Manuscript

Abstract: Forests play a crucial role in the global carbon cycle. Extrapolating the current knowledge of the relationship between tree growth and climate conditions to unobserved regions or past/future periods is essential for mitigating and adapting to global change. Tree growth extrapolations typically rely on either a temporal model using long-term time series from a single site or a spatial model using multi-year averages across climatic gradients. Both models have inherent strengths and limitations. Integrating temporal-spatial variations to quantify drivers of system variations is a key issue in macrosystems ecology studies, but is rare in tree growth extrapolation models. Based on the random forest algorithm, we used *Picea mariana* tree-ring data as a case study to explore the potential of the integrating temporal-spatial (TS) model, which integrates both temporal and spatial variations. To examine the prediction skills of models, we provided isolated validation datasets for rigorous testing. Results revealed that (1) for extrapolating temporal variations, both temporal and TS models performed well, while the spatial model performed extremely poorly; (2) For extrapolating spatial variations, both spatial and TS models performed well, while the temporal model performed extremely poorly; and (3) for extrapolating TS variations, only the TS model performed acceptably, while both temporal and spatial models performed extremely poorly. In contrast, the TS model showed significantly higher predictive skill. In summary, this study provides a rare empirical evaluation, demonstrating that the TS model exhibits notably higher predictive skill compared to the temporal and spatial models. Furthermore, our evaluation studies emphasize the necessity of testing the extrapolation capacity of models using independent external validation data, which has

significant implications for the field of tree growth extrapolation. The TS model offers valuable insights for research planning and informed management decisions, particularly in regions with a limited number of observations.

Keywords: climate change, forest, growth extrapolations, random forest, temporal-spatial model, tree-ring

1 Introduction

Forests play a crucial role in the global carbon cycle by absorbing substantial atmospheric CO_2 from human activities(Friedlingstein *et al.*, 2020) and storing it in woody biomass for decades to centuries(Körner, 2017). Understanding and simulating forest dynamics are essential for guiding environmental policies and management strategies aimed at mitigating and adapting to global change(Chausson *et al.*, 2020; Beaulne *et al.*, 2021). Tree-ring observations provide long-term, annual records of individual tree growth with unparalleled spatiotemporal coverage and resolution, making them valuable for evaluating and simulating forest dynamics(Babst *et al.*, 2017; Zhao *et al.*, 2019). Extrapolating the current knowledge of the relationship between tree growth and climate conditions to unobserved regions or past/future periods is a key priority for tree-ring research. Ideally, effective modeling requires dense coverage of tree-ring observations across climatic gradients(Babst *et al.*, 2018), yet logistical, financial and technical constraints often limit data collection(Miller *et al.*, 2004). Consequently, scientists face the challenge of making reliable extrapolations across both spatial and temporal domains using available data.

Temporal model, fitted using long-term time series to describe how growth varies over time experiencing different climatic gradients, is one approach for tree growth extrapolation. Due to the limited overlapping period of tree-ring and climate observations, typically spanning only a few decades, it captures the fast processes operating on interannual time-scales while disregarding the slower processes(Adler *et al.*, 2020; Rodríguez-Morata *et al.*, 2020). The temporal model usually densely covers a limited climatic gradient, as the climate at a given location typically exhibits relatively small-amplitude oscillatory variations(Blois *et al.*, 2013).

As well as temporal variation, tree growth also exhibits spatial variation(Wu *et al.*, 2022). Thus, spatial model, fitted using multi-year averages to describe how growth varies across sites under different climate conditions, represents a complementary approach for simulating tree growth. It captures the interactions between fast and slow processes over long-term periods but provides no information about the transition speed between different growth states(Adler *et al.*, 2020; Bradter *et al.*, 2022). Spatial model usually sparsely covers a broad climatic gradient, as existing tree-ring observations, primarily used for dendroclimatology, focus on marginal growth conditions(Babst *et al.*, 2018). However, some studies have reported that the spatial and temporal model yield very different extrapolations(Oedekoven *et al.*, 2017; Adler *et al.*, 2020).

Integrating temporal-spatial variations to quantify drivers of system variations is a key issue in macrosystems ecology studies(Levy *et al.*, 2014). Several studies have shown that a deeper understanding of the mechanisms driving community assembly and biodiversity dynamics can be achieved by integrating temporal-spatial variations(White *et al.*, 2010; Rull, 2014; Engels *et al.*, 2020). However, research using integrated temporal-spatial variations in tree growth extrapolation model (i.e., TS model) is even rarer than studies focusing solely on spatial variation.

The performance of tree growth extrapolation models is evaluated during validation by examining the correlation between predictions and observations. This often relies on a re-substitution process, where data used for model training are also used for testing, potentially leading to an overestimation of model performance(Araujo *et al.*, 2005). Some studies warn that models optimized to handle data noise may introduce bias in estimates of prediction errors, potentially losing generality outside the original data(Olden & Jackson, 2000; Olden *et al.*, 2002). To address this, methods such as cross-validation and jack-knifing have been proposed. Among these, cross-validation is frequently employed as it can reduce the need for additional observations(Đorđević *et al.*, 2019; Bodesheim *et al.*, 2022). Nonetheless, training and validation data are often tested under identical climate conditions, which is basically unrealistic for extrapolating tree growth. In contrast, independent external validation, which comprises data not used during extrapolation model developing, serves as a more rigorous test of model robustness and generalizability(Kothari *et al.*, 2023). However, external validation of tree growth extrapolation model under climate change remains poorly explored.

Here, we used *Picea mariana* in the Québec, Canada as a case study for exploring the application potential of temporal, spatial and TS growth extrapolation models. To capture the complex nonlinear relationship between growth and climate conditions, we apply the random forest algorithm to build extrapolation models. We also evaluate the robustness and generalizability of models using independent external validation data.

2 Data materials and methods

2.1 Study area

We use the *Picea mariana* (Mill.) B.S.P in the northern temperate and boreal forest zones of Québec, Canada as a case study. The climate spans from humid continental in the south, characterized by hot humid summers and long cold winters, to subarctic in the north with cooler summers and extended colder winters. The elevation of the study area increases from about 0 m a.s.l. in the boundary to about 1233 m a.s.l. in the central area (Fig 1b). The study area experienced relative climatic stability for most of the 20th century followed by a consistent warming began in the early 1990s (Fig 1c). *Picea mariana* is commonly selected for dendrochronology studies due to its distinctly marked annual growth rings, climatic sensitivity and abundance throughout the North American boreal forest(Beaudoin *et al.*, 2014; D'orangeville *et al.*, 2016).

