Received: 31 July 2020 | Accepted: 30 March 2021

DOI: 10.1111/2041-210X.13616

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

A new algorithm for reconstructing tree height growth with stem analysis data

Christian Salas-Eljatib[1,2,3]

[1]Centro de Modelación y Monitoreo de Ecosistemas, Facultad de Ciencias, Universidad Mayor, Santiago, Chile

[2]Vicerrectoría de Investigación y Postgrado, Universidad de La Frontera, Temuco, Chile

[3]Departamento de Silvicultura y Conservación de la Naturaleza, Universidad de Chile, Santiago, Chile

Correspondence
Christian Salas-Eljatib
Email: cseljatib@gmail.com

Funding information
Chilean Government research grants Fondecyt No. 1191816 and FONDEF No. ID19 10421.

Handling Editor: Ryan Chisholm

Abstract

1. Stem analysis allows us to obtain an abundant amount of information on tree growth. A couple of algorithms exist to utilize section height and growth ring data for reconstructing height and age time-series information.
2. I evaluated two alternatives, a well-known and a newly proposed algorithm using stem analysis data of four species, including deciduous and evergreen broadleaves and a conifer. I reconstructed height–age pairs by both algorithms. I fit height growth equations in a mixed-effects model framework for each species, using the generated data with the respective algorithm. Comparisons considered confidence intervals of the estimated parameters, as well as regression-based equivalence tests.
3. Results showed that the fitted growth models obtained from both stem analysis algorithms were statistically equivalent. However, the proposed algorithm is simpler and thus provides a useful alternative to current methods.
4. Based on the findings, I recommend using this new stem analysis algorithm to reconstruct tree height growth with stem analysis data.

KEYWORDS

equivalence testing, forest ecology, growth rates, height growth models, mixed-effects models

1 | INTRODUCTION

Tree growth is essential for forest management and forest ecosystem research. Growth is the product of several abiotic and biotic factors, interacting with a tree and its surroundings through time. Tree growth is expressed morphologically and is vital for forest management because tree size determines timber production capability (Monserud, 2003). In the same vein, tree growth is the primary driver of forest dynamics because it determines tree crown class differentiation, mortality and regeneration, and therefore it is of pivotal importance for understanding how trees interact in a dynamic ecosystem (Kimmins, 2004).

Stem analysis is a methodology for measuring tree growth. As pointed out by Tesch (1980), the study of tree growth is not a recent endeavour, and the first recorded observations of tree growth are generally credited to Theophrastus (370–285 B.C.), a student of Aristotle. Most trees growing in temperate climates produce a distinct layer of wood every year, indicating the particular tree's age and growth. One layer is formed each year between the bark and the previously formed wood. This layer looks like a ring in a cross-section cut (or ‘crosscut’) of a trunk. In stem analysis (SA), several wood discs are extracted along the stem of a tree. Usually, height is measured, and the number of rings present at each disc is counted (*incomplete SA*), although sometimes the radial growth at each disc is also measured (*complete SA*). SA provides a record of the height growth of a tree and its diameter growth at many points along the stem. SA is the measuring technique that offers the most outstanding detail on the past growth of several tree-level variables, therefore offering crucial useful data for tree growth studies.

Height–age pairs must be derived from the stem analysis data to model height growth. The study of tree height growth is crucial for understanding competition capabilities and its relationship with stand structure and composition (Holste et al., 2011), as well as for assessing forest productivity (Salas-Eljatib, 2021a). Growth can be depicted as a series of envelopes with different taper and a decreasing number of annual rings as stem height increases (Figure 1). A count of the number of rings on a given cross-section gives the tree's age above the section. Thus, it indicates the age of the tree at that point. If the count is made on a cross-section at ground level, it gives the total tree age. A section's height is attained within a period equal to the age at the sampling time minus the section ring count. If we could obtain a disc exactly at the point in the stem where the growth ends for a corresponding time, we would directly know the corresponding height, in this case, equal to the section height (Figure 1a). Nevertheless, due to the trees' conical growth pattern, the actual height at the age corresponding to the cross-section ring count will almost always be located above the measured cross-section (Dyer & Bailey, 1987). Let me define the cross-section age as the difference between the total number of rings at the base of a tree and the cross-sectional number of rings. (Figure 1b). Carmean (1972) pointed out that the height at the point of sectioning underestimates actual height at the presumed age because the section will almost always occur at some intermediate point along with the annual leader rather than at the terminal bud itself. The actual height for the corresponding age at that cross-section will be at some point above that cross-section. The exact height or the height of the tip for a corresponding age is known as the 'hidden tip' (Dyer & Bailey, 1987), as shown in Figure 1b.

A couple of algorithms exist for reconstructing tree height–age data pairs with stem analysis data. From the graphical one of Graves (1906) up to the linear programming-based one of Lappi (2006). I classify the stem analysis algorithms into two groups: (a) 'age-based', those that estimate the height corresponding to the age of each cross-section, and (b) 'height-based', those that estimate the age corresponding to the height of each cross-section. The age-based algorithms have received the most attention in the literature, and they are subdivided into those that use only ring counts (e.g. Carmean, 1972; Fabbio et al., 1994; Lappi, 2006; Lenhart, 1972; Newberry, 1991) and those that use also ring width (e.g. Kariuki, 2002). Overall, and based on the comparisons of Dyer and Bailey (1987), Fabbio et al. (1994), Kariuki (2002), Lappi (2006), Machado et al. (2010), Rayner (1991) and Subedi and Sharma (2010), Carmean's height interpolation algorithm has performed well in almost all the published research on the topic. Regarding the height-based algorithms, Milner (1992) proposed one in words, but no further work has been done on this type of algorithm.

