“Application of machine learning to improve tree growth-climate assessments”

Dissertation

zur Erlangung des Grades

„Doktor der Naturwissenschaften“

im Promotionsfach Geographie

am Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

der Johannes Gutenberg-Universität Mainz

Eileen Sonja Kuhl

Geb. in Köln

Mainz, 2024

Gutachter der Dissertation

Tag der mündlichen Prüfung: 06. November 2024

D77 (Dissertation Universität Mainz)

Summary

The late 20th and early 21st centuries have witnessed rapid climate changes, which are marked by exceptional temperature increases and caused by anthropogenic greenhouse gas emissions. This reinforces the need to comprehend the dynamics of climate drivers, to analyze the impacts on ecosystems, and to place the recent developments into the context of natural climate variability. Since meteorological measurements have only been available from 1850 to present, paleoclimate archives are utilized to reconstruct climate variability of the past centuries to millennia. Dendrochronological studies offer significant potential for deciphering these fluctuations using proxies like tree-ring width (TRW) and maximum latewood density (MXD) and for assessing the impacts of recent climate change on different tree species, forest ecosystems and urban trees. Besides these annually resolved parameters, dendrometer measurements on tree stems can provide detailed insights into the hourly to monthly effects of meteorological variables on the water household and growth of a tree. Recently, machine learning (ML) methods have been increasingly applied in climatological disciplines, offering the potential to be utilized in dendroclimatology as well.

This dissertation investigates the integration of ML into dendroclimatology, emphasizing its potential to enhance the assessment of climate-growth relationships. The study focuses specifically on the application of ML in two pivotal areas of dendroclimatology: refining climate reconstructions and improving the imputation of gaps in dendrometer data. Initially, opportunities to improve dendroclimatological analyses using ML are introduced. In chapter 2, an extreme gradient boosting (XGB) classification model is applied to predict the elevational origin of historical construction timber based on different tree-ring parameters of European larch trees (*Larix decidua* Mill.) from the Simplon valley, Switzerland. Historical samples are successfully categorized into six distinct living-tree sites along an elevational transect ranging from 1400 to 2200 m asl. This approach is subsequently used in chapter 3 to support the development of a new millennium-length Alpine summer temperature reconstruction. The provenance method is applied to datasets from two European larch (*Larix decidua* Mill.) populations in the Simplon and Matter valleys, Switzerland. High correlations between the resulting tree-ring chronology and instrumental temperature records facilitate the use of a linear regression model to reconstruct summer temperatures from 881 to 2017 CE. The study evaluates the impact of utilizing high-elevation historical series versus traditional methods, comparing the new temperature reconstruction with previous efforts. Chapter 4 further investigates the application of ML to impute gaps of up to 30 days in intra-annual dendrometer

data from maples (*Acer platanoides* L.) and plane trees (*Platanus x hispanica* Münchh.) in Mainz, Germany. XGB is successfully utilized to fill data gaps of individual trees. It is distinguished from other methods by its independence from neighbouring trees and climatic variables.

The results of this thesis emphasise the capacity of ML to classify tree-ring series, refine climate reconstructions and enhance data imputation. While ML presents promising advancements in dendroclimatology, ongoing research is crucial to address limitations and methodological uncertainties arising from, for example, the costs of data collection, overfitting problems or nonlinear relationships between tree growth and climate. The findings of this dissertation contribute significantly to enhance our understanding of past climate variability and provides a basis for future method development in dendroclimatological and -ecological research in order to embed ML as an integral component of tree-ring research.

Zusammenfassung

Im späten 20. und frühen 21. Jahrhundert wurde ein rapider Klimawandel verzeichnet, welcher mit einem außergewöhnlich schnellen Temperaturanstieg einher geht und durch anthropogene Treibhausgasemissionen verursacht wird. Dies unterstreicht die Notwendigkeit, die Dynamik der Klimatreiber zu verstehen, die Auswirkungen auf Ökosysteme zu analysieren und die jüngsten Entwicklungen in den Kontext der natürlichen Klimavariabilität einzuordnen. Da meteorologische Messungen im Durchschnitt nur von 1850 bis heute zur Verfügung stehen, werden paläoklimatische Archive verwendet, um die Klimavariabilität der vergangenen Jahrhunderte bis Jahrtausende zu rekonstruieren. Dendrochronologische Studien bieten ein erhebliches Potenzial für die Entschlüsselung dieser Schwankungen durch Proxys wie der Jahrringbreite (TRW) oder der maximalen Spätholzdichte (MXD) und für die Bewertung der Auswirkungen des Klimawandels auf verschiedene Baumarten, Waldökosysteme und Stadtbäume. Neben diesen jährlich aufgelösten Parametern können Dendrometermessungen an Baumstämmen detaillierte Einblicke in die stündlichen bis monatlichen Auswirkungen von meteorologischen Größen auf den Wasserhaushalt und das Wachstum eines Baumes liefern. Seit kurzem haben Methoden des maschinellen Lernens (ML) zunehmend Einzug in klimatologische Disziplinen gehalten bieten das Potenzial, auch in der Dendroklimatologie eingesetzt zu werden.

Diese Dissertation untersucht die Integration von ML in die Dendroklimatologie und betont dessen Potenzial zur Verbesserung der Bewertung von Klima-Baumwachstums-Beziehungen. Die Studie konzentriert sich speziell auf die Anwendung von ML in zwei zentralen Bereichen der Dendroklimatologie: der Verfeinerung von Klimarekonstruktionen und der Verbesserung der Imputation von Lücken in Dendrometerdaten. Zunächst werden die Möglichkeiten zur Verbesserung dendroklimatologischer Analysen mittels ML hervorgehoben und in den nachfolgenden Kapiteln erkundet. In Kapitel 2 wird ein Extreme Gradient Boosting (XGB) Klassifikationsmodell angewendet, um den Ursprung historischer Holzproben basierend auf Jahrringparametern von Europäischen Lärchen (*Larix decidua Mill.*) aus dem Simplon-Tal in der Schweiz vorherzusagen. Die historischen Proben werden dabei erfolgreich den Attributen von sechs verschiedenen lebenden Baumstandorten entlang eines Höhentransekts zwischen 1400 und 2200 m ü. NN. zugeordnet. Diese Herkunftsanalyse wird anschließend in Kapitel 3 genutzt, um die Entwicklung einer neuen 1000-jährigen Temperaturrekonstruktion in den europäischen Alpen zu unterstützen. Die Methode wird dabei auf Datensätze von zwei Lärchenpopulationen (*Larix decidua Mill.*) im Simplon- und Matter-Tal in der Schweiz angewendet. Hohe Korrelationen zwischen der resultierenden Jahrringchronologie und

instrumentellen Temperaturaufzeichnungen ermöglichten die Verwendung eines linearen Regressionsmodells zur Rekonstruktion der Sommertemperaturen von 881 bis 2017 n. Chr.. Die Studie bewertet die Auswirkungen der Nutzung historischer Serien aus hohen Lagen im Vergleich zu traditionellen Ansätzen und vergleicht die neue Rekonstruktion mit früheren Veröffentlichungen. In Kapitel 4 wird zudem die Anwendung von ML zum Füllen von Datenlücken bis zu 30 Tagen in intra-annualen Dendrometerdaten von Spitzahorn (*Acer platanoides L.*) und Ahornblättrigen Platanen (*Platanus x hispanica Münchh.*) in Mainz, Deutschland, untersucht. XGB wird erfolgreich eingesetzt, um Datenlücken einzelner Bäume zu schließen und zeichnet sich durch seine Unabhängigkeit von benachbarten Bäumen und klimatischen Variablen aus.

Die Ergebnisse dieser Dissertation betonen die Fähigkeit von ML, Baumringserien zu klassifizieren, Klimarekonstruktionen zu verfeinern und Imputation bei Datenlücken zu verbessern. Während ML vielversprechende Fortschritte in der Dendroklimatologie zeigt, ist fortlaufende Forschung unerlässlich, um Einschränkungen und methodische Unsicherheiten zu adressieren, die sich beispielsweise aus den Kosten der Datenerhebung, Problemen der Überanpassung oder nichtlinearen Beziehungen zwischen Baumwachstum und Klima ergeben. Die Ergebnisse dieser Dissertation tragen wesentlich zu einem besseren Verständnis der Klimavariabilität in der Vergangenheit bei und bieten eine Grundlage für die zukünftige Methodenentwicklung in der dendroklimatologischen und -ökologischen Forschung, um ML als integralen Bestandteil der Baumringforschung zu verankern.

Acknowledgements

Not available online.

Table of content

1 Introduction

The world's climate has been rapidly changing in the last 70 decades, with temperatures rising at an exceptionally pace since the 21st century, mainly forced by anthropogenic greenhouse gas emissions (Phillips et al. 2023; Esper et al. 2024). These developments highlight the need to enhance our understanding of the drivers of temperature variations and set the recent warming into the context of natural climate variability. To this end, it is vital to pursue the development of new methods for analyzing climatic issues and to reduce the uncertainties inherent in existing methods and climate models (Eyring et al. 2019, Esper and Büntgen 2021). Paleoclimate reconstructions are crucial for contextualizing recent climate changes within the framework of long-term climate history, as instrumental climate observations are limited to 1850 CE. Studies on tree growth have a considerable potential for analyzing the climatic past and the effects of climate change on ecosystems (Anchukaitis 2017; Martinez del Castillo et al. 2022; Treydte et al. 2024). Tree-rings, as paleoclimatic archive, thus offer a valuable opportunity to reconstruct past temperature, precipitation, or drought variability far beyond instrumentally covered periods (Briffa et al. 1998; Wilson et al. 2016; Lara et al. 2020; Torbenson et al. 2023; Büntgen et al. 2024; Esper et al. 2024). Additionally, the effects of a warming climate on tree growth in forest ecotones and urban areas as well as the impact of the trees on their environment in the light of different climate scenarios can be modelled and analyzed (Netherer and Schopf 2010; Martinez del Castillo et al. 2018; He et al. 2021; Schwaab et al. 2021; Pattnaik et al. 2024).

For these purposes, tree-ring proxies obtained from wood core samples have an advantage to proxies from other climate archives, namely in their annual resolution, which allows precise dating certainty and more detailed insights into climate variability (Fritts 1976). Tree growth is, however, influenced by various external parameters. To reconstruct climatic parameters, trees must be sampled in regions, where these parameters are the most limiting factors of growth. For instance, tree growth in high latitudes and altitudes is restricted by temperature and is therefore sensitive to fluctuations in this parameter (Speer 2010). The most commonly used proxies to reconstruct past climate variability are tree-ring width (TRW), maximum latewood density (MXD) and stable isotopes (SI) (Anchukaitis 2017). To achieve millennium length and more, a chronology of measured and cross-dated living and dead tree-ring series is calibrated against instrumental climate data. This relationship is then used to build a regression model or to scale tree-growth to climate variables (Esper and Gärtner 2001; Cherubini et al. 2004). Correlations between tree-growth and climate can also be calculated for analyzing the current states of forests (Rennenberg et al. 2006; Babst et al. 2013; Hartl-Meier et al. 2014; Hartl et al. 2021; Zhang et al. 2024) and for predicting future developments of forest and urban tree

species (Mickler et al. 2002; Rötzer et al. 2021; Martinez del Castillo et al. 2022). In addition to the annually recorded radial tree-ring parameters, intra-annual information on growth can be derived from dendrometers (Fritts 1976; Deslauriers et al. 2007; Speer 2010; Pandey 2021). Dendrometers can measure radial stem growth changes up to a resolution of 1 minute and thus provide a continuous, high-resolution growth monitoring throughout the year (Deslauriers et al. 2007; Oberhuber and Gruber 2010; Miller et al. 2022; Rocha and Holzkämper 2023). This allows, for example, for a detailed examination of alterations of the beginning, end and length of the growing season (Deslauriers et al. 2007) or the impact of drought stress to the tree water deficit and growth (Ziaco and Biondi 2018). Still, limitations in this field can occur from the uncertainty of the climatic signal in the tree-ring proxy or from the use of error-prone devices. ML can help to overcome these limitations due to its capability to learn patterns from existing data to predict missing information.

1.1 Current challenges and limitations in dendroclimatological methods

Millennium-length temperature reconstructions are vital to set recent anthropogenic warming into the context of natural climate variability. Dead wood used to extend living tree-ring chronologies into the past is often sampled from snags, dead stems, or fossil wood, ideally near the corresponding stands (Luckman and Wilson 2005; Briffa et al. 2013; Esper et al. 2014; Klippel et al. 2018; Lara et al. 2020). When in situ wood is unavailable, construction wood from old buildings of unknown elevational origin is utilized (Wilson et al. 2005; Büntgen et al. 2005; Liu et al. 2009; Klippel et al. 2020). Consequently, the strength of the temperature signals stored in these trees is unclear and the inclusion of this wood adds great uncertainty to temperature reconstructions (Babst et al. 2013; Zhang et al. 2015; Hartl et al. 2022).

