Regional height growth models for Scots pine in Poland

Jarosław Socha¹*, Luiza Tymińska-Czabańska¹, Karol Bronisz², Stanisław Zięba¹ & Paweł Hawryło²

Site productivity remains a fundamental concern in forestry as a significant driver of resource availability for tree growth. The site index (SI) reflects the overall impact of all environmental factors that determine tree height growth and is the most commonly used indirect proxy for forest site productivity estimated using stand age and height. The SI concept challenges are local variations in climate, soil, and genotype-environmental interactions that lead to variable height growth patterns among ecoregions and cause inappropriate estimation of site productivity. Developing regional models allow us to determine forest growth and SI more appropriately. This study aimed to develop height growth models for the Scots pine in Poland, considering the natural forest region effect. For height growth modelling, we used the growth trajectory data of 855 sample trees, representing the Scots pine entire range of geographic locations and site conditions in Poland. We compared the development of regional height growth models using nonlinear-fixed-effects (NFE) and nonlinear-mixed-effects (NME) modelling approaches. Our results indicate a slightly better fit to the data of the model built using NFE approach. The results showed significant differences between Scots pine growth in natural forest regions I, II, and III located in northern Poland and natural forest regions IV, V, and VI in southern Poland. We compared the development of regional height growth models using NFE and NME modelling approaches. Our results indicate a slightly better fit to the data of the model built using the NFE approach. The developed models show differences in height growth patterns of Scots pines in Poland and revealed that acknowledgement of region as the independent variable could improve the growth prediction and quality of the SI estimation. Differences in climate and soil conditions that distinguish natural forest regions affect Scots pine height growth patterns. Therefore, extending this research to models that directly describe height growth interactions with site variables, such as climate, soil properties, and topography, can provide valuable forest management information.

Site productivity remains a fundamental concern in forestry as a significant driver of resource availability. Appropriate estimation of site productivity is critical for forest management decisions. Additionally, information on site productivity helps monitor and assess climate change impact on forest ecosystems and provide information to support forest management to take effective adaptation measures. The site index (SI) reflects the overall impact of all environmental factors that determine tree height growth and is the most commonly used indirect proxy for forest site productivity estimated using stand age and height. An appropriate height growth model is needed to determine the SI accurately.

One of the most critical challenges in the SI concept is a local variation of height growth patterns that result from a variety of factors and could cause the under- or over-estimation of the growth potential and productivity of a site. Species variability caused by sensitivity to environmental factors is the primary reason why a more flexible approach to developing height growth models is recommended to include the impact of local conditions and to accurately reflect height growth variation on a regional scale.

Climate, soil, and genotype-environmental interactions can favour species adaptation to local ecological conditions, causing variable height growth patterns for the same species among ecoregions. In a study conducted in Sweden, Johansson found significant variation in forest height growth curves due to mineral soil type and geographical location. The effect of genetic variability on the height development of trees was reported by Adams et al. Genetic variability has also been identified as the primary factor modifying the trajectory of height growth by Buford and Burkhart, who found that the height growth models developed for different provenances differed in their parameters describing the asymptote. Significant differences in height growth...
patterns can also occur for the same species, even within a given SI class. García reported that three sources of height growth development variability could be identified: (a) between sites, (b) within sites, and (c) observation error. Because of the inter-regional variability of growth patterns, regional height growth models (RMs) are derived to improve the model's predictive ability. Studies on Pinus pinaster in Spain indicated significant differences in growth patterns observed in different areas because of the effect of genotype-environment interaction. Research describing regional height growth differences highlights the regional application of the SI concept. The role of regional models may be crucial for sustainable forest management, particularly in solving problems that require accurate information on local growing conditions, where climate and soil type lead to inter-regional variation.

The Scots pine is one of the most important species in Europe and ranges from the boreal region in Northern and Eastern Europe to the mountains of the Mediterranean in Southeastern Europe (Fig. 1). It is also the most important tree species in Poland, covering approximately 58% of the total forested area of the country. The Scots pine grows under very variable ecological conditions, which could affect the difference in the shape of the growth curves among ecoregions. The effective management of such a considerable forest resource requires a thorough assessment of the productivity of stands at a regional level. Growth modelling is crucial in assessing the consequences of forest management activities in long term planning in forestry. The use of inappropriate height growth models may result in erroneous estimation of site productivity. This problem could be solved by the development of models, which reflect the difference in growth trajectories resulting from specific growth conditions in a given area. Regional models reflect the impact of regional variability on height growth dynamics resulting from specific climate conditions, soil properties, and differences in local forest management.

In Poland, yield tables with discrete scales are still used in forest management for forest growth and site productivity estimation. The forestry practice reports the incorrect determination of site productivity and incremented volume using yield tables, probably because of the inadequacy of Schwappach's tables from differences in growth conditions, which have changed considerably during the 100 years since the yield tables were developed. Pretzsch et al. estimated that in Central Europe, site productivity increased by approximately 30–70%. Additionally, the height growth of observed stands differs significantly from that observed 100 years ago and included in the original tables. Moreover, the tables produce errors by not considering regional growth variability on a countrywide scale.

