduced" model structures. The model structure referred to as the "reduced model" uses the same parameter values for all ecoregions examined, whereas the "full model" uses different parameter values for each ecoregion (Pillsbury et al., 1995). Full model structures are usually presented using dummy variables. However, in this study, a new approach was adopted: mixed-effects modeling was used to reveal the full model structures. In order to reveal ecoregional differences, ecoregions were introduced into the models as "random effects", that is, "subject" in the mixed-effects modeling approach. The parameters with the most successful random effects on the models were estimated for three ecoregions with the help of the "nlmeStruct" and "random.effects" functions in the "nlme" package in the R. This enabled the full model structure to be obtained, which was used to test the differences between the ecoregions (Mehtätalo, 2013; Pinheiro & Bates, 2000). The test statistic for the nonlinear sum of extra squares employed to assess the ecoregional differences is as follows (Eq. 5) (Neter et al., 1996; Pillsbury et al., 1995):

$$F == \frac{\left(SSE_R - SSE_F\right)/\left(DF_R - DF_F\right)}{\left(SSE_F/DF_F\right)} \tag{5}$$

In the equation, SSE_R and SSE_F represent the sum of squares of error of reduced and full site index models, respectively, and DF_RandDF_F represent the degrees of freedom of reduced and full site index models, respectively.

F-distribution was used to interpret the F-value in revealing ecoregional differences. It is concluded that dominant heights for ecoregions are statistically different when the p-value is less than 0.05 (Pillsbury et al., 1995). As a result, in case of differences between ecoregions, full model structures with different parameters for each ecoregion should be used instead of reduced models.

Results and discussion

In order to develop site index models, a series of combinations of random effects parameters were tested in order to determine the most appropriate mixed-effects modeling approach. The model's convergence was provided for the initial five models and random effects parameter combinations were evaluated. In the last two models (M6 and M7), some combinations have not been resolved, and some combinations have non-significant parameter values. In this context, 3 different random effects parameter combinations were evaluated for M1 and M5, and 7 different combinations of random effects parameters for M2, M3, and M4. Finally, three random effects parameter combinations for M1, M4, and M5 and 6 random effects parameter combinations for M2 and M3 are provided for the solution. Random effects parameter combinations for the models are given in Table 3. When the relevant table is examined, the "b" parameter for M1 and M2, the "a" parameter for M3 and M5, and the "c" parameter for M4 gave the most satisfactory outcomes.

Once the most successful random effects parameter combinations for the five models were identified, the models were evaluated to determine the best site index

Table 3 Random effects parameter combinations for site index models

Model	Random effect parameters	*AIC*	*BIC*	Relative ranking		Total	Total ranking
				AIC	*BIC*		
M1	a	5523.245	5544.891	3.00	2.08	5.08	3.00
	b	5520.360	5542.006	2.31	1.00	3.31	1.00
	a, b	5514.902	5547.372	1.00	3.00	4.00	1.78
M2	a	5392.268	5419.326	1.01	1.01	2.02	1,01
	b	5391.841	5418.899	1.00	1.00	2.00	1.00
	c	5568.439	5595.497	6.00	6.00	12.00	6.00
	a, b	5394.626	5432.507	1.08	1.39	2.46	1.23
	a, c	5394.504	5432.385	1.08	1.38	2.46	1.23
	a, b, c	5399.734	5453.849	1.22	1.99	3.21	1.61
M3	a	5369.412	5396.470	1.00	1.00	2.00	1.00
	b	5375.015	5402.073	1.14	1.14	2.28	1.14
	c	5570.874	5597.932	6.00	6.00	12.00	6.00
	a, b	5373.412	5411.293	1.10	1.37	2.47	1.23
	a, c	5373.099	5410.979	1.09	1.36	2.45	1.23
	b, c	5376.145	5414.026	1.17	1.44	2.60	1.30
M4	a	5563.279	5590.337	3.00	3.00	6.00	3.00
	b	5372.174	5399.232	1.03	1.03	2.06	1.03
	c	5369.397	5396.455	1.00	1.00	2.00	1.00
M5	a	6624.847	6646.494	2.13	1.00	3.13	1.00
	b	6626.806	6648.452	3.00	1.47	4.47	3.00
	a, b	6622.316	6654.785	1.00	3.00	4.00	2.30

Table 4 The rankings for the site index models

Model	Random effects parameters	*AIC*	*BIC*	*RMSE*	R^2	*ME*	Relative ranking					Total	Total ranking
							AIC	*BIC*	*RMSE*	R^2	*ME*		
M1	b	5520.360	5542.006	1.275	0.975	−0.037	1.48	1.47	1.41	1.34	1.25	6.95	1.38
M2	b	5391.841	5418.899	1.225	0.976	−0.013	1.07	1.07	1.06	1.05	1.00	5.25	1.04
M3	a	5369.412	5396.470	1.217	0.977	−0.020	1.00	1.00	1.00	1.00	1.07	5.07	1.01
M4	c	5369.397	5396.455	1.217	0.977	−0.016	1.00	1.00	1.00	1.00	1.03	5.03	1.00
M5	a	6624.847	6646.494	1.778	0.950	−0.385	5.00	5.00	5.00	5.00	5.00	25.00	5.00

model (Table 4). When Table 4 is examined, it is seen that the models generally have close relative rankings, while the M5 model is the model with the highest ranking. According to the statistical success criteria, it is seen that the M3 and M4 models had almost equal rankings. However, the statistical success criteria of the M4 model had relatively lower values. Consequently, the M4 model was selected as the best model.

