# Trees don’t always act their age: size-deterministic tree-ring standardization for long-term trend estimation in shade-tolerant trees

Rachel Dietrich[1], Madhur Anand[1]
[1]School of Environmental Sciences, University of Guelph, Guelph, N1G2W1, Canada
*Correspondence to*: Madhur Anand (manand@uoguelph.ca)

**Response to reviewers**

We thank both reviewers for their insightful and thoughtful comments on our manuscript. We have implemented a vast majority of the suggestions highlighted in their reviews and in doing so believe the results are more statistically robust and the reasoning is clearer.

To assist comprehensibility of our response this document is structured as follows: 1) Reviewer comment 2) Author response 3) changes to manuscript (if applicable). We will begin our response by addressing the comments of reviewer 2 as they are more extensive and similar to many comments from reviewer 1.

Review 2:

Main comments:

Introduction:

1. *In the current manuscript it is not clear how the proposed methods solve the problem presented in Figure 1. Probably, the size-based solution should also be illustrated in Figure 1.*
    - **a.** This is addressed in the introduction lines 113-116.
2. *The C-method is mentioned but not referenced in the discussion. It should be mentioned in the introduction and cite the paper that describes it - Biondi and Qeadan 2008.*
    - **a.** We agree with the reviewer that mentioning the C-method is important given its prevalent use in tree ring studies. We have amended our analysis to include C-method as one of the tested standardization methods. As requested, Biondi & Qeadan 2008 is referenced in the introduction (line 73) as well as methods (line 202).
3. *The use of similar mixed-effect modelling approaches for tree-ring standardization should also be mentioned in the introduction. It is mentioned only in the discussion in Lines 361 and 402.*
    - **a.** The introduction has been amended to include a more thorough account of the use of explicit BAI models in the literature and their purpose (line 67-70).

Methods

1. *It is not clear if the standardization using the proposed models is applied based on individual series or based on a model fitted to the cloud of all data and then subtracted from each series (as in Fig 4). Please explain it more clearly.*

    a. As with traditional RCS the model is indeed fit to the cloud of data NOT individual series. We have changed the explanation to make it clearer that the function is derived communally (Line 153-155)

2. *For the sake of reproducibility, I recommend the authors to present a worked example with the corresponding R code as supplementary material.*
    a. A sample R code has been included in the supplementary materials S5
3. *It should be explained in the main text how SORTIE simulates tree-ring widths, what is the underlying formulation and the environmental drivers.*
    a. A brief explanation of the calculation is provided on line 211-212. "In SORTIE annual radial tree growth is calculated as an asymptotic function of light availability and previous tree diameter."
4. *As a sensitivity test, the authors should repeat the analysis of Figure 2 for an imposed growth decline and vary the shape of the growth increase to linear and present it as supplementary material. It seems that in Figure 3 the standardization models get a more linear-like increase in growth instead of the sigmoid saturating trend imposed on the synthetic data. To clarify this apparent issue it would help if the mean chronologies of each method are shown as an inset for the last 100 years along with the imposed signal. This would make easier to evaluate if the fitted models suffer from end effects.*
    a. We have added 3 trends to our analysis in both simulated and real tree ring data, the first, a logistic declining trend, is investigated in the main body of the manuscript, while a positive and negative linear trend are interpreted in the supplementary materials. However, we chose not to change Fig 3 as adding the mean chronologies (100 for each method) would decrease from legibility of the figure. We believe the 95% confidence intervals of the resampled mean chronologies adequately show the models' capabilities to reproduce the trends.
5. *Compare the same methods for real world data and not just RCS as currently done.*
    a. Previously the CD and BAI methods were not included in the real tree ring data as they were difficult to evaluate on the basis of parsimony (AIC); BAI because its not an explicit model that allows for AIC calculation and CD because variance explained by the model would be artificially inflated (leading to low AIC) due to inappropriate removal of the long-term trends we are attempting to maintain (and reconstruct in the chronologies). Accordingly, in order to include analysis of BAI, CD (and Cmethod) in real ring data we have adjusted our statistical methodology to be more similar to that performed in the simulated data. Lines (272:289) highlight this methodology. This change in methodology both 1) allows for evaluation of all standardization methods in tree ring data and 2) allows for stronger conclusions regarding the implications of each method for long term trend reconstruction.

Results

1. *Isn't it more logical to start with Figure 3 instead of Figure 2? In this way the reader sees first how the chronologies look like and on what the comparison is based.*
    a. We agree with the reviewer. The figure order has been switched.

2. *In Figure 4 it is clear that the GAM fitting is very noisy at large sizes or ages when there are fewer data points. How much does this noise affect the overall fit? What is the frequency response of the underlying spline in the GAM if any? Melvin et al. 2007 solved this problem by using a time-varying-response smoothing spline, which gets stiffer with age as the data availability declines. Can a similar solution work for this case in the GAM?*
    - **a.** We agree with the reviewer that these are valid concerns and good discussion points regarding the usage of splines in dendrochronological models. However, we believe the assessment of the implications of regression spline parameters for the SDS, RCS and COMB models to be beyond the scope of this paper. Problems with end-fitting and spline frequency choice are not unique to the models presented in this study. To appease the reviewer, we have amended the methods to provide more details on the regression splines used in this study and to provide interested readers with other alternative techniques. (Line 155-158)
3. *What are the different curves in Figure 4 and what are the gray points? It is not stated in the caption*.
    - **a.** We have removed the previous Fig 4 as we do not believe the results presented in it added significantly relevant information.
4. *Why the resulting chronologies are not shown in the current results? I recommend adding a figure with the resulting mean chronologies for each method.*
    - **a.** We have added Fig 4 which presents confidence intervals for the site-wise chronologies produced by each standardization method for both species. As above we present C.I.s not mean chronologies as it eases in interpretation of the figure.
5. *What is COMBred? This comes out of the blue.*
    - **a.** This has been removed.
6. *It is not clear what Figure 5 tells. What does the Rsq mean?*
    - **a.** This figure was removed and replaced with a figure that shows correlations of real tree ring chronologies with imposed trends (similar to Fig 3).

Discussion

1. *The finding that BAI works for recovering mid-frequency growth signals when only large dominant trees are sampled is interesting because it suggests that this method should be less sensitive to the typical big-tree sampling bias of traditional dendrochronological collections.*
    - **a.** We do not believe our results suggest that BAI is less susceptible to big-tree selection bias. In the case of SORTIE simulated data it is less likely that contemporaneous differences in growth rates are significant. As such the probability of big-tree selection bias occurring is low. Further, mortality is stochastic, so slow-grower survivorship bias is unlikely. Accordingly, the only interesting interpretation of this result is that BAI performs poorly when young/ small trees are included in the sample. Lines (462:468) in the discussion highlight that our results should not be used to make conclusions regarding sampling biases.

2. *The discussion should touch on the potential advantages and shared shortcomings of the proposed methods with RCS and BAI in terms of data requirements and biases. How sensitive are the proposed methods to the proportion of aged/unaged trees in the sample and the number of trees in a site?*
    **a.** Biases and data requirements of RCS and BAI are discussed briefly in Lines (462:468) and (477:481). We do not test the sensitivity of COMB method to unaged trees as we believe it to be beyond the scope of the study. The goal of this study was not to provide a review of conventional standardization methods but instead to evaluate new ones in a concise manner, as such we direct the interested reader to an appropriate reference for a systematic review of the use of other standardization methods for long-term trend estimation (line 458). We have added discussion regarding the motivation for this and call for future research in lines (428-433).

Response to reviewer 1:

1. *The method should be better explained:*
   *(a) The reason why the observations are enriched with simulated trees for evaluating the method is only mentioned in the discussion. Move this explanation forward as it may avoid that readers loose attention because they wonder why the dataset is not enough to present the result.*
   **A**: It is not clear to us what explanation or dataset the reviewer is referring to here. Perhaps it is addressed in lines (212-216), but if not we ask to reviewer to provide clarification regarding this comment.

   *(b) It is written that chronologies from different methods were tested with logistic growth-trend for the correlation (L213). To my understanding, a growth trend and chronologies from detrending are contradicting factors because after applying the method, the chronology would be interannual variations remained after removing long-term trend from tree-ring widths.*
   **A:** This is a slight misunderstanding by the reviewer. In this case, the goal of tree ring standardization procedures is to remove age/size related trends from the series but maintain medium and high-frequency time-related variance. RCS, BAI and our proposed models work under the assumption, that by sampling trees from a variety of age/size classes, size/age related variance can be estimated (and removed) independent of time-related variance. This is explained in the methods line (146).

   *(c) For the simulated trees, both shade-tolerant and shade-intolerant species were tested for different methods, but for the data, only shade-tolerant species were selected. What is the reason for this approach? How could it affect the results?*
   **A:** Real tree ring data from shade intolerant trees were not included for simplicity, as the objective of the study was only to test the proposed model in shade tolerant species. Given the physiological justification of the model it is unlikely that the proposed models would produce less accurate results in shade intolerant species relative to tolerant ones.

Evaluated the proposed models is in more tree species is beyond the scope of our study but, nonetheless, we invite future studies to explore the topic (476-480). The justification for using shade intolerant species in simulated data, however, is provided in line (211-213).

2. *Some figures fall short of bringing a visual message.*
   *(a) Figure 2 is difficult to understand. Why is there no CD for the category 'All'? And in figure (a), it looks like BAI has the highest mean for all sampling thresholds but the text lists SDS has having the highest mean correlation (L257). Please, explain this apparent inconsistency.*
   **A:** We failed to explain why CD could not be included in the 'All' category. Line 254 in the methods amends this. When averaged across all sampling thresholds BAI does not produce higher correlations than SDS, mainly because of its unreliability when "all" trees are sampled. This is explained in lines 440+ of the discussion.