Methods to determine the biological provenance of such tree-ring series have primarily relied on correlating the series to reference chronologies, and the technique is therefore dependent on a common period between the reference and the series (Bonde et al. 1997; Bridge 2012; Linderholm et al. 2021). Some studies have tested more advanced methods such as multiple regression models, principal component (gradient) analysis (PC(G)A) or k-nearest neighbor (kNN) (Wilson et al. 2004; Buras et al. 2016; Gut 2018; Akhmetzyanov et al. 2020). Most of these still rely on a common period between the reference and the series and many approaches lack standard ML procedures. When the provenances of construction timbers can be determined, model uncertainty in early reconstruction periods can be significantly reduced. This allows for analysis of the effects of provenance consideration on the outcome of a reconstruction model and for revision of past temperature variability.

Methodological advancements are also achievable for other dendroclimatic fields. Working with dendrometer data is generally challenging, as dendrometer devices are prone to data errors and technical failures (Drew and Downes 2009; Vilenski et al. 2019). It is therefore vital to accurately control and prepare the measured stem radius changes before applying any data analysis. One reoccurring problem are data gaps. Especially when regular maintenance of the devices is challenging (e.g., in very remote places or during pandemic restrictions), long data gaps may occur due to technical failures like battery or logger failure, or technical damage caused by moisture, animals, or humans.

So far, gap-filling methods for correcting data gaps have focused on addressing short gaps under 24h with linear or spline interpolation (Aryal et al. 2020; Knüsel et al. 2021), and few methods exist to fill data gaps exceeding this period (Aryal et al. 2020; Lukovic et al. 2022). ML algorithms offer the potential to develop new methods for gap filling and compare the methods across various tree species and environments, ranging from rural to urban settings.

1.2 Machine learning in dendroclimatology

ML is a rapidly evolving tool for classifying, clustering or predicting data, which involves computers learning from data and experience (Samuel 1959). Learning is mainly differentiated between supervised, non-supervised, semisupervised and reinforcement learning. In supervised learning, the algorithms are trained on specific features to predict a target variable, using the target variable itself to learn the relationship between inputs and outputs. In contrast, algorithms of unsupervised learning are designed to distinguish between noise and patterns in data without any target information. Semisupervised learning algorithms can handle data, when the target data has missing values and reinforcement learning is based on the concept of reward and penalty, and the aim of learning system is to achieve the most reward (Géron 2019).

Supervised learning has already been proven useful across many research fields, from medical to ecological applications (Jordan and Mitchell 2015; Rahmani et al. 2021; Keitt and Abelson 2021). For example, it was successfully used to train models to classify medical images and identify cancer (Rahmani et al. 2021). In dendroclimatology, supervised learning techniques are the most prevalent. A simple example is the well-established use of linear regression models to reconstruct past climate (Fritts 1976). A linear regression algorithm is trained with tree-ring parameters to predict the target climate variable, which are the instrumental data. Recently, more advanced algorithms have been applied for this purpose. For instance, Jevšenak et al. (2018) tested various ML algorithms to reconstruct past climate

variability using different tree-ring proxies, finding that model-tree bagging was the most effective. Jiang et al. (2024) used an ensemble technique of different ML algorithms to reconstruct minimum temperatures. Additionally, artificial neural networks (ANNs) have been employed to learn nonlinear patterns in climate growth relationships or have been trained to predict multiple target variables, such as temperature, precipitation, and aridity, at the same time (Zhang et al. 2000; Gholami et al. 2017). Moreover, ML can be utilized for other research questions in this field, for example for clustering series by basal area increment characteristics or predicting radial stem growth through classification methods (Mendes et al. 2019; Aryal et al. 2023). However, many opportunities to test and apply ML algorithms to assess climate-growth relationships in comparison to standard methods have not been explored yet.

1.3 Objectives of this dissertation

This dissertation aims to apply and evaluate ML as a tool in dendroclimatology assessing tree-growth climate relationships. Classic ML procedures are employed, and various algorithms are tested to address two dendroclimatological challenges: paleoclimate reconstructions and imputation for dendrometer data. The results are evaluated and critically compared with existing approaches in both fields.

Chapter 2, “Using machine learning on tree-ring data to determine the geographical provenance of historical construction timbers”, explores the possibility to predict the origin of construction timber from different elevations using classification algorithms. A dataset of various tree-ring parameters, like mean MXD and earlywood to latewood ratio, is established using living tree-ring series from European larch trees (*Larix decidua* Mill.) in the Simplon valley, Switzerland. Different parameter combinations and algorithms are tested for their ability to predict the elevational origin of construction timber. Standard ML practices, including model training, tuning, and testing of the models, are applied. The best-performing algorithm is chosen and tested with different datasets to evaluate the predictive power of the parameters. The final model is then applied to the historical series from buildings. With this ML approach, series with a high probability of being temperature sensitive can be identified. The author significantly supported the concept development of this work and was responsible for the development of the methodology, analysis of the results and interpretation. She created the visualizations and wrote the manuscript, which were revised by the co-authors. The paper was published in Ecosphere 44 (3) in March 2023 (Kuhl et al. 2023).

Chapter 3, “Revising Alpine summer temperatures since 881 CE,” presents the development of a temperature reconstruction for the European Alps, using only temperature-sensitive historical tree-ring series identified through the provenance approach introduced in Chapter 2. This classification approach is applied to two European larch (*Larix decidua* Mill.) datasets from the Simplon and Matter valleys, Switzerland. High correlations between the MXD chronology and instrumental temperature data allowed for the application of a linear regression model to reconstruct summer temperatures in the European Alps spanning the period 881-2017 CE. This chapter assesses and discusses the effects of using high-elevation historical series compared to the traditional approach and compares the new record to previous reconstructions. The author significantly supported the concept development of this work and was responsible for the development of the methodology, analysis of the results and interpretation. She created the visualizations and wrote the manuscript, which were revised by the co-authors. The paper was published in Climate Dynamics 44 (3) in March 2024 (Kuhl et al. 2024).

Chapter 4, “A machine learning approach to fill gaps in dendrometer data”, shifts from paleoclimate reconstructions to the application of ML algorithms for data imputation on intra-annual dendrometer data from Norway maples (*Acer platanoides* L.), and London plane trees (*Platanus x hispanica* Münchh.) in Mainz, Germany. This study aims to test various dataset combinations and algorithms to find the best solution for filling long data gaps. This is necessary as methods for imputing gaps are limited (Aryal et al. 2020; Lukovic et al. 2022), and only five ML architectures have been tested for this purpose yet. This chapter details how eight ML algorithms, that have not yet been tested, and various combinations of data are analyzed for their performance and applicability to develop a practical application scheme. This scheme provides users with easy access to ML-based gap filling. The approach is tested on both artificial and real data gaps and compared to existing methods. The author significantly supported the concept development of this work and was responsible for the development of the methodology, analysis of the results and interpretation. She created the visualizations and wrote the manuscript, which were revised by the co-authors. The paper is under review in Trees.

2 Using machine learning on tree-ring data to determine the geographical provenance of historical construction timbers

Eileen Kuhl[1], Christian Zang[2], Jan Esper[1,3], Dana F. C. Riechelmann[4], Ulf Büntgen[3,5-7], Martin Briesch[8], Frederick Reinig[1], Philipp Römer[1], Oliver Konter[1], Martin Schmidhalter[9], and Claudia Hartl[10]

[1] Department of Geography, Johannes Gutenberg University, 55099 Mainz, Germany

[2] Department of Forestry, University of Applied Science Weihenstephan-Triesdorf, 85354 Freising, Germany

[3]Global Change Research Centre (CzechGlobe), 60300 Brno, Czech Republic

[4] Institute for Geosciences, Johannes Gutenberg University, 55099 Mainz, Germany

[5] Department of Geography, University of Cambridge, Cambridge CB2 3EN, UK

[6]Swiss Federal Research Institute (WSL), 8903 Birmensdorf, Switzerland

[7]Department of Geography, Masaryk University, 61300 Brno, Czech Republic

[8] Department of Information Systems and Business Administration, Johannes Gutenberg University, 55099 Mainz, Germany

[9] DENDROSUISSE- Labor für Dendrochronologie, 3900 Brig, Switzerland

[10] Nature Rings - Environmental Research and Education, 55118 Mainz, Germany

Published in May 2023: Ecosphere 14 (2023): e4453. DOI: 10.1002/ecs2.4453

Summary

Dendroclimatology offers the unique opportunity to reconstruct past climate at annual resolution and wood from historical buildings can be used to extend such information back in time up to several millennia. However, the varying and often unclear origin of timbers affects the climate sensitivity of individual tree-ring samples. Here, we compare tree-ring width and density of 143 living European larch (*Larix decidua* Mill.) trees at seven sites along an elevational transect from 1400-2200 m asl and 99 historical tree-ring series to parametrize state-of-the-art classification models for the European Alps. To achieve geographical provenance of the historical series nine different supervised machine learning algorithms are trained and tested in their capability to solve our classification problem. Based on this assessment, we consider a tree-ring density-based and a tree-ring width-based dataset for model building. For each of these datasets, a general not species-related model as well as a larch-specific model including the cyclic larch budmoth influence are built. From the nine tested machine learning algorithms, Extreme Gradient Boosting showed the best performance. The density-based models outperform the ring-width models with the larch-specific density model

reaching the highest skill (f1 score = 0.8). The performance metrics reveal that the larch-specific density model also performs best within individual sites and particularly in sites above 2000 m asl which show the highest temperature sensitivities. The application of the specific density model for larch allows the historical series to be assigned with high confidence to a particular elevation within the valley. The procedure can be applied to other provenance studies using multiple tree growth characteristics. The novel approach of building machine learning models based on tree-ring density features allows to omit a common period between reference and historical data for finding the provenance of relict wood and will therefore help to improve millennium-length climate reconstructions.

2.1 Introduction

The ability of computers to learn on the basis of existing data (machine learning (ML)) bears great potential to improve various scientific fields including bio- and geoscience (Jordan and Mitchell 2015; Keitt and Abelson 2021). In tree-ring research, ML has recently been applied for modelling stem diameter growth and vessel lumen or for climate reconstruction purposes (Ou, Lei, and Shen 2019; Jevšenak and Skudnik 2021; Bodesheim et al. 2022; Salehnia and Ahn 2022). In the (paleo)climatological context, tree-rings are an essential source to reconstruct past climate fluctuations beyond the instrumental period. Classical approaches consider the relation between climate elements (e.g., temperature or precipitation) and a tree-ring proxy, for example, tree-ring width (TRW) or maximum latewood density (MXD), by scaling or building linear regression models (Cook and Kairiukstis 1990; Briffa et al. 1992; Esper et al. 2005; 2012; Wilson and Luckman 2003; Gurskaya et al. 2012; Li et al. 2012; Cook et al. 2019; Lara et al. 2020). ML algorithms are tested as transfer functions for this relationship by training artificial neural networks, random forests or boosted regression trees (Jevšenak et al. 2018; Gu et al. 2019; Jevšenak and Skudnik 2021; Salehnia and Ahn 2022).

Besides a robust growth-climate model, multi-centennial reconstructions rely on dead wood to extend living chronologies into the past. These samples are often collected from local historical construction wood of unknown origin (Schweingruber 1988; Büntgen et al. 2005; Wilson et al. 2005; Liu et al. 2009; Tegel et al. 2010; Labuhn et al. 2016; Klippel et al. 2020; Hartl et al. 2022). The determination of the origin of ancient wood, so-called dendroprovenancing, is a frequently applied tool to reconstruct trade and transportation routes (Bonde et al. 1997; Wazny 2002; Shindo and Claude 2019; Linderholm et al. 2021; Daly and Tyers 2022), to uncover illegal logging (Kagawa and Leavitt 2010) and determine the origin of artwork or shipwrecks (Haneca et al. 2005; Bridge 2011; Domínguez-Delmás et al. 2020; Brookhouse et

al. 2021). Classic approaches in dendroclimatology and -archaeology consider the correlation of series from unknown origin to a set of existing reference tree-ring chronologies (Bonde 1992; Bridge 2012). It is argued, however, that this classical dendroprovenancing with chronologies not always serves the complexity in the relationship between tree-growth parameters and regionality (Bridge 2000; 2012; Haneca et al. 2005; Drake 2018; Domínguez-Delmás et al. 2020).

The application of ML is likely suitable for a probable higher complexity in this relationship. Approaches have been tested with different tree-ring proxies using multiple regression models (Wilson et al. 2004; Dittmar et al. 2012), or the ML algorithms k-Nearest Neighbor (kNN) (Gut 2018), Principal Component Analysis (PCA) (Wilson and Hopfmueller 2001) or Principal Component Gradient Analysis (PCGA) (Buras et al. 2016; Akhmetzyanov et al. 2020). While PCGA and PCA rely on a common period of overlap between living and historical series for determining the provenance, Dittmar et al. (2012) used features of individual tree series to build a non-linear regression model. However, many published approaches in dendroprovenancing generally lack basic ML model development steps like testing different algorithms and hyperparameter combinations before opting for a fitting algorithm. The a priori requirement of a common period between historical samples and a reference for effective dendroprovenancing remains the greatest challenge. The potential of additional parameters such as MXD or species-specific disturbance features (such as insect outbreaks) has so far not been tested in ML provenance models.