To eliminate the effect of autocorrelation on model estimation, various error structures were included in the model. Autoregressive (AR(1)) and moving average (ARMA(1, 1)) were included in the model and the error structure to be included in the model was decided based on the improvement in statistical criteria. For this purpose, the M4 model (M4c), which includes the "c" parameter with random effects; the M4c.ar model, which tests the AR(1) error structure; and the M4c.arma model, which tests the ARMA(1,1) error structure added to the M4c model were evaluated in terms of statistical criteria (Table 5).

Upon examining Table 5, it was observed that adding ARMA error structure to the mixed-effects

Table 5 The rankings of error structures related to site index models

Model	*AIC*	*BIC*	*RMSE*	R^2	*ME*	Relative ranking					Total	Total ranking
						AIC	*BIC*	*RMSE*	R^2	*ME*		
M4c.arma	4931.052	4968.933	1.218	0.977	−0.036	1.00	1.01	3.00	2.00	2.82	9.83	1.00
M4c.ar	4934.136	4966.605	1.218	0.977	−0.038	1.01	1.00	3.00	2.00	3.00	10.01	3.00
M4c	5369.397	5396.455	1.217	0.977	−0.016	3.00	3.00	1.00	2.00	1.00	10.00	2.85

Table 6 Parameters and statistical criteria for the M4c.arma model

Parameter	Parameter value	*AIC*	*BIC*	*RMSE*	R^2	*ME*
a	1.3368	4931.052	4968.933	1.218	0.977	−0.036
b	−25.9291					
c	1658.4054					
$\sigma^2_{u_1}$	310.3310					
σ^2	1.2167					

Table 7 Parameters of ecoregions from M4c.arma model and reduced model parameters

Parameter	Parameter value	Ecoregion 1	Ecoregion 2	Ecoregion 3	Reduced
a	1.3368	1.3368	1.3368	1.3368	1.2237
b	−25.9291	−25.9291	−25.9291	−25.9291	−11.6073
c	1658.4054	1307.9770	2057.6080	1609.6310	956.6950

Table 8 *F*-test results for the M4c.arma model

Model	Ecoregions	Reduced model		Full model		*n*	*F*-values	*p*
		SSE	*DF*	*SSE*	*DF*			
M4c.arma	1–2-3	2858.643	1652	2445.149	1648	1655	69.673	< 0.001
	1–2	1585.205	796	1222.164	793	799	78.520	< 0.001
	1–3	2048.313	1289	1813.274	1286	1292	55.564	< 0.001
	2–3	2083.769	1216	1854.860	1213	1219	49.899	< 0.001

dynamic site index models resulted in a significant increase in statistical criteria success. Ithe M4c.arma model is superior to the M4c.ar and M4c models. Consequently, the M4c.arma model was selected as the final model. The mixed-effects model structure of the final model is presented in Eq. 6. The parameters and statistical criteria are provided in Table 6. Table 7 provides the parameters for each ecoregion and reduced model parameters estimates with common parameters for the ecoregions.

$$h = \frac{t^{a}}{b + (c + u_1)X_0 + X_0 t^{a}} X_0 = \frac{\left(\frac{t_0{}^{a}}{h_0}\right) - b}{(c + u_1) + t_0{}^{a}} \quad (6)$$

The *F*-values presented in Table 8 were calculated to evaluate the efficacy of ecoregional models and to determine whether an ecoregional modelling approach is necessary. Examination of the table revealed that the dominant height development of Scots pine stands in all ecoregions within the Kastamonu and Sinop regions showed statistically significant differences from one another ($p < 0.05$). It was thus demonstrated that the ecoregion-based dynamic site index model yielded superior predictions in comparison to the reduced model.

The residuals of the developed dynamic site index model (M4c.arma) are presented in Fig. 3. Upon examination of Fig. 3, it becomes evident that the error values associated with the predictions generated