   *(b) Figure 4 needs to be improved, or the caption needs to be rewritten. What do individual lines represent? I need more explanations for the figures for COMB. I guess the right-hand figures were redrawn on the same X-axis as the left hand figures so they could be better compared.*
   **A:** This figure has been removed, justification is provided above in *Methods 3*.

   *(c) Are the boxplots left of the dashed line of Figure 5 needed? It seems that the difference between COMB and COMB.red or SDS and SDS.red are not dealt in the discussion.*
   **A:** This figure has been removed. Justification is provided in response to *Methods 5* above.

3. *The authors seem to push for the COMB method but*
   *(a) The better performance of the COMB method is not prominent in the result (See figure 2 and 5). The fact that the figures are difficult to understand may have added to this conclusion.*
   **A:** As explained in *Results 5* above, we have changed our analysis of real tree ring data and adjusted our discussion accordingly. Presently in our discussion we advocate for the SDS and COMB methods for two main reasons 1) they work as well as RCS, and are more reliable than BAI, CD ,CM and 2) allow for inclusion of unaged trees.

*(b) The title says 'trees don't act their age', which is a conflict with the best resulted method that used both age and size to estimate the growth trend. When compared against data, RCS and SDS didn't show much difference. The main point of the title is confusing me.*

**A:** We have amended the title to "Trees don't always act their age"

*(c) It would enhance the readability of the paper a lot if the same set of detrending methods were shown throughout. Now some methods presented in the results are not discussed.*

**A:** We adjusted our statistical methodology to allow for the inclusion of all detrending methods throughout. More details are provided in *Methods 5* above.

*4. To use COMB or SDS, the diameter of the tree at the time of sampling is needed. Hence, I doubt about the applicability of the method for existing huge datasets such as ITRDB because in this data set it is not indicated whether or not the record contains the pith. Could you elaborate on this issue in the discussion?*

**A:** We agree! Unfortunately, this is the case for a large number of standardization methods. Without pith offset information RCS, BAI and C-method cannot be reliably used. So, as it stands none of the methods are applicable to tree ring data from ITRDB (This is why we don't use data from there in this study). A line has been added in discussion (482) to push for more stringent requirements in large databases.

*5. A few times, I felt the first sentence of the paragraph seems to be out of phase with the rest of the paragraph. See for example, L298 and L373.*

**A:** These have both been addressed by reworking paragraph structure and we have proofread with this problem in mind.

# ~~When trees~~Trees don't always act their age: size-deterministic tree-ring standardization for long-tern trend estimation in shade-tolerant trees

Rachel Dietrich[1], Madhur Anand[1]
[1]School of Environmental Sciences, University of Guelph, Guelph, N1G2W1, Canada
*Correspondence to*: Madhur Anand (manand@uoguelph.ca)

## Abstract

With increasing awareness of the consequences of climate change for global ecosystems, the focus and application of tree-ring research has shifted to reconstruction of long-term climate-related trends in tree growth. Contemporary methods for removing the biological growth-trend from tree-ring series (standardization) are ill-adapted to shade-tolerant species, leading to biases in the resultant chronology. Further, many methods, including regional curve standardization (RCS), encounter significant limitations for species in which accurate age estimation is difficult. In this study we present and test two tree-ring standardization models that integrate tree size in the year of ring formation into the estimation of the biological growth-trend. The first method, dubbed size deterministic standardization (SDS), uses tree diameter as the sole predictor of the growth-trend. The second method includes the combined (COMB) effects of age and diameter. We show that both the SDS and COMB methods reproduce long-term trends in simulated tree-ring data better than conventional methods – this result is consistent across multiple species. Further, when applied to real tree-ring data, the SDS and COMB models reproduce long-term, time-related trends as reliably as traditional RCS and more so than common standardization methods (i.e. C-method, BAI, conservative detrending). ~~Further, when applied to real tree-ring data, the COMB method is more parsimonious than its than RCS.~~ We recommend the inclusion of tree size in the year of ring formation in future tree-ring standardization models, particularly when dealing with shade-tolerant species, as it does not compromise model ~~parsimon~~accuracy~~y~~ and allows for the inclusion of unaged trees.


## 1 Introduction

Tree-rings have long-served as a record of environmental change in forest ecosystems. Early dendrochronological studies used tree-ring chronologies from climate sensitive species to elucidate the dynamics of growth-climate relationships and reconstruct climate anomalies from periods before the existence of instrumental records. However, with increasing awareness of the consequences of climate change for global ecosystems, the focus and application of tree-ring research has shifted to reconstruction of low-frequency climate related trends in tree growth (Gedalof and Berg 2010, Boisvenue and Running 2006, Jacoby and D'Arrigo 1997). As it stands, previous optimism regarding the benefits of carbon fertilization for forest growth (Battipaglia et al. 2012, Norby et al. 2005) has been quelled by a lack of consistent evidence in real forests. While many studies have noted increases in long-term growth rates over time in temperate forests (Gedalof and Berg 2010, Huang et al. 2007, Martinelli 2004) others suggest no change (Giguère-Croteau et al. 2019, Camarero et al. 2015, Granda et al. 2014, Silva et al. 2010, Peñuelas et al. 2011). Further, in boreal and drought prone species, growth decline (Chen et al. 2017, Dietrich et al. 2016, Girardin et al. 2012, Silva and Anand 2013) and increased mortality (Herguido et al. 2016, Liang et al. 2016) in response to climate stress have been prevalent. Central to all these studies is the assumption that long-term growth-trends can be accurately and unbiasedly estimated from tree-ring data.

Modern standardization methods are designed to estimate biological age/ size-related effects on tree growth independent of time-related variance, thus theoretically, maintaining long-term trends in the final chronologies. Among these, the conversion of tree ring widths to basal area increments (BAI), and the closely related C-method (Biondi and Qeadan 2008), as well as the use of regional curve standardization (RCS), (Briffa et al. 1992), and its many variants (See Helama et al. 2017), have become commonplace (Peters et al. 2015). Traditional RCS relies on the assumption that the species-specific biological growth trend of local trees can be estimated, and thus removed, from a sufficiently large sample of trees using tree age alone. Alternatively, the BAI method assumes that the biological growth trend is sufficiently related to basal area accrued in a given year and, as such, chronologies presented as BAI (instead of raw ring width) contain minimal biological effects. In practice, it is unlikely that this strict relationship accounts for all the variation in ring width that is related to biological size/ age effects. As such, some studies have proposed explicit models of BAI that attempt to include variables related to tree age/ size or environmental conditions (i.e. tree density, soil fertility etc.), (e.g. Linares et al. 2008, Nock et al. 2011). Similarly, the C-method (CM) assumes that tree-wise basal area increment (tree ring area) distributed over a growing surface in time is constant and as such, annual deviations from this trend can represent the standardized chronology (free from biological trend), (Biondi and Quadan 2008). Both BAI and CM are best suited to open-growth, shade-intolerant trees where the strict relationship between annual growth and expected BAI is not impeded by early competition for light.

~~As it stands, accurate estimation of long-term growth-trends in forests may be limited by poorly adapted tree-ring standardization (age-trend removal) methods (Briffa et al. 1996) and inappropriate sampling methods (Nehrbass-Ahles et al. 2014, Brienen et al. 2012). Early standardization methods (i.e. conservative detrending) were designed to maintain high-frequency variation in tree-ring series and discard long-term, low-frequency variation. It is accepted that these methods are inappropriate for estimating long-term climate related growth-trends (Briffa 1992); however, they are still used in situations where contemporary standardization methods are not applicable due to restrictive data requirements (e.g. Villalba et al. 2012, Gedalof and Berg 2010, Geoff Wang et al. 2006). More recently, the use of regional curve standardization (RCS), and its many variants, as well as the conversion of tree-ring widths to basal area increments (BAI) have become commonplace (Peters et al. 2015). But~~However, due to the difficulties in separating climate related trends that vary on long time scales from those related to biological tree growth and/or succession-related environmental change, neither of these methods are likely to produce accurate estimates of external forcing when trees from only a single age/size class are sampled (Brienen et al. 2012, Briffa and Melvin 2011). While increased awareness of sample biases has led to better prescriptions for study design (see Nehrbass-Ahles et al. 2014, Brienen et al. 2012), systematic tests of the ability of these models to accurately reproduce long-term trends are limited (e.g. Sullivan et al. 2016, Peters et al. 2015, Esper 2010).

Despite these limitations, RCS remains the standard method for estimating long-term growth-trends in tree-ring data (Helama et al. 2017). However, the standard RCS approach encounters large limitations for many species in which accurate age estimation is difficult. Additionally, we suggest the inherent assumption of RCS that biological growth-trends are sufficiently determined by tree age may not be appropriate in all species. More specifically, this assumption is problematic for shade-tolerant trees. Shade-tolerant species exhibit relatively low low-light mortality and thus can persist in forest understories for variable amounts of time before release from overstory light suppression. In these cases, traditional age-deterministic models exhibit high variance, and thus low precision, in the period following tree establishment and leading up to the age when most trees have been released from suppression (Fig. 1). This period of ill-fit means that trees which are released relatively early (or late) from light suppression will exhibit inflated (or deflated) growth relative to the chronology. As a result, the final chronology will show less agreement than would be expected in a shade-intolerant species. Even more problematic, if trees are sampled according to minimum size thresholds, the youngest trees in the chronology are likely to be early-release trees leading to an artificial inflation of modern growth rates in the final chronology. While modifications to traditional RCS that address variance in contemporaneous growth rates and regional environmental conditions have been prevalent in shade-intolerant species (see Helama et al. 2017) there has been little to no focus on the improvement of standardization techniques specific to shade-tolerant tree species.

