



# Modeling forest site productivity using climate data and topographic imagery in *Pinus elliottii* plantations of central Argentina

Santiago Fiandino<sup>1,2</sup> · Jose Plevich<sup>1</sup> · Juan Tarico<sup>1</sup> · Marco Utello<sup>1</sup> · Marcela Demaestri<sup>1</sup> · Javier Gyenge<sup>2,3</sup>

Received: 1 July 2019 / Accepted: 22 September 2020  
© INRAE and Springer-Verlag France SAS, part of Springer Nature 2020

## Abstract

• **Key message** To be useful for silvicultural and forest management practices, the models of Site Index (SI) should be based on accessible predictor variables. In this study, we used spatially explicit data obtained from digital elevation models and climate data to develop SI prediction models with high local precision.

• **Context** Predicting tree growth and yield is a key component to sustainable forest management and depends on accurate measures of site quality.

• **Aims** The aim of this study was to develop both empirical models to predict site index (SI) from biophysical variables and a dynamic model of top height growth for plantations of *Pinus elliottii* Engelm. in Córdoba, Argentina.

• **Methods** Site productivity described by SI was related to environmental characteristics, including topographic and climatic variables. Separate models were created from only topographic data and the combination of topographic and climate data.

• **Results** Although SI can be adequately predicted through both types of models, the best results were obtained when combining topographic and climate variables ( $R^2 = 0.83$ ,  $RMSE\% = 7.02\%$ , for the best-fitting model). The key factors affecting site productivity were the landscape position and the mean precipitation of the last 5 years before the reference age, both related to the amount of plant-available water in the soils. Furthermore, the top height growth models developed are fairly accurate, considering the proportion of variance explained ( $R^2 = 98\%$ ) and the precision of the estimates ( $RMSE\% < 8\%$ ).

• **Conclusion** The models developed here are likely to have considerable application in forestry, since they are based on accessible predictor variables, which make them useful for silvicultural and forest management practices, particularly for non-forest areas and for the young or uneven-aged stands.

**Keywords** Site index · Biophysical factors · Digital elevation models · Prediction models · Forest productivity · Córdoba

**Handling Editor:** Jean-Michel Leban

**Contribution of co-authors** Jose Plevich: Contributed analysis tools.

Juan Tarico: Collected the data.

Marco Utello: Collected the data.

Marcela Demaestri: Contributed data.

Javier Gyenge: Participated in the design of the analysis and the writing of the paper.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13595-020-01006-3>) contains supplementary material, which is available to authorized users.

✉ Santiago Fiandino  
santifiandino@gmail.com

<sup>1</sup> Department of plant production, National University of Río Cuarto, Río Cuarto, X5804BYA Córdoba, Argentina

<sup>2</sup> CONICET, Godoy Cruz 2290, C142FQB CABA, Argentina

<sup>3</sup> AER Tandil, INTA EEA Balcarce, 7000 Buenos Aires, Argentina

## 1 Introduction

Predicting tree growth and yield is a key component to sustainable forest management and depends on accurate measures of site quality (Bontemps and Bouriaud 2013; Sharma et al. 2014). Biophysical factors (climate, edaphic, and physiographic factors) are important drivers of forest productivity; thus, many researchers attempted to establish quantitative relationships between site productivity and explicit biophysical indicators (Albert and Schmidt 2010; Sabatia and Burkhart 2014; Sharma et al. 2014; Paulo et al. 2015; Watt et al. 2015; Parresol et al. 2017; Brandl et al. 2018).

One of the most frequent methods for assessing site productivity is to develop stand height-age relationships to estimate the site index (SI), which refers to the dominant height of a stand at some specified reference age or base age, commonly selected to lie close to the average rotation age (Sharma et al.

2014; Paulo et al. 2015). Site index is a useful measurement of forest productivity as it integrates the combined effects of the most important determinants of tree growth and is relatively unaffected by stand density (Watt et al. 2015). Several investigations focused on the biophysical effects on site productivity and developed models to estimate SI from environmental variables (average annual temperature, annual rainfall, soil characteristics, and topography) (Bontemps and Bouriaud 2013; Sabatia and Burkhart 2014; Sharma et al. 2014; Paulo et al. 2015; Watt et al. 2015; Parresol et al. 2017).

Various methods of height modeling have been explored over the years. Traditional multiple linear/nonlinear regression models have proven to be adequate for predicting SI in many studies (Schafer 1988; Socha 2008). Authors such as Schafer (1988) developed fairly precise multivariate linear models for predicting SI in relation to site factors in *Pinus elliottii* stands, using effective soil depth, slope angle, terrain position, and rainfall as predictor variables. Although this traditional method is simple, its assumptions are usually not met. In this sense, mixed-effects models are particularly used when there is no independence in the data, such as those arising from a hierarchical structure (Lappi and Malinen 1994; Sharma et al. 2014). Breidenbach et al. (2007) used datasets from study sites in the USA and Germany, in which several sample plots were located within a stand and the stands were located within two study sites, resulting in hierarchical datasets. In the aforementioned study, the use of mixed-effects models significantly improved the estimates and especially reduced the bias which was present for numerous stands in the estimates of the fixed-effects models. Another approach that is been used increasingly in the last years is the Generalized Additive Models (GAMs). The principal advantages of GAM are as follows: (1) its ability to model highly complex nonlinear relationships and (2) no assumptions are made about the distributions of the predictor variables (Albert and Schmidt 2010; Brandl et al. 2018). In that sense, Albert and Schmidt (2010) fitted GAMs using large yield datasets as well as soil and climate variables from different data sources, to describe the site productivity relationships for *Picea abies* L. and *Fagus sylvatica* L. Although these authors indicate that the GAMs fitted may be improved, they were capable of predicting changes in site index under non-constant site conditions in a large and diverse geographic area.