The research adopted the hypothesis that height growth is regionally differentiated. We assumed that height growth patterns are influenced by the variability in site conditions, which can be acknowledged by considering the specific effect of natural forest regions (NFE), distinguished by environmental conditions. Therefore, the research objective was to analyse the variability of height growth patterns between natural forest regions in Poland and develop height growth models for the Scots pine in Poland taking into account the specific effect of the region.

**Materials and methods**

**Collection of the research material.** The research material consists of the height growth trajectories of 855 dominant and codominant Scots pines (Table 1). The most substantial part of growth trajectories was collected from the sample plots established within the project "Actual and potential site productivity in Poland for..."
main forest-forming tree species\textsuperscript{3} from 2015 to 2016 and representing the whole variability of growth conditions in Poland for the Scots pine trees. Other tree growth trajectories have been collected in Poland over the last 2 decades as part of research projects conducted by the Department of Forest Resources Management, University of Agriculture, Krakow, the Forest Research Institute in Sękocin Stary, and the Faculty of Forestry, Warsaw University of Life Sciences. The final set of sample plots represent the area of Poland, with a whole range of geographic distributions and site conditions (Fig. 1). The age of the analysed sample trees ranged from 12 to 175 years, with an average of 76 years and heights ranging from 3.50 to 36.75 m (Table 1).

Natural forest regions\textsuperscript{37} were adopted as the regionalisation unit in Poland. Natural forest regions were distinguished by the natural diversity of the country, in particular: variability of climatic and geological conditions, naturally occurring ranges of principal forest-forming tree species, presence and distribution of natural landscapes classes, and distribution of the primary units of potential natural vegetation\textsuperscript{37}.

**Stem analysis.** Stem analysis was used to reconstruct the past height growth of the trees. The trees for stem analysis were selected in the direct neighbourhood of the permanent sample plots established in Scots pine stands. For each plot, tree diameter and height close to the top height (TH)—defined as the average height of the 100 thickest trees per 1 ha—were selected and cut for stem analysis (SA). Only trees with typically developed crowns and no visible damage or growth anomalies were selected for SA. After cutting, the total lengths of the trees were measured, and discs were collected for SA. Discs were taken from the base of the felled tree at the height of 0.5 m, and the breast height diameter (DBH) was measured; for trees over 15 m high, subsequent discs were collected at heights of 2.0, 4.0, and successive 2-m intervals to the top of the tree. For trees shorter than 15 m, subsequent discs were collected at heights of 1.5, 2.5, 3.5, and at 1-m intervals further up the tree. The course of growth of the individual trees was reconstructed based on the height of the discs and the number of annual discs. The growth trajectories of all collected trees were visually examined for suppression and release patterns and growth anomalies. As cross-section lengths do not coincide with periodic height growth, the height-age data from the SA were corrected using Carmean's method\textsuperscript{38}. The collected growth series for six natural forest regions were used to develop regional height growth models for Scots pine.

**Models and parameters estimation.** To develop regional height growth models: nonlinear fixed-effects and nonlinear mixed-effects modelling approaches were used.

A **nonlinear mixed-effects modelling approach.** The growth series are characterised by a hierarchical structure, which contains information about individual trees and natural forest regions. The mixed-effects modelling approach is a possible solution to handle such data\textsuperscript{39–42}. To develop the height growth model using a mixed-effects modelling approach, we tested three functions well known in forest growth modelling: Chapman–Richards, Schumacher and Hossfeld:

\[
\text{Chapman-Richards: } H = a_1 \left(1 - e^{a_2 t}\right)^{a_3}
\]

\[
\text{Schumacher: } H = e^{a_1 + a_2 t^{a_3}}
\]

\[
\text{Hossfeld: } H = \left(\frac{a_1 t^{a_2}}{a_3 + t^{a_3}}\right)^{1/3}
\]

where $H$ is the top height (m), $t$ is an age, and $a_1$–$a_3$ are estimated parameters.

Given the value of the mean absolute error (MAE), adjusted coefficient of determination ($R_{adj}^2$), and the Akaike information criterion (AIC), the possibility of defining different random parameters was achieved, with a maximum of all (three) parameters found in the tested functions on both: trees and natural forest regions at grouping levels. In the case of heteroscedastic data, analysing likelihood ratio test and solutions containing powertype variance function with age as a predictor were applied\textsuperscript{38}. Furthermore, given within-tree autocorrelation, the autoregressive correlation structure of order 1—AR(1)\textsuperscript{44} and autoregressive–moving average models for the

| Natural forest region | Number of growth series | Age (years) | Height (m) |
|----------------------|-------------------------|-------------|------------|
|                      |                         | Mean       | Min        | Max        |
| I                    | 146                     | 74         | 20         | 140        | 21.2       | 13.2       | 33.4       |
| II                   | 125                     | 88         | 12         | 175        | 28.2       | 7.1        | 36.7       |
| III                  | 174                     | 75         | 14         | 175        | 20.6       | 8.2        | 33.8       |
| IV                   | 60                      | 68         | 17         | 112        | 23.9       | 9.9        | 33.2       |
| V                    | 39                      | 63         | 12         | 145        | 21.8       | 3.5        | 33.1       |
| VI                   | 311                     | 75         | 23         | 141        | 23.3       | 11.0       | 35.6       |

Table 1. The number and basic characteristics of growth series for the Scots pine trees, obtained from stem analysis and used for the development of the site index models.
Development of the models using nonlinear fixed-effect modelling approach with the functions developed using the algebraic difference approach. The algebraic difference approach (ADA) formulated by Bailey and Clutter allows for the derivation of models in which the SI depends only on one parameter. Therefore it allows obtaining polymorphic models with single asymptotes or anamorphic models with multiple asymptotes. The generalised algebraic difference approach (GADA) developed by Cieszewski and Bailey allows the use of multiple site parameters and models characterised by polymorphism and variable asymptotes for different sites. Height growth models should also show equality of the SI and height at base age and use the same function as the height growth and SI models.