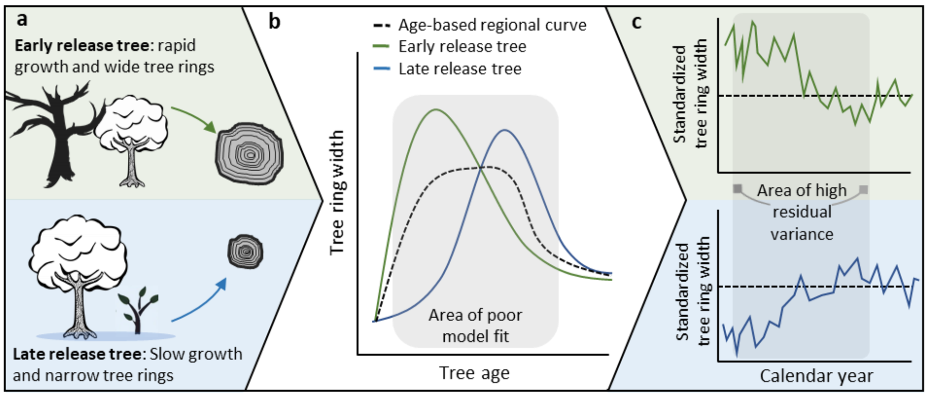


**Figure 1: (a) In shade-tolerant species young trees are stochastically released from low-light suppression in the understory. (b) Since release from suppression is not strictly related to tree age, widely used communal age-trend models (RCS) poorly model tree growth in the period following establishment and leading up to the age when most trees have been released from suppression. (c) Poor model-fit in this period implies that the biological growth-trend is not entirely removed from individual series and leads to high residual variance when standardized tree-ring series are aligned according to calendar year.**

Alternatively, in the field of forest growth and yield modelling, size, rather than age, deterministic predictive growth models are ubiquitous. It is well understood that tree size regulates the capacity for resource acquisition, namely, light (Canham et al. 2004), water and nutrients (Homann et al. 2000), resource allocation (Lehnebach et al. 2018) and metabolic costs (West et al. 2001). As such, the notion of radial growth being deterministic according to size rather than age is logical from both a physiological and ecological perspective. Tree size in a given year is dependent on its previous size and annual growth, so shade-tolerant trees that have yet to be released from overstory light suppression remain small as they grow older. This relaxes the period of ‘ill-fit’ that would be observed in an age-based model. Accordingly, ~~W~~we propose that a size-deterministic model for tree-ring standardization may be more appropriate than traditional RCS for shade-tolerant tree species. The application of size-deterministic models has been limited, with few examples of tree size in a given year being incorporated into BAI models (e.g. Marqués et al. 2016, Camarero et al. 2015, Nock et al. 2011, Martínez-Vilalta et al. 2008) and even fewer of uniquely size-based tree-ring models (e.g. Bontemps and Esper 2011). Further, there have been no systematic evaluations of the ability of size-based models to accurately estimate long-term trends in tree-ring series.

We present two tree-ring standardization models that integrate tree size in the year of ring formation into estimation of the biological growth-trend. The first model uses tree diameter as the sole predictor of the communal growth-trend while the second includes the combined effects of both age and diameter. It follows that the objective of this study is to determine the efficacy of both models in estimating long-term growth-trends in their resultant tree-ring chronologies. First, we use modelled tree-ring data from shade-tolerant and intolerant species to make explicit the inappropriateness of age-based models for shade-tolerant trees. Further, we investigate the performance of size-based models relative to contemporary standardization methods in the presence of size thresholds in tree sampling. Last, we apply the developed models to tree-ring data from shade-tolerant temperate species to evaluate model performance relative to contemporary methods on the basis of accurate reconstruction of known long-term, time-related trends in the series~~model-fit and chronology quality statistics~~.

## 2 Methods

### 2.1 Model formulation

Traditional RCS makes two assumptions about tree growth. First that trees of the same species in a given region exhibit a common growth-trend as they age, and second, that growth of an individual tree in a given year is thus a product of its age and common climatic or environmental forcing in that year (Esper et al. 2003, Briffa et al. 1992). We present a variant of the RCS method that uses tree size, measured by diameter at breast height (DBH), in the year of ring formation as the primary determinant of the common biological growth-trend. As with RCS we assume that the relationship between expected growth and tree size is non-linear and can be approximated for a region from a sufficiently large sample of trees from the species in question. Further, we assume that using a sample of trees from a range of size/age classes ensures estimation of the common trend is not confounded by underlying low-frequency climate or environmental forcing in the chronology (Brienen et al. 2012). The size-based regional curve model, hereafter referred to as the **size deterministic standardization (SDS)** model, takes the following form:

$$(1)\ E(RW_{y,i}) = B_o + f_1(DBH_{y,i}) + e_{yi}$$

Where $E(RW_{yi})$ represents the expected ring width of a given tree (i) in year (y), and $f_1$ represents a non-linear function relating DBH of a given tree (i) in year (y) to $E(RW_{yi})$. As in RCS, the communal non-linear relationship is estimated communally for all local trees of interest. In our study we estimate $f_1$ with a penalized thin plate regression spline in a generalized additive model (GAM), however this relationship could be estimated by a number of different spline fitting or non-linear regression techniques (i.e. *ffcsaps* function in dplR (Bunn et al. 2018), time-varying splines (Melvin et al. 2007)).~~The non-linear relationship is estimated using penalized spline fitting techniques in a generalized additive model (GAM).~~ Under this paradigm the model residuals ($e_{yi}$) represent individual standardized ring width indices and, by extension, individual tree response to climatic or environmental forcing. Annual model residuals subject to a robust mean, thus, represent the final standardized chronology. This approach differs slightly from traditional RCS, whereby occasionally standardized ring width indices are produced by division of raw data by the expected value. Calculation of standardized ring width indices by subtraction from the expected value, as in the case of residuals, is now commonly used as it tends to reduce bias in the resultant chronology (Helama et al. 2004) and eases in the formulation of more complex tree-ring standardization models. However, unlike division methods, the subtraction method does not provide any stabilization of variance in the resulting residuals; as such, it may be necessary to use a stabilization procedure (i.e. log transformation, power transformation) on raw ring width data beforehand.

Tree size in a given year can be estimated by outside-in or inside-out techniques. If the pith of a tree is present in the core (or reasonably close to) $DBH_y$ is a simple summation of all previous ring widths since the year of origin, multiplied by two. Alternatively, if the pith is missed, $DBH_y$ can be calculated via subtraction of more modern ring widths (multiplied by two) from the inside-bark diameter. In this case inside-bark diameter is calculated as the measured DBH minus bark thickness (multiplied by two), where bark thickness can be directly measured or estimated using species-specific allometric equations (e.g. Stayton and Hoffman 1970).

Similar to the model formulation for SDS, RCS models were estimated with GAMs of the following form:

$$(2)\ E(RW_{yi}) = B_o + f_1(Age_{yi}) + e_{yi}$$

Where $Age_{yi}$ is the age of an individual tree in a given year and the resultant standardized tree-ring indices are derived from model residuals ($e_{yi}$).

In addition, a more complex model that integrated independent size and age effects was also evaluated for comparison. This model, hereafter referred to as the **combined model (COMB)**, took the following form:

$$(3)\ E(RW_{yi}) = B_o + f_1(Age_{yia}) + f_2(DBH_{yi}) + e_{iy}$$

In a large variety of long-lived tree species, accurate age estimation (pith sampling) is difficult or impossible; rendering traditional RCS or combined models inappropriate for all trees sampled. To address this issue, the above model incorporates unaged trees. Here $f_1$ represents the non-linear function relating age to expected ring width for the subset of all trees that are aged (ia). In this model, ring widths from unaged trees are assigned arbitrary ages which do not contribute to the linear approximation of the smooth term for Age (i.e. $f_1(Age_{yia})$ but these trees still contribute to the smooth term for size $f_2$ ($DBH_{yi}$). Syntax for missing data in GAMs followed the protocol provided in mgcv (Wood 2011). In this study all GAMs were fit using the mgcv package (Wood 2011) in the R statistical program (v.3.5.0).

In addition to the models presented above we investigated ~~two more~~three additional contemporary standardization methods; conservative detrending (CD), CM and BAI. Conservative detrending describes functions (i.e. negative exponentials, straight lines) or flexible splines fit to individual tree-ring series. In this study we use spline fitting techniques rather than modified negative exponentials as they are more appropriate for shade-tolerant tree species. The C-method estimates tree-specific expected ring widths by assuming constant annual basal area increment (tree ring area) over the life span of the tree (See Biondi and Qeadan 2008). Annual deviations from expected values thus represent standardized ring width indices. For consistency the standard CM approach in dplR (Bunn et al. 2018) was modified in order to calculate indices via subtraction (residuals) instead of division (R code available in Suppl. Materials (S1). Tree ring widths were converted to BAI using the dplR package in R (Bunn et al. 2018).~~Alternatively, BAI attempts to remove biological growth-trends by converting ring widths from individual trees to estimates of annual basal area growth. For simplicity, untransformed BAI was used to compile~~

~~chronologies for this study. Both CD and BAI methods were applied using the dplR package (Bunn 2008) in R.~~

**2.2 Simulated tree-ring data**

~~To evaluate the efficacy of each standardization method in detecting long-term trends, ~~W~~w~~e simulated tree-ring data using a well-established gap-phase model. The SORTIE-ND model was chosen over other similar gap-phase models as it better emulates understory light conditions and low-light mortality both of which are central to the notion of age being an inappropriate determinant of growth in shade-tolerant species. In SORTIE annual radial tree growth is calculated as an asymptotic function of light availability and previous tree diameter. As such, the underlying growth-trend in SORTIE simulated data should be well-approximated by a flexible curve estimated on the basis of tree size (SDS). As such, we use this analysis solely to elucidate the problematic nature of age-based standardization methods for shade-tolerant species not to confirm the efficacy of size-based standardization methods.