Regardless of the statistical approach, in order to be useful for silvicultural and forest management practices, the models of SI should be based on accessible predictor variables (Socha 2008). Topography influences various environmental factors affecting tree growth (Hairston and Grigal 1991). Site topography influences (regional and local) climate, which together with geological substratum affects soil-forming processes, thus determining the types of soil (Socha 2008). In this respect, Hairston and Grigal (1991) found differences in the amount of available soil water among slope positions, which is linked with the surface soil hydraulic properties (SHP). SHP

also varies with landscape position and influences infiltration rates, determining plant-available water in the soil (Sauer et al. 2005), which is usually highly correlated with site productivity (Iverson et al. 1996).

Over the last two decades, topographic information has become more accessible for management of natural resources through the flourishing use of geographic information systems (GIS) and digital elevation models (DEMs) (De Reu et al. 2013). The most commonly used topographic indices are elevation, aspect, slope, and slope position (landscape position), and they all can be derived from DEMs (Hairston and Grigal 1991). For example, Coelho Pereira Neto and Limberger (2004) generated fairly precise slope maps of a representative area of the “Ribeirão Cafezal” basin in Brazil using DEMs. Another example is the use of the Topographic Position Index (TPI) to classify landscapes in slope position classes, which has been applied to several fields (De Reu et al. 2013).

Most of the models developed to describe slash pine dominant height growth are based on flexible height growth function as the Chapman-Richards or Richards models, since they have a form that combines dominant height growth and site index into one model form (Pienaar and Shiver 1984; Pienaar and Harrison 1989; Pienaar and Rheney 1993; Lee and Hong 1999; Fang and Bailey 2001). Traditional estimation techniques for dominant height projection equations involve the arbitrary choice of observed growth intervals and base ages. In order to be base-age invariant, the parameters of these models should be estimated using base-age-invariant techniques (BAI), which avoid the arbitrary choice of observed growth intervals and base ages (Cieszewski et al. 2000). BAI techniques deal with the problems of parameter dependence on the selected base ages, reducing the bias in the parameter estimation (Diéguez-Aranda et al. 2009).

Slash pine (*Pinus elliottii* Engelm.) plantations form a significant proportion of forest land in the province of Córdoba (Fiandino et al. 2018a), and they are mainly distributed in the hills and the piedmont (foothills) of the Sierra de Comechingones. The management of these pine plantations aims for the growing of sawtimber during a 30-year rotation, usually without applying thinning or pruning. Authors such as Chauchard et al. (2013) and Fiandino et al. (2018b) have developed state and transition models to predict the stand volume of these pine plantations, using stand density and top height growth as predictor variables. Nonetheless, there are no models of top height growth or site quality available for slash pine plantations in the region. For that reason, the objectives of the current study were two: (1) develop empirical models to predict site index from biophysical variables and (2) develop a dynamic model of top height growth (which includes SI) for plantations of slash pine. This approach will permit forest managers to exploit the increasing availability of spatially explicit data (created by remote sensing or digital soil mapping) by using SI prediction models with high local precision.

## 2 Methods

### 2.1 Study area

The investigations were carried out in the hills and the piedmont (foothills) of the southern sector of the Sierra de Comechingones, Córdoba, Argentina. Slash pine plantations are located in *Las Guindas* (LG) (32° 35' 16.8" S, 64° 43' 45.42" W, hills, 1000 m above sea level) and *Pozo del Carril* (PZ) (32° 58' 19.52" S, 64° 39' 4.86" W, foothills, 650 m above sea level) rural establishments, which belong to the National University of Rio Cuarto (Appendix 2). The climate of the region is monsoon type, with hot and wet summers and cold and dry winters. The average annual rainfall is higher in LG (1007 mm) than in PZ (760 mm), but its distribution is similar in both sites (with approximately 80% of the rainfall concentrated in spring–summer). The maximum, minimum, and mean annual temperatures are lower in LG (22.3 °C, 8.4 °C, and 14.7 °C, respectively) than in PZ (23 °C, 9.1 °C, and 16 °C, respectively).

The topography of LG site corresponds to a mountain environment, where areas of hills and slopes (50%), valleys and concavities (30%), and rocky outcrops (20%) are found. The areas of hills and slopes occupy a greater proportion of the surface, and it is on these areas where the pine plantations were established. The average slope in these units is around 10%, and the predominant soils are lithic, shallow (50 cm deep), excessively drained, and very stony Hapludols (Jarsún et al. 2006). On the other hand, the topography of the PZ site corresponds to a rolling plain environment, where the average slope is around 2% and the predominant soils are deep, well drained, and sandy Haplustolls (Plevich et al. 2019).

### 2.2 Field measurements

Field data were collected between 2015 and 2018 from 68 sample plots located in even-aged *P. elliotti* stands aged 15–26 years (Table 1). Of the sample plots, 18 were established in silvopastoral systems (SPS; 150–450 trees per hectare), nine in both LG and PZ (further details in Fiandino et al. (2018b) and Plevich et al. (2019)), and 50 were established in denser plantations (42 in LG and 8 in PZ; 600–1100 trees per hectare). The sample plots consisted in fixed rectangles of 300–450 m<sup>2</sup> randomly distributed across the pine plantations. A global positioning system (GPS) was used to record the coordinates at the center of each plot. Within each field plot, the stem diameter was measured at a height of 1.30 m (dbh) using a dendrometric tape. Then, the total height ( $H_0$ ) of the five largest trees (by dbh) of each plot (dominants) was measured using a Suunto hypsometer.