After initial preselection, three ADA models (M2, M4 and M5) and two GADA models (M1, M3) were chosen for further analysis (Table 2). These models meet the criteria mentioned above and have been successfully applied in SI modelling, making them potential candidates for reliable SI models.

Local (site-specific) parameters of individual growth series and global parameters of SI models (M1-M5) were simultaneously fitted using the dummy variable approach (DVA), also known as the varying parameter method. In the DVA, global and site-specific parameters are simultaneously estimated, resulting in the same parameter estimates regardless of the selected base age; therefore, it is a base-age invariant method and produces a unique site parameter for each tree. Due to the longitudinal nature of growth series data, subsequent height estimation, an expression containing simulated variables and a site-specific parameter allows the calculation of height at age 0, and $a_1$, $a_2$, $a_3$, $a_4$, $a_5$, $b_1$, $b_2$, and $b_3$ are model parameters.

The potential improvement of height growth prediction was analysed by expanding the height growth model equations tested in height growth model development. Where $H$ is the tree height, $S$ site index, $X$ and $X_0$ theoretical variable representing site conditions, $T$ tree age, $H_i$ is the measured height at age $T_i$, $H_0$ is a site parameter denoting stand height at age $T_0$, and $a_1$, $a_2$, $a_3$, $a_4$, $a_5$, $b_1$, $b_2$, and $b_3$ are model parameters.

within-group errors was analysed using `corAR1` and `corARMA(1,0)` function from `nlme` R package respectively with RStudio software.

Table 2. The dynamic generalised algebraic difference approach (GADA) and algebraic difference approach (ADA) formulation of growth functions tested in height growth model development. Where $H$ is the tree height, $S$ site index, $X$ and $X_0$ theoretical variable representing site conditions, $T$ tree age, $H_i$ is the measured height at age $T_i$, $H_0$ is a site parameter denoting stand height at age $T_0$, and $a_1$, $a_2$, $a_3$, $a_4$, $a_5$, $b_1$, $b_2$, and $b_3$ are model parameters.

| Model no | Base model forms | Parameter related to site | Solution for theoretical variable $X$ | Dynamic ADA or GADA formulation and reference |
|----------|------------------|--------------------------|----------------------------------------|---------------------------------------------|
| M1       | $H = \frac{a_3X^3}{1+e^{b_1+b_2X}}$ | $b = b_1 + X$ | $a = a_1/X$ | $R = Z_0 + \frac{Z_1}{t_0} + \frac{\log(b_1)}{t_0}$ | $H_i = H_0 + \frac{\log(b_1)}{t_0}$ |
| M2       | $H = a_1 \left[1 - \exp\left(-a_2 T\right)\right]$ | $a_1 = X$ | $X_0 = H_0 / (1 - \exp(-b_2 T_0))$ | $H_i = H_0 \left(\frac{\exp(-b_2 T_i)}{\exp(-b_2 T_0)}\right)$ |
| M3       | $H = a_1 \left[1 - \exp\left(-a_2 T\right)\right]$ | $a_1 = \exp(X)$ | $a_2 = b_1, b_2/X$ | $X_0 = \frac{1}{2} \ln Y_0 - b_2 L_0 + \sqrt{(\ln Y_0 - b_2 L_0)^2 - 4 b_2 L_0}$ | $H_i = H_0 \left(\frac{\exp(-b_2 T_i)}{\exp(-b_2 T_0)}\right)$ |
| M4       | $H = \frac{X_0}{\frac{X_0}{X_0} + \frac{X_0}{X_0}}$ | $a_2 = X$ | $b_i = X_0 - b_2 T_0$ | $H_i = \left(\frac{X_0}{X_0} + \frac{X_0}{X_0}\right)$ |
| M5       | $H = a_1 \exp\left(-a_2 T^{a_3}\right)$ | $a_2 = X$ | $X_0 = -\ln\left(\frac{b_3}{b_2}\right) T_0^{a_3}$ | $H_i = b_1 \left(\frac{b_3}{b_2}\right) T_i^{a_3}$ |
were used to simultaneous fitting the expansion terms specific for each ARMA(10,0) solution allows to reach the best goodness-of-fit measures (Table 4).

Within-group errors ARMA(10,0) has significant influence on model fit. Alike autoregressive correlation structure of order 1—AR(1) autoregressive–moving average models for the well. Likewise, the likelihood test indicated that the variance function has a nonsignificant influence on model fit. The range (0.359–0.710, Table 3) of the mean absolute error indicates specific for the given jth region (\(R_j\)) and ith global parameter (\(i\)). Six dummy variables (\(R_j\)) representing regions were used to simultaneous fitting the expansion terms specific for each ith parameter (\(b_i\)) and jth region.