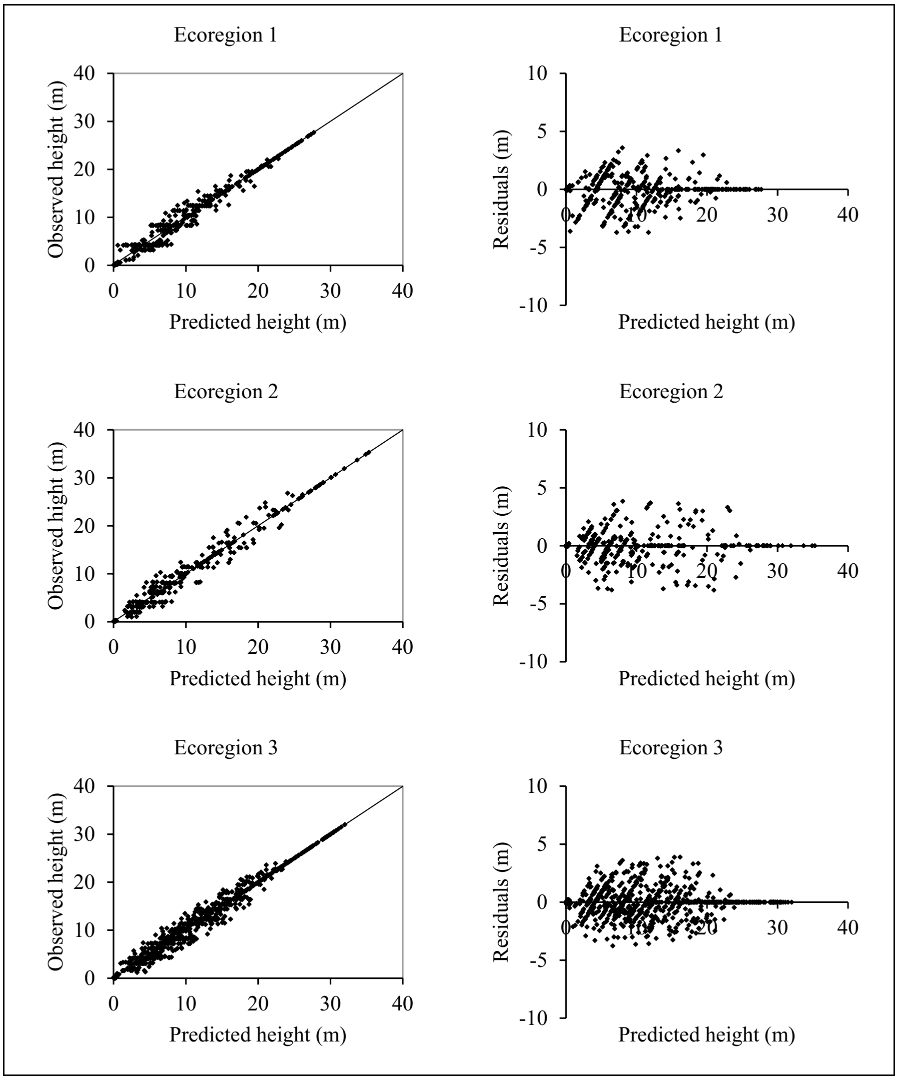

Fig. 3 Residuals for ecoregion-based dynamic site index models

by the M4c.arma model exhibit a random distribution and are devoid of any discernible trend. This is a crucial because the absence of a discernible trend in residuals aligns with the assumptions of regression analysis. This is an essential criterion for the success and suitability of the model.

In this study, seven models were evaluated in order to develop ecoregion-based dynamic site index models for Scots pine stands in the Kastamonu and Sinop regions. Among these models, M4c emerged as the most promising. In order to eliminate the effect of autocorrelation in model predictions, error structures were added to the model and examined. The results indicated that M4c.arma yielded the most satisfactory outcomes.

Figure 4 was created by considering the relative errors for different ages in order to determine the base age. Upon examination of Fig. 4, it becomes evident that the lowest relative errors are observed after the age of 90. As stated in the study by Diéguez-Aranda et al. (2006), the base age should not exceed the rotation age of the subject tree species for it to be a useful indicator of the silvicultural treatments to be applied to the stands. In light of the aforementioned considerations, it is appropriate to determine the base age as 100, taking into account the age range with the lowest relative errors.

The compatibility of the ecoregion-based mixed-effects site index models with the desirable biological attributes was examined through the use of

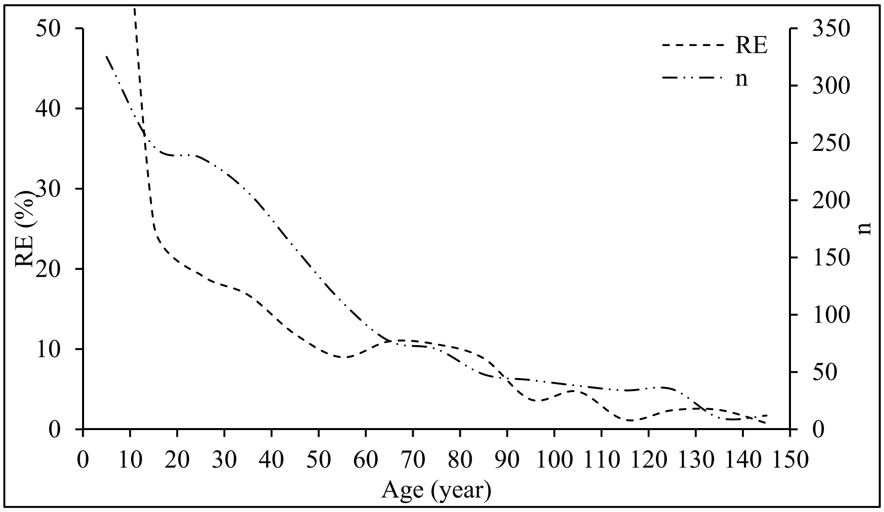

Fig. 4 Relative error for different base ages. RE is the relative error; *n* is the number of data

graphics. In order to achieve this objective, the base age (100 years) and three site index values (30.5, 23.5, and 16.5 m; site class I, II, and III, respectively) were employed to plot curves through the models. In order to ascertain the change in the mean annual increment (*MAI*) and the change in the site index curves, the 200-year change was analyzed graphically, separately for the three ecoregions, taking into account the three site index values (Fig. 5).

The compatibility of the M4c.arma model with desirable biological presumptions was evaluated. This evaluation included examining how the M4c.arma model varies with age and ecoregions, whether it exhibits a sigmoid growth pattern in age-height relationships depending on site productivity, demonstrates polymorphism, features multiple asymptotes, and shows a decrease in age required to reach the maximum mean annual increment (*MAI*) as site productivity increases. In this context, the examination of Fig. 5 reveals that the site index curves of the M4c.arma model for each ecoregion exhibit satisfactory biological characteristics, including a sigmoid growth pattern, polymorphism, and multiple asymptotes. Furthermore, it is evident that the M4c.arma model also provides the attribute of "the maximum value of the *MAI* and the decrease in the age to reach this maximum value depending on the improvement in the site productivity" for all ecoregions (Cieszewski & Bailey, 2000).