We selected 58 trees for stem analysis, 32 in LG and 26 in PZ. Some of the sampling plots in LG were also used in other research projects. For that reason, we could not fell any of its

**Table 1** Descriptive statistics for the inventory data and the biophysical variables associated with the research plot locations. The table includes only the variables that were most important predicting site index

Variable	Mean	Min	Max	SD
Age (years)	18.99	15.00	26.00	3.97
Mean square dbh (cm)	25.33	16.80	31.60	3.61
Top height (m)	16.12	9.75	25.00	3.32
Site index (m)	13.13	9.17	18.42	2.10
Stand volume (m <sup>3</sup> h <sup>-1</sup> )	261.58	45.58	729.97	152.17
Elevation (m)	896.60	636.50	1022.00	148.35
Slope (%)	11.37	1.28	25.79	6.77
Topographic Position Index (m)	-0.08	-2.88	4.13	1.08
Precipitation (0–16)	938.50	766.36	1007.73	100.64
Precipitation (12–16)	934.03	742.80	1060.29	122.93

dbh diameter at breast height

trees. From the plots where it was possible to fell trees, one or two sample trees per plot were selected for stem analysis. However, we tested the sample size and it offered sufficient statistical power (confidence level, 95%; maximum error, 6.25% of the mean SI). The selected trees were felled, and the logs crosscut to obtain discs of 5 cm thick, at 1.5-m sampling interval. The annual height growth was estimated using the tree annual radial growth (TARG) method (Kariuki 2002).

### 2.3 Auxiliary variables used in the modeling: climate and topographic data

Daily precipitation and minimum, mean, and maximum temperature data were obtained for each site for the years 1993–2018 from the weather stations placed in both LG and PZ. The precipitation and temperature data were processed to provide annual and seasonal climate variable figures for each year in the range 1993–2018. The 26-year climate variable normal for site was then computed as the arithmetic mean of the yearly values over the 26-year period. For every sample plot, the mean precipitation values for the period between planting and the reference age (0–16 years;  $P_{16}$ ) and for the last 5 years before reaching the reference age were calculated ( $P_5$ ). The reason for including  $P_5$  as a predictor variable is that in most of the unthinned slash pine plantations of the region, there is an increase in competition after age 10 and an increase in tree mortality after age 15 (Fernandez Corradi, Plevich, and Fiandino. UNRC—undergraduate thesis: “Forest management tables for stands of *Pinus elliottii* of the Sierra de Comechingones, Córdoba”. Unpublished data). Thus, it is reasonable to assume that deviations from the norm in the mean precipitation values during that period could have significant effects in height growth.

The elevation data of the hills and the piedmont of the southern sector of the Sierra de Comechingones were downloaded from the Argentina National Geographic Institute (IGN) Digital Elevation Model of Argentina (MDE-AR) available on the website <http://www.ign.gob.ar/category/tem%C3%A1tica/geodesia/mde-ar> (last accessed on 17th December 2018). The data were obtained in a grid format with a pixel size of 30 m. The elevation grid was cut according to the limits of each forest site. Mean values of metrics from the pixels that were spatially coincident with the plots were extracted for the analyses.

For this study, the elevation ( $E$ ) data was used to calculate topographic properties that might affect the amount of soil water available for tree growth, using QGIS software (QGIS Development Team 2018). The relative landscape position (LP) indicates the position of a sample plot in the toposquence of each forest site (LG or PZ) and was calculated using Eq. 1.

$$LP_{ij} = \frac{E_{ij} - E_{MINj}}{E_{MAXj} - E_{MINj}} \quad (1)$$

where  $LP_{ij}$  is the relative landscape position of the  $i$ th sample plot of the  $j$  site,  $E_{ij}$  is the elevation of the  $i$ th sample plot of the  $j$  site,  $E_{MINj}$  is the minimum elevation value of a sample plot in the  $j$  site, and  $E_{MAXj}$  is the maximum elevation value of a sample plot in the  $j$  site.

Position on the slope was also described through Topographic Position Index (TPI). TPI is the difference between the elevation at a cell and the average elevation in a neighborhood surrounding that cell. In the QGIS software (QGIS Development Team 2018), TPI is calculated as the difference between a central pixel and the mean of its surrounding cells (eight pixels). Positive values indicate that the cell is higher than its neighbors while negative values indicate the cell is lower. TPI values provide a simple and powerful means to classify the landscape into morphological classes (Tagil and Jenness 2008). Logically, high TPI values would be found near the tops of hills while low TPI values would be found in valley bottoms. TPI values near 0 would be found on either flat ground or somewhere mid-slope, and slope values are used to distinguish between these two possibilities. In the current study, TPI values were scaled to positive values (between 0 and 1) through Eq. 2.

$$STPI_i = \frac{TPI_i - TPI_{MIN}}{TPI_{MAX} - TPI_{MIN}} \quad (2)$$

where  $STPI_i$  is the Scaled Topographic Position Index of the  $i$ th sample plot,  $TPI_i$  is the Topographic Position Index of the  $i$ th sample plot,  $E_{MINj}$  is the minimum value of the Topographic Position Index of the sample plots, and  $E_{MAXj}$  is the maximum value of the Topographic Position Index of the sample plots.

Slope ( $S$ ) is the angle of inclination to the horizontal. It was calculated as percent slope using Horn method (Horn 1981), since it is suitable for rough surfaces.

Finally, we also explored the inclusion of stand density as a predictor variable, but it caused intercorrelation problems ( $VIF > 5$ ) and did not improve the models.

## 2.4 Data management and model development

The 68 observations that made up the sample plot data and the 58 observations that made up the stem analysis data were each divided into a model development (66% of the total number of observations) and a model validation (33%) data set. The validation observations were selected by random sampling in both cases.