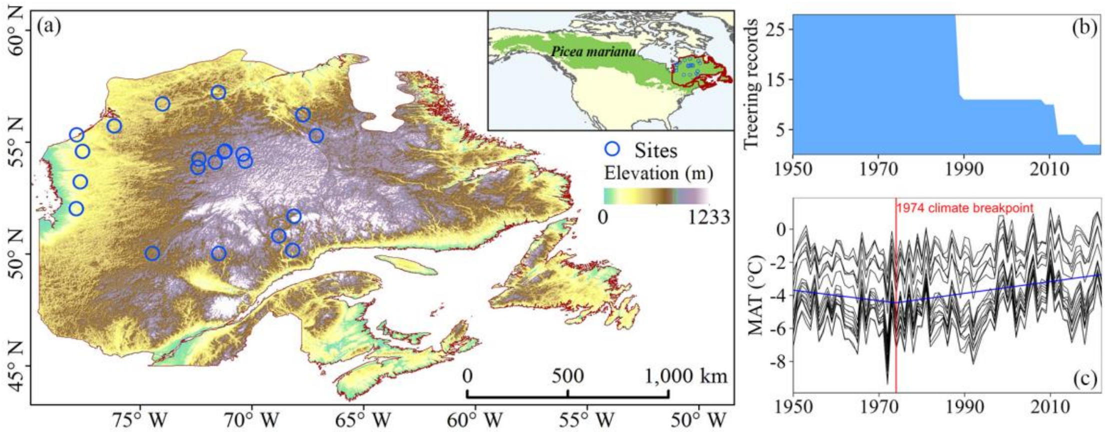

Fig 1 Spatial-temporal distribution of tree-ring sample sites. (a) Spatial range of *Picea mariana* (colored area) and Sample sites (blue circles), (b) Number of tree-ring sample sites and (c) Temporal variation of mean annual temperature at each of tree-ring sample sites in the study

2.2 Tree-ring data

Tree-ring width data were retrieved from the International Tree-Ring Data Bank (https://www.ncei.noaa.gov/pub/data/paleo/treering/measurements/), a public database known for its utility in investigating long-term tree growth patterns(Zhao *et al.*, 2019; Pearl *et al.*, 2020). First, we downloaded all available tree-ring width data of *Picea mariana* for the study area. We then thoroughly cleaned and filtered the tree-ring width data to ensure that all cores met the following criteria: the radius was greater than 5 cm and less than 100 cm, and the average ring width and 95% of annual ring widths were less than 1 cm. We further subset the tree-ring width to include only measurements after 1950, as the significant improvement in the quality of climate data after 1949(Harris *et al.*, 2020). Finally, we retained 38,170 ring width measurements from 733 cores across 28 sites. The number of sites available for each year peaks before 1988 and declines dramatically thereafter: at 57% of sites the last measurements was before 1988, at 86% of sites before 2011, and only two sites have measurements in 2022 (Fig 1c).

We transformed the tree-ring width to annual basal area increments (BAI, mm^2/yr), which are closely related to tree productivity(Mirabel *et al.*, 2023) using the formula:

$$\mathrm{BAI} = \pi(R_t^2 - R_{t-1}^2)$$

where, R represents the tree radius and t represents the year of tree-ring formation. To estimate tree radius, we assumed that the ring width was uniform in a circular cross-section of the tree, and the oldest record corresponded to the pith (i.e., radius at any given year is the sum of all previous ring widths). Tree age was estimated as the sum of recorded rings. Single BAI and age series for the same site were averaged by Tukey's

Biweight Robust Mean method into a site BAI and age chronology. All these procedures were performed using the dplR package(Bunn *et al.*, 2021) in R software(Team, 2018).

2.3 Climatic data

As driving tree growth variables, we used monthly temperature (°C), precipitation (mm), cloud cover (%), potential evapotranspiration (PET, mm/d), and diurnal temperature range (DTR, °C) spanning 1949-2022 with a spatial resolution of 0.5° from the Climate Research Unit Time Series (CRU TS) V 4.07 dataset(Harris *et al.,* 2020). To obtain these variables for the individual tree-ring sample sites, values were extracted from the gridded dataset and the corresponding grid pixels. Considering the influence of both previous and current year on tree growth for a single year, we incorporated climatic variables from May of previous year to October of current year.

2.4 Experimental design

In this study, three major steps were taken to evaluate the potential of models in simulating tree growth variations (Fig 2). First, we established the foundational database. Second, we constructed the temporal and spatial variation datasets for tree growth and their driving factors. Finally, we fitted and evaluated three tree growth variation extrapolation models.