Dendroclimatological studies focus on regions, where tree growth is limited by a dominating factor e.g., the temperature at latitudinal or elevational treeline sites (Briffa et al. 1988; Liu et al. 2009; Babst et al. 2013; Schneider et al. 2015; Esper et al. 2016; Wilson et al. 2016; Ljungqvist et al. 2020; Hartl et al. 2021; 2022). Consequently, the quality of a tree-ring based climate reconstruction derived from living and historic wood, hinges on the consistency of the signal strength across the proxy sources, as the temperature signal of trees fades with decreasing elevation (Wilson et al. 2004; Babst et al. 2013; Hartl-Meier, Dittmar, et al. 2014; Salzer et al. 2014; Konter, Rosner, et al. 2015; P. Zhang et al. 2015; Riechelmann et al. 2020; Hartl et al. 2021; 2022). At lower elevations, other biotic factors e.g., intra- and interspecific competition or trophic interactions with insects, can influence tree growth (Coomes and Allen 2007; Hartl-Meier, Zang, et al. 2014; Konter, Rosner, et al. 2015; Hartl-Meier et al. 2017; Saulnier et al. 2017; Harr et al. 2021). Climate signals in tree-rings of European larch (Larix decidua Mill.), for example, are superimposed by growth disturbances resulting from larch budmoth (Zeiraphera griseana Hübner, LBM) mass outbreaks in the Alps (Baltensweiler and Rubli 1999; Esper et al. 2007; Baltensweiler et al. 2008; Hartl-Meier et al. 2017; Hartl et al.

2022). LBM larvae feed on the needles of larch trees, lowering photosynthetic capacity and altering growth rates (Baltensweiler and Rubli 1984). Various studies on LBM mass outbreak cyclicity and effects on larch growth conducted in the European Alps agree on an occurrence rate of mass outbreaks of eight to ten years and constrain outbreak locations largely to elevations between 1700 and 2000 m asl (Baltensweiler and Rubli 1999; Rolland et al. 2001; Esper et al. 2007; Baltensweiler et al. 2008; Büntgen et al. 2009; Daux et al. 2011; Konter, Esper, et al. 2015; Hartl-Meier et al. 2016; Hartl-Meier et al. 2017; Saulnier et al. 2017; Hartl et al. 2022). Consequently, elevational classification of larch wood from such regions will be determined by the potentially inherent LBM signals of historical series.

For palaeoclimatological studies, the knowledge of the sample origin of historical material is likewise important to adequately remove age-depended growth trends (Bräker 1981). The commonly applied regional curve standardization (Briffa et al. 1992, RCS) must be performed site-by-site or on mean-adjusted series (P. Zhang et al. 2015; Römer et al. 2021), because altitude-dependent offsets among MXD (and TRW) series are observed for European larch (King et al. 2013; P. Zhang et al. 2015; Riechelmann et al. 2020; Rozenberg et al. 2020; Hartl et al. 2022). Neglecting the elevational discrepancies of the regional curves could bias the amplitudes and long-term trends of a subsequent climate reconstruction severely, thus leading to a misinterpretation of past climate variability (P. Zhang et al. 2015; Riechelmann et al. 2020; Hartl et al. 2022).

In this study, we aim at improving dendroprovenancing by applying state-of-the-art ML procedures to eventually strengthen millennium-length climate reconstructions. We use 149 samples of living larch trees from an elevational transect ranging from 1400 to 2200 m asl in the Simplon Valley of the Swiss Alps and test nine different ML algorithms. We fit the best-performing algorithm to different sets of tree-ring parameters and, for the first time, include X-ray measurements and species-specific parameters in a provenance model.

2.2 Material and methods

2.2.1 Tree-ring datasets and study area

Eight larch datasets were collected in the Simplon Valley, Switzerland. One of these contains 99 historical series from different buildings in the Simplon Village (~ 1470 m asl) (introduced in Riechelmann et al. 2013; Riechelmann et al. 2020). The seven living-tree sites span an elevational transect from 1400 m to 2200 m asl, with four sites south exposed (S14, S17, S20 and S22) and three north exposed (N16, N17 and N19) (**Fig. 2-1**, **Tab. S 2-1**). Each site consists of up to 24 series from twelve trees (see Hartl et al. 2022). The two sites at 1700 m

asl (S17 and N17) are merged to one dataset SN17 including twelve series from each site to represent this elevation. Both, the living and historical series, have been accurately dated to build a robust chronology (**Fig. 2-2**).

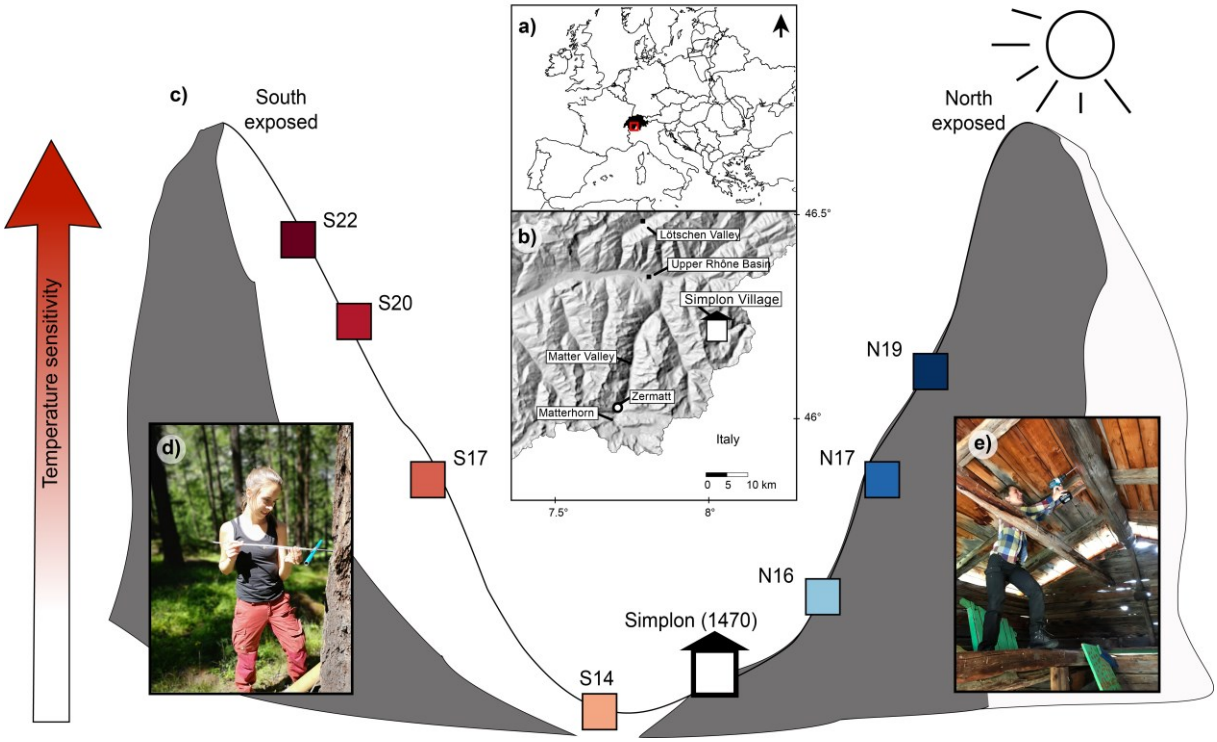


Fig. 2-1 Study area in the Simplon Valley, Switzerland in **a)**, **b)** and sampling scheme with site codes in **c)**, exact elevations are shown in Tab. S 2-1. Sampling of high elevation living trees with a high temperature sensitivity **d)** and of historical construction timber in the lower elevated Simplon Village **e)** (Photo Credits: P. Schulz, C. Hartl).

2.2.2 Dendrochronological measurements and tree-ring parameters

242 high-resolution tree-ring density profiles were measured using a Walesch2003 (WALESCH, Electronic GmbH, Switzerland) following the X-ray densitometry procedure described in Björklund et al. (2019). We considered six different tree-ring parameters: TRW, MXD, earlywood ring width (EWW), earlywood density (EWD), latewood ring width (LWW) and latewood density (LWD) (Tab. S2-2). Descriptive statistics were calculated and forwarded into the ML models including the arithmetic mean, standard deviation (std), skewness (skew), Gini coefficient (gini) and maximum/minimum values (max/min). Additional parameters include the age (including the pith offset), the first-order autocorrelation of TRW (A1 TRW) and the ratio between EWD and LWD (ED/LD ratio). To address the LBM mass outbreak mean cyclicity of 9 years, the spectrum value at 1.11 frequency (9 years) was determined by applying Lomb-Scargle Fourier transformation to 30-year spline detrended TRW series. A dataset D of all living series containing their parameters, from here on referred to as features, and

corresponding elevations was created using R 4.1.0 (R Core Team 2021) and the dplR package (Bunn 2010).

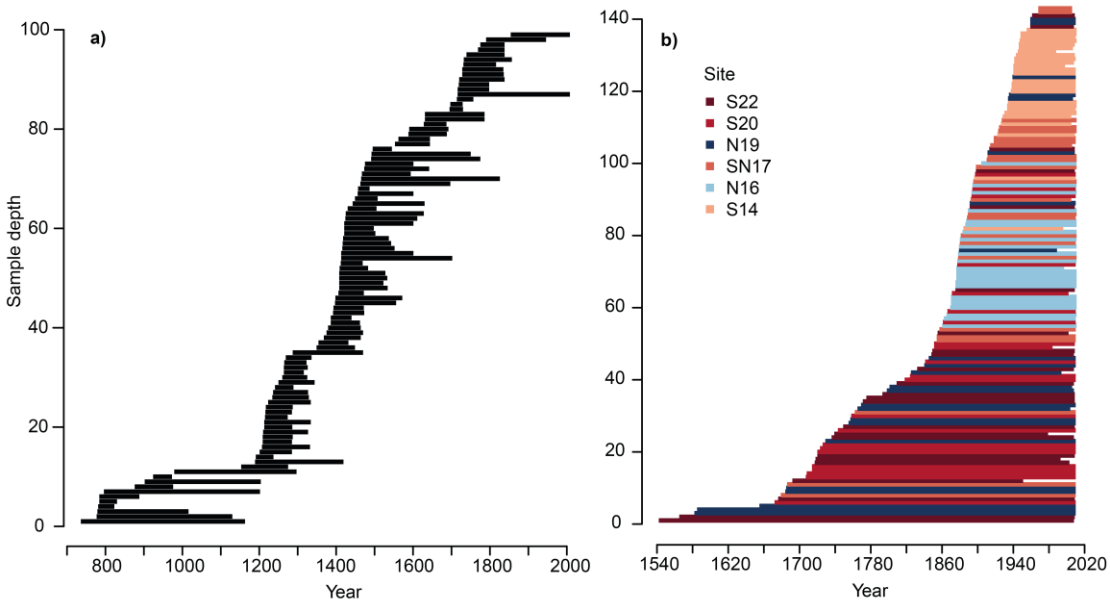


Fig. 2-2 Segment plots of the historic **a)** and the living tree samples **b)** aligned by start date. Colours in b) denote to the different sites of the living material (**Fig. 2-1 c)**).

2.2.3 Site-specific chronology building and climate signals

To illustrate the fading of the temperature signal with decreasing elevation, the living tree-ring data (TRW and MXD) were site-wise power transformed and RCS detrended (Briffa et al. 1992; Esper et al. 2003) using the software ARSTAN (Cook 1985). Site chronologies were constructed by averaging single series with a robust mean and the chronology variance was stabilized using the Rbar weighted method (Osborn et al. 1997). Temperature correlations between site chronologies and gridded temperature data (EOBS 0.25° v23.1e; Cornes et al. 2018) were calculated via classical bootstrap of Pearson's correlation coefficients for the summer seasons June-August (JJA) over the 1928-2009 common period using the R packages treeclim (Zang and Biondi 2015) and dplR (Bunn 2010).