For SI prediction, separate models were created from only topographic data, and the combination of topographic and climate data. Linear and nonlinear multiple regression techniques were used to develop the models relating SI to biophysical variables. The general formulation of the parametric models is as follows:

$$SI = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_n * X_n + \varepsilon \quad (3)$$

where SI is the site index (dominant height at a reference age of 16 years),  $X_1, \dots, X_n$  are the explanatory variables,  $\beta_1, \dots, \beta_n$  are the unknown parameters, and  $\varepsilon$  is the error term.

First, the variables that were most important predicting SI were identified through a correlation analysis. Then, these variables were introduced sequentially into the model, starting with the variable that exhibited the strongest correlation, until further additions were not significant, or did not substantially improve the model coefficient of determination and AIC. Variable selection was undertaken manually, one variable at a time (Watt et al. 2015). Before including the independent variable in the model, its redundancy was estimated by calculating its variance inflation factor (VIF). If  $VIF = 1$ , no intercorrelation exists between the independent variables; if VIF stays within the range 1–5, the corresponding model is acceptable; if  $VIF > 10$ , the corresponding model is unstable (Choubin et al. 2016). Since in our dataset several sample plots were located within two types of systems (silvopastoral systems and denser plantations) and the stands were located within two study sites (LG and PZ), the samples trees on each plot might not be independent. For that reason, mixed-effects models with a random intercept on the location and type of system were also fitted (Breidenbach et al. 2007).

For developing site index curves, data were derived from reconstructed height/age development patterns for individual trees obtained using stem analysis (from the 58 trees felled). Three dynamic models (Eqs. 4–6) have been considered for

developing the site index system for *P. elliottii* (Diéguez-Aranda et al. 2009).

$$H_0 = SI * \left( \frac{1 - \exp(-\alpha * t_0)}{1 - \exp(-\alpha * t_r)} \right)^\beta \quad (4)$$

$$H_0 = \frac{\alpha}{1 - (1 - \alpha / SI)^{*(t_r/t_0)^\beta}} \quad (5)$$

$$H_0 = \alpha * \left( SI / \alpha \right)^{(t_r/t_0)^\beta} \quad (6)$$

where  $H_0$  is the mean dominant height (m) at  $t_0$ ,  $t_0$  is the age of the stand (years), SI is the site index at the reference age  $t_r$  (16 years), and  $\alpha$  and  $\beta$  are the model parameters.

The three dynamic models (Eqs. 4–6) were fitted using both the traditional method and the iterative BAI technique proposed by Cieszewski et al. (2000). In this case, BAI technique did not improve model goodness of fit or precision. The model parameters stabilized after 8–9 interactions (Eq. 4, 8; Eq. 5, 9; Eq. 6, 8), and mean SI values changed by 0.15 m (less than 1.5% change). For that reason, we decided to present the traditional method results.

## 2.5 Statistical analysis

The goodness of fit and the prediction capacity of the models fitted were evaluated through the coefficient of determination ( $R^2$ ), which indicates goodness of fit; the root mean square error (RMSE), which analyzes the accuracy of the estimates; and the bias ( $\text{bias} = \frac{\sum(y_i - \hat{y}_i)}{n}$ , where  $y_i$  and  $\hat{y}_i$  are the measured and estimated values (Mayer and Butler 1993). For model comparison, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were also used.

For validation purposes, the goodness of fit and the prediction capacity of the models were evaluated through the coefficient of determination ( $R^2$ ), the mean absolute error (MAE), the observed vs. predicted plot (with the line  $y = \hat{y}$  marked), and the bias (Mayer and Butler 1993).

The adjustment of all the regression models was performed with Infostat statistical software (Di Rienzo et al. 2017). For linear regression, the parameters were estimated using the least squares method. Instead, the parameter estimation of nonlinear models was done using Nelder-Mead simplex algorithm (Nelder and Mead 1965).

## 3 Results

### 3.1 Inventory and biophysical data

The stand age averaged 19 years, ranging from 15 to 26 years (Table 1). Stand volume exhibited the greatest variation,

ranging from 45.58 m<sup>3</sup> ha<sup>-1</sup> in young low-density silvopastoral systems to 729.97 m<sup>3</sup> ha<sup>-1</sup> in older and denser stands. Mean square diameter at breast height, site index, and top height variations were similar, with coefficients of variation of 14%, 16%, and 20%, respectively. The variation in biophysical variables was relatively wide across the study site. Precipitation for the period between planting and the reference age ( $P_{16}$ ) averaged 938.5 mm ranging from 766.36 to 1007.73 mm, while precipitation for the last 5 years before reaching the reference age ( $P_5$ ) averaged 934.03 mm ranging from 742.8 to 1060.29 mm. The variation in precipitation for the 12–16 period was 23% higher than for the 0–16 period (Table 1). Elevation averaged 896.6 m.a.s.l., ranging from 636.5 m.a.s.l. in the foothills to 1022 m.a.s.l. in the hills (Table 1). The majority of the terrain over which measurements were made was relatively undulated with slope averaging 11.37%. The slope ranged from 1.28 to 25.79% (Table 1) indicating that some measurements were taken on flat sites, and some others on steeper sites. The mean value for Topographic Position Index was -0.08 m, ranging from -2.88 to 4.13 m.

### 3.2 Site index prediction models fitting and validation

Mixed-effects models did not improve models goodness of fit or precision. Therefore, the traditional fixed-effects models were preferred. SI model 3 (Eq. 7; Table 2) had the highest goodness of fit and prediction capacity among the models created using only metrics derived from topographic imagery (Supplementary material: Appendix 1 Table 1).

$$SI = \alpha + \beta * STPI^S + \gamma * LP^\lambda + \xi * E \quad (7)$$

where SI is the site index at the reference age of 16 years (m), STPI is the Scaled Topographic Position Index,  $S$  is the slope (1/100 fraction),  $E$  is the elevation (km), LP is the relative landscape position, and  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\lambda$ , and  $\xi$  are the model parameters.