For the dynamic site index models, the "c" parameter had random effects and took different parameter values for the ecoregions. In order to examine the differences in dominant height development between ecoregions, the M4c.arma model was used to graphically examine 200 years of height development for each ecoregion (Fig. 6). The graphs also include estimations of the reduced model, which provides estimations without consideration of ecoregional differences. It is understood that there are general differences between ecoregions in terms of dominant height. It is seen that these differences are more evident due to the stronger competition for site class I. Comparisons of dominant height differences between ecoregions for site class III show smaller differences than those observed for site classes I and II. Although these differences were relatively minor until the base age, they became more pronounced after this point.

In Ecoregion 1, the dominant heights are observed to be higher than those of the other ecoregions until the base age. However, after this point, the dominant heights are found to be less than those of the other ecoregions. Conversely, dominant heights for Ecoregion 2 are lower than those for other ecoregions up to the base age but higher thereafter. To summarize briefly, in terms of dominant height estimations, the rankings are as follows: Ecoregion 1 is the highest, followed by Ecoregion 3, and Ecoregion 2 before the base age, and Ecoregion 2, Ecoregion 3, and Ecoregion 1 from high to low after the base age. The dominant height values from the reduced model, which does not account for ecoregion differences, are similar to those for Ecoregion 3 but quite different from those for Ecoregion 1 and Ecoregion 2. Using ecoregion-based mixed-effects dynamic site index models,

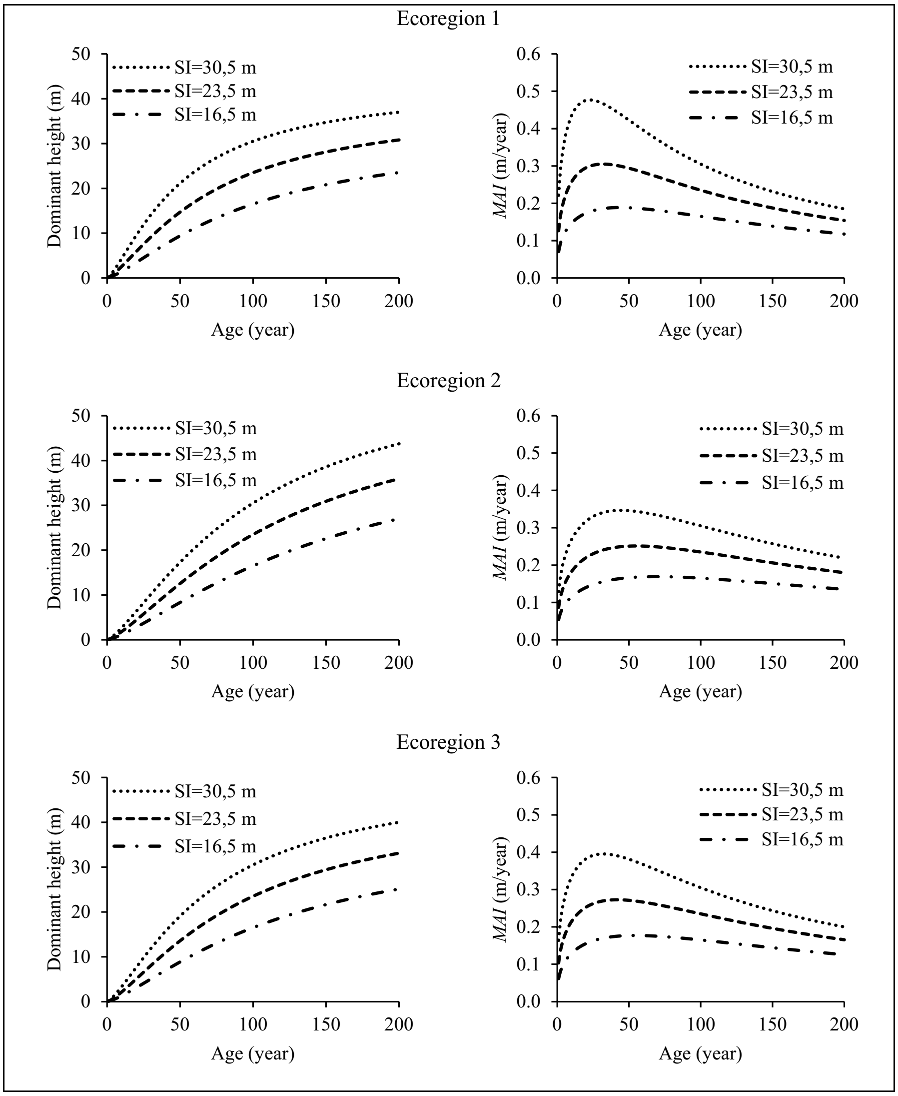

Fig. 5 Variation of site index (SI) and mean annual increment (*MAI*) for ecoregions

the dominant heights were predicted for each ecoregion and Bias and Bias(%) values were calculated (Table 9).

Upon examination of Table 9, it can be observed that the bias is positive and relatively minor when the ecoregion-based mixed-effects dynamic site index models are employed in the appropriate ecoregion. The bias for the predictions made using reduced site index models that do not take into account the differences between ecoregions and have common parameter values for ecoregions is low for Ecoregion 1, whereas it is rather high for Ecoregion 1 and Ecoregion 2. The utilization of a mixed-effects dynamic site index model specific to a given ecoregion in a different ecoregion results in the generation of highly biased values. The highest degree of bias was observed when the models of Ecoregion 1 and Ecoregion 2 were used interchangeably. This can be attributed to the fact that the dominant height developments in these two ecoregions are markedly disparate from one another.