Dyer and Bailey (1987) pointed out that Carmean's algorithm has two assumptions: First, annual height increment is constant for each year for which height growth is contained within the section. Second, the hidden tip will occur in the middle of a year's height increment. While neither assumption is likely to be upheld, it is germane to examine an alternative approach that does not rely on either. I aim to describe the basis of a new stem analysis algorithm and compare it against Carmean's algorithm.

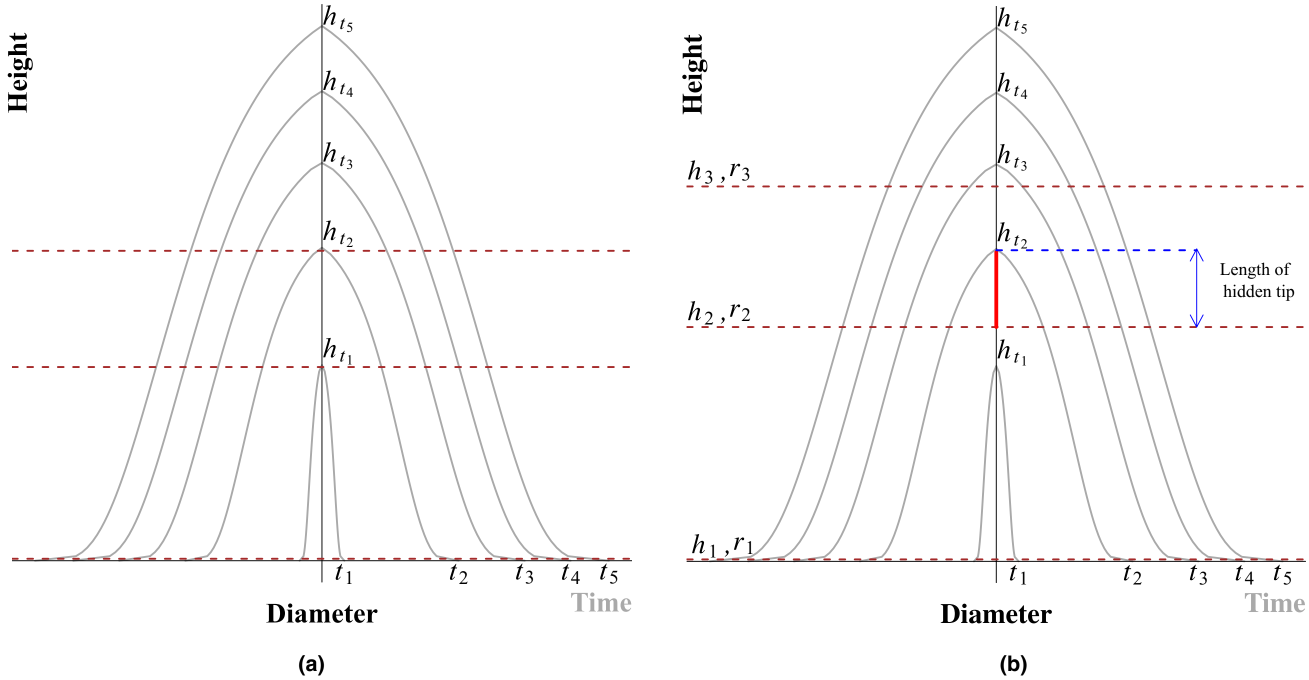


FIGURE 1 Stem analysis basis on a longitudinal split showing the progressive taper development (wood layers) for a 5-year-old tree. (a) Horizontal dashed lines represent the ideal location of two cross-sections to accurately reconstruct the tree heights at time 1 (t_1) and time 2 (t_2), that is, h_{t_1} and h_{t_2} respectively. (b) In practice, the locations of cross-sections will not coincide exactly with tips. Besides, the height of three cross-section (h_1, h_2 and h_3) and the ring count on each cross-section (r_1, r_2 and r_3) are represented. Finally, the length of the hidden tip for t_2 is also depicted

2 | MATERIALS AND METHODS

2.1 | Data

I used stem analysis data of sample trees selected from fixed-area sample plots for different tree species. Data were representative of broadleaved (two deciduous and one evergreen) and coniferous tree species. The species were as follows: (a) *Nothofagus dombeyi* (Coigüe), an evergreen tree species with a wide distribution in Chile, being very abundant, especially in the country's southern part. (b) *Pseudotsuga menziesii* (Douglas fir), a conifer native from the Pacific Northwest of North America, extensively planted as an exotic elsewhere. (c) *Nothofagus alpina* (Raulí), a deciduous tree species that has great potential for forest management and high timber value. (d) *Nothofagus obliqua* (Roble), a deciduous tree species as well. The *Nothofagus* species data are obtained from several studies described by Salas-Eljatib (2020) and cover a latitudinal range between 35°50′ and 41°30′S in south-central Chile (Figure S1a). In contrast, the Douglas fir data are obtained from the study of Monserud (1984) and cover between 42°10' and 49°50' N in the Inland Northwest region, northern Rocky Mountains, of the United States (Figure S1b).

All the tree-level data were collected from sample plots. For the *Nothofagus* data, these plots were established in mixed-species, even-aged, secondary stands dominated by Coigüe, Raulí or Roble. Veblen et al. (1996) provide additional ecological features of these forests. Meanwhile, for the Douglas fir data, the plots were established in even- and uneven-aged stands (Monserud, 1984). For each dataset, dominant trees were selected from fixed-area plots, provided they were healthy and of good form, of seed origin and belonging to the upper canopy. The selected trees were felled, and after measuring DBH (d) and total height (h), cross-sectional discs were obtained at several heights along the stem. There were, on average, 10 sections per tree. Rings were counted in the laboratory. The number of plots by species is 30, 181, 53 and 62 for Coigüe, Douglas fir, Raulí and Roble, respectively, and have between three and four sample trees per plot.