The significance of the expansion terms was tested using t-statistics and was assumed as a statistical test for the difference between a given global parameter \(b_i\) and the regional parameter values. Regional parameter values are the sum of the global parameters \(b_i\) and expansion term \((\delta_{iR_j})\). Moreover, to compare the fits of the full (regional) and reduced (global) model, we used the anova() function with the regression objects as two separate arguments. The anova() function takes model objects as arguments and returns an ANOVA to check that a more complex model is much better at capturing data than a simpler model

The full model trajectories were directly compared by plotting the differences between observed and model trajectories. Parameter estimation was carried out in R using the procedures and suitably defined model form.

### Results

**Nonlinear mixed-effects models.** The range (0.359–0.710, Table 3) of the mean absolute error indicates that the application for the mixed-effects modelling approach when elaborating the height growth models allows obtaining high accuracy. The obtained results indicate that a better fit characterises the functions with defined two random parameters than one random parameter. The Chapman–Richards function with parameters \(a\) and \(b\), which refer respectively to the asymptote and the growth rate, defined as random effects, was characterised by two random parameters than one random parameter. The Chapman–Richards function with parameters \(a\) and \(b\), which refer respectively to the asymptote and the growth rate, defined as random effects, was characterised by the best fit (Table 3). Note the shape of the function is expressed as a parameter \(c\) is estimated as a fixed-effect.

Moreover, analysis of the residuals for the best fitted Chapman–Richards model (estimated parameters shown in Table 4) indicated heteroscedasticity (Fig. 4). The assumed variance function did not model heteroscedasticity well. Likewise, the likelihood test indicated that the variance function has a nonsignificant influence on model fit. Alike autoregressive correlation structure of order 1—AR(1) autoregressive–moving average models for the within-group errors ARMA(10,0) has significant influence on model fit (\(p < 0.0001\) for likelihood test). Whereas ARMA(10,0) solution allows to reach the best goodness-of-fit measures (Table 4).

**Nonlinear fixed-effects models.** All preselected functions showed a good fit to the data and explained (in most cases) close to 99% (adjusted \(R^2\)) of the height growth variation (Table 5). The error structured modelling with the 10-order autoregressive error was used in parameter estimation. All developed models prediction errors were relatively low, with root mean absolute error (MAE) varying from 0.59 to 0.69 m.

Considering the largest part of the explained variance, the smallest average absolute error and the AIC value, the function M1 was selected as best fitted to the data. Next function M1 was used to develop the model taking into account the regional variability. For fitting of the full model, taking into account the specific effect of region, the parameters of the model M1 were expanded as a function of regions (Eq. 6).

| Model name                  | Model type               | \(R^2_{adj}\) | AIC  | MAE   |
|-----------------------------|--------------------------|---------------|------|-------|
| Chapman–Richards            | \(H = a_i(1 - e^{b_i})^c\) | 0.985         | 43.037 | 0.710 |
|                             | \(H = a_i(1 - e^{b_i})^c\) | 0.988         | 40.323 | 0.643 |
|                             | \(H = a_i(1 - e^{b_i})^c\) | 0.995         | 32.205 | 0.419 |
|                             | \(H = a_i(1 - e^{b_i})^c\) | 0.996         | 28.711 | 0.359 |
|                             | \(H = a_i(1 - e^{b_i})^c\) | 0.995         | 31.589 | 0.406 |

Table 3. Adjusted coefficient of determination (\(R^2_{adj}\)), the Akaike information criterion (AIC) and mean absolute error (MAE) for Chapman–Richards function with different random parameters on both: trees and natural forest regions as grouping levels (random effects are expressed with index \(i\)).

| Parameter                  | Chapman–Richards | Chapman–Richards AR(1) | Chapman–Richards ARMA(10,0) |
|----------------------------|------------------|-------------------------|-----------------------------|
| a                          | 33.838 (1.051)   | 33.786 (1.025)          | 33.905 (1.114)              |
| b                          | -0.023 (0.001)   | -0.023 (0.001)          | -0.022 (0.001)              |
| c                          | 1.432 (0.004)    | 1.414 (0.007)           | 1.408 (0.008)               |
| Natural region level       |                  |                         |                             |
| sd(\(\beta_1\))           | 2.444            | 2.403                   | 2.621                       |
| sd(\(\beta_2\))           | 0.003            | 0.003                   | 0.003                       |
| corr(\(\beta_1, \beta_2\))| 0.432            | 0.511                   | 0.510                       |
| Tree level                 |                  |                         |                             |
| sd(\(\beta_1\))           | 7.095            | 5.601                   | 5.344                       |
| sd(\(\beta_2\))           | 0.007            | 0.006                   | 0.006                       |
| corr(\(\beta_1, \beta_2\))| 0.422            | 0.342                   | 0.254                       |
| Residual                   | 0.516            | 0.705                   | 0.617                       |

Table 4. The parameter estimates and standard errors (in brackets) of the best fitted mixed-effects Chapman–Richards model with (Chapman–Richards AR(1), AR(10) and without (Chapman–Richards)) autoregressive correlation structure.
where \( R = Z_0 + \left( Z_0^2 + \frac{2(\delta_1 + \delta_2 R_j \times R_j)}{H_0} \right)^{0.5} \) ; \( Z_0 = H_0 + \left( b_3 + \delta_3 R_j \times R_j \right) \).