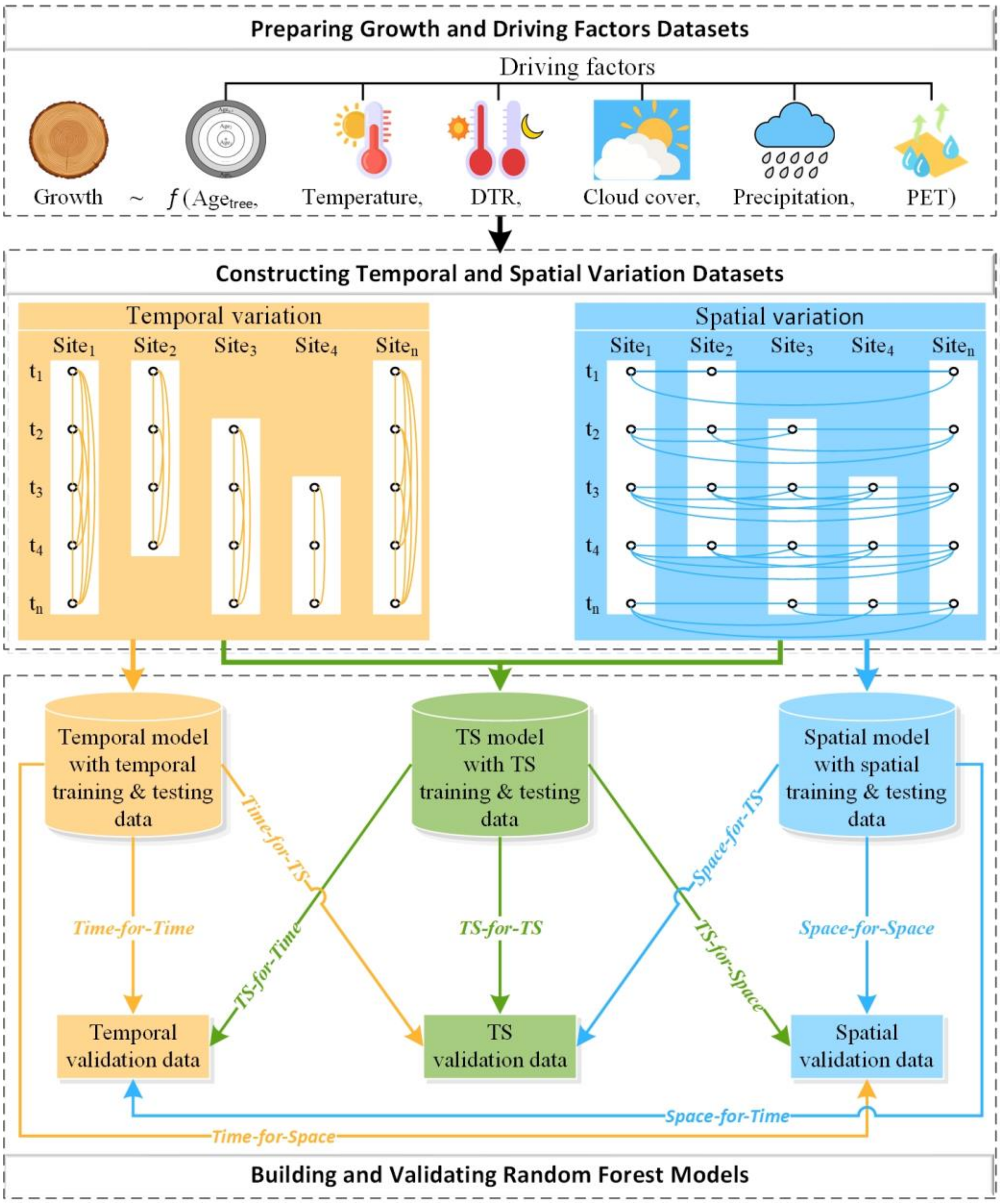

Fig 2 The experimental design and workflow to evaluate the performance of extrapolation models in tree growth variation. PET and DTR represent the potential evapotranspiration and diurnal temperature range, respectively. TS, integrating temporal-spatial.

2.4.1 Constructing the temporal and spatial variations

For temporal variation data, the dissimilarities of BAI, age and climate conditions between all sample pairs at a single tree-ring site over time were calculated then combined with temporal comparisons obtained from all other sites. For example, within site1, the temporal dissimilarities in tree growth were calculated by a subtraction expression (e.g., growth in 2000 minus growth in 2001). For spatial variation data, the dissimilarities between all site pairs for a single year across space were calculated and then combined with spatial comparisons obtained from all other years. For example, in the year 2000, the spatial dissimilarities in tree growth were calculated by a subtraction expression (e.g., growth at site1 minus growth at site2).

To build and evaluate extrapolation models, we split the temporal and spatial dissimilarities data into training, testing (i.e., internal validation) and validation (i.e., external validation) partitions (Fig 3). We randomly selected 70% of sites, and then (1) assigned 70% of their temporal dissimilarities to construct temporal training and testing data (16,984 pairwise dissimilarities) and the remaining 30% to construct temporal validation data (7,279 pairwise); (2) assigned 70% of their spatial dissimilarities to construct spatial training and testing data (5,116 pairwise) and the remaining 30% to construct spatial validation data (2,192 pairwise); (3) pooled their temporal and spatial dissimilarities together to construct TS training and testing data (22,100 pairwise). For the remaining 30% of sites, we pooled their temporal and spatial dissimilarities together to construct TS validation data (12,655 pairwise). Through the above processes, we combined tree growth and climate data across space and time and thus integrates shifting niches, no-analog climatic conditions, and other factors to capture diverse manifestations of growth-climate relationships(Nogués-Bravo, 2009; Blois *et al.*, 2013). Moreover, the integration created large datasets, enhancing analytical power.

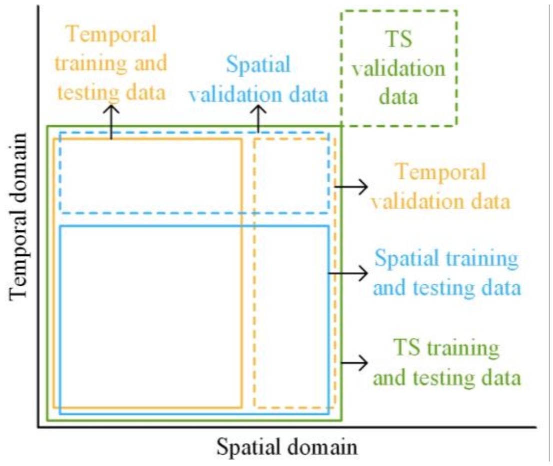

Fig 3 Data composition used for model building and evaluation. TS, integrating temporal-spatial

2.4.2 Building and evaluating extrapolation models

We used the random forest (RF) algorithm as a multivariate non-parametric regression method(Bhuyan *et al.*, 2017) to explore the effects of age and climate on tree growth variations. The RF algorithm, an ensemble learning method developed by Breiman (2001), constructs several random decision trees during training phase, with the ultimate output being the average of all decision trees results(Liaw & Wiener, 2002). It is widely applied in extrapolations and well-suited for high-dimensional non-linear modeling of tree growth(Mirabel *et al.*, 2023; Jevšenak *et al.*, 2024). Incorporating Tobler' First Law of Geography(Tobler, 1970), which emphasizes the importance of spatial proximity, we incorporate space distance as an explanatory variable. Recognizing the influence of adaptive plasticity on growth response to climate, i.e., the effects of rapid and slow climate changes may vary(DeSoto *et al.*, 2014; Wilmking *et al.*, 2020), we also incorporate time distance as an explanatory variable. Ecologically, the relationship between growth and climate often reflects non-linearities, resembling reaction norms in evolutionary biology, and featuring physiological optima and either threshold or saturation effects(Wilmking *et al.*, 2020). Identical climate variations under various climate conditions may result in diverse impacts on growth, potentially even exhibiting opposite signs. Thus, we further estimate interactions of climate conditions between sample pairs at a single site over time, or between site pairs for a single year across space. Finally, the fitting function has the form:

$$\Delta BAI \sim f\left(\Delta age, \frac{\Delta climate_{i,k}}{(1 + dist_T) * (1 + dist_S)}, climate_{onset,i,k} * climate_{end,i,k}\right)$$

where, ΔBAI, Δage and $\Delta climate$ represent the dissimilarities of BAI, age and climate between sample pairs at a single site over time, or between site pairs for a single year across space, respectively. The index i specifies five climate variables: temperature, precipitation, cloud cover, potential evapotranspiration and diurnal temperature range. The index k specifies six 2-month seasons, starting from the previous year's July and August, September and October, …, until the current year's September and October (noting that the winter season spans from previous November to current April). The index onset and end specify the onset and end climate condition for tree growth variations. The terms $dist_T$ and $dist_S$ represent the time and space distance between sample pairs at a single site over time, or between site pairs for a single year across space. To prevent division by zero, a bias term of 1 is added to both $dist_T$ and $dist_S$. The RF model was fitted with the 'sklearn' package from Python(Pedregosa *et al.*, 2011).