The age of a tree is the length of time that has elapsed since the germination of the seed or the sprout's budding (Husch et al., 2003). The most common dating method for mature trees is to count the number of annual growth rings on a transverse section or an increment core sampled at ground level, that is, at what appears to be the root collar (DesRochers & Gagnon, 1997). I computed total age (*tot.age*), at the time of sampling by linear extrapolation of ring counts of the lowest two sections down to ground level, rounding to the nearest integer. This age was used as the time variable for reconstructing past height growth. As a reference, I also computed the breast height age (*bha*) as the number of rings present at 1.3 m on the stem (Table 1).

I tried to include more species in the analysis; however, pertinent data from other species were unavailable. The data that could be used have some of the following issues: small sample size; does not have the minimum information needed, such as the number of rings per cross-section; and trees concentrated in a single location, among other issues. Nonetheless, obtaining data of this sort is not as easy as a user might expect, but is the reality of most research disciplines (Poisot et al., 2019).

2.2 | Algorithms for cross-section height–age adjustment

Before describing the algorithms, I will introduce notation for the raw data in a given tree. Notice that I will omit a tree's subscript to simplify details: h_k, r_k and t_k are height, number of rings and age at the kth cross-section. For these data pairs, the Carmean algorithm reconstructs corrected pairs $\left(t_k, h_{t_k}\right)$ and the proposed one reconstructs $\left(t_{h_k}, h_k\right)$. That is to say, the main difference between the two algorithms is either reconstructing height at a given rounded time (denoted by h_{t_k}), or reconstructing time at a given height (represented by t_{h_k}). From this point forward, I will refer to Carmean's algorithm as ABA (age-based algorithm), because based on age, estimate height; and to the proposed algorithm as HBA (height-based algorithm), because based on height, estimate time. How to compute t_k, h_{t_k} and t_{h_k} is explained in the following sections.

2.2.1 | Carmean (ABA)

Carmean (1972) did not offer a mathematical expression for his algorithm, but Lenhart (1972) did. It can be expressed as follows.

$$h_{t_k} = h_k + \frac{h_{(k+1)} - h_k}{2\left[r_k - r_{(k+1)}\right]} = h_k + \frac{\Delta h_k}{2\Delta r_k}, \tag{1}$$

where h_{t_k} is the estimated total tree height at age t_k and h_k is the height of the kth cross-section. The rest of the terms were described above. Therefore, Δh_k is the length of the $(k+1)$th section and Δr_k is the difference in the number of rings between the k and $(k+1)$th sections, provided that $r_k > r_{(k+1)}$. The age of the tree (associated with the innermost ring) at the kth cross-section, t_k, is computed as

$$t_k = t_b - r_k + 1, \tag{2}$$

where t_b is the reference age for height reconstruction (e.g. either total age or breast height age). Other mathematical details of the algorithm are in Supporting Information Appendix 1. In this regard, it is important to point out that it is common practice to interpolate the heights to every missing year between sections (using the third component of Equation S1). This practice is unsuited because it inflates the sample size and smooths the data before fitting a model. The latter's problem is those fit statistics measuring deviations about smoothed height increment data are misleading and strongly biased (Hasenauer & Monserud, 1997). Hereafter Equation (1) defines the ABA, where the length of the hidden tip is estimated as half of the periodic annual height increment for the kth section.

TABLE 1 Tree sample variables' summary by species. The variables are *d* is the DBH, *h* is the total height, *tot.age* is the total age and *bha* is the breast height age

		Variable			
		d	*h*	*tot.age*	*bha*
Species	**Statistic**	**(cm)**	**(m)**	**(year)**	**(year)**
Nothofagus dombeyi (Coigüe)					
(*n* = 107)	Minimum	5.3	9.9	21	15
	Maximum	60.2	33.7	71	68
	Mean	26.6	19.9	41.1	37.7
	CV(%)	19	16.6	15.6	17.3
Pseudotsuga menziesii (Douglas fir)					
(*n* = 312)	Minimum	30.5	24.2	54	51
	Maximum	95.2	49.4	198	195
	Mean	52.3	33.3	120.6	116.4
	CV(%)	11.7	9.6	19.9	19.8
Nothofagus alpina (Raulí)					
(*n* = 169)	Minimum	5.3	7.1	19	17
	Maximum	49.9	31.2	81	76
	Mean	24.9	20.3	45.9	42.2
	CV(%)	16.2	15.1	17	18.4
Nothofagus obliqua (Roble)					
(*n* = 155)	Minimum	7.3	7.9	14	13
	Maximum	59.1	37	103	94
	Mean	28.8	22.6	45.2	41
	CV(%)	19.4	18.3	19.2	19.3

2.2.2 | Proposed algorithm (HBA)

The HBA seeks to estimate the age of each cross-section based on the height of that section. The idea is to reduce the calculated cross-section age by half a year, instead of increasing the section height by half the height increment as in Carmean's algorithm. This approach is more straightforward than the ABA; however, no studies have focused on its effectiveness or explained its basis. To the best of my knowledge, the first reasoning of this kind was suggested by Milner (1992). Therefore, in the following, I refine Milner's wording by introducing new concepts and proposing a formal notation and mathematical expressions.