The extended full (regional) model explains 0.47% more residual variance, while all expansion terms are statistically significant (\( t < 0.001 \), Table 6). Besides, the ANOVA result indicates that more complex, the regional model is significantly better (\( p < 0.01 \)) than the reduced (global) model and thus favour the regional model.

The residuals of the developed regional growth model are homoscedastic in the entire range of values predicted for individual trees and in individual regions (Fig. 2).

### Table 5. Parameter estimates and goodness-of-fit statistics for the individual functions. \( R^2_{\text{adj}} \) adjusted coefficient of determination, AIC akaike information criterion, MAE mean absolute error.

| Function | Parameter | Estimate | \( R^2_{\text{adj}} \) | AIC | MAE |
|----------|-----------|----------|------------------------|-----|-----|
| M1       | \( b_1 \)  | 1.363    |                        | 34.961 | 0.59 |
|          | \( b_2 \)  | 5920.904 |                        | 38,784 | 0.69 |
|          | \( b_3 \)  | 30.443   |                        |        |     |
| M2       | \( b_2 \)  | 0.017    |                        | 36,536 | 0.63 |
|          | \( b_3 \)  | 1.236    |                        |        |     |
| M3       | \( b_1 \)  | 0.019    |                        | 36,356 | 0.63 |
|          | \( b_2 \)  | 0.864    |                        |        |     |
|          | \( b_3 \)  | 8.293    |                        |        |     |
| M4       | \( b_1 \)  | 10.854   |                        | 35,275 | 0.60 |
|          | \( b_2 \)  | 0.018    |                        |        |     |
| M5       | \( b_0 \)  | 102.076  |                        | 36,155 | 0.63 |
|          | \( b_2 \)  | 0.434    |                        |        |     |

### Table 6. Parameters of the regional model developed using a nonlinear fixed-effects modelling approach. Expansion terms (\( \delta_{ij} \)) expressing the difference in the values of \( i \)th parameters (\( b_i, i = 1, 2, 3 \)) of the Eq. (5) in \( j \)th regions (\( R_j, j = 1, 2, \ldots, 6 \)).

| Global parameter | Regional expansion term | Estimate | Standard error (m) | t-statistic | p value |
|-------------------|-------------------------|----------|--------------------|-------------|---------|
| 1.363             | \( \delta_{a1} \)     | −0.13666 | 0.004              | −33.266     | <0.001  |
|                   | \( \delta_{a2} \)     | 0.06231  | 0.003              | 21.171      | <0.001  |
|                   | \( \delta_{a3} \)     | −0.09902 | 0.004              | −24.061     | <0.001  |
|                   | \( \delta_{a4} \)     | 0.02643  | 0.006              | 4.134       | <0.001  |
|                   | \( \delta_{a5} \)     | 0.02700  | 0.008              | 3.332       | <0.001  |
|                   | \( \delta_{a6} \)     | 0.08085  | 0.003              | 28.843      | <0.001  |
| 5920.904          | \( \delta_{b1} \)     | 25,305.5 | 1747.08            | 14.485      | <0.001  |
|                   | \( \delta_{b2} \)     | 7828.3   | 441.195            | 17.743      | <0.001  |
|                   | \( \delta_{b3} \)     | 31,820.2 | 1648.675           | 19.300      | <0.001  |
|                   | \( \delta_{b4} \)     | 21,248.2 | 2898.633           | 7.330       | <0.001  |
|                   | \( \delta_{b5} \)     | −2595.9  | 339.176            | −7.653      | <0.001  |
|                   | \( \delta_{b6} \)     | 15,746.3 | 617.851            | 25.485      | <0.001  |
| 30.443            | \( \delta_{c1} \)     | −29.0509 | 2.711              | −10.7145    | <0.001  |
|                   | \( \delta_{c2} \)     | −2.8214  | 0.56               | −5.0352     | <0.001  |
|                   | \( \delta_{c3} \)     | −45.0508 | 2.555              | −17.6339    | <0.001  |
|                   | \( \delta_{c4} \)     | −50.8217 | 6.588              | −7.7144     | <0.001  |
|                   | \( \delta_{c5} \)     | 6.2408   | 0.766              | 8.1472      | <0.001  |
|                   | \( \delta_{c6} \)     | −39.6201 | 1.263              | −31.3782    | <0.001  |
The developed regional model does not cause systematic errors in SI classes, age classes and tree height classes (Fig. 3). The mean residuals slightly deviate from 0 only in the least common SI below 15 m, and the equally rare age class above 120 years and the height class above 35 m (Fig. 3).

Figure 2. Relationship between residual and predicted height using the site index model for Scots pine in Poland developed using nonlinear fixed-effect modelling approach with the parameters expanded as a function of regions (I–VI).
Comparison of regional models developed with the use of nonlinear mixed-effects and nonlinear fixed-effects modelling approach. Based on the model fit statistics and graphical error analysis, it can be concluded that the regional model developed using the fixed-effects approach is characterised by better fit statistics (Table 7) and the distribution of errors (Fig. 4).