For simplicity, a 100% sugar maple (*Acer saccharum)* dominated stand was simulated as sugar maple is a model shade-tolerant species that grows in self-replacing stands. All living trees (>5 cm dbh), (n=3657) in the final year of the model run were used for further analysis. Additionally, to elucidate our claim that age-deterministic growth estimation is more problematic in shade-tolerant species, we completed a similar SORTIE simulation for the shade-intolerant species white pine (*Pinus strobus*). Again, the stand was 100% white pine, standard model parameters were used, and the simulation was run for 1000 years. All living trees (>5 cm dbh), (n=7362) in the final year of the model run were used for further analysis. Additional details regarding model parameters for the SORTIE simulations are provided in the supplementary materials (~~Suppl.~~ S2~~1~~).

**Formatted:** Space After: 0 pt, Don't hyphenate

To simulate a low-frequency climate related growth-trend, a logistic trend was added to raw tree-ring width of individual trees produced by both SORTIE simulations. The logistic trend simulated an initial rapid increase in growth and subsequent levelling off that aimed to represent a period of carbon fertilization and eventual acclimation. The logistic model was applied to the last 100 years of growth and took the following form, where *RWt* represents ring widths with the simulated long-term trend and *RWr* are raw ring widths:

$$(4)\quad RWt_{yi} = RWr_{yi}\left(\frac{k}{1+ae^{-ry}} + 1\right)$$

The logistic trend parameters ($r=0.12$, $k=0.629$, $a=20$~~r, a~~) were chosen such that increases in growth did not exceed 5% of individual average tree growth per decade. Additionally, we tested the standardization models in their ability to detect simulated negative trends in tree growth as previous studies have noted a failure of contemporary methods to accurately reproduce declining growth trends (Peters et al. 2015). The simulated negative logistic trend took the form of eq (4) with parameters ($r=0.12$, $k=-0.421$, $a=20$) chosen such that decreases in growth averaged 5% per decade. For completeness, we also simulated positive and negative linear trends. Results of those analyses are provided in the supplementary materials (S3).

Sixty trees were randomly selected, without replacement, from the simulated tree populations and subject to each of the five standardization methods (SDS, RCS, COMB, CD, BAI). Model residuals (in the case of RCS, SDS and COMB), and standardized (CD) or transformed (BAI) tree-ring widths were compiled into an annual mean chronology using Tukey's biweight robust mean. The resultant chronologies were then tested for significant correlation with the logistic growth-trend using Spearman's rank correlation coefficient. This process was bootstrap resampled 100 times to produce confidence intervals for correlation coefficients.

To examine the effect of minimum size sampling thresholds on the accuracy of long-term trend reconstruction by each of the standardization methods, we completed the same analysis on trees from the simulated populations that exceeded certain size thresholds. The thresholds employed were 10 cm DBH, which represented a practical minimum size threshold for sampling, and 30 and 50 cm DBH which represented thresholds for mature and dominant trees, respectively. The CD method was only applied when size thresholds exceeded 10cm DBH due to the troublesome nature of fitting splines to excessively short timeseries. The mean Spearman's rho for all detrending methods and sampling thresholds were compared using two-way ANOVA and post-hoc tests. ~~Further, two-way ANOVA compared the effect of model choice on Spearman's rho between species (sugar maple and white pine).~~

**2.3 Real tree-ring data**

Additionally, ~~W~~we evaluated the appropriateness of the SDS, COMB and RCS models for use in real tree-ring data from shade-tolerant species. We collected tree-ring data from seven mature sugar maple dominated stands in Ontario, Canada (Table 1). Further, tree-ring data sets from the shade-tolerant species red spruce (*Picea rubens*) were obtained from the DendroEcological Network database (https://www.uvm.edu/femc/dendro), (Table 1). Red spruce was chosen as it had sufficient replication across studies in the database. Descriptions of the sampling strategies and data processing methods for all sites considered are provided in either the supplementary materials (~~Suppl.~~ S4~~2~~) or in their respective references (i.e. Kosiba 2013, Kosiba 2017). Data was considered suitable for this study if age and DBH estimates were provided and if a minimum 20 trees per site and species were sampled. All cores in which pith offset was estimated to be greater than 10 years were considered unaged. ~~The SDS, RCS, and COMB models were fit to tree-ring data from all site-species combinations and the resultant chronologies were compiled~~

~~with a robust mean. In all cases models were fit to log-transformed ring widths, as it increased residual homoskedasticity. For simplicity and ease of model comparison we did not fit CD or BAI models to the real tree-ring data set.~~ To simplify comparisons of the resultant chronologies unaged tree were not included in the models.

~~Model fits from the SDS, RCS and COMB methods were compared according to Akaike information criterion (AIC) and percent variance explained ($R^2$). Since model comparison via AIC requires equal sample sizes, reduced data SDS ($SDS_{red}$) and COMB ($COMB_{red}$) models,~~

**Table 1:**
Location, sample size, chronology length and source of tree ring data sets used in this study.

| Species | Site (code) | Longitude (°) | Latitude (°) | N. trees total | N. trees aged | Length of chronology | Source |
|---|---|---|---|---|---|---|---|
| Sugar maple (*A. saccharum*) | Toobee Lake (TB) | 46.7459 | -82.8668 | 79 | 67 | 1750-2015 | This study |
| | Wolf Mtn. (WM) | 46.7390 | -82.8467 | 22 | 18 | 1827-2015 | … |
| | Roosevelt Road (RS) | 47.2852 | -79.7063 | 20 | 11 | 1792-2016 | … |
| | Raven Lake (RL) | 45.3309 | -78.6339 | 31 | 19 | 1864-2015 | … |
| | Freezy Lake (FR) | 45.2998 | -78.4329 | 20 | 11 | 1887-2015 | … |
| | Kakakise Lake (KK) | 46.0554 | -81.3317 | 22 | 7 | 1773-2016 | … |
| | Mt. Zion Road (MT) | 46.4000 | -83.7004 | 29 | 15 | 1777-2015 | … |
| Red spruce (*P. rubens*) | Mt. Mansfield (MTM) | 44.3750 | -73.8750 | 111 | 109 | 1769-2011 | Kosiba et al. (2016) |
| | Burnt Mtn. (BNT) | 44.2068 | -72.3515 | 40 | 40 | 1891-2010 | Kosiba et al. (2013) |
| | Mt. Carmel (CAR) | 43.7709 | -72.9205 | 41 | 41 | 1795-2010 | … |
| | Mt. Ellen (ELL) | 44.1656 | -72.9221 | 42 | 42 | 1824-2010 | … |
| | Mt. Equinox (EQU) | 43.1487 | -73.1273 | 89 | 89 | 1857-2010 | … |
| | Mt. Greylock (GRY) | 42.6738 | -73.1575 | 44 | 44 | 1911-2010 | … |
| | Hubbard Brook (HUB) | 43.9577 | -71.7350 | 89 | 89 | 1885-2010 | … |
| | Killington Mtn. (KIL) | 43.6146 | -72.8088 | 104 | 103 | 1742-2010 | … |
| | Mt. Mansfield (MAN) | 44.5106 | -72.8297 | 57 | 57 | 1767-2010 | … |
| | Mt. Moosilauke (MOO) | 44.0056 | -71.8215 | 54 | 54 | 1760-2010 | … |
| | Mad River Glen (MRG) | 44.1932 | -72.9232 | 36 | 36 | 1927-2010 | … |

~~which only included aged trees, were also fit. These reduced data models have no practical application but allow for direct AIC comparison between the RCS, $COMB_{red}$ and $SDS_{red}$ models.~~

~~Further, we calculated chronology quality statistics including: mean interseries-correlation, expressed population signal (EPS) and signal-to-noise ratio (SNR), for all chronologies.~~

~~Differences between model fit statistics and quality indices among models were tested using a linear mixed-effect modelling (LME) approach whereby, model error was specified according to site. This approach is analogous to traditional repeated-measures ANOVA but allows for contrast analysis between models.~~

Prior to model application a time-deterministic thin plate regression spline was applied to all raw ring widths from each site. This ensured there was no underlying time-trend present in the data. Since trees of multiple ages/sizes were sampling in each study we assume the removed time-trend is therefore independent of biological trends in the series. For each site residuals from the regression spline were centred according to the site-wise mean and standard deviation of raw ring widths prior to analysis.

Again, increasing and decreasing logistic trends (Eq 4) as well as linear trends (Suppl. S3) were added to the (re-centered) tree ring residuals. Trend parameters were chosen such that the increase (or decrease) in tree growth averaged 5% per decade over the last 50 years of growth ($r=0.12$, $k=0.276$, $a=20$ (positive trend)), ($r=0.12$, $k=-0.226$, $a=20$ (negative trend)). For each site all trees were subject to each of the six standardization methods (SDS, RCS, COMB, CD, BAI, CM). Model residuals (in the case of RCS, SDS, COMB, CD and CM) or transformed (BAI) tree ring widths were compiled into an annual mean chronology using Tukey's biweight robust mean. The resultant chronologies were then tested for significant correlation with the imposed trends using Spearman's rank correlation coefficient. In both species (sugar maple and red spruce) one-way ANOVA and Tukey post-hoc comparisons were used to test for significant differences in model performance- as estimated by chronology correlation with the imposed trend.

## 3 Results

Commented [rd1]: Note: To accommodate new analysis all results have been rewritten and figures redrafted.

### 3.1 Comparisons of methods in simulated data

In order to evaluate the efficacy of each standardization method we calculated correlations between chronologies produced by each method and a variety of imposed trends in simulated sugar maple and white pine tree ring data. Bootstrapped confidence intervals for chronologies from each of the standardization methods are provided in Figure 2a and 2b for sugar maple and red pine, respectively. Distributions of the respective spearman's rank correlation coefficients between the chronologies and the imposed trends are provided in Figure 3a for sugar maple and 3b for white pine.