SI model 11 (Eq. 8; Table 2) had the highest goodness of fit and prediction capacity among the models created using metrics derived from topographic imagery and climate variables (Supplementary material: Appendix 1 Table 2).

$$SI = \alpha + \beta * STPI^S + \gamma * LP^\lambda + \xi * P_5^\rho \quad (8)$$

where SI is the site index at the reference age 16 years (m),  $P_5$  is the mean precipitation for the last 5 years before reaching the reference age (m), LP is the relative landscape position, STPI is the Scaled Topographic Position Index,  $S$  is the slope (1/100 fraction), and  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\lambda$ ,  $\xi$ , and  $\rho$  are the model parameters.

In both cases, all model parameters were statistically significant ( $p < 0.05$ ). Site index is positively correlated with the mean annual precipitation of the last 5 years before reaching the reference age ( $P_5$ ;  $\beta$  in M11) and the elevation ( $E$ ;  $\gamma$  in M3)

**Table 2** Estimation and statistical significance of the parameters of the selected regression models

Model	Estimate	S.E.	P value	AIC	BIC	
[3] $SI = \alpha + \beta * STPI^S + \gamma * LP^\lambda + \xi * E$	$\alpha$	16.47	1.84	< 0.0001	149.5	160.3
	$\beta$	- 4.73	1.26	0.0006		
	$\gamma$	- 3.99	0.57	< 0.0001		
	$\lambda$	1.72	0.64	0.0095		
	$\xi$	2.36	1.12	0.0411		
[11] $SI = \alpha + \beta * STPI^S + \gamma * LP^\lambda + \xi * P_5^\rho$	$\alpha$	15.88	1.06	< 0.0001	126.4	139.0
	$\beta$	- 3.41	0.99	0.0014		
	$\gamma$	- 2.90	0.51	< 0.0001		
	$\lambda$	2.65	1.03	0.0139		
	$\xi$	1.15	0.41	0.0078		
	$\rho$	14.80	5.04	0.0056		

S.E. standard error, AIC Akaike information criteria, BIC Bayesian information criterion

and negatively with the relative landscape position (LP;  $\lambda$  in both M3 and M11) and the interaction effect of the Scaled Topographic Position Index (STPI) and the Slope ( $S$ ;  $\beta$  in M3 and  $\rho$  in M11).

There were substantial increases in precision for models that used both topographic properties and climate data, compared to those that used only topographic properties (Table 2). Addition of precipitation ( $P_5$ ) markedly improved the models with RMSE values reducing by 0.27 m (23%), and  $R^2$  increasing by 0.19. Mean prediction errors (RMSE and MAE) were similar for both modeling and validation datasets (Table 3). The magnitude of these errors (M3: RMSE% and MA%E < 10.5%; M11: RMSE% and MA%E < 8.4%) is half the variability found in the data from inventories (Table 1). A slight bias toward underestimation (bias < 3.00%) was identified when analyzing the results obtained from the validation dataset for both models (higher density of points below the  $y = \hat{y}$  line; Fig. 1). However, no bias was identified when analyzing the results obtained from the fitting dataset (Table 3).

### 3.3 Site index curves fitting and validation

The three proposed models adequately expressed the patterns of height growth. The percentage of total variation in  $H_0$  accounted for in the three models is 98%. Nevertheless, the modified version of “the Chapman-Richards” model (Eq. 4)

**Table 3** Goodness of fit and prediction capacity of the selected models

Model	RMSE (f)	MAE (v)	$R^2$ (f)	$R^2$ (v)	Bias (f)	Bias (v)
[3]	1.18	1.31	0.71	0.64	- 0.000067	- 0.32
[11]	0.91	1.08	0.83	0.81	0.00010	- 0.39

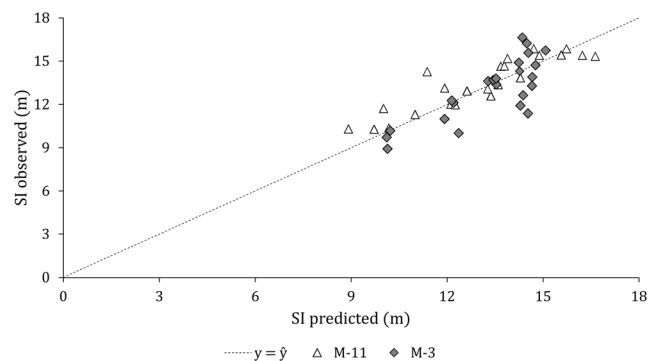
f fitting, v validation, RMSE root mean square error (m), MAE mean absolute error (m),  $R^2$  coefficient of determination

fitted the data best, since it presented the lowest RMSE, AIC, and BIC values (Tables 4 and 5), and its residual plot shows a fairly random pattern (Fig. 2). All model parameters were statistically significant ( $p < 0.005$ ; Table 4).

For all cases, mean prediction errors for modeling dataset (RMSE) were smaller (20–25%) than mean prediction errors (MAE) for validation dataset (Table 3). The RMSE% and MA%E for Eqs. 4, 5, and 6 were, respectively, 7.50/9.12%, 7.78/9.34%, and 7.78/9.79%. A slight bias toward overestimation (2.15–2.85%) was identified when analyzing the results obtained from the validation dataset for the three models. However, no bias was identified when analyzing the results obtained from the fitting dataset (Table 5). Finally, Fig. 3 shows the calculated site index curves for the range of site productivity in slash pine plantations of Southwestern Córdoba.