Alemdağ (1967) prepared site index tables for Scots pine, while Şenyurt and Ercanlı (2013) developed a dynamic site index model for Kastamonu province. The ecoregion-based mixed-effects dynamic site index model and the other site index

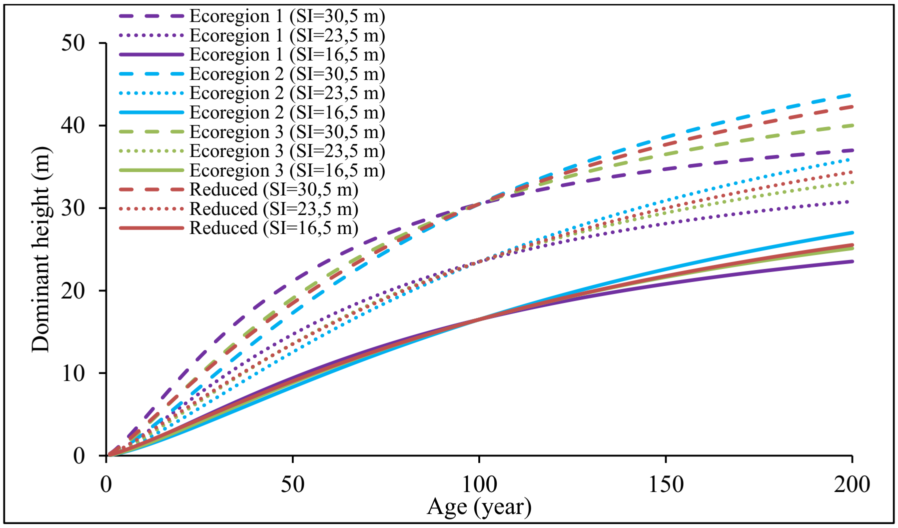

Fig. 6 Site index curves according to ecoregions

Table 9 Prediction errors for site index models

Model form	Applied ecoregion	*n*	Average observed heights (m)	Average predicted heights (m)	Bias (m)	Bias (%)
Reduced	E1	436	9.5	9.0	0.47	5.01
E1				9.4	0.09	0.91
E2				8.5	1.01	10.64
E3				8.9	0.58	6.17
Reduced	E2	363	9.5	10.0	−0.50	−5.27
E1				10.6	−1.16	−12.25
E2				9.4	0.05	0.55
E3				10.0	−0.50	−5.25
Reduced	E3	856	10.4	10.5	−0.06	−0.58
E1				11.1	−0.72	−6.92
E2				9.8	0.60	5.77
E3				10.4	0.00	0.03

models developed in the aforementioned studies were subjected to a graphical comparison, with a view to examining the extent of the differences in predictions between ecoregions (Fig. 7).

As illustrated in Fig. 7a, it is seen that the differences in the change of site index curves for Ecoregion 1 are less in the study by Şenyurt and Ercanlı (2013) in comparison to the study by Alemdağ (1967). Differences after the base age are smaller with Şenyurt and Ercanlı (2013) for site class I, but Alemdağ (1967) shows less variation for site class III. In Fig. 7b, the estimations of Ecoregion 2 are examined. It is observed that while the difference between the estimations produced by M4c.arma and those produced by Alemdağ (1967) is relatively minor, the differences increase significantly when Şenyurt and Ercanlı (2013) are considered, particularly after the base age. For Ecoregion 3, discrepancies between the M4c.arma model and literature models are less before the base age but become more pronounced afterward, particularly in comparison to the study by Şenyurt and Ercanlı (2013) (Fig. 7c). Following a graphical examination, it was observed that there were differences in the predictions of the models for each ecoregion.