3.11 Simulated sugar maple tree ring data

In the simulated sugar maple data, two-way ANOVA suggested a significant effect of both standardization model ($p<0.001$) and minimum size sampling threshold ($p<0.001$) on average correlation with the positive logistic trend. Alternatively, for the negative logistic trend there was a significant effect of standardization model ($p<0.001$) but not of size sampling threshold. For both positive and negative logistic trends SDS ($\overline{r_s}$=0.974±0.037, $\overline{r_s}$=0.954±0.068, respectively) and COMB ($\overline{r_s}$=0.965±0.039, $\overline{r_s}$=0.894±0.123, respectively) models produced chronologies with significantly higher correlations than all other models ($p<0.001$ for all) but not significantly different from each other (p=0.998, p=1.000, respectively). For the positive imposed trend BAI ($\overline{r_s}$=0.864±0.236) and RCS ($\overline{r_s}$=0.900±0.162) produced chronologies with correlations significantly higher than CD ($\overline{r_s}$=-0.503±0.329) and CM ($\overline{r_s}$=0.746±0.306), ($p<0.001$ for all) but not significantly different then each other (p=0.996). Notably, correlations exhibited by BAI chronologies were dependent on size sampling thresholds with BAI chronologies performing best when size thresholds exceeded 50 cm DBH (Fig 3a). At this threshold BAI chronologies produced significantly higher correlations than when all trees were sampled (p=0.003) and when trees >10 cm DBH were sampled (p<0.001). The CD method produced chronologies that

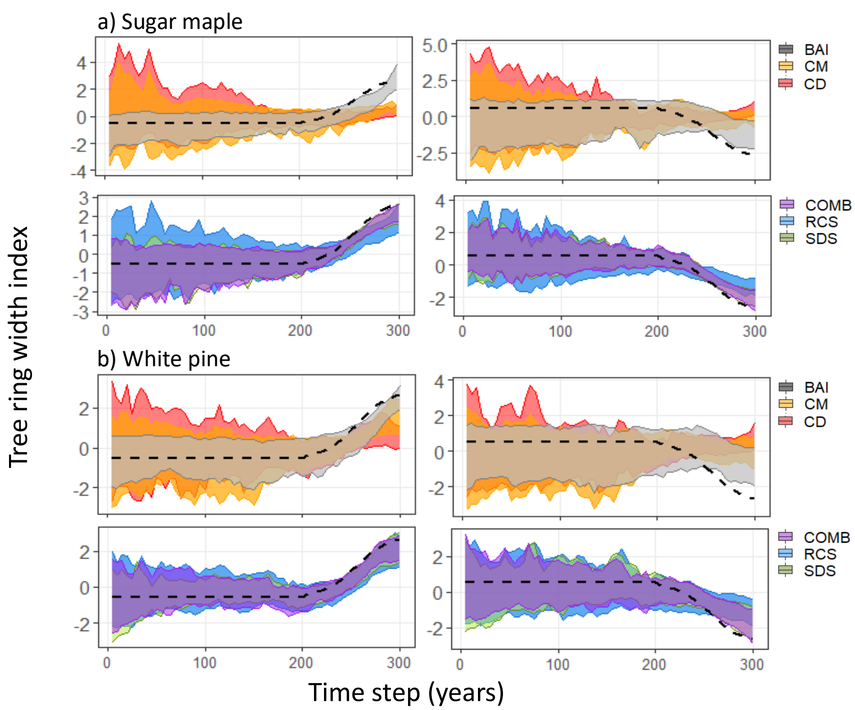


**Figure 2***:* 95% confidence intervals for standardized chronologies produced by each standardization method (legend right side) applied SORTIE simulated sugar maple and white pine tree ring data. Confidence intervals obtained via bootstrap resampling (rep=100) of 60 trees (>10 cm DBH) from the SORTIE simulated populations. Dotted lines indicate the standardized positive (left side) or negative (right side) logistic trend that was added to the raw tree ring data.

exhibited the lowest average correlation with the imposed positive trend of all models (p<0.001 for all).

Alternatively when considering negative imposed trends, BAI ($\overline{r_s}$=0.745±0.426) chronologies performed significantly worse than RCS ($\overline{r_s}$=0.706±0.281, p<0.001) but still better than CD ($\overline{r_s}$=-0.609±0.291) and CM ($\overline{r_s}$=0.666±0.364), (p<0.001 for both). Again, CD chronologies exhibited significantly lower correlations than all other models (p<0.001 for all). Notably, RCS chronologies produced at the 50 cm DBH sampling threshold exhibited significantly lower

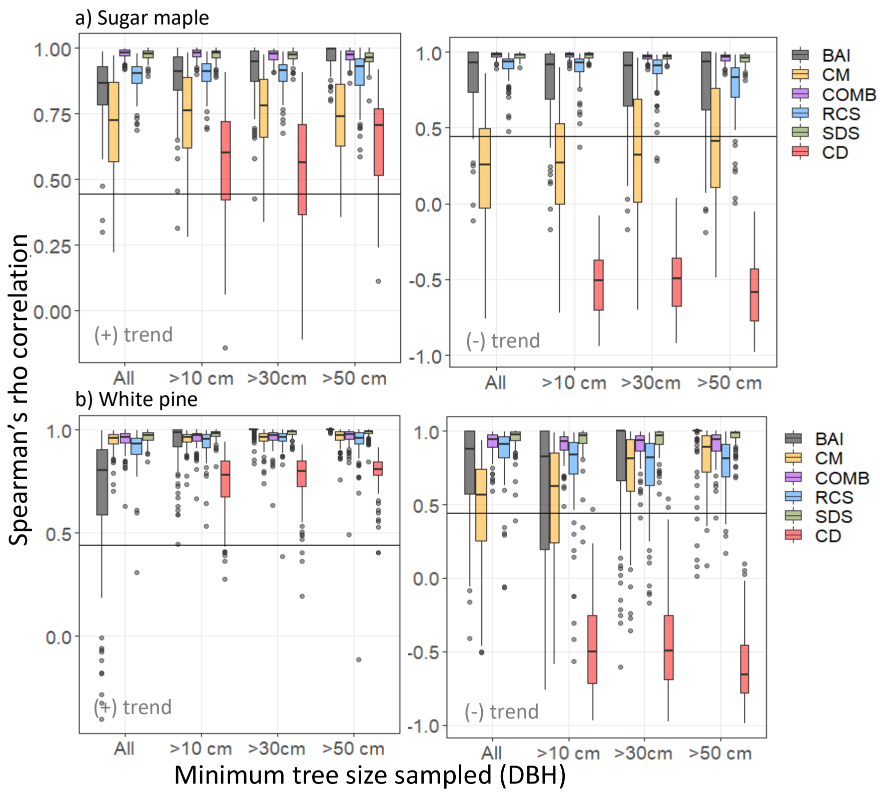

**Figure 3:** Spearman's correlation between chronologies produced by each of the five standardization methods and the imposed positive (left column) or negative (right column) logistic trend in SORTIE simulated (a) sugar maple and (b) white pine tree-ring data. Correlation distribution created by bootstrap resampling 60 trees (rep=100) from SORTIE simulated tree populations.  Horizontal axis denotes minimum tree size (DBH) thresholds for sampling from the population. Horizontal lines indicate threshold for significant Spearman's rho (a=0.05) for correlation between chronologies and the imposed trend.

correlations than all other sampling thresholds ($p<0.001$), (Fig 3a). All other models exhibited similar correlation distributions across the various size thresholds for sampling.

### 3.12 Simulated white pine tree ring data

In simulated white pine data, two-way ANOVA suggested a significant effect of both standardization model ($p<0.001$) and minimum size sampling threshold ($p<0.001$) on average correlations for both the positive and negative logistic trend analyses. For the positive trend, chronologies produced by SDS ($\bar{r}_s$=0.977±0.026), RCS ($\bar{r}_s$=0.932±0.091), COMB ($\bar{r}_s$=0.956±0.052) and CM ($\bar{r}_s$=0.953±0.045) produced high correlations across all sampling thresholds with SDS performing significantly better than CM ($p=0.006$) and RCS ($p=0.001$). All four models produced significantly higher correlations than those produced by BAI ($\bar{r}_s$=0.899±0.222) or CD ($\bar{r}_s$=0.767±0.126) chronologies, with CD producing the lowest correlations of all models. Contrasts suggested that the significant effect of minimum size threshold was driven by significant differences in correlations from BAI chronologies across sample thresholds, whereby BAI chronologies exhibited significantly lower correlations when no minimum size thresholds (i.e. all trees sampled) were employed ($p<0.001$ in all cases), (Fig. 3b). When examining negative imposed trends, SDS ($\bar{r}_s$=0.942±0.090) and COMB ($\bar{r}_s$=0.904±0.0.97) models produced chronologies with significantly higher correlations than all the other models, but not significantly different from each other ($p=0.594$). BAI ($\bar{r}_s$=0.750±0.390) and RCS ($\bar{r}_s$=0.772±0.245) produced chronologies with correlations significantly higher than CD ($\bar{r}_s$=-0.505±0.316) and CM ($\bar{r}_s$=0.623±0.362), ($p<0.001$ for all) but not significantly different then each other ($p=1.00$). CD chronologies exhibited significantly lower correlations than all other models ($p<0.001$ for all). Contrasts suggested that the significant effect of minimum size threshold was driven by significant difference in correlations of chronologies produced by BAI and CM among sampling thresholds. As evident in Figure 3b, BAI chronologies performed significantly better when sampling thresholds exceeded 50 cm DBH and CM chronologies performed best when sampling thresholds exceeded 30 cm DBH.