## 4 Discussion

In spite of that it is possible to determine slash pine site index based on topography features and climate data, there was



**Fig. 1** Predicted vs. observed site index (SI) of the validation dataset. M3, model 3; M11, Model 11. The dotted line represents  $y = \hat{y}$

**Table 4** Estimation and statistical significance of the parameters of the selected regression models

Model	Estimate		S.E.	P value	AIC	BIC
[Eq. 4] $H_0 = SI * \left( \frac{1 - \exp(-\alpha * t_0)}{1 - \exp(-\alpha * t_r)} \right)^\beta$	$\alpha$	0.045	0.003	< 0.0001	1454.9	1468.5
	$\beta$	1.134	0.025	< 0.0001		
[Eq. 5] $H_0 = \frac{\alpha}{1 - (1 - \alpha / IS)} * (t_r / t_0)^\beta$	$\alpha$	51.289	3.552	< 0.0001	1506.1	1519.7
	$\beta$	1.065	0.021	< 0.0001		
[Eq. 6] $H_0 = \alpha * (IS / \alpha)^{(t_r / t_0)^\beta}$	$\alpha$	600.58	210.74	0.0045	1506.7	1520.4
	$\beta$	0.205	0.018	< 0.0001		

S.E. standard error, AIC Akaike information criteria, BIC Bayesian information criterion

substantial variation in precision for SI models developed from only topographic data and the combination of topographic and climate data. Models that included only variables from topographic data were the least precise. Use of precipitation data for the last 5 years before reaching the reference age ( $P_5$ ) resulted in substantial gains in model precision. Nevertheless, models created using only topographic data provided reasonably precise estimates of SI. The precision range for these models ( $R^2 = 0.61-0.71$ ,  $RMSE\% = 9.01-11.63\%$ ; Supplementary material: Appendix 1 Table 1) was similar to that of other models of Site Index created from topographic data and soil properties (Chen et al. 2002; Socha 2008; Pinno et al. 2009; Pietrzykowski et al. 2015).

As with other studies, our results clearly show that forest productivity is directly influenced by changes in precipitation (Sharma et al. 2014). Models created using precipitation data provided a significant improvement in precision ( $R^2 = 0.70-0.83$ ,  $RMSE\% = 7.02-10.26\%$ ; Supplementary material: Appendix 1 Table 2). Under local conditions of the hills and the piedmont of the Sierra de Comechingones, the increase in elevation causes an increase in precipitation (Table 1). In this regard, elevation was used as a proxy of precipitation to improve the precision in the models that included only variables from topographic data (Eq. 7). However, as in most hilly environments, soil quality varies greatly among landscape positions, which affects soil water holding capacity (Hairston and Grigal 1991; Sauer et al. 2005; Socha 2008). Thus, plant-available water in the soil depends on both landscape position and precipitation.

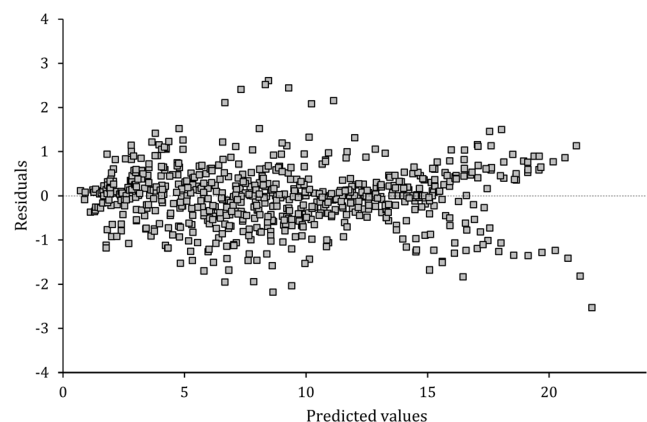
**Table 5** Goodness of fit and prediction capacity of the selected models

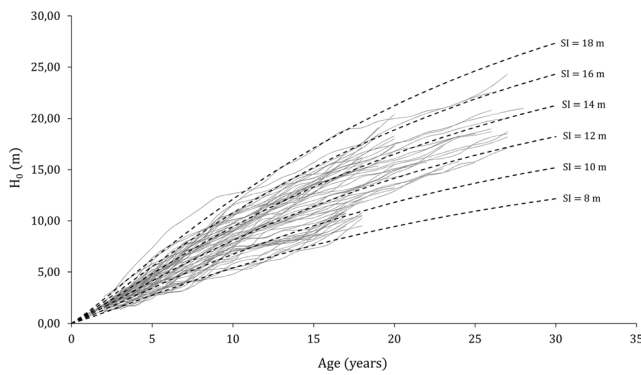
Model	RMSE (f)	MAE (v)	$R^2$ (f)	$R^2$ (v)	Bias (f)	Bias (v)
Eq. 4	0.68	0.87	0.98	0.98	-0.043	0.206
Eq. 5	0.71	0.89	0.98	0.98	-0.011	0.250
Eq. 6	0.71	0.93	0.98	0.97	0.001	0.278

f fitting, v validation, RMSE root mean square error (m), MAE mean absolute error,  $R^2$  coefficient of determination

In accordance with recent research, all the variables included in the models and the parameter values associated with each one suggest that water availability lays a key role for slash pine productivity (Brandl et al. 2014; Menéndez-Miguélez et al. 2015; Paulo et al. 2015). The three topographic variables included in the best fit model (landscape position (LP), slope (S), and Topographic Position Index (TPI)) are related to soil hydraulic properties and infiltration rates (Sauer et al. 2005). Decreases in LP and TPI are associated with increases in SI, but the influence of LP on SI prediction was greater. Nevertheless, including the interaction between TPI and S resulted in an increase in the explained variance, despite SI not affected by S when it was treated as one of the explanatory variables included in the model in the additive form. The observed relationship of site index with the interaction between TPI and S is probably because both variables (TPI and S) bring complementary information in respect to the landforms (Weiss 2001).