Analysis of regional variability of height growth patterns based on a model developed using the fixed-effects modelling approach. The height growth patterns in individual regions show that it
changes directionally from the north and northwest to south and southeast Poland (Fig. 5). In the NFR located in the northwestern (I, III) and northern (II) lowlands of Poland, the height increment rate is also maintained in older stands. In the case of NFR located in the highlands of southern and southeastern Poland, the height increment in old age is inhibited, and therefore the growth curves are more asymptotic (Fig. 5).

Growth trajectories of models developed for individual regions were directly compared using the height at base age 40 as the reference level (Fig. 6). A direct comparison of the height growth trajectories for different SI values determined for the base age of 40 years shows a significant diversification of the height growth patterns between regions, especially in more fertile sites. Up to the base age of 40, the height growth in individual regions is very similar. However, above the base age, a considerable diversification of the growth pattern is observed, especially for most fertile sites with SI equal to 23 or 17 m. Intensive height increment persists for the longest in older stands in the lowlands of northwest and northern Poland, especially in NFR I and III, which are under the influence of the Atlantic climate. Scots pine NFR II located in northeastern Poland grows only slightly more slowly. In the stands growing in regions VI and IV located in the highlands of southeastern Poland, where the climate is continental, the growth is slowed down in older stands. Scots pines from the NRF V located in southern Poland lowlands show weaker growth in old age in the most fertile sites, but in the poorer, their growth dynamics persist until old age (Fig. 6).

Discussion
We compared the development of regional height growth models using nonlinear fixed-effects and nonlinear mixed-effects modelling approaches. Our results indicate a better fit to the data of the model developed using the nonlinear fixed-effects modelling approach. Comparison of fixed-effects and mixed-effects models was conducted by Nigh40 who also found a better fit of models built with GADA. Similar results were also obtained by Cieszewski55, who demonstrated that self-referencing GADA models based on the nonlinear fixed-effects approach possess all the desirable properties associated with logical behaviour of the model and estimation statistics, while the nonlinear mixed-effects models lack the well-behaved model properties. The reason for the better fit of the models may be that the nonlinear fixed-effect models are by definition more flexible than the nonlinear mixed-effect models55. Moreover, models can significantly overestimate variances based on untenable assumptions and have more restrictions imposed regarding the site effects’ distributional properties. Cieszewski55 points out that this is because the NME model fitting process imposes an additional constraint in that the random effect must conform to the arbitrary assumption of randomness, meaning that it must always have a greater or equal sum of square errors when the same models are fit into the same dataset as site models. Our results also confirm well-behaved properties of the nonlinear fixed-effects self-referencing models and their better suitability for modelling TH growth and SI compared to mixed-effects models.

The developed models allowed us to compare the differences in Scots pine height growth dynamics between natural forest regions in Poland. We found that height growth is regionally differentiated. The most relevant

| Modelling approach | RMSE [m] | MAE   | $R^2_{adj}$ | AIC   |
|--------------------|---------|-------|-------------|-------|
| Mixed-effects      | 0.5596  | 0.4186| 0.9954      | 15,765|
| Fixed-effects      | 0.3698  | 0.2690| 0.9987      | 12,721|

Table 7. Characterisation of the fit statistics of the regional models developed using mixed-effects and fixed-effects modelling approaches.
differences in growth patterns are observed from the northwest to the southeast gradient. The shape of the growth trajectory for stands located in natural forest regions from the north and northwestern Poland (I–III) differs for younger and older trees. In these natural forest regions, the stands also reach higher absolute heights than those of southern regions. Our results show that the use of the regional model Poland reduces errors in the height growth prediction.

Regional variation in growth patterns was previously reported; these variations have usually been interpreted as the result of variability in climate, geology, soil type, geographical location, and genotype13–15.

We found significant differences between the height growth patterns from NRF located in lowlands of northern- northwestern and northeastern Poland and highlands of southern- southeastern Poland. This may be because natural forest regions in northern Poland are under the influence of Atlantic climate and have deep podsolic, brown podsolic, and brown glacial soils, whereas natural forest regions VI and IV are under the influence of continental climate and grow on upland soils with less depth. The depth and physical properties of soil and the sand and clay contents lead to varying soil fertility and productivity56,57. Additionally, the availability of water is strongly correlated with the growth of stands and impacts the productivity of sites58. A linear increase in net primary production with increased water availability was observed at dry sites59. The amount of annual

Figure 5. Regional nonlinear fixed-effects model height trajectories against the observed growth series from stem analysis collected in natural forest regions I-VI.
precipitation in Poland changes from south to north and, with soil fertility and depth, can explain the difference in height growth variability between the northern and southern regions.

We found that growth trajectories for individual regions vary with age. The growth of stands in southern and southeastern Poland is slowing rapidly, whereas northern and northeastern stands have more continuous rapid height growth. The difference in growth pattern in regions IV, V, and VI observed in older age can be explained as the impact of soil depth. When soil depth is not a limiting factor in the youngest age classes, the impacts of climate and other factors are more important for growth dynamics, but with the transition to older age classes and increased interspecific competition, soil depth likely begins to limit the growth dynamic. An evident slowdown of growth dynamics in oldest age classes in natural forest regions IV and VI may be the result of a combination of shallower and less fertile podsol and brown podsol soils and low annual rainfall. In regions, I, II, and III, geological and climatic conditions are more favourable for Scots pines and likely result in a larger height increment in the older stands.