2.2.4 Developing a machine learning model for dendroprovenancing

In supervised learning, a model is fit for a classification or regression task to a labeled dataset (e.g., D) (**Fig. 2-3 a)**). Using a training dataset D_{train} of an input matrix X, the model must predict the corresponding target vector y with a prediction ŷ. During training the settings of the given model are adjusted by computing and minimizing the total loss

$$L = \sum_{i=1}^{m} l(y_i, \hat{y}_i)$$

(eqn. 1)

where m equals the number of entries (number of tree-ring series) in D_{train} and l is a loss function (e.g., cross-entropy). To assure that the model has not only learned D_{train} by heart but has adopted a meaningful representation that generalises to unseen data, it is applied to a test dataset D_{test}. Potential hyperparameters of a model can be adjusted using a validation dataset D_{val} (Vapnik 1991; Ying 2019). Here, the input matrix X of the 143 living-tree series from varying elevational sites between 1400 and 2200 m asl was split into D_{train} and D_{test} using a stratified sampling by elevation to ensure a representation of all sites in both sets (80:20 split). In a second split, D_{train} was divided by stratified sampling into the final D_{train} and D_{val} (80:20 split). We tested nine different ML classification algorithms on D_{train}: kNN (Fix and Hodges 1951), Ridge Regression (Hoerl and Kennard 2000), Logistic Regression (here Softmax Regression) (Berkson 1944), Support Vector Machines (Vapnik and Chervonenkis 1974), Stochastic Gradient Decent (Kiefer and Wolfowitz 1952), Gaussian Naïve Bayes (H. Zhang 2004), Random Forest (Breiman 2001), Linear Discriminant Analysis (Fisher 1936) and Extreme Gradient Boosting (XGBoost) (Chen and Guestrin 2016). The hyperparameters of the algorithms (e.g., maximum tree depth or learning rate) were fine-tuned using grid search and stratified k-fold cross-validation (k = 10) on D_{train} ensuring the consideration of the specified combinations (chosen hyperparameters are listed in **Tab. S 2-3**). Afterwards, the models' performances were tested again on D_{train} using a repeated stratified k-fold cross validation (k = 10 and repeats = 100) to choose the best performing ML algorithm.

We measured the performance of the models using f1 score, precision and recall:

$$f1\ score = 2 \times \frac{precision \times recall}{precision + recall} \quad \text{(eqn. 2)}$$

with

$$precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad \text{(eqn. 3)}$$

and,

$$recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad \text{(eqn. 4).}$$

Precision describes the correctly sorted series per elevational class, while recall counts the number of series, which belong to a certain class but are not predicted into it (Géron 2019). The f1 score over all classes per model was calculated as arithmetic mean.

We find that highest f1 scores are reached by the XGBoost algorithm (**Tab. S 2-4**) after stratified repeated k-fold cross-validation with a mean f1 score of 0.7. XGBoost (Chen and Guestrin 2016) is based on the gradient boosting algorithm by Friedmann (2001), which iteratively builds an ensemble model consisting of multiple decision trees by minimizing the loss function in each iteration (see **Fig. 2-3 b)** as simplified classification scheme with two target classes (filled circle or unfilled diamond)). The algorithm proceeds until a final model N is found specified by a stopping criterion (Ying 2019).

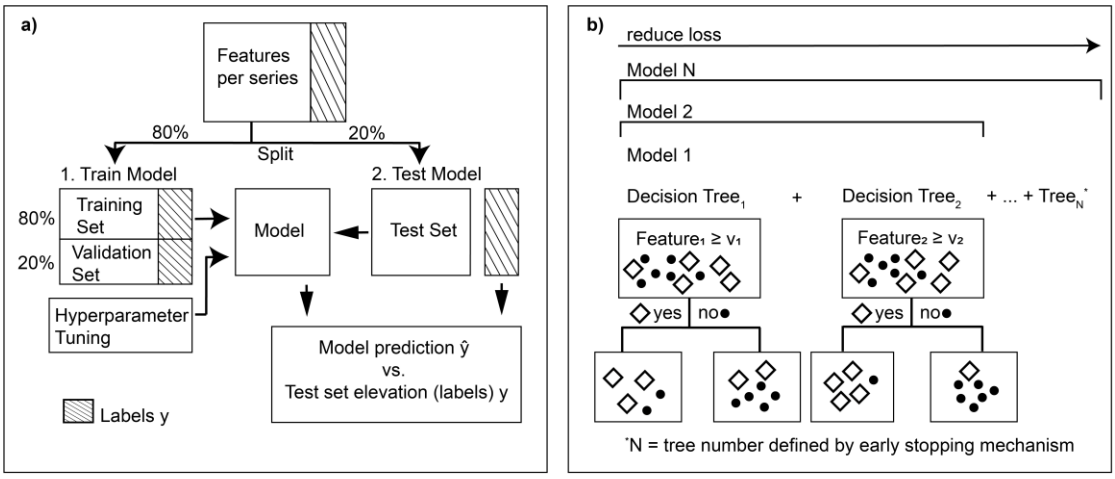


Fig. 2-3 a) Basic Machine Learning Scheme and **b)** a simplified Gradient Boosting Scheme with two classes (filled circle or unfilled diamond).

Based on the applicability in tree-ring science, four different XGBoost models were trained on different combinations of the features. Two models were trained on all density and ring-width features: a general not species-related (39 features, DM_{gen}) and a larch-specific model including the 9-year spectrum (40 features, DM_{sp}). Two additional models excluding densitometric measured features were built: a general cross-species ring-width model (RWM_{gen}, 7 features) and a larch-specific ring-width model (RWM_{sp}, 8 features). The average and site-wise performances of these models were assessed on D_{test}. Finally, a feature matrix for the historical timber was built and fed to the four XGBoost Models. All models were implemented in Python 3.8.5 (Van Rossum and Drake 2009) with the packages Scikit-Learn (Pedregosa et al. 2011) and XGBoost (Chen and Guestrin 2016).

For comparison with existing methods for dendroprovenancing, we additionally tested the performance of these approaches with our living series of D_{train} and D_{test}. We built a regression model from the D_{train} tree-ring series (following the approach of Wilson et al. (2004)) based on

the correlation between the individual site chronologies and the highest elevation chronology S22 (2200 m asl). We also tested a PCGA approach (Buras et al. 2016) on our tree-ring parameters (TRW, MXD, EWW, EWD, LWW and LWD) to check for differences in PCGA loadings. The tree-ring series of D_{test} were then sorted to a provenance following the steps described in Akhmetzyanov et al. (2019; 2020). We used our LWD densitometry measurements as equivalent substitutes for their use of latewood blue intensity (Campbell et al. 2007). Precision, recall and f1 score were calculated for these approaches to compare them to our XGBoost models (equations 2-4).

2.3 Results and discussion

2.3.1 Model performances on the test dataset

Testing the models revealed that DM_{sp} and RWM_{sp} (**Tab. 2-1**) perform better than their species-independent equivalents DM_{gen} and RWM_{gen} (**Tab. S 2-5**). While DM_{gen} and RWM_{gen} reach average f1 scores of 0.69 and 0.25, the inclusion of the larch-specific LBM feature increases the scores to 0.80 (DM_{sp}) and 0.31 (RWM_{sp}), respectively. In DM_{sp}, the highest site-wise f1 scores are observed for the sites S22, N16, and S20 and it executes better than DM_{gen} in almost all sites, except for the highest and the lowest elevation. This performance gain is most pronounced in SN17 and N19. The provenance model results (**Fig. 2-4**) imply that the LBM signals appear to be stronger at these sites and serve as an important feature for site distinctions in DM_{sp} (**Fig. 2-4 c)**). This is also reflected by the high feature importance of the LBM spectrum (**Fig. 2-4 d)**) and the increased performance of RWM_{sp} compared to RWM_{gen} at site N19 (**Tab. 2-1**, **Tab. S 2-5**). Our results indicate that future applications of likewise models should be tested with and without tree species-related characteristics (e.g., interactions with insects) when specific influences on growth are known and observed. Even though the ring-width models themselves are not very reliable, including site- or species-specific information improved the model's ability to distinguish sites (mean f1 scores: 0.25 in RWM_{gen} and 0.31 in RWM_{sp}). The comparison of models with and without disturbance effects can offer the chance to detect site specific influences such as LBM mass outbreaks and support the assessment of past disturbances in historical series. However, the increased performances of the larch-specific models in our example reflect the distinct impact of LBM outbreaks on larch growth and raise the question, if general not species-related models will have increased performances when applied to undisturbed non-host species like the Swiss pine (*Pinus cembra* L.). In the Simplon valley, the inherent LBM signal strength on certain elevations has a strong influence on the performance of provenance models for larch trees. When working with species, which react similarly to a disturbance, a feature that describes the strength of this disturbance on each individual tree-ring series should be considered.

Tab. 2-1 Classification Report: performance metrics of D_{test}[1] on DM_{sp}[2] and RWM_{sp}[3] (in brackets) in each individual class.

Site	Precision	Recall	f1 score
S14	0.57 (0.57)	0.8 (0.8)	0.67 (0.67)
N16	1 (0.40)	1 (0.4)	1 (0.4)
SN17	0.75 (0.0)	0.6 (0.0)	0.67 (0.0)
N19	0.6 (0.6)	0.6 (0.6)	0.6 (0.6)
S20	1 (0.0)	0.75 (0.0)	0.86 (0.0)
S22	1 (0.17)	1 (0.2)	1 (0.18)
Average	0.82 (0.29)	0.79 (0.33)	0.8 (0.31)

[1] test dataset

[2] larch specified density and ring-width parameter model

[3] larch specified ring-width parameter model

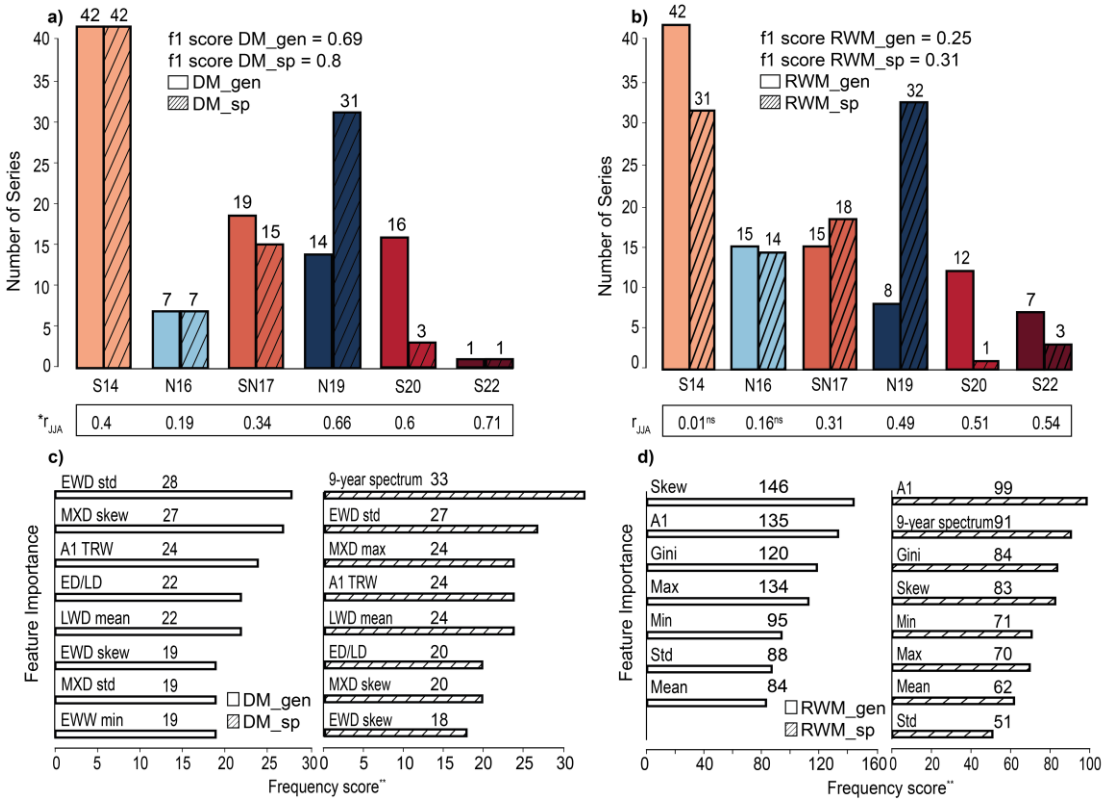


Fig. 2-4 a) Classified historic timber by DM_{sp} (larch specified density and ring-width parameter model) and DM_{gen} (general unspecified density and ring-width parameter model) **b)** Accordingly, RWM_{sp} (larch specified ring-width parameter model) and RWM_{gen} (general u unspecified ring-width parameter model), including the f1 score of D_{test} (test dataset) and correlations between EOBS 0.25° mean temperature data for June-August ($*r_{JJA}$) and MXD (maximum latewood density) a) and TRW (tree-ring width) b) chronologies, respectively. Correlations tagged with ^{ns} have no significant correlation. **c)** and **d)** illustrate the most important features of the models above.

**Frequency score refers to the number of decision tree nodes, a feature was used for (see methods for detailed information on feature abbreviations).