### 3.2 Comparisons of methods in real tree-ring data

Standardization methods were evaluated on the basis of correlations between their resultant chronologies and known time-related trends in tree ring series from shade-tolerant species. Confidence intervals surrounding chronologies produced from each of the standardization methods applied to the tree ring series from six sugar maple stands are provided in Figure 4a for both positive and negative logistic trends. The corresponding distributions of Spearman's rank correlation coefficients are provided in Figure 5a with significant differences ($p<0.05$) being denoted by letters. Chronologies and corresponding correlation coefficients for the identical analysis performed on 12 red spruce stands are provided in Figure 4b and 5b.

Regardless of trend direction RCS, COMB and SDS chronologies exhibited comparable and consistent results across both species (Fig. 5). In general chronologies produced by all three methods exhibited conservative, but reliable, estimations of the imposed trends (Fig. 4). SDS produced chronologies with correlations as high or higher (Fig. 5b (negative trend)) than traditional RCS chronologies. Notably, the BAI and CM methods produced strong positive correlations between chronologies and the imposed trend only when the imposed trend was increasing (Fig. 4, 5) but both consistently failed to reproduce negative trends (Fig. 4). Finally, across both species, CD chronologies exhibited low correlations with the imposed trend regardless of direction (Fig. 4,5).

## 4 Discussion

### 4.1 Size-vs age-deterministic models for long-term trend reconstruction

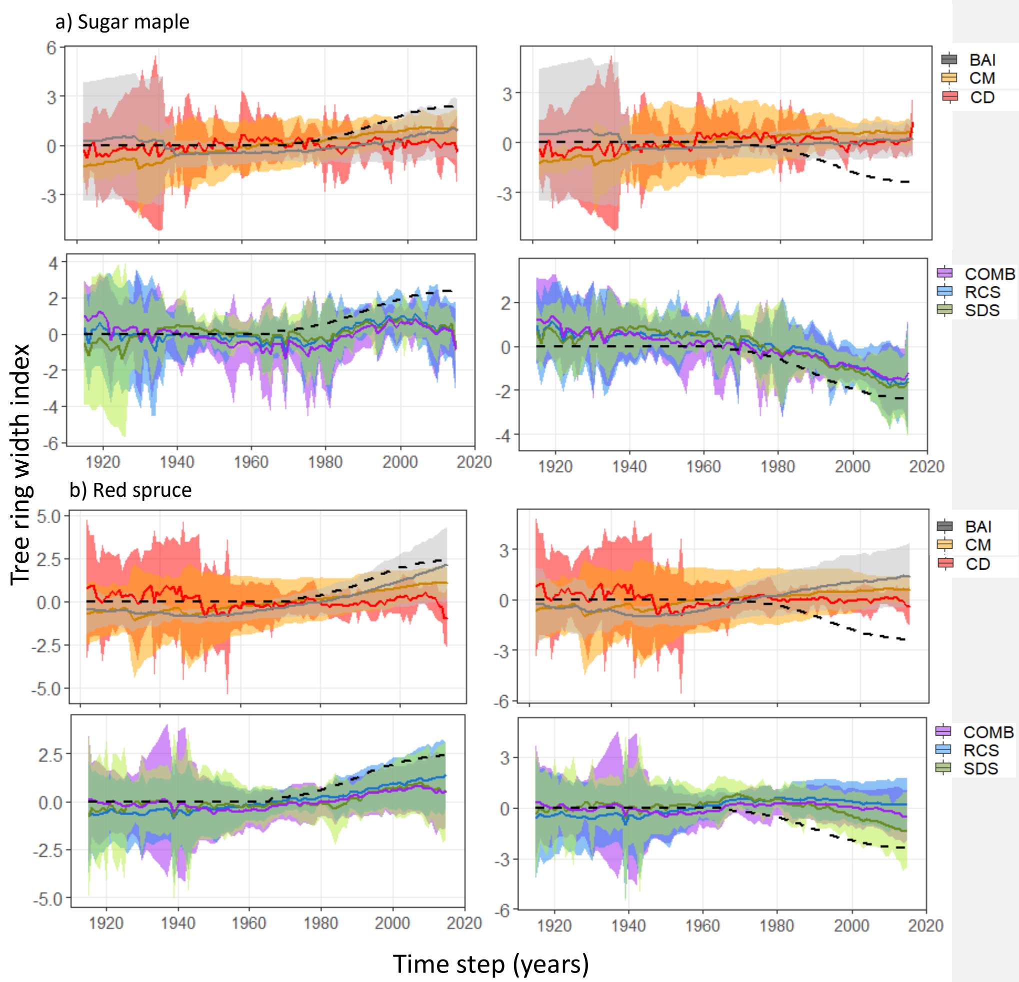


**Figure 4:** Standardized chronologies produced by each standardization method (legend right side) applied to tree ring series from a) sugar maple (n=6) and b) red spruce (n=12) stands. Solid lines represent the resultant model-wise mean chronologies across all stands considered while ribbons represent respective 95% confidence intervals. Dotted lines indicate the standardized positive (left side) or negative (right side) logistic trend that was added to the raw tree ring data.

Using simulated tree-ring data, from the shade-tolerant species sugar maple, we have shown that standardization models which include tree size in the year of ring formation (SDS, COMB) produced chronologies that retain long-term/low-frequency variation better than those produced by models that only include age as a predictor (RCS). Alternatively, in the shade-intolerant species white pine, chronologies produced by the ~~SDS,~~ RCS and COMB models showed no significant difference in their estimation of long-term trends, though SDS chronologies slightly outperformed RCS chronologies.~~.~~ ~~Further, our analysis suggests that the traditional RCS method performed significantly worse in the shade-tolerant species, sugar maple, than in shade-intolerant white pine.~~ As discusses previously, ~~T~~the finding that size-based standardization models performed well in simulated tree-ring data is not surprising given that the SORTIE model calculates annual tree growth as function of tree size. Thus, the underlying growth-trend would be well-approximated by a flexible curve estimated on the basis of tree size. As such, we use these results solely to elucidate the problematic nature of age-based standardization methods for shade-tolerant species. SORTIE's use of diameter, rather than age, as a determinant of tree growth is not arbitrary; it is well established that tree metabolic processes are directly related to size (West et al. 2001). Additionally, there is little evidence for a unique effect of age on tree growth that is independent

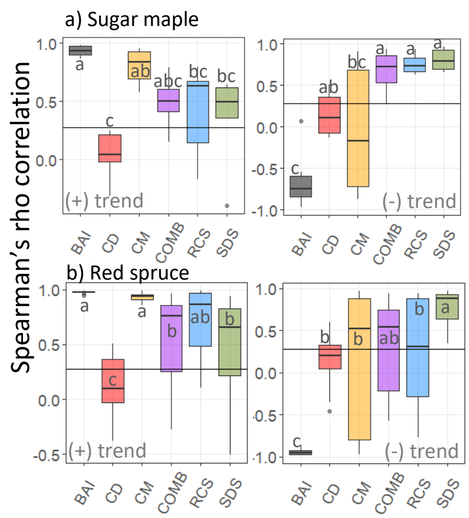


**Figure 5:** Spearman's correlation between chronologies produced by each of the five standardization methods and the imposed positive (left column) or negative (right column) logistic trend in tree ring series from (a) sugar maple and (b) red spruce stands. Horizontal lines indicate threshold for significant Spearman's rho (a=0.05) for correlation between chronologies and the imposed trend. Letters indicate significant differences among samples as estimated by Tukey honest significant differences (a=0.05).

of size (Munné-Bosch 2007 (and within)). With the exception of dendrochronological models, the vast majority of individual tree growth and process models are indeed size-based. It follows that the ubiquitous use of age or calendar year in tree-ring standardization methods (RCS, signal-free standardization, C-method, CD, Hugershoff curves) is a practice born out of convenience rather than physiological consideration. As such, we agree with previous accounts that this assumption may be especially problematic in shade-tolerant trees where age and size may not be perfectly correlated (Peters et al. 2015, Bontemps and Esper 2011).

Unfortunately, all systematic comparisons of tree-ring standardization methods in real tree-ring data (e.g. Sullivan et al. 2016) are limited by their inability to validate long-term trends estimated by chronologies. In this study we evaluate standardization methods on their ability to reconstruct artificial trends in tree ring data. ~~Instead, we evaluate standardization models on the basis of model parsimony. We have shown that in the shade-tolerant species red spruce, COMB models are significantly more parsimonious (estimated by AIC) than simpler models (RCS, SDS). Further, the COMB models explain more variance (estimated by $R^2$) in tree-ring data regardless of differences in underlying sample sizes~~. ~~Overall, our results are conservative relative to similar comparisons performed by Nock et al. (2011) in tropical tree species of varying shade-tolerance. Nock et al. (2011) note that LMEs of BAI that included tree diameter had more support than those that included age. In line with discussion above, Nock et al. (2011) attribute this finding to size being a more important determinant of light capture as it relates to tree height and crown size (King et al. 2005). Further, in both red spruce and sugar maple we have shown that tree size and age exhibit stronger relationships with average growth when their unique effects are estimated simultaneously in COMB models rather than alone in SDS and RCS models, respectively. This result is interesting given the high correlation expected between these~~

~~variables and it may explain why COMB models explained significantly more variance than each of the simpler models. Given the relatively weak trends shown in predictors from both the SDS and RCS, models we suggest that low-frequency variance related to the underlying biological growth-trend may be retained in these chronologies.~~

~~Regardless of differences in model fits, the implications for the resultant chronologies remain conservative (Fig. S.2). Similarly, in comparison of RCS and SDS chronologies in common beech (*Fagus sylvatica* L.) Bontemps and Esper (2011) note both chronologies exhibit similar annual variations~~. We show that SDS and COMB models are as reliable as the traditional RCS method in accurately detecting long-term trends in shade-tolerant species. Further, SDS appears to provide more reliable reconstructions when the underlying trend is negative. To our knowledge, only one other study has evaluated size-deterministic models on the basis of long-term trend reconstruction in chronologies. Bontemps and Esper (2011) compared RCS and SDS chronologies in common beech (*Fagus sylvatica* L.)) and conclude that both exhibit similar variations, with the magnitude of difference varying between 3-7%. However, other studies have examined the influence tree size in explicit models of BAI. In tropical tree species of varying shade-tolerance Nock et al. (2011) note that linear mixed models of BAI that included tree diameter had more support than those that included age. This result is corroborated by analyses of mixed models of BAI in Mediterranean pine species which suggest that the effect of DBH on BAI is more important than the effect of tree age (Marqués et al. 2016). In line with discussion above, Nock et al. (2011) attribute this finding to size being a more important determinant of light capture as it relates to tree height and crown size (King et al. 2005).