When having stem analysis data, we observe the height of the cross-sections. In this algorithm, I use that height as such, without modification. Because of that, I refer to this algorithm as a height-based one. Instead, I only alter the computation of the age for each cross-section. Equation (2) allows obtaining that age for the ABA, symbolized by t_k. However, given that for the proposed algorithm, h_k is directly treated as the height of interest, we know that the tree's age when was h_k tall (symbolized by t_{h_k}) must be lower than t_k. Besides, inasmuch as the age difference must be between 0 and 1, I think that a uniform distribution is a suitable choice for accommodating its uncertainty, as follows.

$$E\left[t_{h_k}\right] = t_b - r_k + 1 - E\left[u_k\right] = t_b - r_k + 0.5, \tag{3}$$

where: t_{h_k} is the age of the tree when was as tall as the height of the kth cross-section, u_k is a random number from a uniform distribution, $u_k \sim U\left[0,1\right]$, E is the expected value operator and the other terms were already defined. Equation (3) is the algorithm I propose, and hereafter refer to as 'HBA' (height-based algorithm). Other mathematical details of this algorithm are in Supporting Information Appendix 2. It is important to argue that I do not interpolate the heights for a given year but the age for a given cross-section height only. In summary, although from both ABA and HBA algorithms we reconstruct height–age data, they are built from different perspectives. The structure of the constructed data from both algorithms is in Table S1. Furthermore, Figure 2 illustrates both algorithms' differences where two cross-sections are measured. In the ABA, the estimation acts in the vertical axis (i.e. height); in contrast, the HBA acts in the horizontal axis (i.e. age). The reconstructed height growth series using both algorithms are represented by species, showing a wide variety of shapes, indicating large variability in tree height growth (Figure S2).

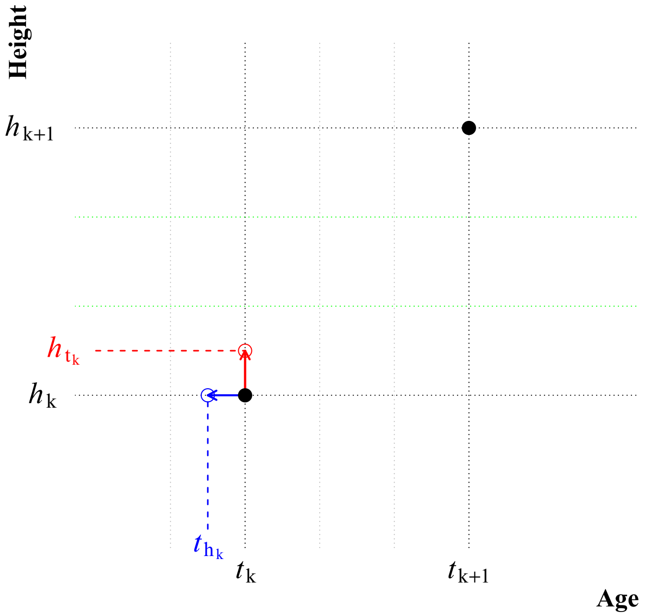


FIGURE 2 Graphical representation of how the algorithms work. Here two measurements for a given tree are represented with black dots at heights h_k and h_{k+1}. Ages at sections k and $k+1$ are represented by t_k and t_{k+1}, respectively, and computed as in Equation (2). To illustrate the differences, each year (apart from the ones where the cross-sections are obtained) is marked with grey-dashed vertical lines, and the annual height increments between each cross-sections are marked with green-dashed horizontal lines. In the ABA (age-based, Carmean), the corresponding height at the age t_k, that is, h_{t_k}, is estimated, therefore the pair of reconstructed data is $\left(t_k, h_{t_k}\right)$, which is represented in red. In contrast, in the HBA (height-based), the corresponding age at height h_k, that is, t_{h_k}, is estimated, therefore the pair of reconstructed data is $\left(t_{h_k}, h_k\right)$, which is represented in blue

2.3 | Comparisons

To compare the algorithms' performance for correcting the section height–age data from stem analysis, we fit the same model (i.e. base model) using the two different datasets, one obtained with the ABA adjustments and the other with the HBA algorithm. I do so to show the effects of the data in the resulting fitted growth models. I repeat the following analyses for each species.

2.3.1 | Growth law

I used the following baseline growth function, which as a statistical model is represented by.

$$h_{iz} = \alpha \left\{ 1 - e^{-\beta(t_{iz})} \right\}^{1/\gamma} + \epsilon_{iz}, \tag{4}$$

where h_{iz} is the height for the ith tree at the zth time t_{iz}; while α, β and γ are parameters and ϵ_{iz} is the random error term. Notice that $i = 1, \ldots, n$ and $z = 1, \ldots, T_i$, where $T_i = K_i + 1$, and K_i is the total number of discs on the ith tree. Therefore, h_{iT_i} is the tree's total height at sampling. The parameters have the following interpretations (Salas-Eljatib, 2020): α is the upper asymptote or maximum level of growth for the state variable h, β is a parameter which governs the rate of change or scale parameter and γ is a shape parameter determining the height of the growth curve inflection point. This model is among the most widely used growth function to study tree growth (Salas-Eljatib, 2020).

As I have several observations of h over time for the same tree, the data have a temporal correlation. To account for this correlation, I fit a mixed-effects model. I consider the tree (in each plot) as the random factor (i.e. group). I added random effects to one of the parameters of Equation (4), therefore having different model variants (Supporting Information Appendix 3). The random individual effects induce an intra-individual correlation structure that accounts for the lack of independence among the same tree observations. I fitted the mixed-effects models by maximum likelihood and used variance functions to model the within-stratum errors' variance structure. Since the *Nothofagus* data come from different studies, I used this information as a stratum; meanwhile, the national forest information was used as a stratum for the Douglas-fir data. All models were fitted using the NLME package (Pinheiro et al., 2021) implemented in R (R Core Team, 2020). All models were compared with the corrected Akaike information criterion (AICc). From here, I obtained two models, one fitted using the 'ABA-generated data' and another using the 'HBA-generated data'.

2.3.2 | Parameter estimates

To compare the algorithms' difference, I compared the parameter estimates obtained for the same model using the generated data. This comparison involves comparing the 95% asymptotic confidence intervals for each parameter estimate by the type of algorithm used to build the height growth series. In this way, I was able to examine the sensitivity of parameter estimates to the algorithm for reconstructing height–age data.