Before now in Córdoba, there has been no development of dominant height growth models for slash pine; hence, direct comparison of results is not possible. Nevertheless, the obtained results seem accurate with any of the equations tested, considering the proportion of variance explained ( $R^2 = 98\%$ ) and the precision of the estimates ( $RMSE\% < 8\%$ ) and compared to the results obtained by other authors who used similar

**Fig. 2** Relationship between residuals and predicted values of the site index model (Eq. 4)



**Fig. 3** Dominant height ( $H_0$ )/Site Index (SI) profiles generated using Eq. 4 for slash pine trees in Southwestern Córdoba for the range of site productivity. Continuous lines correspond to the values obtained for each tree in the stem analysis

approaches (Fang and Bailey 2001; Diéguez-Aranda et al. 2009). Furthermore, these dynamic site models permit to directly estimate height and site index from any other height and age and are base-age invariant (Cieszewski et al. 2000). Therefore, by using these equations, it is possible to adequately predict height growth from climate and topographic data, since dominant height growth and site index are explicitly combined in the site equations. In addition, these site equations can be used to predict the stand volume of silvopastoral systems using the state and transition models developed by Fiandino et al. (2018b) for the same region.

## 5 Conclusion

The models developed here are likely to have considerable application in forestry. Our results clearly show that Site Index can be accurately predicted from topographic and climate data. As an alternative to climate data, the model that used only topographic variables derived from a digital elevation model was fairly precise and unbiased. The main advantage of these models is that they are based on accessible predictor variables, which makes them useful for silvicultural and forest management practices, particularly for non-forest areas and for the young or uneven-aged stands.

The results obtained in the present study provide further knowledge related to the ecology of slash pine stands in southwest Córdoba. According to the regression models, precipitation data for the last 5 years before reaching the reference age and landscape position were the most important factors in explaining site productivity. These results suggest that water availability lays a key role for slash pine productivity in the region.

Finally, we propose the modified version of “the Chapman-Richards” model (Eq. 4) fitted in this study for predicting the height growth of slash pine stands within the range of altitude (600–1100 m.a.s.l) and mean precipitation (750–1050 mm) of

the hills and piedmont of southwest Córdoba. The main advantage of this dynamic model is that it permits to estimate directly height and site index from any other height and age and is base-age invariant.

**Acknowledgements** The authors thank the FAV-UNRC Agrometeorology research group for providing the meteorological data of the study area, and Mr. Ignacio Fernandez Corradi and Mr. Franco Banchemo for their help in the forest inventories. We also thank to the people in charge of “Las Guindas” and “Pozo del Carril” rural establishments for their help and support. Finally, we thank Dr. Fernando Casanoves for his advice on modeling methods.

**Funding** This research was funded through a PPI (SECyT-UNRC) and CONICET (postdoctoral position).

**Data availability** The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Albert M, Schmidt M (2010) Climate-sensitive modelling of site-productivity relationships for Norway spruce (*Picea abies* (L.) Karst.) and common beech (*Fagus sylvatica* L.). For Ecol Manag 259:739–749
- Brandl S, Mette T, Falk W, Vallet P, Rötzer T, Pretzsch H (2018) Static site indices from different national forest inventories: harmonization and prediction from site conditions. Ann For Sci 75:56
- Bontemps JD, Bouriaud O (2013) Predictive approaches to forest site productivity: recent trends, challenges and future perspectives. Forestry 87:109–128
- Brandl S, Falk W, Klemmt HJ, Stricker G, Bender A, Rötzer T, Pretzsch H (2014) Possibilities and limitations of spatially explicit site index modelling for spruce based on National Forest Inventory data and digital maps of soil and climate in Bavaria (SE Germany). Forests 5: 2626–2646
- Breidenbach J, McGaughey RJ, Andersen HE, Kändler G, Reutebach SE (2007) A mixed effects model to estimate stand volume by means of small footprint airborne lidar data for an American and a German study site. In Proceedings of ISPRS workshop laser scanning, pp 12–14
- Chauchard L, Andenmatten E, Letourneau F (2013) Modelización del crecimiento y la producción de los rodales a través de Índices de Densidad. In Aplicaciones de modelos ecológicos a la gestión de recursos naturales, 1ra ed. OmniaScience Monographs, pp 109–124.
- Chen HYH, Krestov PV, Klinka K (2002) Trembling aspen site index in relation to environmental measures of site quality at two spatial scales. Can J For Res 32:102–119
- Choubin B, Khalighi-Sigaroodi S, Malekian A, Kisi Ö (2016) Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals. Hydrol Sci J 61:1001–1009
- Cieszewski CJ, Harrison M, Martin SW (2000) Practical methods for estimating non-biased parameters in self-referencing growth and yield models. University of Georgia PMRC-TR 7