The resultant chronology is more likely to be influenced by sample size of the underlying tree population than by choice of standardization model. Tree age can be difficult or impossible to accurately estimate for some trees. In contrast, annual tree size can be reliability estimated from DBH and tree-ring measurements more ubiquitously. We note that in this study only 66% of sugar maple trees could be accurately aged. Since unaged trees are likely to be the oldest trees in the chronology, it follows that RCS chronologies may exhibit poor sample replication (especially in early years) and may be significantly shorter than those produced by SDS or COMB models. This has obvious implications for data quality and suitability. Considerably problematic is the "segment length curse" whereby, almost all standardization methods are ill-equipped to estimate long-term trends on time scales greater than or equal to the length of the chronology itself (Cook et al. 2005). Excessively short RCS chronologies are therefore limited in their application. A large advantage of SDS and COMB models is that they can incorporate otherwise inadmissible tree-ring data.

This study does not explicitly test the efficacy of COMB models relative to SDS in the presence of unaged trees. Nor have we provided evidence to suggest that the added complexity of COMB models relative to SDS is beneficial to accurate reconstruction of trends in the resultant chronologies. Given, the merit the of size-deterministic models presented here, we suggest future research explore the implications of the trade-off between model information and complexity in the presence of unaged trees.

### 4.2 BAI and CD methods for long-term trend reconstruction

The finding that CD did not produce accurate long-term trends in simulated tree-ring data is consistent with our expectations (Peters et al. 2015, Briffa et al. 1992). We maintain CD should be avoided if the goal is long-term reconstruction from tree-ring data. More interestingly, we have shown that CM and BAI, although designed for shade-intolerant open growth trees, do not reliably reconstruct negative long-term trends in simulated white pine tree ring data.

~~BAI chronologies accurately reproduced long-term trends in simulated tree-ring data.~~

~~However~~Further, our analysis suggests BAI is less reliable when small/young trees are sampled. This result is corroborated in our study by a failure of both methods to reconstruct negative trends in shade-tolerant, sugar maple and red spruce, tree ring data.Further, t~~T~~his finding ~~wa~~is in line with Peters et al. (2015) who note low reliability of BAI ~~chronologies to imposed long-term trends, but~~and that BAI is likely to produce erroneous trends when the underlying trend is of low signal, as would be the case for young/small trees that have low BAI rates and low climate sensitivity. ~~As presented here, the~~

Both BAI ~~method~~ and CM impart~~s~~ a strict relationship between tree size and growth. It has been suggested that this relationship may not account for the entire biological growth-trend, leading to the maintenance of erroneous long-term trends in the resultant chronologies (Peters et al. 2015). Erroneous increasing trends are indeed noted in both sugar maple (Fig 4a) and red spruce (Fig 4b) chronologies produced by BAI and CM in our study. Accordingly, we caution future studies in their interpretation of BAI trends in low-signal tree-ring series. ~~Alternatively, O~~Other studies have explicitly modelled size and/or age effects on BAI using a mixed-effect modelling approach (e.g. Marqués et al. 2016, Camarero et al. 2015, Nock et al. 2011, Martínez-Vilalta et al. 2008). We suggest this approach may better account for species- and site-specific factors that influence expected growth rates, leading to more accurate estimates of long-term trends in the resultant chronology. While our findings regarding the importance of inclusion of size in tree-ring standardization models are presented in the context of raw tree-ring width models, they are also directly relevant to explicit models of BAI. A more thorough discussion of the limitations of CD, BAI and CM method as relevant to reconstruction of long-term trends is beyond the scope of this study. The interested reader is directed to Peters et al. (2015).

**4.3 Other considerations and future research**

It is important to note that the goal of this study was not to explicitly test the effect of sample biases (i.e. modern sample bias, selection bias, etc.) on trend reconstruction, but instead to assess reliability across different underlying sampling distributions. Accordingly, our results do not suggest that any of the discussed standardization methods are immune to sample biases (i.e. big tree selection bias, slow grower survivorship bias) as our study is not designed to detect, and isolate, the effects of contemporaneous differences in growth among trees that lead to these biases. There is now considerable evidence to suggest that the long-standing practice of sampling only dominant trees or trees exceeding a minimum size threshold within a stand leads to considerable bias in the resultant chronology (Nehrbass-Ahles et al. 2014, Brienen et al. 2012, Briffa and Melvin 2011). This bias is consistent across standardization methods (Nehrbass-Ahles et al. 2014). We maintain that in cases of long-term trend reconstruction, stands should be sampled according to the underlying stand age/size distribution, either through use of fixed-plots or random tree selection, regardless of the standardization procedure used.

~~Our study has suggested that the choice of standardization model (SDS, RCS, COMB) has no discernable effect on indices of chronology quality (EPS, SNR, interseries-correlation). We suggest this finding is a result of the chosen species exhibiting low climate sensitivity (Phipps 1982) and thus, low common signal in the chronology. As such we do not regard this finding as failure of any of the standardization models. We suspect more conclusive results would be found in climate sensitive species.~~ Given the underlying physiological justification of the models presented here, we have no reason to suggest they are not broadly applicable to species of all shade-tolerance levels. We recommend future studies investigate the applicability of SDS and COMB models to both raw tree-ring width and BAI data in wider range of species. That said, shade-tolerant and broadleaf species, and their applicable standardization procedures, are underrepresented in dendrochronological studies (Zhao et al. 2019). Further, the applicability of enhanced tree ring standardization models (including traditional RCS and BAI) to global tree ring data sets is limited by widely unavailable metadata (i.e. tree age and DBH) in tree ring databases. Accordingly, we recommend more stringent requirements on the inclusion of applicable metadata in global databases in order to accommodate more complicated standardization models. We advocate for continued refinement of tree-ring standardization procedures that are relevant to the ecological questions they aim to address.

**Acknowledgements**

We are grateful to two anonymous reviewers whose thoughtful and thorough comments greatly improved the impact and intelligibility of this paper. We thank the National Science and Engineering Research of Council of Canada for scholarship support for R. Dietrich and research funds to M. Anand (NSERC Discovery) and the MNRF Climate Science program for funding for field work related to this study. We are grateful to staff of the Ontario Forest Research Institute that supported field work for this study, namely F. Wayne Bell. Further, we thank Ontario Parks and the Haliburton Forest Reserve for providing access to field sites. All tree-ring data used in this study are available in the DendroEcological Network database (https://www.uvm.edu/femc/dendro). All SORTIE-ND simulation data are available by request from the corresponding author.

**References**

Battipaglia, G., Saurer, M., Cherubini, P., Calfapietra, C., McCarthy, H. R., Norby, R. J., & Francesca Cotrufo, M.: Elevated CO 2 increases tree-level intrinsic water use efficiency: insights from carbon and oxygen isotope analyses in tree-rings across three forest FACE sites, New Phytol., 197(2), 544-554, 2013.

Boisvenue, C., & Running, S. W.: Impacts of climate change on natural forest productivity–evidence since the middle of the 20th century. Global Change Biol., 12(5), 862-882, 2006

Bontemps, J. D., & Esper, J.: Statistical modelling and RCS detrending methods provide similar estimates of long-term trend in radial growth of common beech in north-eastern France. Dendrochronologia, 29(2), 99-107, 2011.

Brienen, R. J., Gloor, E., & Zuidema, P. A.: Detecting evidence for CO2 fertilization from tree-ring studies: The potential role of sampling biases. Global Biogeochem. Cycles, 26(1). 2012.

Biondi, F. & Qeadan, F.: A theory-driven approach to tree-ring standardization: defining the biological trend from expected basal area increment. Tree-Ring Research, 64(2), pp.81-97, 2008.

Briffa, K. R., Jones, P. D., Bartholin, T. S., Eckstein, D., Schweingruber, F. H., Karlen, W., ... & Eronen, M.: Fennoscandian summers from AD 500: temperature changes on short and long timescales. Climate Dyn., 7(3), 111-119, 1992.

Briffa, K. R., Jones, P. D., Schweingruber, F. H., Karlén, W., & Shiyatov, S. G.: Tree-ring variables as proxy-climate indicators: problems with low-frequency signals. In Climatic variations and forcing mechanisms of the last 2000 years (pp. 9-41). Springer, Berlin, Heidelberg, 1996.

Briffa, K. R., & Melvin, T. M.: A closer look at regional curve standardization of tree-ring records: justification of the need, a warning of some pitfalls, and suggested improvements in its application. In Dendroclimatology (pp. 113-145). Springer, Dordrecht. 2011.

Bunn, A. G.: A dendrochronology program library in R (dplR). Dendrochronologia, 26(2), 115-124, 2008.

Canham, C. D., LePage, P. T., & Coates, K. D.: A neighborhood analysis of canopy tree competition: effects of shading versus crowding, Can. J. For. Res., 34(4), 778-787, 2004.