2.3.3 | Equivalence testing

I conducted a regression-based equivalence test within a nonparametric bootstrapping framework. Traditionally, equivalence tests are approached similarly as a traditional significant test for the same parameters, that is, one statistic is tested versus one computed from another treatment. However, this approach is not suitable for assessing the prediction performance of models. Robinson and Froese (2004) and Robinson et al. (2005) proposed to apply the equivalence test in a validation context that has data (i.e. validation sample) that are independent of the fitting process (as explained below) and assessing the model in the prediction of the response variable for the validation sample. Robinson et al. (2005) proposed a nonparametric bootstrap approach which we used in the present research. It is essential to clarify that I cannot directly assess equivalence among the algorithms because the Carmean one modifies the height data while the HBA modifies the age data. Therefore, the two algorithms neither share the same value of age nor height, on which a simple comparison can be established. I instead compared the predictions obtained from the same height growth model but fitted using the generated data from each algorithm to the observed height growth. All the details on the application of the equivalence testing are given in Supporting Information Appendix 4. This testing's main outcome is to compute the proportion on which the predictions are within the equivalent region; if it is greater than 0.95, I will reject the corresponding null hypothesis of dissimilarity at 5% of the significance level. That is to say, the predictions of the height growth model fitted with the generated data using the corresponding stem analysis algorithm are statistically equivalent to the observed height growth. Equivalence testing was carried out using the EQUIVALENCE R package (Robinson, 2010).

2.3.4 | Bias

Apart from testing the equivalence between the algorithms, I also compared their absolute bias. For doing so, I have proceeded using the following four steps for each species: (a) To create simulated data using the growth laws. I generated 1,000 sample trees having six cross-sections. I predicted heights for randomly generated rounded ages from a uniform probability density function, ranging between 10 and 200. I predicted using the baseline model and its parameter estimates (Table S2). Later, I added to each predicted heights a white noise coming from a Gaussian distribution with an expected value of 0 and a standard deviation equal to the residual model's standard deviation, that is, $\hat{\sigma}_e$. (b) To derive data of height-ring counts using both algorithms. (c) To fit the baseline model using the derived data.

(d) To compute the absolute bias of both fitted models in predicting the fundamental growth law.

3 | RESULTS

The polymorphic mixed-effect model (Equation S6) had a better fit, using as reference the AICc statistic, for both the data generated with the ABA and the HBA, as well as for all species (Table 2). Notice that comparisons must be only established among fitted models using the same data because maximum likelihood-based statistics alone are meaningless. As the polymorphic model had a better fit for all species, it was used for subsequent analysis.

Fitting the polymorphic model (Equation S6) to height–age pairs from the two algorithms using each species stem analysis data resulted in parameter estimates significantly different from zero in each case (Figure 3). Confidence intervals (CIs) of all the parameter estimates of the height growth models fitted with data generated from the two stem analysis algorithms overlapped to a large extent (Figure 3). Overlap of CIs of all parameter estimates is almost total for Douglas fir and Coigüe. The smaller overlap of CIs occurs for $\hat{\gamma}$ of roble, however it is still greater than 80%.

The null hypothesis (*H*0) of the equivalence test is equal to the statistics being not equivalent to the defined equivalence region. As explained above, if the bootstrap CIs are contained within the equivalent region, we reject *H*0; that is, we have strong evidence against dissimilarity, favouring equivalence. The equivalence testing results are shown in Table 3 for all species. In 100% of the bootstrap replicates, both estimated parameters, $\hat{\phi}_0$ and $\hat{\phi}_1$, of model (S8) were contained within their respective equivalence regions, I_0 and I_1 respectively. Although it might seem strange to have 100%, similar results had also been reported (e.g. Robinson et al., 2005), but more than the value, what is important is to emphasize the strong evidence against *H*0. The bootstrap CI for a 95% confidence level for both parameters and all species is contained within their respective equivalence regions. Therefore, I reject the null hypothesis of not equivalent for both $\hat{\phi}_0$ and $\hat{\phi}_1$, and for both all species and the two algorithms. Predicted values obtained from both a growth model fitted with data generated with the ABA and the same model but fitted with data generated with the HBA are statistically equivalent to their respective measured height growth values.

The average bias, an indicator of accuracy, clearly shows that the HBA unbiasedly estimated height growth (Table 4). On the other hand, the ABA is biased, with a slight (1%) trend towards the overestimation. The described pattern is consistent across all the species. Regarding the absolute bias, the HBA has a lower value for all species as well. These results enhance the difference between the two algorithms and favour the HBA.

TABLE 2 Maximum likelihood-based statistic of model variants. Each variant represents adding random effects either into the asymptote (as in Eq. S5) or into the shape parameter (as in Equation S6) of the baseline height growth model (Equation 4). AICc is the corrected Akaike information criterion. Notice that comparisons based on AICc are only valid among models using the same data. That is to say, we can compare the AICs between the anamorphic and polymorphic variants but only for a given data generated by a specific algorithm

		Algorithm	
		ABA	HBA
Species	Model variant	AICc	AICc
Coigüe			
	$\alpha + a_i$	4,191.7	4,286.0
	$\beta + b_i$	3,954.1	4,066.0
Douglas fir			
	$\alpha + a_i$	14,590.0	14,638
	$\beta + b_i$	14,071.0	14,118
Raulí			
	$\alpha + a_i$	5,771.0	5,900.7
	$\beta + b_i$	5,506.6	5,649.6
Roble			
	$\alpha + a_i$	6,925.2	7,040.4
	$\beta + b_i$	6,618.4	6,741.1