Including tree-ring density features generally improves model results. The ring-width models are not able to distinguish SN17 and S20 from the other sites (f1 scores = 0; **Tab. 2-1**, **Tab. S 2-5**), produce more classification errors on D_{test} than the density models (sum of errors for RWM_{gen} = 21, RWM_{sp} = 19, DM_{gen} = 9 and DM_{sp} = 6) and misclassify series belonging to SN17 to S22. With respect to detrending, mistakenly handling low elevation series as high elevation series can impact a mean chronology, particularly when such misplaced series are clustered during certain periods. Additionally, RWM_{gen} sorts series from N19, characterized by higher temperature sensitivities, to the lowest elevated site, which contains no temperature signal (see rJJA in **Fig. 2-4 b)**). These errors will, however, not influence a temperature reconstruction, as sites with a low climate response should be excluded from a final chronology. Nonetheless, the sample replication of useful historical series would be reduced, and uncertainties increased. In the Simplon valley, the ring-width models appear to lack the ability to correctly determine the provenance and distinguish minor elevational differences (≤ 200 m), although small-scale elevation steps can result in differences in a tree species growth response to environmental factors (rJJA in **Fig. 2-4 a, b)**) (Salzer et al. 2009; 2014; Bunn et al. 2011; Hartl et al. 2022).

The applied XGBoost algorithm is not able to find a well-fitting model, when trained on our ring-width features and, as Akhmetzyanov et al. (2019) has already shown for dendroprovenancing with PCGA, performs better when tree-ring density measurements are included. We suggest using DM_{sp} for finding the provenance of our historical tree-ring series, as it is the most reliable model indicated by the highest f1 score and lowest number of prediction errors. This unique dataset including density measurements and LBM features is, however, tailored to larch trees from the Simplon valley. Datasets of varying species and other regions might perform better with a different algorithm.

2.3.2 Extreme Gradient Boosting models compared to previous dendroprovenancing approaches

The performance of previous approaches on D_{test} indicates that these approaches are not suitable for our data as they show lower performances scores and higher error numbers than our DM_{sp}. The application of PCGA to find the provenance of the D_{test} series resulted in f1 scores of 0.28 (LWD) and 0.16 (TRW), respectively (**Fig. S 2-1 a-c)**). PCGA has demonstrated the ability to distinguish between high and low elevation sites (Akhmetzyanov et al. 2020) but does not appear to be able to determine the provenance of the series as precise as the XGBoost model in this study. Utilizing the method described in Wilson, Esper, and Luckman

(2004) on the dataset indicates a non-linear correlation between the highest site and the other elevational chronologies (**Fig. S 2-1 a, d)**). The striking low correlation between the highest site and the N19 chronology might result from differing LBM signals between the sites, which is supported by the observed performances decrease in DM_{gen} compared to DM_{sp} at N19 (**Tab. 2-1, Tab. S 2-5**). Thus, fitting a linear regression model on the series of D_{train} and testing this model using D_{test} reveals a lower f1 score of 0.25 (**Fig. S 2-1 a)**).

Most existing approaches depend on a common period between the historical and living tree-ring series, but even when a common period is given, DM_{sp}, DM_{gen} and RWM_{sp} outperform traditional dendroprovenancing approaches on our dataset. The XGBoost models (but also all other tested ML models, **Tab. S 2-3, S 2-4**) do not require a common period. Our approach has the advantage of including specifications on certain regions or tree-species (here European larch) if needed and could likewise be applied to other dendroprovenancing objectives, such as shipwrecks or art provenance as well as trade or transportation route studies (Wazny 2002; Bridge 2011; Shindo and Claude 2019; Linderholm et al. 2021; Daly and Tyers 2022). Testing different ML algorithms, which are independent of a common period, might enlarge the pool of useable sites for these wood provenance studies and could thereby improve detecting the geographical origin.

2.3.3 Dendroprovenancing of historical timber using Extreme Gradient Boosting

The classification of historical material reveals differences between our four models (**Fig. 2-4 a, b)**). DM_{sp}, DM_{gen} and RWM_{gen} classify most series to the lowest site S14, while RWM_{gen} sorts most series to N19. In contrast to their general not species-related equivalents the larch-specific models classify more series to the LBM-influenced site N19 but less series to the highest elevation sites. Comparison of the individual series predictions reveals that most of the series assigned to a higher elevation by DM_{gen} were sorted to N19 by the larch-specific DM_{sp}. This might again indicate that information on the intensity of the LBM mass outbreaks cyclicity influences a model's differentiability of these sites and thus the model outcomes. An assessment of the predictions reveals that 26% of the historical series are sorted to different sites by DM_{sp} and DM_{gen}. With respect to the better performance and lower number of errors of DM_{sp} compared to DM_{gen} during the model validations, DM_{sp} likely performs more confident with the historic series as well. RWM_{gen} indicates a similar problem outlining a tendency to sort historical series to S20, while the larch-specific RWM_{sp} attributes them to N19. Both ring-width models have very low performance scores as well as high error numbers on D_{test} and, compared to the density models, classify 49% (RWM_{gen} to DM_{gen}) and 41% (RWM_{sp} to DM_{sp})

of the historical series differently. The discrepancy between RWM_{sp} and DM_{sp} classifications persist throughout time (**Fig. 2-5**), culminating between 1250-1600 CE. This peak period is not covered by living material (of known origin) and erroneous predictions (e.g., from RWM_{sp}) could lead to incorrect variability or mean levels in an RCS detrended chronology, when historical series sorted to wrong elevations. As varying allocations of historical series may bias long composite chronologies differently in certain time periods, these chronologies depend on a reliable provenance of historical series.

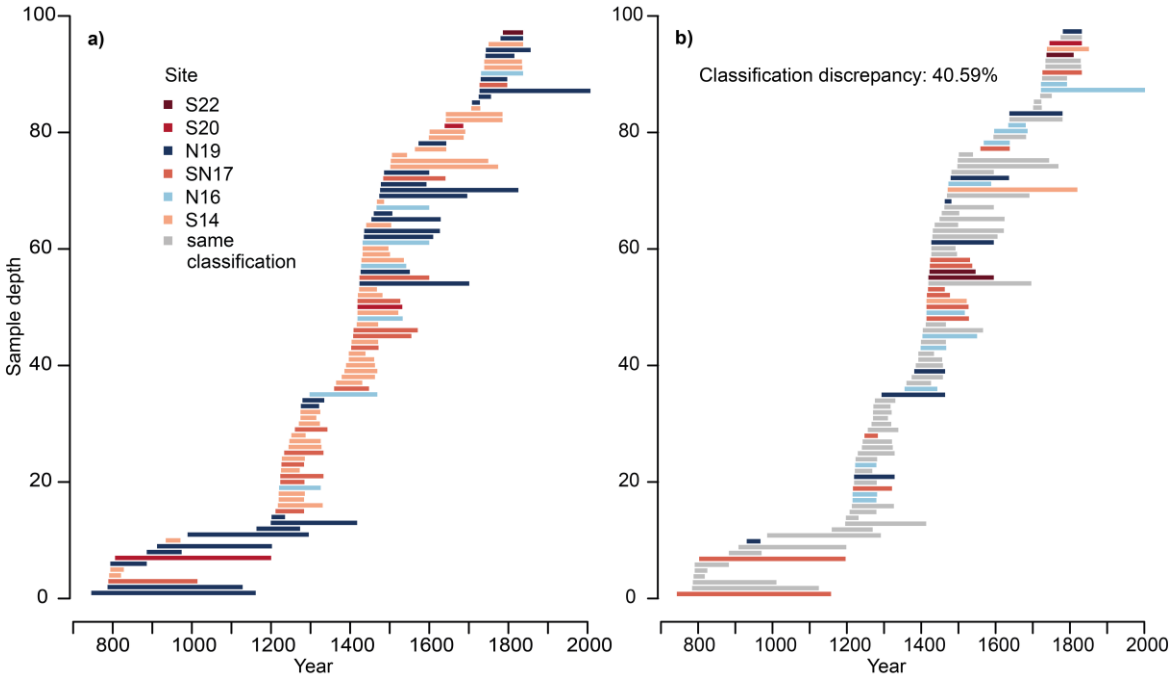


Fig. 2-5 Segment plots for the classified historic timber by DM_{sp} (larch specified density and ring-width parameter model) in **a)** and RWM_{sp} (larch specified ring-width parameter model) in **b)**. Colours in a) denote to the different elevational classes, the historical material was grouped to by DM_{sp}. In comparison to a), identical classifications of DM_{sp} and RWM_{sp} are greyed in b). Colours in b) pronounce classification differences and give the respective class.

As the analyses on D_{test} imply more accurate predictions by the DM_{sp} in contrast to the other models, especially to the ring-width models, the results of the classification of historical timber by DM_{sp} should be considered for further proceedings. For a reconstruction, we suggest using series allocated to the temperature sensitive high elevation sites N19, S20 and S22. This will result in using only 35 of 99 historical series but will improve the robustness of the signal strength of these series. In the Simplon valley, these high elevated sites show distinct offsets in their regional curves (Hartl et al. 2022), which must be considered when merging samples and sites into one RCS run (Esper et al. 2014). Excluding 54 historical series from a chronology would, however, massively reduce the sample replication. The period between 742-1450 CE

will almost never exceed five series per year and puts the development of a continuous millennium-length climate reconstruction at risk. Treeline shifts during the last millennium can alter the elevational temperature signal strength, and bias climate reconstructions (Büntgen et al. 2022). The provenance model cannot account for temporal changes in treeline elevation, but it helps reducing the bias by excluding historical series classified to recent lower elevations and reveals the series from close treeline sites with a very high likeliness of high temperature sensitivity.

2.3.4 Outlook and future applications of machine learning algorithms in dendroprovenancing

Our new approach for dendroprovenancing using ML shows considerable skill to differentiate tree-ring samples over short distances and among different elevations. While this application is limited to the Simplon Valley in the Swiss Alps and European larch, the proposed scheme (**Fig. 2-3 a)**) could similarly be applied in other provenance studies. If suitable, existing multi-centennial to -millennial long chronologies, based on ex situ historical or relict wood might be improved using ML techniques. Improved provenance determination of dead wood will increase the temporal stability of the climate signal of a chronology and enable a more reliable reconstruction of past climate.

Algorithms may also be trained with geographical coordinates as target y and a matrix X of series features from different chronologies to detect wood trade routes and origins of art or ship timber. It is mentionable that for different study areas, the chosen final algorithm might not match our best performing one, since tested algorithms might outperform each other differently depending on the region or species (Wolpert 1996). We acknowledge that the available features in this study are unique and are often not available in this quantity, hindering the exact reproduction of the models. We consider the selected basic features of our study as a good starting point, which can be extended with other tree-ring parameters, e.g., blue intensity, wood anatomical features, dendrochemical parameters, biomarkers or isotopic signatures, that are already tested in other provenance studies (e.g., Hajj et al. 2017; Traoré et al. 2018; Akhmetzyanov et al. 2020; Domínguez-Delmás et al. 2020). In the presented study, the LMB influence strongly impacts the performance of our models. Therefore, we suggest testing our approach with a non-host species, i.e., with the tree-ring series of previous approaches, to compare general cross-species models with each other. We hypothesis that the reduced performance of a general cross-species model (DM_{sp}, RWM_{sp}) likely results from the LBM manifestations in the larch samples. Using an undisturbed tree species might also result in

better performances of the ring-width models. More data from living trees would likely improve training the presented models. The DM_{sp} should especially be tested in future studies using corresponding blue intensity features, because it is a less labour and cost intensive approach for gaining information on tree-ring density (McCarroll et al. 2002; Campbell et al. 2007; Björklund et al. 2014).

2.4 Conclusion

Our novel approach using the ML algorithm XGBoost with tree-ring density and width data including species-specific features (DM_{sp}) improved to determine the provenance of wood of unknown origin without relying on a common period with (living-tree) reference data. The origin of 99 historical series was assigned along an elevational transect ranging from 1400 to 2200 m asl. Importantly, series from sites with diverging temperature responses have been identified and were consequentially excluded from a reconstruction. Our approach enables the user to include multiple parameters of individual trees and species and test various ML algorithms. It reveals how model performances are impacted by tree growth disturbances and how these performances can be used to detect the strength of growth disturbances. Our novel approach may serve as a framework for future applications of ML and dendroprovenancing in tree-ring research.

2.5 Acknowledgements

We are grateful to the forest district Simplon-South for the sampling permission. We thank Markus Kochbeck, Christian Gnaneswaran, Benedikt Lang, Philipp Schulz, Lara Meurer, Sophie Spelsberg and Jannes Fischer for help in the laboratory. CZ acknowledges support by the Bavarian Ministry of Science and the Arts in the context of the Bavarian Climate Research Network (BayKliF), JE by the Gutenberg Research College, JE and UB by SustES (CZ.02.1.01/0.0/0.0/16_019/0000797) and ERC (AdG 882727), PR by the German Research Foundation (ES 161/12-1), and CH by the German Research Foundation (HA 8048/1-1).

2.6 References

Akhmetzyanov L, Buras A, Sass-Klaassen U, den Ouden J, Mohren F, Groenendijk P, García-González I (2019) Multi-variable Approach Pinpoints Origin of Oak Wood with Higher Precision. J Biogeogr 46 (6):1163–77. https://doi.org/10.1111/jbi.13576.