The significant height growth differences in different regions indicate that the use of one generalised height growth model for the entire country may result in the inappropriate estimation of site productivity and have direct forest management consequences. The importance of regional models may increase because the projections of climate change in Poland predict accentuated differences in site conditions between the northwest and southeast regions. The obtained results concerning the spatial variability of Scots pines’ height growth dynamics should be considered when making forest management decisions concerning the rotation age or intensity of silvicultural treatments.

The amount of carbon stored in forests is more significant than in any other terrestrial ecosystem, and tree growth data play a significant role in carbon sequestration and biomass production calculations. Several conversion methods have been developed for different types of forest to obtain reliable estimates. However, our results show that not considering specific regional growth conditions and the resulting differentiated growth patterns by using only global models for these calculations may result in an over- or under-estimation of regional height growth, biomass production, and carbon sequestration. Many studies on the estimation of biomass production and carbon sequestration emphasise that the regional approach is promising; therefore, the use of models developed on a regional scale shows particular importance for effective forest management.

With the development of geographical information systems and technologies, tools such as the mapping SI are increasingly common. Many problems regarding the stability of forest stands and disturbance of forests in Europe concern relatively small areas. Numerous studies found that SI mapping and analyses, (particularly on a small scale) provide information for a more dynamic forest growth assessment and should be an essential component of regional planning and management. Therefore the use of advanced forestry solutions focused on smaller areas, regional growth models, and the analysis of local trends have great potential for the fast detection of problems and the precise formulation of answers to stand instability and forest disturbances.

We used the height growth trajectories of dominant Scots pines collected using the unified methodology to develop a regional model; however, collection methods and data processing in different regions should be carefully executed. Differences in methodology or data sources can lead to significant inconsistencies between growth models that are not based on ecological or biological causes. A given species growth patterns in different ecological or geographical conditions are significant and should be determined before developing local models for small areas. The differences in local model predictions may not be large enough in small geographical regions to justify the development costs. In this situation, smaller regions with low climatic and geomorphological variability should be grouped to develop a single model for larger areas.

Conclusions

We compared the development of regional height growth models using nonlinear-fixed-effects and nonlinear-mixed-effects modelling approaches. Our results indicate a slightly better fit to the data of the model using the nonlinear fixed-effects approach. The developed models show differences in height growth patterns of Scots
pines in Poland and revealed that acknowledgement of region as the independent variable could improve the growth prediction and quality of the SI estimation. The presented study showed regional differences in height growth patterns of Scots pines in Poland. Differences in climate and soil conditions that distinguish natural forest regions affect Scots pine height growth patterns. Therefore, extending this research to models that directly describe height growth interactions with site variables, such as climate, soil properties, and topography, can provide valuable forest management information.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 25 August 2020; Accepted: 30 April 2021
Published online: 14 May 2021