We then built three extrapolation models: (1) the temporal model fitted using temporal training and testing data, (2) the spatial model fitted using spatial training and testing data, and (3) the TS model fitted using TS training and testing data. To mitigate overfitting, we used a 10-fold cross-validation approach to search for the best parameters(Hastie *et al.*, 2009). In this approach, the input data was randomly split each time by assigning 90% train the model and the remaining 10% to test the model (i.e., the internal validation). The determination coefficients (R^2) were used to evaluate model accuracy, with a higher value closer to 1 indicating a more accurate model(Alavi *et al.*, 2010). The mean absolute error (MAE) and root mean square error (RMSE) were adopted to depict the average magnitude of errors. MAE equal to 0 indicates predictions completely coincides with observations, and the model is deemed ideal. Although RMSE is sensitive to outliers, it avoids using absolute values and is deemed more suitable than MAE when model errors conform to a normal distribution(Chai & Draxler, 2014). The value of R^2, MAE and RMSE are in the ranges [0, 1], [0, +∞] and [0, +∞], respectively, and were calculated using the following formulas:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(Y_i^{model} - Y_i)^2}{\sum_{i=1}^{N}(Y_i - Y_{mean})^2}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}\left|Y_i^{model} - Y_i\right|}{N}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i^{model} - Y_i)^2}{N}}$$

where, N is the number of samples, Y_i^{model} is model predictions, Y_i is observations, Y_{mean} is the mean of observations.

The Time-for-Time, Time-for-Space and Time-for-TS methods were applied to evaluate the predictive skill of temporal model on the temporal, spatial, and integrating temporal-spatial (TS) dissimilarities, respectively. Similarly, the Space-for-Time, Space-for-Space and Space-for-TS were applied to evaluate the predictive skill of spatial model on the temporal, spatial, and TS dissimilarities, respectively. Finally, the

TS-for-Time, TS-for-Space and TS-for-TS were applied to evaluate the predictive skill of TS model on the temporal, spatial, and TS dissimilarities, respectively.

3 Results

The error metrices calculated for the training and testing data of three extrapolation models are shown in the Table 1. R^2 values of all models exceed 0.95 for the training data, indicating that more than 95% of the variance in the training data can be explained by the explanatory variables. Additionally, the low MAE and RMSE values demonstrate that all extrapolation models can effectively fit the training data. Compared to the training phase, the relatively low R^2 values and high MAE and RMSE values during testing phase are understandable, given the greater challenge of predicting unseen data during fitting process. Although overfitting might be present, all models exhibit an R^2 greater than 0.71 for the testing data, along with low MAE and RMSE values, indicating their good generalization ability for the testing data.

Table 1 The statistical evaluation of three extrapolation models

Model	Training			Testing		
	R^2	MAE	RMSE	R^2	MAE	RMSE
Temporal	0.952	0.07	0.12	0.710	0.19	0.29
Spatial	0.973	0.16	0.24	0.865	0.42	0.57
TS	0.965	0.09	0.16	0.798	0.23	0.38

Certain differences in the relative importance of explanatory variables were observed among three extrapolation models (Fig 4). For the temporal model, diurnal temperature range factors (26.57%) were identified as the most important climatic variables in explaining tree growth variations, followed by the potential evapotranspiration factors (19.34%) and age (16.14%). Diurnal temperature range factors were also the most important climatic variables for both spatial (32.55%) and TS model (31.65%), while the order of importance thereafter changes to cloud cover factors (28.83% and 17.08%) and temperature factors (10.38% and 15.40%). Among the 61 predictor variables (Table S1), Age_dif (16.14%), Dtr*_Winter (13.06%) and Pet*_Summer (3.83%) were the top three important variables for the temporal model. For the spatial model, Cld*_Winter (15.61%), Age_dif (9.03%) and Cld_dif_Winter (8.17%) were the top three important variables. In contrast, for the TS model, Age_dif (10.38%) was the most important variable, with Cld*_Autumn (7.80%) and Dtr*_spring (6.41%) being the second and third most important. In general, Age_dif was a very important variable in explaining tree growth variations, consistently ranking within the top two in all extrapolation models.

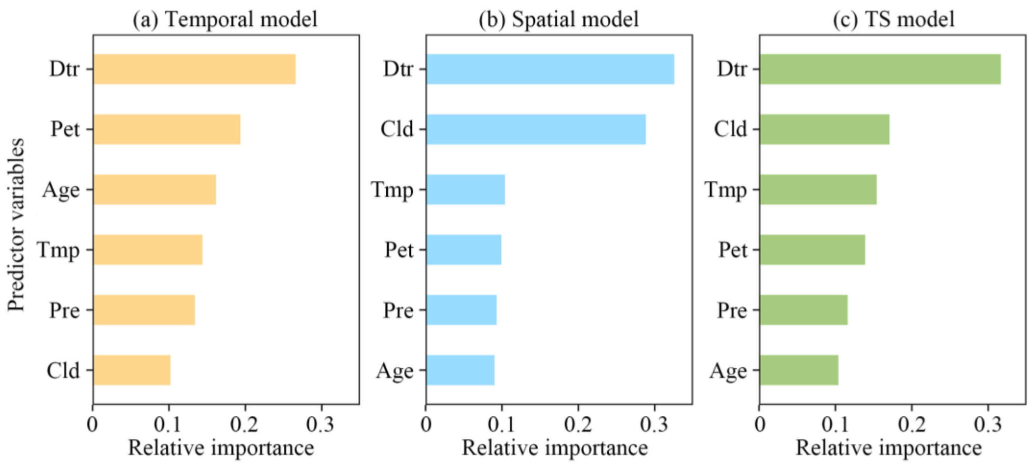

Fig 4 The relative importance of 6 types explanatory variables contributing to tree growth variations as determined from the (a) Temporal, (b) Spatial and (c) TS models. Cld, cloud cover; Dtr, diurnal temperature range; Pet, potential evapotranspiration; Pre, precipitation; Tmp, temperature.