Camarero, J. J., Gazol, A., Tardif, J. C., & Conciatori, F.: Attributing forest responses to global-change drivers: limited evidence of a CO 2-fertilization effect in Iberian pine growth, J. Biogeogr., 42(11), 2220-2233, 2015

Cook, E. R., Briffa, K. R., Meko, D. M., Graybill, D. A., & Funkhouser, G.: The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies, Holocene, 5(2), 229-237, 1995.

Chen, L., Huang, J. G., Dawson, A., Zhai, L., Stadt, K. J., Comeau, P. G., & Whitehouse, C.: Contributions of insects and droughts to growth decline of trembling aspen mixed boreal forest of western Canada, Global Change Biol., 24(2), 655-667, 2018.

Dietrich, R., Bell, F. W., Silva, L. C., Cecile, A., Horwath, W. R., & Anand, M.: Climatic sensitivity, water-use efficiency, and growth decline in boreal jack pine (Pinus banksiana) forests in Northern Ontario, J. Geophys. Res. Biogeosci, 121(10), 2761-2774, 2016.



**Formatted:** Space After: 0 pt, Don't hyphenate

Duchesne, L., Houle, D., Ouimet, R., Caldwell, L., Gloor, M., & Brienen, R.: Large apparent growth increases in boreal forests inferred from tree-rings are an artefact of sampling biases. Scientific reports, 9(1), 6832, 2019.

Esper, J., Cook, E. R., Krusic, P. J., Peters, K., & Schweingruber, F. H.: Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies, Tree-Ring Res., 59(2), 81-98, 2003.

Gedalof, Z. E., & Berg, A. A.: Tree-ring evidence for limited direct CO2 fertilization of forests over the 20th century, Global Biogeochem. Cycles, 24(3), 2010.

Geoff Wang, G., Chhin, S., & Bauerle, W. L.: Effect of natural atmospheric CO2 fertilization suggested by open-grown white spruce in a dry environment, Global Change Biol., 12(3), 601-610, 2006.

Giguère-Croteau, C., Boucher, É., Bergeron, Y., Girardin, M. P., Drobyshev, I., Silva, L. C., ... & Garneau, M.: North America's oldest boreal trees are more efficient water users due to increased [CO2], but do not grow faster, Proc. Nat. Acad. Sci., 201816686, 2019.

Girardin, M. P., Bernier, P. Y., Raulier, F., Tardif, J. C., Conciatori, F., & Guo, X. J.: Testing for a CO2 fertilization effect on growth of Canadian boreal forests, J. Geophys. Res. Biogeosci, 116(G1), 2011.

Granda, E., Rossatto, D. R., Camarero, J. J., Voltas, J., & Valladares, F.: Growth and carbon isotopes of Mediterranean trees reveal contrasting responses to increased carbon dioxide and drought, Oecologia, 174(1), 307-317, 2014.

Helama, S., Lindholm, M., Timonen, M., & Eronen, M.: Detection of climate signal in dendrochronological data analysis: a comparison of tree-ring standardization methods, Theor. Appl. Climatol., 79(3-4), 239-254, 2004.

Helama, S., Melvin, T. M., & Briffa, K. R.: Regional curve standardization: State of the art, Holocene, 27(1), 172-177, 2017.

Herguido, E., Granda, E., Benavides, R., García-Cervigón, A. I., Camarero, J. J., & Valladares, F.: Contrasting growth and mortality responses to climate warming of two pine species in a continental Mediterranean ecosystem, For. Ecol. Manage., 363, 149-158, 2016.

Homann, P. S., McKane, R. B., & Sollins, P.: Belowground processes in forest-ecosystem biogeochemical simulation models, For. Ecol. Manage., 138(1-3), 3-18, 2000.

Huang, J. G., Bergeron, Y., Denneler, B., Berninger, F., & Tardif, J.: Response of forest trees to increased atmospheric CO2, CRC Crit. Rev. Plant Sci., 26(5-6), 265-283, 2007.

Jacoby, G. C., & D'Arrigo, R. D.: Tree-rings, carbon dioxide, and climatic change, Proc. Nat. Acad. Sci., 94(16), 8350-8353, 1997.

King, D. A., Davies, S. J., Supardi, M. N., & Tan, S.: Tree growth is related to light interception and wood density in two mixed dipterocarp forests of Malaysia, Funct. Ecol., 19(3), 445-453, 2005.

Formatted: Space After: 0 pt, Don't hyphenate

Kosiba, A. M., Schaberg, P. G., Hawley, G. J., & Hansen, C. F.: Quantifying the legacy of foliar winter injury on woody aboveground carbon sequestration of red spruce trees, For. Ecol. Manage., 302, 363-371, 2013.

Kosiba, A. M., Schaberg, P. G., Rayback, S. A., & Hawley, G. J.: Comparative growth-trends of five northern hardwood and montane tree species reveal divergent trajectories and response to climate, Can. J. For. Res., 47(6), 743-754, 2017.

Lehnebach, R., Beyer, R., Letort, V., & Heuret, P.: The pipe model theory half a century on: a review, Ann. Bot., 121(5), 773-795, 2018.

Liang, E., Leuschner, C., Dulamsuren, C., Wagner, B., & Hauck, M.: Global warming-related tree growth decline and mortality on the north-eastern Tibetan plateau, Clim. Change, 134(1-2), 163-176, 2016.

Linares, J.C., Delgado-Huertas, A., Julio Camarero, J., Merino, J., & Carreira, J. A. Competition and drought limit the response of water-use efficiency to rising atmospheric carbon dioxide in the Mediterranean fir *Abies pinsapo*. Oecologia, 161(3), 611–624, 2009.

Marqués, L., Camarero, J. J., Gazol, A., & Zavala, M. A.: Drought impacts on tree growth of two pine species along an altitudinal gradient and their use as early-warning signals of potential shifts in tree species distributions, For. Ecol. Manage., 381, 157-167, 2016.

Martinelli, N.: Climate from dendrochronology: latest developments and results, Glob. Planet. Change, 40(1-2), 129-139, 2004.

Martínez-Vilalta, J., López, B. C., Adell, N., Badiella, L., & Ninyerola, M.: Twentieth century increase of Scots pine radial growth in NE Spain shows strong climate interactions, Global Change Biol., 14(12), 2868-2881, 2008.

Melvin, T. M., Briffa, K. R., Nicolussi, K., & Grabner, M.: Time-varying-response smoothing. Dendrochronologia, 25(1), 65-69, 2007.

Munné-Bosch, S.: Aging in perennials, Crit. Rev. Plant Sci., 26(3), 123-138, 2007.

Nehrbass-Ahles, C., Babst, F., Klesse, S., Nötzli, M., Bouriaud, O., Neukom, R., ... & Frank, D.: The influence of sampling design on tree-ring-based quantification of forest growth, Global Change Biol., 20(9), 2867-2885, 2014.

Nock, C. A., Baker, P. J., Wanek, W., Leis, A., Grabner, M., Bunyavejchewin, S., & Hietz, P.: Long-term increases in intrinsic water-use efficiency do not lead to increased stem growth in a tropical monsoon forest in western Thailand, Global Change Biol., 17(2), 1049-1063, 2011.

Norby, R. J., DeLucia, E. H., Gielen, B., Calfapietra, C., Giardina, C. P., King, J. S., ... & De Angelis, P.: Forest response to elevated CO2 is conserved across a broad range of productivity, Proc. Nat. Acad. Sci., 102(50), 18052-18056, 2005.

Peñuelas, J., Canadell, J. G., & Ogaya, R.: Increased water-use efficiency during the 20th century did not translate into enhanced tree growth, Glob. Ecol. Biogeograph., 20(4), 597-608, 2011.

Peters, R. L., Groenendijk, P., Vlam, M., & Zuidema, P. A.: Detecting long-term growth-trends using tree-rings: a critical evaluation of methods, Global Change Biol., 21(5), 2040-2054, 2015.

Formatted: Space After: 0 pt, Don't hyphenate

Formatted: Space After: 0 pt, Don't hyphenate

~~Phipps, R. L.: Comments on the interpretation of climatic information from tree rinds, eastern North America, Tree-Ring Bulletin, 42,11-22, 1982.~~

Silva, L. C., & Anand, M.: Probing for the influence of atmospheric CO2 and climate change on forest ecosystems across biomes, Glob. Ecol. Biogeograph., 22(1), 83-92, 2013.

Silva, L. C., Anand, M., & Leithead, M. D: Recent widespread tree growth decline despite increasing atmospheric CO2, PloS one, 5(7), e11543, 2010.

Stayton, C. L., & Hoffman, M.: Estimating sugar maple bark thickness and volume (USDA Forest Service Research Paper NC-38), St. Paul, Minnesota: U.S., 1970.

Sullivan, P. F., Pattison, R. R., Brownlee, A. H., Cahoon, S. M., & Hollingsworth, T. N.: Effect of tree-ring detrending method on apparent growth-trends of black and white spruce in interior Alaska, Environ. Res. Lett., 11(11), 114007, 2016.

Villalba, R., Lara, A., Masiokas, M. H., Urrutia, R., Luckman, B. H., Marshall, G. J., ... & Allen, K.: Unusual Southern Hemisphere tree growth patterns induced by changes in the Southern Annular Mode, Nat. Geosci., 5(11), 793, 2012.

West, G. B., Brown, J. H., & Enquist, B. J.: A general model for ontogenetic growth, Nature, 413(6856), 628, 2001.

Wood, S. N.: Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models, J. R. Stat. Soc. Series B Stat. Methodol., 73(1), 3-36, 2011.

Zhao, S., Pederson, N., D'Orangeville, L., HilleRisLambers, J., Boose, E., Penone, C., ... & Manzanedo, R. D.: The International Tree-Ring Data Bank (ITRDB) revisited: Data availability and global ecological representativity, J. Biogeogr., 46(2), 355-368, 2019.