Akhmetzyanov L, Sánchez-Salguero R, García-González I, Buras A, Dominguez-Delmás M, Mohren F, den Ouden J, Sass-Klaassen U (2020) Towards a New Approach for Dendroprovenancing Pines in the Mediterranean Iberian Peninsula. Dendrochronologia 60: 125688. https://doi.org/10.1016/j.dendro.2020.125688.

Babst F., Poulter B, Trouet V, Tan K, Neuwirth B, Wilson RJS, Carrer M, Grabner M, Tegel W, Levanic T, Panayotov M, Urbinati C, Bouriaud O, Ciais P, Frank D (2013) Site- and Species-Specific Responses of Forest Growth to Climate across the European Continent: Climate Sensitivity of Forest Growth across Europe. Glob Ecol Biogeogr 22 (6): 706–17. https://doi.org/10.1111/geb.12023.

Baltensweiler W, Rubli D (1984) Forstliche Aspekte Der Lärchenwickler-Massenvermehrung Im Oberengadin. Mit 39 Abbildungen Und 21 Tabellen. Zürich: Konkordia, Druck- und Verlags-AG.

Baltensweiler W, Rubli D (1999) Dispersal: An Important Driving Force of the Cyclic 367 Population Dynamics of the Larch Bud Moth, Zeiraphera Diniana Gn. For Snow Landsc Res 74 (1): 3–153.

Baltensweiler W, Weber UM, Cherubini P (2008) Tracing the Influence of Larch-Bud-Moth Insect Outbreaks and Weather Conditions on Larch Tree-Ring Growth in Engadine (Switzerland). Oikos 117 (2): 161–72. https://doi.org/10.1111/j.2007.0030-1299.16117.x.

Berkson, J (1944) Application of the Logistic Function to Bio-Assay. J Am Stat Assoc 39 (227): 357–65. https://doi.org/10.1080/01621459.1944.10500699.

Björklund J, von Arx G, Nievergelt D, Wilson RJS, Van den Bulcke J, Günther B, Loader NJ, Rydval M, Fonti P, Scharnweber T, Andreu-Hayles L, Büntgen U, D'Arrigo R, Davi N, De Mil T, Esper J, Gärtner H, Geary J, Gunnarson BE, Hartl C, Hevia A, Song H, Janecka K, Kaczka RJ, Kirdyanov AV, Kochbeck M, Liu Y, Meko M, Mundo I, Nicolussi K, Oelkers R, Pichler T, Sánchez-Salguero R, Schneider L, Schweingruber F, Timonen M, Trouet V, Van Acker J, Verstege A, Villalba R, Wilmking M, Frank D (2019) Scientific Merits and Analytical Challenges of Tree-Ring Densitometry. Rev Geophys 57(4). https://doi.org/10.1029/2019RG000642

Björklund, J, Gunnarson BE, Seftingen K, Esper J, Linderholm HW (2014) Blue Intensity and Density from Northern Fennoscandian Tree Rings, Exploring the Potential to Improve Summer Temperature Reconstructions with Earlywood Information. Clim Past 10 (2): 877–85. https://doi.org/10.5194/cp-10-877-2014.

Bodesheim P, Babst F, Frank D, Hartl C, Zang CS, Jung M, Reichstein M, Mahecha M (2022) Predicting Spatiotemporal Variability in Radial Tree Growth at the Continental Scale with Machine Learning. Environ Data Sci 1: e9, 1–35. https://doi.org/10.1017/eds.2022.8.

Bonde, N (1992) Dendrochronology and Timber Trade in Norther Europe from the 15th to 17th Century. In: Bartholin TS, Berglund BE, Eckstein D, Schweingruber F, Eggertsson O (eds) Proceedings of the International Dendrochronological Symposium 34:53–55. Ystad, Sweden: Lundqua Report.

Bonde N, Tyers I, Wazny T (1997) Where Does The Timer Come From? Dendrochronological Evidence of the Timber Trade in Northern Europe. In: Sinclair A, Slater E, Gowlett J (eds) Archaeological Sciences 1995. Proceedings of a Conference on the Application of Scientific Techiques to the Study of Archaeology: 201-204. Liverpool: Oxbow Books.

Bräker OU (1981) Der Alterstrend bei Jahrringdichten und Jahrringbreiten von Nadelhölzern und sein Ausgleich. Mitteilungen der forstlichen Bundesversuchsanstalt Wien 142:75–102

Breiman L (2001) Random Forests. Machine Learning 45:5–32. https://doi.org/10.1023/A:1010933404324

Bridge M (2000) Can Dendrochronology Be Used to Indicate the Source of Oak Within Britain? Vernac Archit 31 (1): 67–72. https://doi.org/10.1179/vea.2000.31.1.67.

Bridge M (2011) Resource Exploitation and Wood Mobility in Northern European Oak: Dendroprovenancing of Individual Timbers from the Mary Rose (1510/11-1545). The Int J Naut Archaeol 40 (2): 417–23. https://doi.org/10.1111/j.1095-9270.2010.00309.x.

Bridge M (2012) Locating the Origins of Wood Resources: A Review of Dendroprovenancing. J Archaeol Sci 39 (8): 2828–34. https://doi.org/10.1016/j.jas.2012.04.028.

Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber F, Karlén W, Zetterberg P, Eronen M (1992) Fennoscandian summers from ad 500: temperature changes on short and long timescales. Clim Dyn 7(3):111–119. https://doi.org/10.1007/BF00211153

Briffa KR, Jones PD, Pilcher JR, Hughes MK (1988) Reconstructing Summer Temperatures in Northern Fennoscandinavia Back to A.D. 1700 Using Tree-Ring Data from Scots Pine. Arc Alp Res 20 (4): 385–94. https://doi.org/10.2307/1551336.

Brookhouse M, Ives S, Dredge P, Howard D, Bridge M (2021) Mapping Henry: Dendrochronological Analysis of a Sixteenth-Century Panel Painting Based Upon Synchrotron-Sourced X-Ray Fluorescence Mapping. Stud Conserv 66 (7): 384–96. https://doi.org/10.1080/00393630.2020.1848133.

Bunn AG (2010) Statistical and visual crossdating in R using the dplR library. Dendrochronologia 28(4):251–258. https://doi.org/10.1016/j.dendro.2009.12.001

Bunn AG, Hughes MK, Salzer MW (2011) Topographically modified tree-ring chronologies as a potential means to improve paleoclimate inference: A letter. Climatic Change 105(3–4):627–634. https://doi.org/10.1007/s10584-010-0005-5

Büntgen U, Esper J, Frank D, Nicolussi K, Schmidhalter M (2005) A 1052-year tree-ring proxy for Alpine summer temperatures. Clim Dyn 25(2–3):141–153. https://doi.org/10.1007/s00382-005-0028-1

Büntgen U, Frank D, Liebhold A, Johnson D, Carrer M, Urbinati C, Grabner M, Nicolussi K, Levanic T, Esper J (2009) Three centuries of insect outbreaks across the European Alps. New Phytologist 182(4):929–941. https://doi.org/10.1111/j.1469-8137.2009.02825.x

Büntgen U, Piermattei A, Crivellaro A, Reinig F, Krusic PJ, Trnka M, Torbenson M, Esper J (2022) Common Era treeline fluctuations and their implications for climate reconstructions. Glob Planet Change 219:103979. https://doi.org/10.1016/j.gloplacha.2022.103979

Buras A, van der Maaten-Theunissen M, van der Maaten E, Ahlgrimm S, Hermann P, Simard S, Heinrich I, Helle G, Unterseher M, Schnittler M, Eusemann P, Wilmking M (2016) Tuning the Voices of a Choir: Detecting Ecological Gradients in Time-Series Populations. PLoS ONE 11(7):e0158346. https://doi.org/10.1371/journal.pone.0158346

Campbell R, McCarroll D, Loader NJ, Grudd H, Robertson I, Jalkanen R (2007) Blue intensity in Pinus sylvestris tree-rings: developing a new palaeoclimate proxy. The Holocene 17(6):821–828. https://doi.org/10.1177/0959683607080523

Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, San Francisco California USA, pp 785–794

Cook ER (1985) A time series analysis approach to tree ring standardization. Dissertation, University of Arizona

Cook ER, Kairiukstis LA (eds) (1990) Methods of Dendrochronology. Applications in the Environmental Sciences. Springer-Science+Business Media, B. V., Laxemburg, Austria

Cook ER, Kushnir Y, Smerdon JE, Williams AP, Anchukaitis KJ, Wahl ER (2019) A Euro-Mediterranean tree-ring reconstruction of the winter NAO index since 910 C.E. Clim Dyn 53(3–4):1567–1580. https://doi.org/10.1007/s00382-019-04696-2

Coomes DA, Allen RB (2007) Effects of size, competition and altitude on tree growth. J Ecol 95(5):1084–1097. https://doi.org/10.1111/j.1365-2745.2007.01280.x

Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. J Geophys Res Atmos 123(17):9391–9409. https://doi.org/10.1029/2017JD028200

Daly A, Tyers I (2022) The sources of Baltic oak. Archaeol Sci 139:105550. https://doi.org/10.1016/j.jas.2022.105550

Daux V, Edouard JL, Masson-Delmotte V, Stievenard M, Hoffmann G, Pierre M, Mestre O, Danis PA, Guibal F (2011) Can climate variations be inferred from tree-ring parameters and stable isotopes from Larix decidua? Juvenile effects, budmoth outbreaks, and divergence issue. Earth Planet Sci Lett 309(3–4):221–233. https://doi.org/10.1016/j.epsl.2011.07.003

Dittmar C, Eißing T, Rothe A (2012) Elevation-specific tree-ring chronologies of Norway spruce and Silver fir in Southern Germany. Dendrochronologia 30(2):73–83. https://doi.org/10.1016/j.dendro.2011.01.013

Domínguez-Delmás M, Rich S, Traoré M, Hajj F, Poszwa A, Akhmetzyanov L, García-González I, Groenendijk P (2020) Tree-ring chronologies, stable strontium isotopes and biochemical compounds: Towards reference datasets to provenance Iberian shipwreck timbers. J Archaeol Sci 34:102640. https://doi.org/10.1016/j.jasrep.2020.102640

Drake BL (2018) Source & Sourceability: Towards a probabilistic framework for dendroprovenance based on hypothesis testing and Bayesian inference. Dendrochronologia 47:38–47. https://doi.org/10.1016/j.dendro.2017.12.004

Esper J, Büntgen U, Frank D, Nievergelt D, Liebhold A (2007) 1200 years of regular outbreaks in alpine insects. Proc R Soc B 274(1610):671–679. https://doi.org/10.1098/rspb.2006.0191

Esper J, Büntgen U, Timonen M, Frank D (2012) Variability and extremes of northern Scandinavian summer temperatures over the past two millennia. Glob Planet Change 88–89:1–9. https://doi.org/10.1016/j.gloplacha.2012.01.006

Esper J, Cook ER, Krusic PJ, Peters K, Schweingruber F (2003) Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies. Tree-Ring Res 59(2):81–89

Esper J, Düthorn E, Krusic PJ, Timonen M, Büntgen U (2014) Northern European summer temperature variations over the Common Era from integrated tree-ring density records. J Quat Sci 29(5):487–494. https://doi.org/10.1002/jqs.2726

Esper J, Frank D, Wilson RJS, Briffa KR (2005) Effect of scaling and regression on reconstructed temperature amplitude for the past millennium. Geophys Res Lett 32(7):L07711. https://doi.org/10.1029/2004GL021236

Esper J, Krusic PJ, Ljungqvist FC, Luterbacher J, Carrer M, Cook ER, Davi N, Hartl-Meier C, Kirdyanov AV, Konter O, Myglan VS, Timonen M, Treydte K, Trouet V, Villalba R, Yang B, Büntgen U (2016) Ranking of tree-ring based temperature reconstructions of the past millennium. Quat Sci Rev 145:134–151. https://doi.org/10.1016/j.quascirev.2016.05.009

Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugen 7(2):179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Fix E, Hodges JL (1951) Discriminatory analysis. nonparametric discrimination: consistency properties. USA Air Force School of Aviation Medicine, Randolph Field, Texas, USA

Friedman JH (2001) Greedy function approximation: A gradient boosting machine. Ann Statist 29(5):1189–1232. https://doi.org/10.1214/aos/1013203451

Géron A (2019) Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd edn. O'Reilly, Sebastopol

Gu H, Wang J, Ma L, Shang Z, Zhang Q (2019) Insights into the BRT (Boosted Regression Trees) Method in the Study of the Climate-Growth Relationship of Masson Pine in Subtropical China. Forests 10(3):228. https://doi.org/10.3390/f10030228

Gurskaya M, Hallinger M, Singh J, Agafonov L, Wilmking M (2012) Temperature reconstruction in the Ob River valley based on ring widths of three coniferous tree species. Dendrochronologia 30(4):302–309. https://doi.org/10.1016/j.dendro.2012.04.002