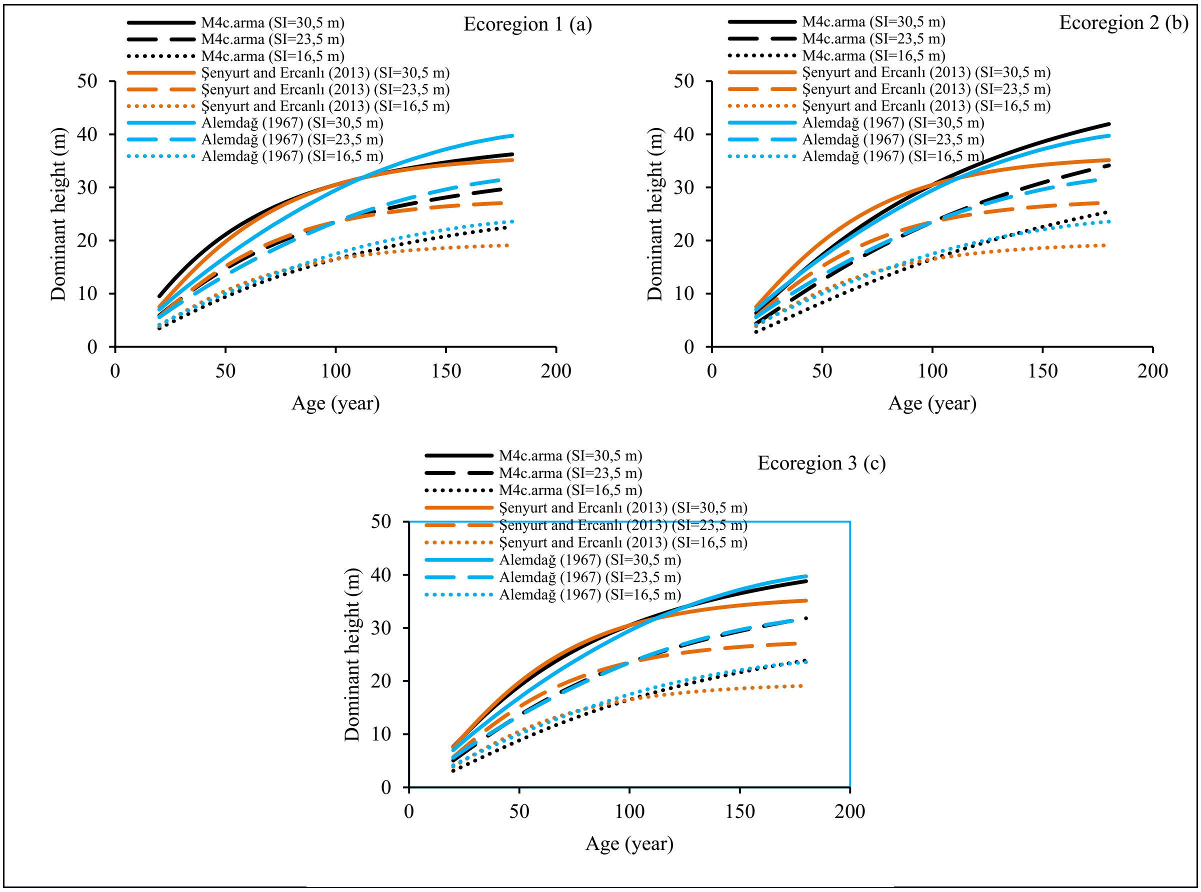

Fig. 7 Predictions of site index models for Ecoregion 1, Ecoregion 2, and Ecoregion 3

These differences may arise from whether the data used in the development of the models take into account the ecoregion differences or the differences in the site index model structures developed, as well as the modeling approach used.

Conclusion

The results of this study indicate that the dominant heights of all ecoregions in the Scots pine (*Pinus sylvestris* L.) stands of the study area exhibited statistically significant differences from one another ($p < 0.05$). Consequently, the use of an ecoregion-based and mixed-effects dynamic site index model has been demonstrated to provide a superior prediction to the reduced model. The models developed for this study yielded reliable predictions for the regions and species from which the data were obtained. Furthermore, the utilization of ecoregion-based site index models in the appropriate ecoregions was of significant importance in terms of the reliability of predictions and the minimization of prediction errors. In contrast, predictions made without considering ecological differences using dynamic site index models tend to show increased errors.

Acknowledgements This study was produced from a PhD thesis prepared by Fadime Sağlam and supervised by Oytun Emre Sakici for the Institute of Natural and Applied Science at Kastamonu University in Türkiye.

Author contribution Conceptualization: Fadime Sağlam, Oytun Emre Sakici; methodology: Fadime Sağlam, Oytun Emre Sakici; formal analysis and investigation: Fadime Sağlam; writing—original draft preparation: Fadime Sağlam; writing—review and editing: Fadime Sağlam, Oytun Emre Sakici; funding acquisition: Oytun Emre Sakici; resources:

Fadime Sağlam; supervision: Oytun Emre Sakici. The authors read and approved the final manuscript.

Funding This study was supported by the Scientific Research Projects Coordination Unit of Kastamonu University, under the grant number KÜ-BAP01/2019–41.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethics approval All authors have read, understood, and complied as applicable with the statement "Ethical responsibilities of Authors" as found in the Instructions for Authors.

Conflict of interest The authors declare no competing interests.

References

Adame, P., Cañellas, I., Roig, S., & Del Rio, M. (2006). Modelling dominant height growth and site index curves for rebollo oak (*Quercus pyrenaica* Willd.). *Annals of Forest Science, 63*(8), 929–940.

Alemdağ, İ.Ş. (1967). *Türkiye'deki sarıçam ormanlarının kuruluşu, verim gücü ve bu ormanların işletilmesinde takip edilecek esaslar*. Ormancılık Araştırma Enstitüsü. (in Turkish)

Álvarez-González, J. G., Ruíz-González, A. D., Rodríguez-Soalleiro, R., & Barrio-Anta, M. (2005). Ecoregional site index models for *Pinus pinaster* in Galicia (northwestern Spain). *Annals of Forest Science, 62*(2), 115–127.

Atalay, İ. (2014). *Türkiye'nin ekolojik bölgeleri. Genişletilmiş 2. Baskı, İzmir: Meta Basım Matbaacılık*. (in Turkish)

Bailey, R. L., & Clutter, J. L. (1974). Base-age invariant polymorphic site curves. *Forest Science, 20*, 155–159.

Barrio-Anta, M., & Diéguez-Aranda, U. (2005). Site quality of pedunculate oak (*Quercus robur* L.) stands in Galicia (northwest Spain). *European Journal of Forest Research, 124*(1), 19–28.

Calama, R., Cañadas, N., & Montero, G. (2003). Inter-regional variability in site index models for even-aged stands of stone pine (*Pinus pinea* L.) in Spain. *Annals of Forest Science, 60*, 259–269.