Aside from the internal validation (e.g., temporal, spatial and TS testing data), we also evaluated the model performance under less ideal scenarios using independent external validation data (e.g., three validation data). Performance of the temporal model experienced a slight decrease, dropping from $R^2 = 0.71$ (internal, see testing R^2 in the Table 1) to 0.69 (external, Fig 5a), representing a 3% decrease. This was followed by the spatial model, which experienced a 6% decrease from 0.86 to 0.81 (Fig 5e). In contrast, the TS model's performance experienced the most significant decrease, dropping substantially by 59% from 0.8 to 0.21 (Fig 5i). This might be attributed to the greater differences between the TS validation data and the TS training and testing data, compared to the other two approaches, as the validation data is comprised of entirely independent tree-ring sites. These findings indicate that all extrapolation models exhibited declines in performance from internal to independent external validation.

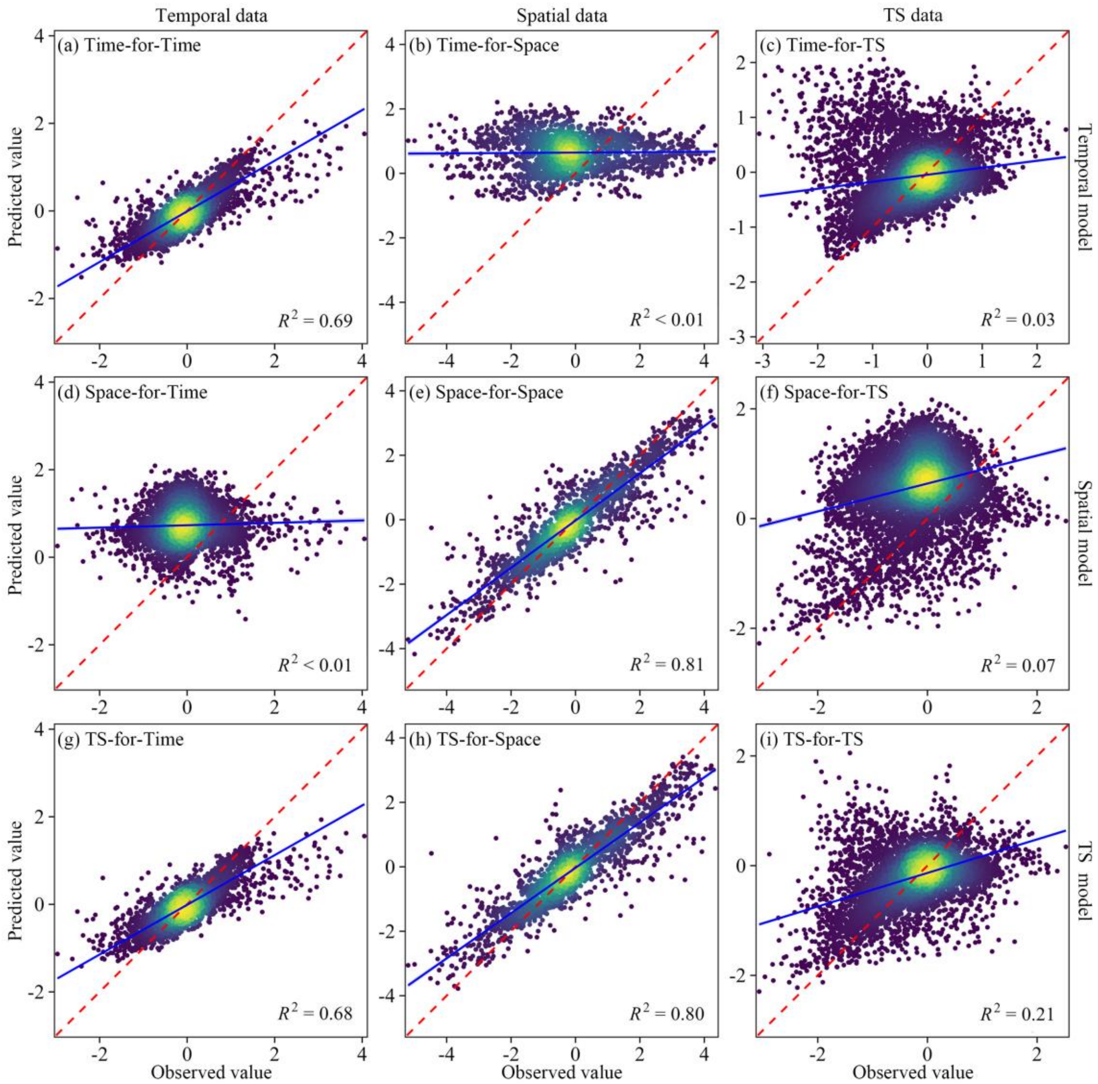

Fig 5 Performance of temporal model (a-c), spatial model (d-f) and TS model (g-i) for simulating the temporal dissimilarities (a, d, g), spatial dissimilarities (b, e, h), and TS dissimilarities (c, f, i) of tree growth. In all cases, color represents point density, with yellow indicating higher density. The blue lines indicate best-fit lines from ordinary least squares regression, and red dashed lines indicate the 1:1 line. R2 denotes the goodness of fit.

The Space-for-Time substitution method exhibited extremely poor performance (Fig 5d). The relationship between temporal dissimilarities predicted by spatial model and observed temporal dissimilarities was notably weak ($R^2 < 0.01$), and the predictive skill using Space-for-Time substitution relative to Time-for-Time substitution dropped to almost zero (0.22%, Table 2). In contrast, predicted temporal dissimilarities were comparable between the TS-for-Time substitution and Time-for-Time substitution (Fig 5g). The association between temporal dissimilarities predicted by TS model and observed temporal dissimilarities was fairly strong ($R^2 = 0.68$), with the TS-for-Time substitution exhibiting 99.82% of the predictive skill using Time-for-Time substitution, resulting in a slight loss of predictive skill (Table 2).