Gut U (2018) Evaluating the key assumptions underlying dendro-provenancing: How to spruce it up with a scissor plot. Dendrochronologia 52:131–145. https://doi.org/10.1016/j.dendro.2018.09.008

Hajj F, Poszwa A, Bouchez J, Guérold F (2017) Radiogenic and “stable” strontium isotopes in provenance studies: A review and first results on archaeological wood from shipwrecks. J Archaeol Sci 86:24–49. https://doi.org/10.1016/j.jas.2017.09.005

Haneca K, Wazny T, Van Acker J, Beeckman H (2005) Provenancing Baltic timber from art historical objects: success and limitations. J Archaeol Sci 32(2):261–271. https://doi.org/10.1016/j.jas.2004.09.005

Harr L, Esper J, Kirchhefer JA, Zhou W, Hartl C (2021) Growth response of Betula pubescens Ehrh. to varying disturbance factors in northern Norway. Trees 35(2):421–431. https://doi.org/10.1007/s00468-020-02043-1

Hartl C, Düthorn E, Tejedor E, Kirchhefer AJ, Timonen M, Holzkämper S, Büntgen U, Esper J (2021) Micro-site conditions affect Fennoscandian forest growth. Dendrochronologia 65:125787. https://doi.org/10.1016/j.dendro.2020.125787

Hartl C, Schneider L, Riechelmann DFC, Kuhl E, Kochbeck M, Klippel L, Büntgen U, Esper J (2022) The temperature sensitivity along elevational gradients is more stable in maximum latewood density than tree-ring width. Dendrochronologia 73:125958. https://doi.org/10.1016/j.dendro.2022.125958

Hartl-Meier C, Büntgen U, Esper J (2016) On the occurrence of cyclic larch budmoth outbreaks beyond its geographical hotspots. In: Hevia A, Sánchez-Salguero R, Linares JC, Olano JM, Camarero JJ, Gutiérrez E, Helle G, Gärtner H (eds) Proceedings of the DENDROSYMPOSIUM 2015: May 20th - 23rd, 2015 in Sevilla, Spain, (Scientific Technical Report ; 16/04), 14th TRACE conference (Tree Rings in Archaeology, Climatology and Ecology). Deutsches GeoForschungsZentrum GFZ, (Sevilla, Spain), Potsdam, pp 86–92

Hartl-Meier C, Dittmar C, Zang C, Rothe A (2014a) Mountain forest growth response to climate change in the Northern Limestone Alps. Trees 28(3):819–829. https://doi.org/10.1007/s00468-014-0994-1

Hartl-Meier C, Esper J, Liebhold A, Konter O, Rothe A, Büntgen U (2017) Effects of host abundance on larch budmoth outbreaks in the European Alps. Agr For Entomol 19(4):376–387. https://doi.org/10.1111/afe.12216

Hartl-Meier C, Zang C, Dittmar C, Esper J, Göttlein A, Rothe A (2014b) Vulnerability of Norway spruce to climate change in mountain forests of the European Alps. Clim Res 60(2):119–132. https://doi.org/10.3354/cr01226

Hoerl AE, Kennard RW (2000) Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 42(1):80–86. https://doi.org/10.1080/00401706.2000.10485983

Jevšenak J, Džeroski S, Zavadlav S, Levanic T (2018) A Machine Learning Approach to Analyzing the Relationship Between Temperatures and Multi-Proxy Tree-Ring Records. Tree-Ring Res 74(2):210–224. https://doi.org/10.3959/1536-1098-74.2.210

Jevšenak J, Skudnik M (2021) A random forest model for basal area increment predictions from national forest inventory data. For Ecol Manag 479:118601. https://doi.org/10.1016/j.foreco.2020.118601

Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. Science 349(6245):255–260. https://doi.org/10.1126/science.aaa8415

Kagawa A, Leavitt SW (2010) Stable carbon isotopes of tree rings as a tool to pinpoint the geographic origin of timber. J Wood Sci 56(3):175–183. https://doi.org/10.1007/s10086-009-1085-6

Keitt TH, Abelson ES (2021) Ecology in the age of automation. Science 373(6557):858–859. https://doi.org/10.1126/science.abi4692

Kiefer J, Wolfowitz J (1952) Stochastic Estimation of the Maximum of a Regression Function. Ann Math Stat 23(3):462–466. http://www.jstor.org/stable/2236690

King GM, Gugerli F, Fonti P, Frank D (2013) Tree growth response along an elevational gradient: climate or genetics? Oecologia 173(4):1587–1600. https://doi.org/10.1007/s00442-013-2696-6

Klippel L, Büntgen U, Konter O, Kyncl T, Esper J (2020) Climate sensitivity of high- and low-elevation Larix decidua MXD chronologies from the Tatra Mountains. Dendrochronologia 60:125674. https://doi.org/10.1016/j.dendro.2020.125674

Konter O, Esper J, Liebhold A, Kyncl T, Schneider L, Düthorn E, Büntgen U (2015a) Tree-ring evidence for the historical absence of cyclic larch budmoth outbreaks in the Tatra Mountains. Trees 29(3):809–814. https://doi.org/10.1007/s00468-015-1160-0

Konter O, Rosner K, Kyncl T, Esper J, Büntgen U (2015b) Spatiotemporal variations in the climatic response of Larix decidua from the Slovakian Tatra Mountains. In: Scientific Technical Report STR 15/06. Deutsches GeoForschungsZentrum GFZ, Sevilla, Spain, pp 62–68

Labuhn I, Daux V, Girardclos O, Stievenard M, Pierre M, Masson-Delmotte V (2016) French summer droughts since 1326 CE: a reconstruction based on tree ring cellulose d18O. Clim Past 12(5):1101–1117. https://doi.org/10.5194/cp-12-1101-2016

Lara A, Villalba R, Urrutia-Jalabert R, González-Reyes A, Aravena JC, Luckman BH, Cuq E, Rodríguez C, Wolodarsky-Franke A (2020) +A 5680-year tree-ring temperature record for southern South America. Quat Sci Rev 228:106087. https://doi.org/10.1016/j.quascirev.2019.106087

Li Z-S, Zhang Q-B, Ma K (2012) Tree-ring reconstruction of summer temperature for A.D. 1475–2003 in the central Hengduan Mountains, Northwestern Yunnan, China. Clim Change 110(1–2):455–467. https://doi.org/10.1007/s10584-011-0111-z

Linderholm HW, Gunnarson BE, Fuentes M, Büntgen U, Hormes A (2021) The origin of driftwood on eastern and south-western Svalbard. Polar Sci 29:100658. https://doi.org/10.1016/j.polar.2021.100658

Liu Y, An ZS, Linderholm HW, Chen DL, Song HM, Cai QF, Sun JY, Tian H (2009) Annual temperatures during the last 2485 years in the mid-eastern Tibetan Plateau inferred from tree rings. Sci China Ser D-Earth Sci 52(3):348–359. https://doi.org/10.1007/s11430-009-0025-z

Ljungqvist FC, Piermattei A, Seim A, Krusic PJ, Büntgen U, He M, Kirdyanov AV, Luterbacher J, Schneider L, Seftigen K, Stahle DW, Villalba R, Yang B, Esper J (2020) Ranking of tree-ring based hydroclimate reconstructions of the past millennium. Quat Sci Rev 230:106074. https://doi.org/10.1016/j.quascirev.2019.106074

McCarroll D, Pettigrew E, Luckman A, Guibal F, Edouard J-L (2002) Blue Reflectance Provides a Surrogate for Latewood Density of High-latitude Pine Tree Rings. Arc Antarct Alp Res 34(4):450–453. https://doi.org/10.1080/15230430.2002.12003516

Osborn TJ, Briffa KR, Jones PD (1997) Adjusting variance for sample-size in tree-ring chronologies and other regional mean timeseries. Dendrochronologia 15:89–99

Ou Q, Lei X, Shen C (2019) Individual Tree Diameter Growth Models of Larch–Spruce–Fir Mixed Forests Based on Machine Learning Algorithms. Forests 10(2):187. https://doi.org/10.3390/f10020187

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine Learning in Python. JMLR 12:2825–2830. https://doi.org/10.48550/arXiv.1201.0490

R Core Team (2021) R: A language and environment for statistical computing. Vienna, Austria Riechelmann DFC, Hartl C, Esper J (2020) The effect of provenance of historical timber on tree-ring based temperature reconstructions in the Western Central Alps. iForest 13(1):351–359. https://doi.org/10.3832/ifor3412-013

Riechelmann DFC, Schmidhalter M, Büntgen U, Esper J (2013) Extending the high elevation larch ring width chronology from the Simplon region in the Swiss Alps over the past millennium. In: Helle G, Gärtner H, Beck W, Heinrich I, Heussner K-U, Müller A, Sanders T (eds) Proceedings of the DENDROSYMPOSIUM 2012: May 8th - 12th, 2012 in Potsdam and Eberswalde, Germany, (Scientific Technical Report ; 13/05), 11th TRACE conference (Tree Rings in Archaeology, Climatology and Ecology). Deutsches GeoForschungsZentrum GFZ, Potsdam, Eberswalde, pp 103–107

Rolland C, Baltensweiler W, Petitcolas V (2001) The potential for using Larix decidua ring widths in reconstructions of larch budmoth (Zeiraphera diniana) outbreak history: dendrochronological estimates compared with insect surveys. Trees 15(7):414–424. https://doi.org/10.1007/s004680100116

Römer P, Hartl C, Schneider L, Bräuning A, Szymczak S, Huneau F, Lebre S, Reinig F, Büntgen U, Esper J (2021) Reduced Temperature Sensitivity of Maximum Latewood Density Formation in High-Elevation Corsican Pines under Recent Warming. Atmosphere 12(7):804. https://doi.org/10.3390/atmos12070804

Rozenberg P, Chauvin T, Escobar-Sandoval M, Huard F, Shishov V, Charpentier J-P, Sergent A-S, Vargas-Hernandez JJ, Martinez-Meier A, Pâques L (2020) Climate warming differently affects Larix decidua ring formation at each end of a French Alps elevational gradient. Ann For Sci 77(2):54. https://doi.org/10.1007/s13595-020-00958-w

Salehnia N, Ahn J (2022) Modelling and reconstructing tree ring growth index with climate variables through artificial intelligence and statistical methods. Ecol Indic 134:108496. https://doi.org/10.1016/j.ecolind.2021.108496

Salzer MW, Hughes MK, Bunn AG, Kipfmueller KF (2009) Recent unprecedented tree-ring growth in bristlecone pine at the highest elevations and possible causes. PNAS 106(48):20348–20353. https://doi.org/10.1073/pnas.0903029106

Salzer MW, Larson ER, Bunn AG, Hughes MK (2014) Changing climate response in near-treeline bristlecone pine with elevation and aspect. Environ Res Lett 9(11):114007. https://doi.org/10.1088/1748-9326/9/11/114007

Saulnier M, Roques A, Guibal F, Rozenberg P, Saracco G, Corona C, Edouard J-L (2017) Spatiotemporal heterogeneity of larch budmoth outbreaks in the French Alps over the last 500 years. Can J For Res 47(5):667–680. https://doi.org/10.1139/cjfr-2016-0211

Schneider L, Smerdon JE, Büntgen U, Wilson RJS, Myglan VS, Kirdyanov AV, Esper J (2015) Revising midlatitude summer temperatures back to A.D. 600 based on a wood density network. Geophys Res Lett 42(11):4556–4562. https://doi.org/10.1002/2015GL063956

Schweingruber F (1988) Tree Rings. Paul Haupt, Bern, Stuttgart

Shindo L, Claude S (2019) Buildings and wood trade in Aix-en-Provence (South of France) during the Modern period. Dendrochronologia :29–36. https://doi.org/10.1016/j.dendro.2019.02.003

Tegel W, Vanmoerkerke J, Büntgen U (2010) Updating historical tree-ring records for climate reconstruction. Quat Sci Rev 29(17–18):1957–1959. https://doi.org/10.1016/j.quascirev.2010.05.018

Traoré M, Kaal J, Martínez Cortizas A (2018) Differentiation between pine woods according to species and growing location using FTIR-ATR. Wood Sci Technol 52(2):487–504. https://doi.org/10.1007/s00226-017-0967-9

Van Rossum G, Drake FL (2009) Python 3 Reference manual. CreateSpace, Scotts Valley, CA

Vapnik V (1991) Principles of Risk Minimization for Learning Theory. In: Moody J, Hanson S, Lippmann RP (eds) NeurIPS Proceedings. p 8

Vapnik V, Chervonenkis A (1974) Theory of Pattern Recognition. Nauka, Moscow, Russia

Wazny T (2002) Baltic timber in Western Europe – an exciting dendrochronological question. Dendrochronologia 20(3):313–320. https://doi.org/10.1078/1125-7865-00024