Chenge, I. B. (2021). Height-diameter relationship of trees in Omo strict nature forest reserve, Nigeria. *Trees, Forests and People, 3*, 100051. https://doi.org/10.1016/j.tfp.2020.100051

Cieszewski, C. J. (2001). Three methods of deriving advanced dynamic site equations demonstrated on inland Douglas-fir site curves. *Canadian Journal of Research, 31*(1), 165–173.

Cieszewski, C. J. (2002). Comparing fixed- and variable-base-age site equations having single versus multiple asymptotes. *Forest Science, 48*(1), 7–23.

Cieszewski, C. J. (2004). *GADA derivation of dynamic site equations with polymorphism and variable asymptotes from Richards to Weibull and other exponential functions* (pp. 2004–2005). Plantation Management Research Cooperative Technical Report.

Cieszewski, C. J., & Bailey, R. L. (2000). Generalized algebraic difference approach: Theory based derivation of dynamic site equations with polymorphism and variable asymptotes. *Forest Science, 46*, 116–126.

Cieszewski, C. J., & Strub, M. (2008). Generalized algebraic difference approach derivation of dynamic site equations with polymorphism and variable asymptotes from exponential and logarithmic functions. *Forest Science, 54*(3), 303–315.

Cieszewski, C. J., Strub, M., & Zasada, M. J. (2007). New dynamic site equation that fits best the schwappach for Scots pine (*Pinus slyvestris* L.) in Central Europe. *Forest Ecology and Management, 23*, 83–93.

Dağdaş, S., Elmas, M., & Timur, L. (2019). Türkiye'de, sarıçamın (*Pinus sylvestris* L.) en güney enlemlerde belirlenen yeni yayılışları. *Doğanın Sesi, 3*, 50–62. (in Turkish).

Diéguez-Aranda, U., Burkhart, H. E., & Amateis, R. L. (2006). Dynamic site model for loblolly pine (*Pinus taeda* L.) plantations in the United States. *Forest Science, 52*(3), 262–272.

Eler, Ü., & Carus, S. (2006). *Orman hasılat bilgisi*. Süleyman Demirel Üniversitesi Orman Fakültesi Yayın No:66, 201s. (in Turkish)

Ercanlı, İ, Kahriman, A., & Yavuz, H. (2014). Dynamic base-age invariant site index models based on generalized algebraic difference approach for mixed Scots pine (*Pinus sylvestris* L.) and Oriental beech (*Fagus orientalis* Lipsky) stands. *Turkish Journal of Agriculture and Forestry, 38*, 134–147.

Ercanlı, İ. (2010). *Trabzon ve Giresun Orman Bölge Müdürlükleri Sınırları İçerisinde Yer Alan Doğu Ladini (Picea orientalis (L.) Link)-Sarıçam (Pinus sylvestris L.) Karışık Meşcerlerine İlişkin Büyüme Modelleri*. Doktora Tezi, Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü. Trabzon. (in Turkish)

Fang, Z., & Bailey, R. L. (2001). Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *Forest Science, 47*(3), 287–300.

García, O. (2008). Visualization of a general family of growth functions and probability distributions - the growth-curve explorer. *Environmental Modelling & Software, 23*(12), 1474–1475.

García-Espinoza, G. G., Aguirre-Calderón, O. A., Quiñonez-Barraza, G., AlanísRodríguez, E., González-Tagle, M. A., & García-Magaña, J. J. (2019). Global-local and fixed random parameters to model dominant height growth of Pinus pseudostrobus Lindley. *Revista Chapingo Serie Ciencias Forestales y Del Ambiente, 25*(1), 141–156.

GDF. (2021). *Türkiye orman varlığı,* 2020. Orman Genel Müdürlüğü Yayınları. (in Turkish)

Goelz, J. C. G., & Burk, T. E. (1992). Development of a well-behaved site index equation: Jack pine in north central Ontario. *Canadian Journal of Forest Research, 22*(6), 776–784.

Grégoire, T. G., Schabenberger, O., & Barrett, J. P. (1995). Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Canadian Journal of Forest Research, 25*(1), 137–156.

Günel, A. (1981). *Orman hasılat bilgisi ders notları*. İstanbul Üniversitesi Orman Fakültesi (Basılmamıştır). (in Turkish)

Huang, S., Price, D., & Titus, S. J. (2000). Development of ecoregion-based height diameter models for white spruce in boreal forests. *Forest Ecology and Management, 129*(1–3), 125–141.

Huang, S. (1994). *Ecologically based individual tree volume estimation for major Alberta tree species.* AB Sustainable Resource Development Public Lands and Forests Division.

Kahriman, A., Sönmez, T., & Gadow, V. K. (2018). Site index models for Calabrian pine in the central Mediterranean region of Turkey. *Journal of Sustainable Forestry, 35*(5), 459–474.

Kalıpsız, A. (1998). *Orman hasılat bilgisi.* İstanbul Üniversitesi Orman Fakültesi Yayınları. (in Turkish)

Klos, R. J., Wang, G. G., Dang, Q. L., & East, E. W. (2007). Taper equations for five major commercial tree species in Manitoba, Canada. *Western Journal of Applied Forestry, 22*(3), 163–170.

Krumland, B., & Eng, H. (2005). *Site index systems for major young-growth forest and woodland species in northern California*. California Department of Forestry and Fire Protection, Report No: 4.

Lee, D., Beuker, E., Vihera-Aarni, A., & Hynynen, J. (2021). Site index models with density effect for hybrid aspen (*Populus tremula* L.×*P. tremuloides* Michx.) plantations in southern Finland. *Forest Ecology and Management, 480*, 118669.