Table 2 Results of ordinary least-squares regression between observed and predicted dissimilarities to explain the predictive performance of each substitution method

Dataset	Method	β	α	R^2	Skill: %
Temporal dissimilarities	Time-for-Time	-0.009	0.578	0.6851	100
	Space-for-Time	0.731	0.027	0.0015	0.22
	TS-for-Time	-0.011	0.567	0.6839	99.82
Spatial dissimilarities	Space-for-Space	-0.027	0.733	0.8055	100
	Time-for-Space	0.645	0.006	0.0003	0.03
	TS-for-Space	-0.032	0.702	0.8037	99.78
TS dissimilarities	TS-for-TS	-0.133	0.306	0.2094	100
	Time-for-TS	-0.046	0.128	0.0312	14.90
	Space-for-TS	0.639	0.255	0.0689	32.90

Note: β, α, and R^2 represent the slope, intercept, and goodness of fit of the ordinary least-squares regression, respectively. Predictive skill for temporal dissimilarities was calculated by dividing the R^2 of each substitution method by the R^2 of Time-for-Time substitution (i.e., $R^2_{\text{Time-for-Time}}/R^2_{\text{Time-for-Time}}$, $R^2_{\text{Space-for-Time}}/R^2_{\text{Time-for-Time}}$, $R^2_{\text{TS-for-Time}}/R^2_{\text{Time-for-Time}}$). Predictive skill for spatial dissimilarities was calculated by dividing the R^2 of each substitution method by the R^2 of Space-for-Space substitution (i.e., $R^2_{\text{Space-for-Space}}/R^2_{\text{Space-for-Space}}$, $R^2_{\text{Time-for-Space}}/R^2_{\text{Space-for-Space}}$, $R^2_{\text{TS-for-Space}}/R^2_{\text{Space-for-Space}}$). Similarly, predictive skill for TS (integrating temporal-spatial) dissimilarities was calculated by dividing the R^2 of each substitution method by the R^2 of TS-for-TS substitution (i.e., $R^2_{\text{TS-for-TS}}/R^2_{\text{TS-for-TS}}$, $R^2_{\text{Time-for-TS}}/R^2_{\text{TS-for-TS}}$, $R^2_{\text{Space-for-TS}}/R^2_{\text{TS-for-TS}}$).

The Time-for-Space substitution method also performed extremely poor, with the fit between spatial dissimilarities predicted by temporal model and observed spatial dissimilarities being notably weak ($R^2 < 0.01$, Fig 5b). Inevitably, the predictive skill using Time-for-Space substitution relative to Space-for-Space substitution dropped to almost zero (0.03%, Table 2). Similar to the TS-for-Time substitution, the TS-for-Space substitution performed well ($R^2 = 0.80$, Fig 5h), achieving 99.78% of the predictive skill using Space-for-Space substitution (Table 2).

For the TS dissimilarities, a fairly weak goodness of fit was observed between the observations and predictions from both temporal model ($R^2 = 0.03$, Fig 5c) and spatial model ($R^2 = 0.07$, Fig 5f). The predictive skill using Time-for-TS substitution and Space-for-TS substitution achieved 14.9% and 32.9% of the predictive skill using TS-for-TS substitution, respectively (Table 2).

4 Discussions

4.1 Generalization performance of extrapolation models

Assessments of accuracy for tree growth extrapolation models under climate changes typically rely on the re-substitution process, where the data used for model training also serve for validation(Klesse *et al.*, 2020; Bodesheim *et al.*, 2022; Zuidema *et al.*, 2022). However, this process may lead to models overfitting the training and validation data, casting doubt on whether high accuracy on internal data reflects good predictive accuracy on independent external data (i.e., data outside the original data)(Olden & Jackson, 2000). To address this, we evaluated model accuracy by applying adjusted models onto external validation data and comparing the consistency between predictions and observations. Results showed that three growth extrapolation models fitted well on training data (R^2 > 0.95), produced accurate predictions on internal validation (testing data, R^2 = 0.71-0.86) and yield good to acceptable predictions on external validation data (R^2 = 0.21-0.81). We also found that the performance of all models declined sequentially from training to internal validation and further to external validation. This is understandable since the external validation data is primarily unseen by models during their building phase and may encompass observations both within and outside the ranges of input variables(Ashraf & Dua, 2023). Of course, this result supports the cautious use of accuracy measurements on internal data as a surrogate for accuracy on external independent data(Araujo *et al.*, 2005).

In addition, the R^2 of 0.21 seems quite low, representing only 21% of the explained variance. However, it is important to consider that environmental factors other than climate, such as soil, biotic, human activities, and various others not incorporated into our models, also affect tree growth and further led to its nonlinear response to climate change(Vaganov *et al.*, 2011; Biermann & Grissino-Mayer, 2018). This makes it impossible to account for all potential factors driving variations in tree growth(Babst *et al.*, 2018). Moreover, with only 28 sites representing a small part of the *Picea mariana* distribution, there is some epistemic uncertainty in capturing tree growth variations. Errors are thus an inherent property of extrapolation models(Stewart, 2000), and it is unrealistic to expect consistently high precision in predictions on external validation data. Nevertheless, the performance analysis on external validation confirmed the adequate modeling and generalization ability of the random forest model for tree growth variations, at least over modest time and space scales.

4.2 Opportunity for extrapolation models on tree growth

Using data from sample sites to drive temporal/spatial models for extrapolation of unobserved regions or past/future periods has always been pivotal in ecological inquiry(Miller *et al.*, 2004; Dormann, 2007; Casalegno *et al.*, 2010). The evaluation results showed that the temporal model performed extremely poorly on the extrapolation of spatial and integrating temporal-spatial (TS) variations, while the spatial model performed extremely poorly on the extrapolation of temporal and TS variations. This discrepancy may be attributed to three main reasons. Firstly, the rates and magnitudes of growth variation along climatic gradients do not entirely similar across time and space, which was further confirmed by the notable discrepancy in the relative importance of explanatory variables between temporal and spatial models. Secondly, the extent of temporal and spatial climatic gradients often mismatch, with spatial climatic variation always larger than temporal variation(Blois *et al.*, 2013; Belle *et al.*, 2022). Finally, ecological processes driving growth variation operate on different timescales across time (short) and space (long)(Adler *et al.*, 2020). For example, temporal dissimilarities are mainly driven by rapid environmental changes(Lovell *et al.*, 2023), while spatial dissimilarities may result from long-term evolutionary adaptation in long-lived sessile organisms(Clark *et al.*, 2001). Additionally, our results demonstrated the effective performance of the TS model on the extrapolation of both temporal and spatial variations. This is unsurprising, as integrating temporal and spatial variations can provide insights that are not possible to get from either alone and may mitigate some inherent weaknesses in each type of variation.