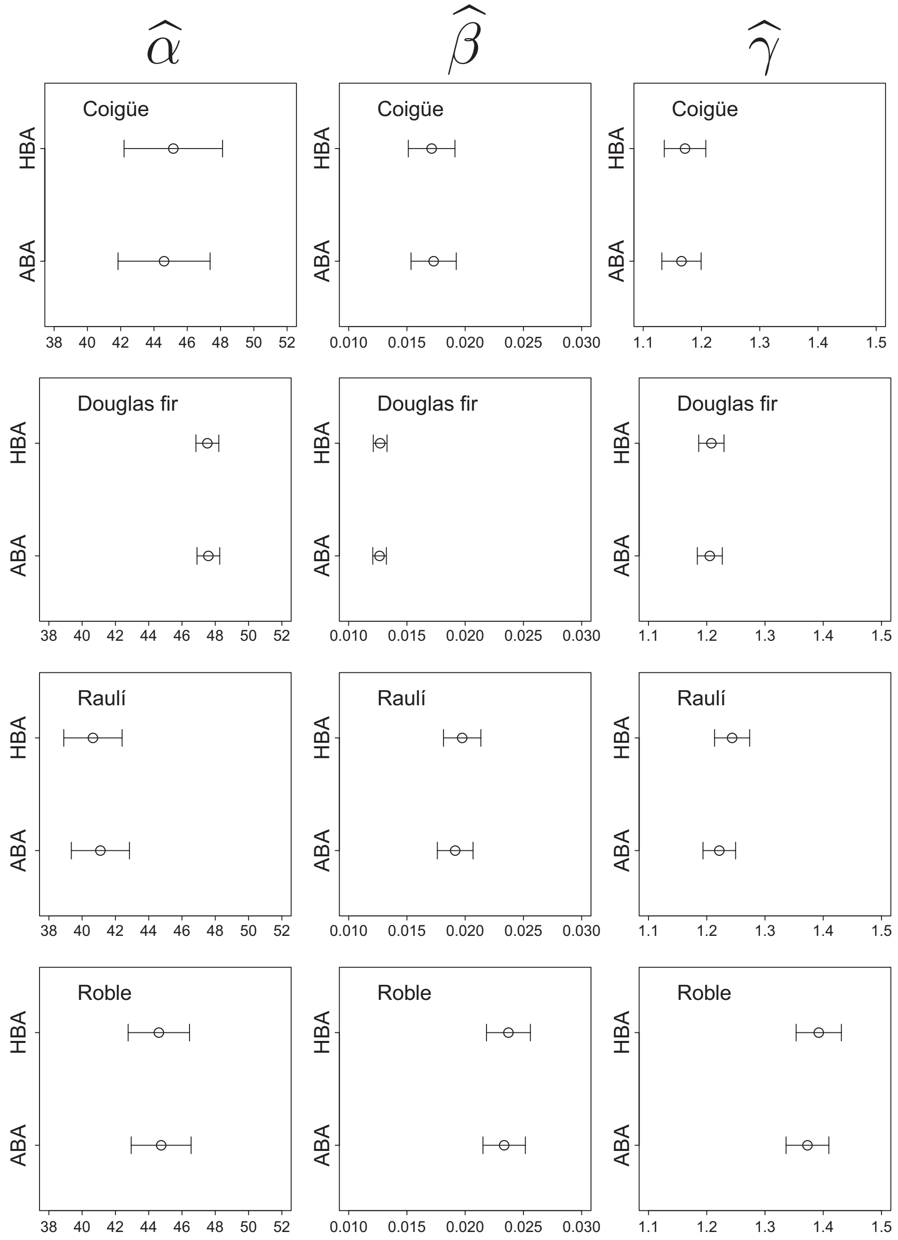
4 | DISCUSSION

I have presented a pleasingly simple novel method for tree height growth reconstruction, which involves preserving the field-measured height data and estimating age by adding 0.5 year to the naive age estimate obtained from ring counts. Though simple, this new method gives results that are nearly equivalent to those of the slightly more complex Carmean algorithm, and thus the new method may be preferred by practitioners.

I fitted the baseline height growth model (Equation 4) in a mixed-effects framework by allocating random effects to different parameters of the model (Equations S5 and S6). The model variant having random effects into the β parameter was always the best for all species, regardless of the algorithm used to generate the data (Table 2). This result agrees with other studies where using the same growth model, the same variant, showed better goodness-of-fit indicators (García, 1983; Hu & García, 2010). Concordantly, using a reliable growth law for modelling the height development of the species under study is secure. Indeed, other models can provide better prediction capabilities. Still, I aim to use a suitable quantitative tool that models the growth rates of the state variable height for further analyses.

Hypothesis testing is not usually taking into account when comparing modelling approaches. For instance, Lappi (2006) focused on only applying a mathematical method to stem analysis data, regardless of the statistical implications. Here, I computed the confidence interval of the estimated parameters by species and the algorithm used to generate the data (Figure 3). Comparing the confidence intervals is essential to assess the null hypothesis of both parameters

FIGURE 3 Asymptotic 95% confidence intervals for the parameter estimates of the polymorphic mixed-effects height growth model (S6) by species, using height–age data built from applying the ABA (age-based, Carmean) and the HBA (height-based) algorithms. Notice that each column panel of the figure represents a parameter estimate, and each row panel represents a species



being equal and to see the drawn conclusion when hypothesis testing is applied. In this regard, the results clearly showed overlapping confidence intervals for the estimated parameters of both data algorithm generated. Therefore, both algorithms would offer almost the same precision on these estimated parameters. The hypothesis to be tested using the variance components of the parameters will reach the same conclusions.