References
1. Bontemps, J. D. & Bouriaud, O. Predictive approaches to forest site productivity: Recent trends, challenges and future perspectives. *Forestry* 87, 109–128 (2014).
2. Skovsgaard, J. P. & Vanclay, J. K. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. *Forestry* 81, 13–31 (2008).
3. Albert, M. & Schmidt, M. 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 (2010).
4. García, O. Dynamical implications of the variability representation in site-index modelling. *For. Ecol. Manag.* 257, 951–959 (2009).
5. Hagglund, B. & Lundmark, J. E. Site index estimation by means of site properties of Scots pine and Norway spruce in Sweden. *Stud. For. Suec.* 138, 5–38 (1977).
6. Johansson, T. Site index curves for common alder and grey alder growing on different types of forest soil in Sweden. *Scand. J. For. Res.* 14, 441–453 (1999).
7. Raulier, F., Lambert, M.-C., Pothier, D. & Ung, C.-H. Impact of dominant tree dynamics on site index curves. *For. Ecol. Manag.* 184, 65–78 (2003).
8. Corral Rivas, J. I., Álvarez González, J. G., Ruiz González, A. D. & von Gadow, K. Compatible height and site index models for five pine species in El Salto, Durango (Mexico). *For. Ecol. Manag.* 201, 145–160 (2004).
9. Tewari, V. P., Rivas, J. I. C., Vilčko, F. & Von Gadow, K. Height-Growth and site index equations for social forestry plantations of acacia nilotica and eucalyptus hybrid in gujarat state of India. *For. Trees Livelihoods* 17, 125–140 (2007).
10. Monserrut, R. A. & Rehfeldt, G. E. Genetic and environmental components of variation of site index in inland douglas-fir. *For. Sci.* 36, 1–9 (1990).
11. Alvarez-González, J. G., Ruiz-Gonzalez, A. D., Rodriguez-Soolleiro, R. & Barrio-Anta, M. Ecoregional site index models for *Pinus pinaster* in Galicia (northwestern Spain). *Ann. For. Sci.* 62, 115–127 (2005).
12. Bravo-Oviedo, A., Tomé, M., Bravo, F., Montero, G. & del Río, M. Dominant height growth equations including site attributes in the generalised algebraic difference approach. *Can. J. For. Res.* 38, 2348–2358 (2008).
13. Johansson, T. Site index curves for Norway spruce plantations on farmland with different soil types. *Stud. For. Suec.* 198, 1–19 (1995).
14. Adams, J. P., Matney, T. G., Land, S. B. Jr., Belli, K. L. & Duzan, H. W. Jr. Incorporating genetic parameters into a loblolly pine growth-and-yield model. *Can. J. For. Res.* 36, 1959–1967 (2006).
15. Buford, M. A. & Burkhart, H. E. Genetic improvement effects on growth and yield of loblolly pine plantations. *For. Sci.* 33, 707–724 (1987).
16. Monserrut, R. A. Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type. *For. Sci.* 30, 943–965 (1984).
17. García, O. Dynamical implications of the variability representation in site-index modelling. *Eur. J. For. Res.* 130, 671–675 (2010).
18. Calama, R., Cañadas, N. & Montero, G. Inter-regional variability in site index models for even-aged stands of stone pine (*Pinus pinea* L.) in Spain. *Ann. For. Sci.* 60, 259–269 (2003).
19. Adame, P., Hynynen, J., Cañellas, I. & del Río, M. Individual-tree diameter growth model for reboilio oak (*Quercus pyrenaica* Willd.) coppices. *For. Ecol. Manag.* 255, 1011–1022 (2008).
20. Bravo-Oviedo, A., del Río, M. & Montero, G. Geographic variation and parameter assessment in generalised algebraic difference site index modelling. *For. Ecol. Manag.* 247, 107–119 (2007).
21. Bontemps, J. D. & Bouriaud, O. Predictive approaches to forest site productivity: recent trends, challenges and future perspectives. *Forestry* 87, 1–20 (2013).
22. Claessens, H., Pauwels, D., Thibaut, A. & Rondeloup, C. J. Site index curves and autecology of ash, sycamore and cherry in Wallonia (Southern Belgium). *Forestry* 72, 171–182 (1999).
23. Karlsson, K. Height growth patterns of Scots pine and Norway spruce in the coastal areas of western Finland. *For. Ecol. Manag.* 135, 205–216 (2000).
24. Martín-Benito, D., Gea-Izquierdo, G., del Río, M. & Cañellas, I. Long-term trends in dominant-height growth of black pine using dynamic models. *For. Ecol. Manag.* 256, 1230–1238 (2008).
25. Dyrekacja Generalna Lasów Państwowych, *Raport o stanie lasów w Polsce* (2018).
26. Socha, J., Orolz, S. Dynamic site index curves for Scots pine (*Pinus sylvestris* L.) stands in the High Ebro Basin (northern Spain) using soil attributes. *For. Ecol. Manag.* 74, 395–406 (2001).
27. Gadow, K., Hui, G. *Modelling Forest Development*, Vol. 57, Forestry Sciences (Springer Netherlands, Dordrecht, 1999). ISBN 978-1-4020-0276-2.
28. Schwappach, A. *Ertragstafeln der Wichtigeren Holzarten* (Druckerei Merkur, Prag, 1943).
29. Szymkiewicz, B. Niekotóre zagadnienia dotyczące tablic zasobności drzewostanów sosnowych. *Pr. IBL* Seria A 67 (1948)
30. Szymkiewicz, B. Tablice Zasobności i Przyrostu Dzewostanów (Państwowe Wydawnictwo Rolnicze i Leśne, Warszawa, 2001). ISBN 83–09–01745–6.
31. Pretzsch, H., Biber, P. *Schütze*, G. Uhl, E. & Rötzer, T. Forest stand growth dynamics in Central Europe have accelerated since 1870. *Nat. Commun.* 5, 4967 (2014).
32. Socha, J., Orzel, S. Site index models for main forest-forming tree species in Poland. *Forests* 11, 8–10 (2020).
Acknowledgements

We would like to thank the Forest Management and Geodesy Bureau from Sekocin Stary for cooperation in implementing the project "Actual and potential site productivity in Poland for the main forest forming tree species", particular for help in collecting the height growth data. We thank Professor Arkadiusz Bruchwald for sharing stem analysis data collected within his research projects.

Author contributions

Conceived and designed the study: J.S. Collected data: J.S. Analysed the data: J.S., K.B., L.T.C., S.Z., P.H. All authors read and approved the final manuscript.
Funding
This work was supported by the project I-MAESTRO. Project Innovative forest MAnagEment STrategies for a Resilient bioEconomy under climate change and disturbances (I-MAESTRO) is supported under the umbrella of ForestValue ERA-NET Cofund by the National Science Centre, Poland and French Ministry of Agriculture, Agrifood, and Forestry; French Ministry of Higher Education, Research and Innovation, German Federal Ministry of Food and Agriculture (BMEL) via Agency for Renewable Resources (FNR), Slovenian Ministry of Education, Science and Sport (MIZS). ForestValue has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 773324. Data collection was supported by the project “Actual and potential site productivity in Poland for the main forest forming tree species”, financed by the General Directorate of State Forests [agreement no. ER-2717–11/14, 15th April 2014].

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to J.S.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021