The almost 100% predictive skill of TS-for-Time and TS-for-Space substitutions, coupled with good performance on the extrapolation of TS variations, collectively demonstrate the TS model's effectiveness in capturing tree growth variation. Here, we constructed the temporal and spatial growth dissimilarities dataset, providing a substantial sample size for growth dissimilarity models. Whereas, the sample size for temporal and spatial growth models is unlikely to be so large. The temporal model is based on the relationship between tree productivity and climate factors over time at a single site, with the sample size determined by the length of time series at that site. The spatial model is based on the relationship between multiyear averages of productivity and climate factors across space, with the sample size determined by the number of sites. While the TS model incorporates all-time series, achieving a notably larger sample size compared to temporal and spatial models, thus enhancing analytical power.

In summary, this study provides a rare empirical evaluation of whether the temporal, spatial, and TS models can be effectively used for extrapolating to unobserved regions or past and future periods. And we suggest that (1) when performing temporal extrapolation at sample sites, both temporal and TS models are suitable, but the TS model should be prioritized unless the time series of that site is extensive; (2) when performing spatial extrapolation on sample sites with partial observations, both spatial and TS models are suitable, but the TS model should be prioritized unless there are numerous sites; and (3) when performing extrapolation on completely unobserved regions, only the TS model can achieve an acceptable level.

Compared to the almost zero accuracy of the temporal and spatial models, the TS model demonstrated notably higher predictive skill but had limited accuracy at 21%, still facing numerous challenges in extrapolating tree growth.

4.3 Implications

Our evaluation studies emphasize the essential of testing the extrapolation capacity of extrapolation models using independent external validation data, which has significant implications for the field of tree growth extrapolation. While three extrapolation models in this study performed well during internal and independent external validation, we remain cautious about their predictive capacities over distant past/future periods or spatial distances. Reaction norms may change, especially when the new climate conditions fall outside the range experienced during modeling(Hodgson *et al.*, 2011), potentially causing unobserved growth patterns to deviate systematically from model projections. This phenomenon holds true even for long-lived sessile organisms with heightened levels of phenotypic plasticity in growth. Nevertheless, quantitative predictions based on the best available science are still preferable to proceeding blindly(Rastetter, 1996; Miller *et al.*, 2004), offering valuable insights for research planning and informed management decisions, especially in regions with only a finite number of observations. Therefore, this study complements studies that have focused on modeling tree growth and provides scientific guidance for selecting methods for extrapolation of tree growth under climate change.

Additionally, we acknowledge that extrapolating the TS model to unobserved regions or past/future periods still faces numerous challenges. In the future, more work should focus on considering more comprehensive environmental factors and improving their spatiotemporal resolution to further enhance the predictive skill of TS model.

5 Conclusions

This study explored the potential of an integrated temporal-spatial (TS) model for extrapolating tree growth variations using tree-ring data of *Picea mariana*. The predictive skills of the TS model were compared with those of the temporal and spatial models using independent external validation data. The results showed that: (1) for the extrapolation of temporal variations, both the temporal and TS model provided good results, while the spatial model was extremely poor; (2) for the extrapolation of spatial variations, both the spatial and TS model provided good results, while the temporal model was extremely poor; (3) for the extrapolation of integrating temporal-spatial variations, only the TS model provided acceptable results, while both the temporal and spatial model were extremely poor. In summary, the TS model demonstrated notably higher predictive skill but had limited accuracy at 21%, still facing numerous challenges in extrapolating tree growth. This study provides scientific guidance for the extrapolation of tree growth under climate change, especially in regions with limited observations. In future work, it will be important to include more comprehensive environmental factors and improve their spatiotemporal resolution.

Acknowledgements

This work was supported by the Supported by China Postdoctoral Science Foundation (2023M740287), the National Natural Science Foundation of China (42301070), and the Postdoctoral Fellowship Program of CPSF (GZB20230072).

CRediT authorship contribution statement

F.W: Writing – original draft, Methodology, Conceptualization. X.C: Writing – review & editing, Project administration, Conceptualization. J.J: Validation, Methodology. Y.J: Writing – review & editing. C.L: Writing – review & editing.

Data availability

All data used in this study from publicly accessible data sources. All tree-ring measurements were obtained from https://www.ncei.noaa.gov/pub/data/paleo/treering/. The CRU climate datasets were obtained from http://www.cru.uea.ac.uk/. R markdown have been uploaded to Zenodo https://doi.org/10.5281/zenodo.12633536.

Manuscript

References

Adler, P. B et al., 2020. Matching the forecast horizon with the relevant spatial and temporal processes and data sources. Ecography, 43(11), 1729-1739.

Alavi, A. H et al., 2010. Multi expression programming: A new approach to formulation of soil classification. Engineering with Computers, 26, 111-118.

Araujo, M. B et al., 2005. Validation of species-climate impact models under climate change. Global Change Biology, 11(9), 1504-1513.

Ashraf, W. M., Dua, V., 2023. Machine learning based modelling and optimization of post-combustion carbon capture process using MEA supporting carbon neutrality. Digital Chemical Engineering, 8, 100115.

Babst, F et al., 2018. When tree rings go global: Challenges and opportunities for retro- and prospective insight. Quaternary Science Reviews, 197, 1-20.

Babst, F et al., 2017. Improved tree-ring archives will support earth-system science. Nature Ecology & Evolution, 1(2), 0008.

Beaudoin A et al., 2014. Mapping attributes of Canada’s forests at moderate resolution through kNN and MODIS imagery. Canadian Journal of Forest Research. 44 (5), 521-532.

Beaulne, J et al., 2021. Peat deposits store more carbon than trees in forested peatlands of the boreal biome. Scientific Reports, 11(1), 2657.

Belle, S et al., 2022. Unravelling chironomid biodiversity response to climate change in subarctic lakes across temporal and spatial scales. Hydrobiologia, 849(11), 2621-2633.

Bhuyan, U et al., 2017. Exploring relationships among tree-ring growth, climate variability, and seasonal leaf activity on varying timescales and spatial resolutions. Remote Sensing, 9(6), 526.

Biermann, C., Grissino-Mayer, H. D., 2018. Shifting Climate Sensitivities, Shifting Paradigms: Tree-Ring Science in a Dynamic World. In the Palgrave Handbook of Critical Physical Geography. Springer.