The equivalence testing framework allowed us to perform a reliable assessment of two confronting approaches (or algorithms in this case). The current research possess difficulties not commonly faced when comparing competing models. The stem analysis algorithms reconstruct height–age series (Figure 2; Figure S2); however, I lack the real or observed height–age series, a fundamental component of any model assessment. Inasmuch as the restriction of the absence of actual height–age data, the equivalence testing offers a suitable alternative to focus on assessing whether the predictions of the fitted models fitted from both height–age series are equivalent. Predictions obtained from a fitted model with both algorithms' data are statistically equivalent (Figure 3). These results are consistent across four species' growth data, one coniferous and three broadleafs (two deciduous and one evergreen), spanning a large geographical area. The calculation of both average bias and absolute bias of the algorithms to reproduce the growth laws (Table 4) offers a reliable but straightforward alternative to assess them. Again, the HBA outperformed the Carmean algorithm.

The HBA is more straightforward than the ABA (Carmean) and makes fewer assumptions. ABA is the most widely used stem analysis algorithm, most likely because of the thorough assessment conducted by Dyer and Bailey (1987). Nonetheless, this algorithm assumes a constant periodic annual increment in height for the bolt between two cross-sections and that the hidden tip is reached at half of that annual increment (Figure 2). Both assumptions are barely

TABLE 3 Summary of equivalence-based regression results by species. The equivalence region (ER) for the intercept $\hat{\phi}_0$, termed I_0, was set to the mean measured height ±10%. The ER for the slope $\hat{\phi}_1$, termed I_1, was set to 1 ± 0.25. The 95% bootstrap-based confidence intervals (CI) were computed for both $\hat{\phi}_0$ and $\hat{\phi}_1$. Notice that the proportion of times that each of these parameter estimates were contained within the respective ER was 1 for all the cases

	$\hat{\phi}_0$				$\hat{\phi}_1$			
	I_0		Bootstrap CI		I_1		Bootstrap CI	
Species	**Lower**	**Upper**	**Lower**	**Upper**	**Lower**	**Upper**	**Lower**	**Upper**
Coigüe								
ABA	7.71	12.86	10.78	11.38	0.75	1.25	0.97	1.07
HBA	7.52	12.54	10.49	11.09	0.75	1.25	0.97	1.06
Douglas fir								
ABA	15.37	25.62	19.25	19.77	0.75	1.25	0.87	0.92
HBA	15.27	25.45	19.12	19.63	0.75	1.25	0.88	0.91
Raulí								
ABA	7.54	12.56	9.54	9.98	0.75	1.25	0.936	1.00
HBA	7.35	12.25	9.28	9.72	0.75	1.25	0.943	1.03
Roble								
ABA	8.51	14.18	10.27	10.78	0.75	1.25	0.86	0.95
HBA	8.28	13.80	9.96	10.46	0.75	1.25	0.87	0.94

TABLE 4 Mean bias and absolute bias of the stem analysis algorithms by species

Species	Algorithm	Bias (%)	Absolute bias (%)
Coigüe	ABA	−0.9063	9.0831
	HBA	−0.0105	9.0535
Douglas fir	ABA	−1.2751	15.1443
	HBA	0.0217	15.0737
Raulí	ABA	−0.9506	8.1791
	HBA	−0.0085	8.1404
Roble	ABA	−0.9783	8.8320
	HBA	−0.0181	8.7921

justified, and they are unreliable. On the other hand, the HBA only requires the measured height at a cross-section, and the age for that section is obtained just by subtracting 0.5 (Equation 3); therefore, it is pretty simple. If a user prefers to try the HBA stochastic version, Equation (S3) must be used instead. Although a stochastic version of the ABA could be built similarly, the probability density function (pdf) for defining the randomness's magnitude and frequency is more challenging to justify than a uniform pdf for the HBA. Based on the results presented here, I recommend using the HBA to reconstruct tree height growth with stem analysis data. I have provided an R code (see Salas-Eljatib, 2021b) implementing a simple example from data of a stem analysis sample tree.

One theoretical mathematical issue of the HBA arises because it uses a non-integer time. On the contrary, in the ABA, only integer times are considered. Therefore, the HBA mimics a continuous-time system, concurrently the ABA mimics a discrete-time system. Height growth occurs mainly in early spring for many species, especially those growing in temperate zones (Kimmins, 2004). In this case, the within-year growth dynamics would be flat most of the year. I cannot define a single time associated with height (i.e. height did not change during most of the year). The HBA, however, does not attempt to represent the seasonal pattern of growth. Despite the baseline height growth model being a solution to a differential equation, in applications, I restrict it to integer time values. Leary (1985) advocated for similar uses of differential equation in forest modelling. In practice, for computation purposes, some available alternatives use corrections depending upon the time within the year that we measure the state variable. Other options are available; regardless, they are well beyond the scope of the present paper.

5 | CONCLUSIONS

The stem analysis algorithm presented here increases the flexibility of using this technique for reconstructing time-series data on height growth. This type of data is essential to many areas of ecology and sustainable forest management, and improving their analysis is undoubtedly useful. The two algorithms are presented and compared using modern and suitable statistical models. The proposed algorithm is simpler, based on fewer assumptions, more straightforward than the classical one, and finally yields equivalent results. I finally recommend its use.

ACKNOWLEDGEMENTS

The author thanks Drs Robert Monserud, Pablo Donoso, Hans Grosse, Alicia Ortega, and Pablo Gajardo and Prof. Patricio Núñez for providing the data used here. He especially thanks Dr Oscar García for discussing aspects that he have tried to organize here. The reviewers and the Associate Editor greatly improved earlier versions of the manuscript by providing valuable comments.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13616.

DATA AVAILABILITY STATEMENT

Data and an R code implementing the proposed algorithm are deposited in the Dryad Digital Repository https://doi.org/10.5061/dryad.qnk98sfgc (Salas-Eljatib, 2021b).

ORCID

Christian Salas-Eljatib https://orcid.org/0000-0002-8468-0829

REFERENCES

Carmean, W. H. (1972). Site index curves for upland oaks in the central states. *Forest Science*, *18*, 109–120.