Mehtätalo, L. (2013). *Forest biometrics with examples in R*. University of Eastern Finland School of Computing.

Meng, S. X., Huang, S., Yang, Y., Trincado, G., & VanderSchaaf, C. L. (2009). Evaluation of population-averaged and subject-specific approaches for modeling the dominant or codominant height of lodgepole pine trees. *Canadian Journal of Forest Research, 39*(6), 1148–1158.

Mensah, A. A., Holmström, E., Nyström, K., & Nilsson, U. (2022). Modelling potential yield capacity in conifers using Swedish long-term experiments. *Forest Ecology and Management, 512*, 120162.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models*. Irwin.

Ni, C., & Liu, C. (2008). Evaluating behaviors of factors affecting the site index estimate on the basis of a single stand using simulation approach. *Canadian Journal of Forest Research, 38*, 2762–2770.

Nigh, G. (2015). Engelmann spruce site index models: A comparison of model functions and parameterizations. *PLoS ONE, 10*(4), 1–19.

Nigh, G. (2017). *Development of a lodgepole pine site ındex model with the grounded generalized algebraic difference approach (g-GADA)*. B.C. Research Report 31. www.for.gov.bc.ca/hfd/pubs/Docs/Rr/Rr31.htm.

Nothdurft, A., Kublin, E., & Lappi, J. (2006). A non-linear hierarchical mixed model to describe tree height growth. *European Journal Forest Research, 125*(281), 289.

Pillsbury, N. H., McDonald, P. M., & Simon, V. (1995). Reliability of tanoak volume equations when applied to different areas. *Western Journal of Applied Forestry, 10*(2), 72–78.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. Springer Verlag.

Poudel, K. P., & Cao, Q. V. (2013). Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. *Forest Science, 59*(2), 243–252.

R Core Team (2021) R version 4.0.5 (2021–03–31). The R foundation for statistical computing platform.

Seki, M., & Sakici, O. E. (2017). Dominant height growth and dynamic site index models for Crimean pine in the Kastamonu-Taşköprü region of Turkey. *Canadian Journal of Forest Research, 47*(11), 1441–1449.

Senilliani, M. G., Bruno, C., & Brassiolo, M. (2019). Site index for *Prosopis alba* plantations in the semi-arid Chaco through mixed models. *Cerne, 25*(2), 195–202.

Şenyurt, M., & Ercanlı, İ. (2013). Dynamic base-age invariant site index models for Scotch pine (*Pinus sylvestris* L.) in Kastamonu Forest Regional Directorate. International Caucasian Forestry Symposium, 179–183, Artvin.

Şenyurt, M. (2011). Increment and Growth in *Pinus sylvestris* Forest in West Black Sea Region. Phd Thesis, Istanbul University. (in Turkish)

Sharma, M., & Reid, D. E. B. (2018). Stand height/site index equations for jack pine and black spruce trees grown in natural stands. *Forest Science, 64*(1), 33–40.

Sharma, M., Subedi, N., Ter-Mikaelian, M., & Parton, J. (2015). Modeling climatic effects on stand height/site index of plantation-grown jack pine and black spruce trees. *Forest Science, 61*(1), 25–34.

Skovsgaard, J. P., & Vanclay, J. K. (2008). Forest site productivity: A review of the evolution of dendrometric concepts for even-aged stands. *Forestry, 81*(1), 13–31.

Trim, K. R., Coble, D. W., Weng, Y., Stovall, J. P., & Hung, I. K. (2020). A new site index model for intensively managed loblolly pine (*Pinus taeda*) plantations in the West Gulf Coastal Plain. *Forest Science, 66*(1), 2–13.

URL-1. (2021). https://kastamonuobm.ogm.gov.tr/Sayfalar/Kurulusumuz/GenelBilgiler.aspx. Accessed 15.11.2021.

Vanclay, J. K. (1994). *Modeling forest growth and yield. Applications to mixed tropical forests*. CAB International.

Wang, Y., LeMay, V. M., & Baker, T. G. (2007). Modelling and prediction of dominant height and site index of *Eucalyptus globulus* plantations using a nonlinear mixed-effects model approach. *Canadian Journal of Forest Research, 37*(8), 1390–1403.

Wang, M., Borders, B. E., & Zhao, D. (2008). An empirical comparison of two subject specific approaches to dominant heights modeling: The dummy variable method and mixed model method. *Forest Ecology and Management, 255*, 2659–2669.

Wang, M., Bhatti, J., Wang, Y., & Varem-Sanders, T. (2011). Examining the gain in model prediction accuracy using serial autocorrelation for dominant height prediction. *Forest Science, 57*(3), 241–251.

Wang, M., Kane, M. B., & Zhao, D. (2017). Correlation-regression analysis for understanding dominant height projection accuracy. *Forest Science, 63*(6), 549–558.

Westfall, J. A., Hatfield, A. A., Sowers, P. A., & O'Connell, B. M. (2017). Site index models for tree species in the Northeastern United States. *Forest Science, 63*(3), 283–290.

Xu, W. (2012). *Modelling jack pine (Pinus banksiana Lamb) and black spruce (Picea mariana (Mill.) BSP) growth and yield in Manitoba.* PhD Thesis, Faculty of Graduate Studies of The University of Manitoba. Manitoba.

Zhu, G., Hu, S., Chhin, S., Zhang, X., & He, P. (2019). Modelling site index of Chinese fir plantations using a random effects model across regional site types in Hunan province, China. *Forest Ecology and Management, 446*, 143–150.