Blois, J. L et al., 2013. Space can substitute for time in predicting climate-change effects on biodiversity. Proceedings of the National Academy of Sciences, 110(23), 9374-9379.

Bodesheim, P et al., 2022. Predicting spatiotemporal variability in radial tree growth at the continental scale with machine learning. Environmental Data Science, 1, e9.

Bradter, U et al., 2022. Decomposing the spatial and temporal effects of climate on bird populations in northern European mountains. Global Change Biology, 28(21), 6209-6227.

Breiman, L., 2001. Random forests. Machine Learning, 45(1), 5-32.

Bunn, A et al., 2021. Package 'dplR.'

Casalegno, S et al., 2010. Vulnerability of Pinus cembra L. in the Alps and the Carpathian mountains under present and future climates. Forest Ecology and Management, 259(4), 750-761.

Chai, T., Draxler, R. R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature. Geoscientific Model Development, 7(3), 1247-1250.

Chausson, A et al., 2020. Mapping the effectiveness of nature‐based solutions for climate change adaptation. Global Change Biology, 26(11), 6134-6155.

Clark, J. S et al., 2001. Ecological forecasts: An emerging imperative. Science, 293(5530), 657-660.

DeSoto, L et al., 2014. Different growth sensitivity to climate of the conifer *Juniperus thurifera* on both sides of the Mediterranean Sea. International Journal of Biometeorology, 58, 2095-2109.

Đorđević, V et al., 2019. Exploring the performance of genomic prediction models for soybean yield using different validation approaches. Molecular Breeding, 39, 1-15.

D'orangeville, L et al., 2016. Northeastern North America as a potential refugium for boreal forests in a warming climate. Science, 352(6292), 1452-1455.

Dormann, C. F., 2007. Promising the future? Global change projections of species distributions. Basic and Applied Ecology, 8(5), 387-397.

Engels, S et al., 2020. Temperature change as a driver of spatial patterns and long‐term trends in chironomid (Insecta: Diptera) diversity. Global Change Biology, 26(3), 1155-1169.

Friedlingstein, P et al., 2020. Global carbon budget 2020. Earth System Science Data, 12(4), 3269-3340.

Harris, I et al., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific Data, 7(1), 1-18.

Hastie, T et al., 2009. The elements of statistical learning: Data mining, inference, and prediction (Vol. 2). New York: Springer.

Hodgson, J. A et al., 2011. Predicting insect phenology across space and time. Global Change Biology, 17(3), 1289-1300.

Jevšenak, J et al., 2024. Incorporating high-resolution climate, remote sensing and topographic data to map annual forest growth in central and eastern Europe. Science of The Total Environment, 913, 169692.

Klesse, S et al., 2020. Continental‐scale tree‐ring‐based projection of Douglas‐fir growth: Testing the limits of space‐for‐time substitution. Global Change Biology, 26(9), 5146-5163.

Körner, C., 2017. A matter of tree longevity. Science, 355(6321), 130-131.

Kothari, S et al., 2023. Predicting leaf traits across functional groups using reflectance spectroscopy. New Phytologist, 238(2), 549-566.

Levy, O et al., 2014. Approaches to advance scientific understanding of macrosystems ecology. Frontiers in Ecology and the Environment, 12(1), 15-23.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News, 2(3), 18-22.

Lovell, R. S et al., 2023. Space‐for‐time substitutions in climate change ecology and evolution. Biological Reviews, 98(6), 2243-2270.

Miller, J. R et al., 2004. Spatial extrapolation: The science of predicting ecological patterns and processes. BioScience, 54(4), 310-320.

Mirabel, A et al., 2023. Increasing atmospheric dryness reduces boreal forest tree growth. Nature Communications, 14(1), 6901.

Nogués-Bravo, D., 2009. Predicting the past distribution of species climatic niches. Global Ecology and Biogeography, 18(5), 521-531.

Oedekoven, C. S et al., 2017. Attributing changes in the distribution of species abundance to weather variables using the example of British breeding birds. Methods in Ecology and Evolution, 8(12), 1690-1702.

Olden, J. D., Jackson, D. A., 2000. Torturing data for the sake of generality: How valid are our regression models? Ecoscience, 7(4), 501-510.

Olden, J. D et al., 2002. Predictive models of fish species distributions: A note on proper validation and chance predictions. Transactions of the American Fisheries Society, 131(2), 329-336.

Pearl, J. K et al., 2020. New frontiers in tree-ring research. The Holocene, 30(6), 923-941.

Pedregosa, F et al., 2011. Scikit-learn: Machine learning in Python. The Journal of Machine Learning Research, 12, 2825-2830.

Rastetter, E. B., 1996. Validating models of ecosystem response to global change. BioScience, 46(3), 190-198.

Rodríguez-Morata, C et al., 2020. Climate impacts on tree growth in a Neotropical high mountain forest of the Peruvian Andes. iForest-Biogeosciences and Forestry, 13(3), 194-201.

Rull, V., 2014. Time continuum and true long-term ecology: From theory to practice. Frontiers in Ecology and Evolution, 2, 75.

Stewart, T. R., 2000. Uncertainty, judgment, and error in prediction. Prediction: Science, Decision Making, and the Future of Nature, 41, 42.

Team, R. C., 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http s. Http://Www. R-Project. Org/.

Tobler, W. R., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46(sup1), 234-240.

Vaganov, E. A et al., 2011. How Well Understood Are the Processes that Create Dendroclimatic Records? A Mechanistic Model of the Climatic Control on Conifer Tree-Ring Growth Dynamics. In M. K. Hughes, T. W. Swetnam, & H. F. Diaz (Eds.), Dendroclimatology: Progress and Prospects. Springer Netherlands.

White, E. P et al., 2010. Integrating spatial and temporal approaches to understanding species richness. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1558), 3633-3643.

Wilmking, M., et al., 2020. Global assessment of relationships between climate and tree growth. Global Change Biology, 26(6), 3212-3220.

Wu, F et al., 2022. Applying space-for-time substitution to infer the growth response to climate may lead to overestimation of tree maladaptation: Evidence from the North American White Spruce Network. Global Change Biology, 28(17), 5172-5184.

Zhao, S et al., 2019. The International Tree-Ring Data Bank (ITRDB) revisited: Data availability and global ecological representativity. Journal of Biogeography, 46(2), 355-368.

Zuidema, P. A et al., 2022. Tropical tree growth driven by dry-season climate variability.