DesRochers, A., & Gagnon, R. (1997). Is ring count at ground level a good estimation of black spruce age? *Canadian Journal of Forest Research*, *27*, 1263–1267.

Dyer, M. E., & Bailey, R. L. (1987). A test of six methods for estimating true heights from stem analysis data. *Forest Science*, *33*, 3–13.

Fabbio, G., Frattegiani, M., & Manetti, M. C. (1994). Height estimation in stem analysis using second differences. *Forest Science*, *40*, 329–340.

García, O. (1983). A stochastic differential equation model for the height growth of forest stands. *Biometrics*, *39*, 1059–1072.

Graves, H. S. (1906). *Forest mensuration* (p. 458). John Willey & Sons.

Hasenauer, H., & Monserud, R. A. (1997). Biased predictions for tree height increment models developed from smoothed 'data'. *Ecological Modelling*, *98*, 13–22. https://doi.org/10.1016/S0304-3800(96)01933-3

Holste, E. K., Kobe, R. K., & Vriesendorp, C. F. (2011). Seedling growth responses to soil resources in the understory of a wet tropical forest. *Ecology*, *92*, 1828–1838. https://doi.org/10.1890/10-1697.1

Hu, Z., & García, O. (2010). A height-growth and site-index model for interior spruce in the Sub-Boreal Spruce biogeoclimatic zone of British Columbia. *Canadian Journal of Forest Research*, *40*, 1175–1183. https://doi.org/10.1139/X10-075

Husch, B., Beers, T. W., & Kershaw, J. A. (2003). *Forest mensuration* (4th ed., p. 443). Wiley.

Kariuki, M. (2002). Height estimation in complete stem analysis using annual radial growth measurements. *Forestry*, *75*, 63–74. https://doi.org/10.1093/forestry/75.1.63

Kimmins, J. P. (2004). *Forest ecology* (3rd ed., p. 611). Pearson Prentice Hall.

Lappi, J. (2006). Smooth height/age curves from stem analysis with linear programming. *Silva Fennica*, *40*, 291–301. https://doi.org/10.14214/sf.344

Leary, R. A. (1985). *Interaction theory in forest ecology and management* (p. 219). Martinus Nijhoff/Dr. W. Junk Publishers.

Lenhart, J. D. (1972). An alternative procedure for improving height/age data from stem analysis. *Forest Science*, *18*, 332.

Machado, S. A., Rodrigues, L. C., Figura, M. A., Teo, S. J., & Mendes, R. G. (2010). Comparison of methods for estimating height from complete stem analysis data for *Pinus taeda*. *Ciência Florestal*, *20*, 45–55.

Milner, K. S. (1992). Site index and height growth curves for ponderosa pine, western larch, lodgepole pine, and Douglar-fir in western Montana. *Western Journal of Applied Forestry*, *7*, 9–14.

Monserud, R. A. (1984). Height growth and site index curves for Inland Douglas-fir based on stem analysis data and forest habitat type. *Forest Science*, *30*, 943–965.

Monserud, R. A. (2003). Modeling stand growth and management. In R. A. Monserud, R. W. Haynes, & A. C. Johnson (Eds.), *Compatible forest management* (pp. 145–175). Kluwer Academic Publishers.

Newberry, J. D. (1991). A note on Carmean's estimate of height from stem analysis data. *Forest Science*, *37*, 368–369.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2021). *nlme: Linear and nonlinear mixed effects models*. R package version 3.1-152.

Poisot, T., Bruneau, A., Gonzalez, A., Gravel, D., & Peres-Neto, P. (2019). Ecological data should not be so hard to find and reuse. *Trends in Ecology & Evolution*, *34*, 494–496. https://doi.org/10.1016/j.tree.2019.04.005

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rayner, M. E. (1991). Estimation of true height from karri (*Eucalyptus diversicolor*) stem analysis. *Australian Forestry*, *54*, 105–108.

Robinson, A. (2010). *equivalence: Provides tests and graphics for assessing tests of equivalence*. R package version 0.5.6.

Robinson, A. P., Duursma, R. A., & Marshall, J. D. (2005). A regression-based equivalence test for model validation: Shifting the burden of proof. *Tree Physiology*, *25*, 903–913. https://doi.org/10.1093/treephys/25.7.903

Robinson, A. P., & Froese, R. E. (2004). Model validation using equivalence tests. *Ecological Modelling*, *176*, 349–358. https://doi.org/10.1016/j.ecolmodel.2004.01.013

Salas-Eljatib, C. (2020). Height growth–rate at a given height: A mathematical perspective for forest productivity. *Ecological Modelling*, *431*, 109198. https://doi.org/10.1016/j.ecolmodel.2020.109198

Salas-Eljatib, C. (2021a). An approach to quantify climate-productivity relationships: An example from a widespread *Nothofagus* forest. *Ecological Applications*. https://doi.org/10.1002/eap.2285

Salas-Eljatib, C. (2021b). Data from: A new algorithm for reconstructing tree height growth with stem analysis data. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.qnk98sfgc

Subedi, N., & Sharma, M. (2010). Evaluating height-age determination methods for jack pine and black spruce plantations using stem analysis data. *Northern Journal of Applied Forestry*, *27*, 50–55. https://doi.org/10.1093/njaf/27.2.50

Tesch, S. D. (1980). The evolution of forest yield determination and site classification. *Forest Ecology and Management*, *3*, 169–182. https://doi.org/10.1016/0378-1127(80)90014-6

Veblen, T. T., Hill, R. S., & Read, J. (Eds.). (1996). *The ecology and biogeography of nothofagus forests* (p. 428). Yale University Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Salas-Eljatib C. A new algorithm for reconstructing tree height growth with stem analysis data. *Methods Ecol Evol*. 2021;00:1411–9. https://doi.org/10.1111/2041-210X.13616