- Coelho Pereira Neto O, Limberger L (2004) Análisis de la precisión de mapas de pendientes utilizando modelos digitales de terreno. *Revista Fac Agron Univ Nac La Plata* 105:97–103
- De Reu J, Bourgeois J, Bats M, Zwertvaegher A, Gelorini V, De Smedt P et al (2013) Application of the topographic position index to heterogeneous landscapes. *Geomorphology* 186:39–49
- Di Rienzo JA, Casanoves F, Balzarini MG, Gonzalez L, Tablada M, Robledo CW (2017) InfoStat versión 2017. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. URL <http://www.infostat.com.ar>
- Diéguez-Aranda U, Rojo Alboreca A, Castedo-Dorado FF, Álvarez Gonzalez JG, Barrio-Anta M, Crecente-Campo JM et al (2009) Herramientas selvícolas para la gestión forestal sostenible en Galicia. Xunta de Galicia, Santiago de Compostela ISBN: 978-84-692-7395-1
- Fang Z, Bailey RL (2001) Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *For Sci* 47:287–300
- Fiandino SI, Plevich JO, Tarico JC, Nuñez C, Rusch V, Gyenge JE (2018a) Effects of low-density *Pinus elliottii* (Slash Pine) afforestation on environmental conditions and native plant diversity, in the mountains of central Argentina. *Appl Veg Sci* 21:442–450
- Fiandino SI, Plevich JO, Tarico JC, Utello M, Gyenge J (2018b) Prediction of post-thinning stem volume in slash pine stands by means of state and transition models. *New For* 50:1–14. <https://doi.org/10.1007/s11056-018-9688-7>
- Hairton AB, Grigal DF (1991) Topographic influences on soils and trees within single mapping units on a sandy outwash landscape. *For Ecol Manag* 43:35–45
- Horn BK (1981) Hill shading and the reflectance map. *Proc IEEE* 69:14–47
- Iverson LR, Scott CT, Dale ME, Prasad A (1996) Development of an integrated moisture index for predicting species composition. In: Kohl M, Gertner GZ (eds) *Caring for the Forest: Research in a Changing World*. Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, pp 101–116
- Jarsún B, Gorgas J, Zamora A, Bosnero H, Lovera E, Ravelo A, Tassile JL (2006) Recursos naturales de la Provincia de Córdoba: Los Suelos. Agencia Córdoba Ambiente, Córdoba
- Kariuki M (2002) Height estimation in complete stem analysis using annual radial growth measurements. *Forestry* 75:63–74
- Lappi J, Malinen J (1994) Random-parameter height/age models when stand parameters and stand age are correlated. *For Sci* 40:715–731
- Lee YJ, Hong SC (1999) Estimation of site index curves for loblolly pine (*Pinus taeda* L.) and slash pine (*Pinus elliottii* Engelm.) Plantations. *J Korean For Soc* 88:285–291
- Mayer DG, Butler DG (1993) Statistical validation. *Ecol Model* 68:21–32
- Menéndez-Miguélez M, Álvarez-Álvarez P, Majada J, Canga E (2015) Effects of soil nutrients and environmental factors on site productivity in *Castanea sativa* Mill. coppice stands in NW Spain. *New For* 46:217–233
- Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput J* 7:308–313
- Paulo JA, Palma JH, Gomes AA, Faias SP, Tomé J, Tomé M (2015) Predicting site index from climate and soil variables for cork oak (*Quercus suber* L.) stands in Portugal. *New For* 46:293–307
- Parresol BR, Scott DA, Zarnoch SJ, Edwards LA, Blake JI (2017) Modeling forest site productivity using mapped geospatial attributes within a South Carolina Landscape, USA. *For Ecol Manag* 406:196–207
- Pienaar LV, Shiver BD (1984) The effect of planting density on dominant height in unthinned slash pine plantations. *For Sci* 30:1059–1066
- Pienaar LV, Harrison WM (1989) Simultaneous growth and yield prediction equations for *Pinus elliottii* plantations in Zululand. *South Afr For J* 149:48–53
- Pienaar LV, Rheney JW (1993) Yield prediction for mechanically site-prepared slash pine plantations in the southeastern coastal plain. *South J Appl For* 17:163–173
- Pietrzykowski M, Socha J, van Doorn NS (2015) Scots pine (*Pinus sylvestris* L.) site index in relation to physico-chemical and biological properties in reclaimed mine soils. *New For* 46:247–266
- Pinno BD, Paré D, Guindon L, Bélanger N (2009) Predicting productivity of trembling aspen in the Boreal Shield ecozone of Quebec using different sources of soil and site information. *For Ecol Manag* 257:782–789
- Plevich JO, Gyenge J, Delgado AS, Tarico JC, Fiandino SI, Utello MJ (2019) Production of fodder in a treeless system and in silvopastoral System in Central Argentina. *Floresta Ambient* 26. <https://doi.org/10.1590/2179-8087.051716>
- QGIS Development Team (2018) QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>
- Sabatia CO, Burkhart HE (2014) Predicting site index of plantation loblolly pine from biophysical variables. *For Ecol Manag* 326:142–156
- Sauer TJ, Logsdon SD, Van Brahana J, Murdoch JF (2005) Variation in infiltration with landscape position: implications for forest productivity and surface water quality. *For Ecol Manag* 220:118–127
- Schafer GN (1988) A site growth model for *Pinus elliottii* in the Southern Cape. *South Afr For J* 146:12–17
- Sharma M, Subedi N, Ter-Mikaelian M, Parton J (2014) Modeling climatic effects on stand height/site index of plantation-grown jack pine and black spruce trees. *For Sci* 61:25–34
- Socha J (2008) Effect of topography and geology on the site index of *Picea abies* in the West Carpathian, Poland. *Scand J For Res* 23:203–213
- Tagil S, Jenness J (2008) GIS-based automated landform classification and topographic, landcover and geologic attributes of landforms around the Yazoren Polje, Turkey. *J Appl Sci* 8:910–921
- Watt MS, Dash JP, Bhandari S, Watt P (2015) Comparing parametric and non-parametric methods of predicting Site Index for radiata pine using combinations of data derived from environmental surfaces, satellite imagery and airborne laser scanning. *For Ecol Manag* 357:1–9
- Weiss A (2001) Topographic position and landforms analysis. In Poster presentation, ESRI user conference, San Diego, CA

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.