Wilson RJS, Anchukaitis KJ, Briffa KR, Büntgen U, Cook ER, D'Arrigo R, Davi N, Esper J, Frank D, Gunnarson BE, Hegerl GC, Helama S, Klesse S, Krusic PJ, Linderholm HW, Myglan V, Osborn TJ, Rydval M, Schneider L, Schurer A, Wiles G, Zhang P, Zorita E (2016) Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. Quat Sci Rev 134:1–18. https://doi.org/10.1016/j.quascirev.2015.12.005

Wilson RJS, Esper J, Luckman BH (2004) Utilising historical tree-ring data for dendroclimatology: A case study from the Bavarian Forest, Germany. Dendrochronologia 21(2):53–68. https://doi.org/10.1078/1125-7865-00041

Wilson RJS, Hopfmueller M (2001) Dendrochronological investigations of Norway spruce along an elevational transect in the Bavarian Forest, Germany. Dendrochronologia 19(1):67–79

Wilson RJS, Luckman BH (2003) Dendroclimatic reconstruction of maximum summer temperatures from upper treeline sites in Interior British Columbia, Canada. The Holocene 13(6):851–861. https://doi.org/10.1191/0959683603hl663rp

Wilson RJS, Luckman BH, Esper J (2005) A 500 year dendroclimatic reconstruction of spring-summer precipitation from the lower Bavarian Forest region, Germany. Int J Climatol 25(5):611–630. https://doi.org/10.1002/joc.1150

Wolpert DH (1996) The Existence of A Priori Distinctions Between Learning Algorithms. Neural Comput 8(7):1391–1420. https://doi.org/10.1162/neco.1996.8.7.1391

Ying X (2019) An Overview of Overfitting and its Solutions. J Phys: Conf Ser 1168:022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Zang C, Biondi F (2015) treeclim: an R package for the numerical calibration of proxy-climate relationships. Ecography 38(4):431–436. https://doi.org/10.1111/ecog.01335

Zhang H (2004) The Optimality of Naive Bayes. Aa 1(2):3

Zhang P, Björklund J, Linderholm HW (2015) The influence of elevational differences in absolute maximum density values on regional climate reconstructions. Trees 29(4):1259–1271. https://doi.org/10.1007/s00468-015-1205-4

2.7 Supplementary material

Tab. S 2-1 Descriptive statistics of the living tree sites and the historical dataset: mean series length (MSL), length of chronology (Range), First-Order Autocorrelation (AR1) and inter-series correlation (Rbar) of tree-ring width (TRW) and maximum latewood density (MXD) data.

Code	Elevation [m asl]	Exposure	MSL	First year (>5)	Last year	Range	AR1 TRW	AR1 MXD	Rbar TRW	Rbar MXD
S14	1400	south	73	1884 (1928)	2011	128	0.68	0.5	0.7	0.37
N16	1600	north	134	1860 (1870)	2011	152	0.83	0.44	0.64	0.43
SN17	1700	south & north	133	1679 (1854)	2011	333	0.77	0.44	0.56	0.49
N19	1900	north	205	1582 (1685)	2010	429	0.67	0.54	0.64	0.56
S20	2000	south	213	1672 (1714)	2010	339	0.77	0.61	0.7	0.62
S22	2200	south	218	1542 (1717)	2009	468	0.7	0.52	0.71	0.64
Historical	**unknown**	**unknown**	115	742 (790)	2002	1261	0.75	0.51	0.47	0.31

Tab. S 2-2 Mean values of tree-ring proxies: average growth rate (AGR), site means of maximum latewood density (MXD), earlywood width (EWW), latewood width (LWW), earlywood density (EWD) and latewood density (LWD).

Code	Elevation [m asl]	AGR	Mean MXD	Mean EWW	Mean LWW	Mean EWD	Mean LWD
S14	**1400**	1.58	1.02	1.08	0.49	0.35	0.91
N16	**1600**	0.94	1.02	0.66	0.29	0.34	0.91
SN17	**1700**	0.91	0.98	0.62	0.29	0.36	0.87
N19	**1900**	0.97	0.9	0.67	0.3	0.34	0.79
S20	**2000**	0.72	0.86	0.52	0.21	0.34	0.76
S22	**2200**	0.72	0.84	0.51	0.21	0.36	0.73
Historical	**unknown**	1.16	0.99	0.84	0.32	0.37	0.88

Tab. S 2-3 Hyperparameters of the fine-tuned models.

ML Algorithm	**Hyperparameter**
Ridge Regression Classifier	{'alpha': 0.3, 'solver': 'cholesky'}
Logistic Regression Classifier (Softmax Regression)	{'multi_class': 'multinomial', 'solver': 'newton-cg', 'C': 15, 'penalty': 'l2', 'max_iter': 9000000}
Gaussian Naïve Bayes	{'var_smoothing': 8.111308307896856e-09}
Random Forest	{'criterion': 'entropy','max_depth': 20, 'min_impurity_decrease': 0.0001, 'min_sample_split': 5, 'n_estimators': 75}
Extreme Gradient Boosting (XGBoost)	{'eval_metric': 'mlogloss', 'eta': 0.3, 'max_depth': 5, 'subsample': 1, 'n_estimators': 250, 'colsample_by_tree': 1, 'gamma': 0.1, 'min_child_weight': 1, earlystopping}
k-Nearest Neighbor	{'algorithm': 'kd_tree', 'leaf_size': 1, 'metric': 'euclidean', 'n_neighbors': 1, 'weights': 'uniform'}
Support Vector Machine	{'C': 30, 'gamma': 0.005}
Stochastic Gradient Decent	{'alpha': 0.1, 'learning_rate': 'optimal', 'loss': 'squared_hinge', 'penalty': 'l1'}
Linear Discriminant Analysis	{'shrinkage': 1, 'solver': 'lsqr'}

Tab. S 2-4 Tested fine-tuned Machine learning algorithms with f1 score means and standard deviations (std) of repeated stratified k-fold (k = 10, repeats = 100) cross-validation.

ML Algorithm	**f1 score mean**	**f1 score std**
Ridge Regression Classifier	0.58	0.12
Logistic Regression Classifier (Softmax Regression)	0.57	0.13
Gaussian Naïve Bayes	0.59	0.12
Random Forest	0.68	0.12
Extreme Gradient Boosting (XGBoost)	0.70	0.12
k-Nearest Neighbor	0.68	0.13
Support Vector Machine	0.61	0.14
Stochastic Gradient Decent	0.16	0.08
Linear Discriminant Analysis	0.35	0.08

Tab. S 2-5 Classification Report: performance metrics of D_{test} (test dataset) on XGBoost in each individual class for DM_{gen} (density and ring width parameter model unspecified and general) and RWM_{gen} (ring width parameter model unspecified and general) (in brackets).

Site	Precision	Recall	f1 score
S14	0.63 (0.5)	1 (0.8)	0.77 (0.62)
N16	1 (0.67)	0.6 (0.4)	0.75 (0.5)
SN17	0.67 (0.0)	0.4 (0.0)	0.5 (0.0)
N19	0.38 (0.17)	0.6 (0.2)	0.46 (0.18)
S20	1 (0.0)	0.5 (0.0)	0.67 (0.0)
S22	1 (0.2)	1 (0.2)	1 (0.2)
Average	0.78 (0.26)	0.68 (0.27)	0.69 (0.25)

a)	Precision	Recall	f1score
b: Akhmetzyianov et al. 2019/2020 (LWD)	0.4	0.22	0.28
c: Akhmetzyianov et al. 2019/2020 (TRW)	0.25	0.13	0.16
d: Wilson et al. (2004)	0.28	0.24	0.25

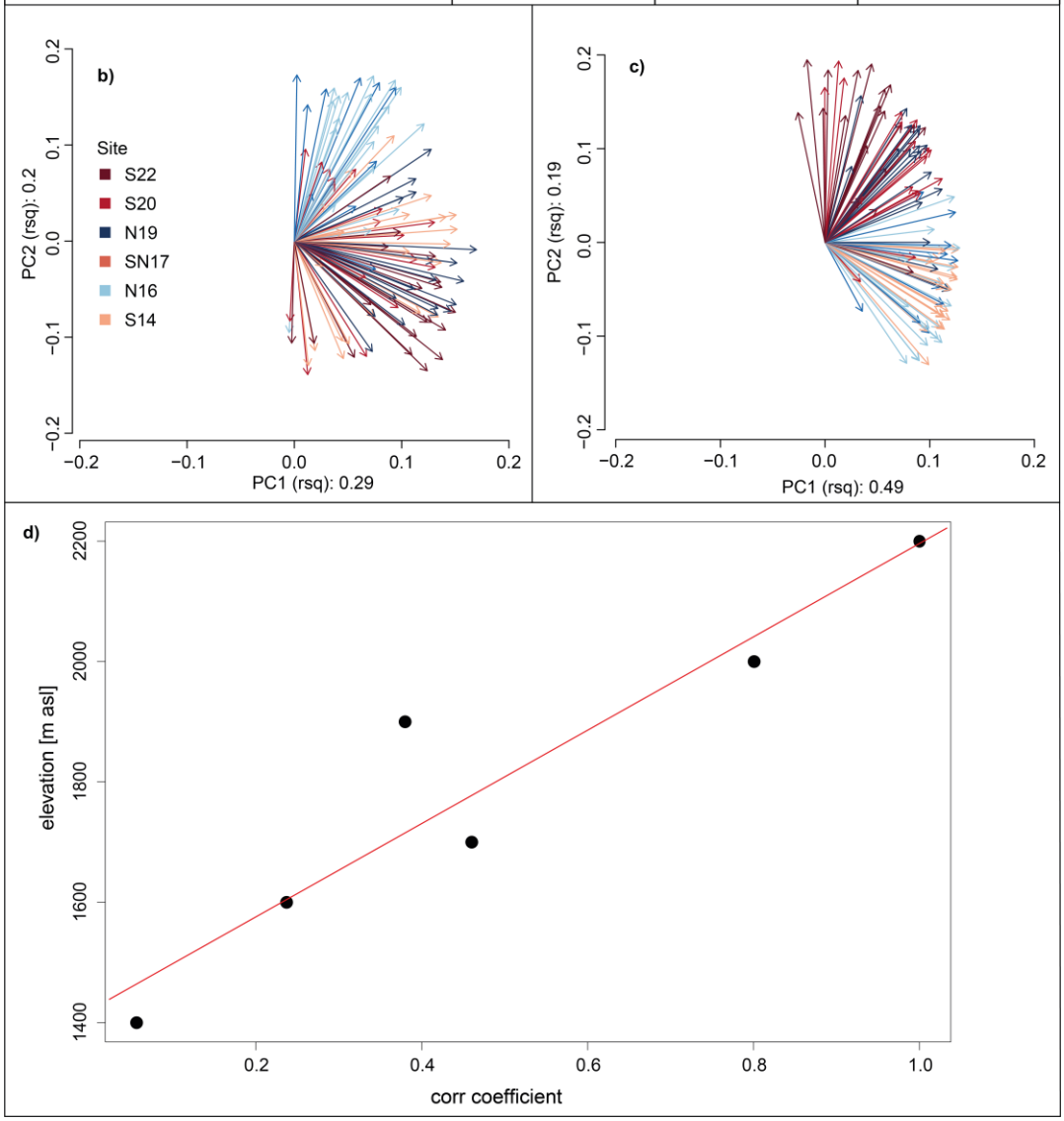


Fig. S 2-1 a) Other tested approaches. Results from the PCGA with **b)** latewood density and **c)** tree-ring width and **d)** the linear regression model. Although ring width in c) seems to be better separated, the density PCGA provenancing b) works better for provenancing (see a)), which is in line with findings from Akhmetzyanov et al. (2020).

3 Revising Alpine summer temperatures since 881 CE

Eileen Kuhl[1], Jan Esper[1,7], Lea Schneider[2], Valerie Trouet[3], Marcel Kunz[1], Lara Klippel[4], Ulf Büntgen[5,7-9] and Claudia Hartl[6]

[1] Department of Geography, Johannes Gutenberg University, Mainz, Germany

[2]Department of Geography, Justus-Liebig-University, Gießen, Germany

[3]Laboratory of Tree-Ring Research, University of Arizona, Tucson, USA

[4]Deutscher Wetterdienst, Offenbach, Germany

[5]Department of Geography, University of Cambridge, Cambridge, UK

[6]Nature Rings - Environmental Research and Education, Mainz, Germany

[7]Global Change Research Centre (CzechGlobe), Brno, Czech Republic

[8]Swiss Federal Research Institute (WSL), Birmensdorf, Switzerland

[9]Department of Geography, Masaryk University, Brno, Czech Republic

Published in March 2024: Climate Dynamics (2024). DOI: 10.1007/s00382-024-07195-1

Summary

Europe experienced severe heat waves during the last decade, which impacted ecological and societal systems and are likely to increase under projected global warming. A better understanding of pre-industrial warm-season changes is needed to contextualize these recent trends and extremes. Here, we introduce a network of 352 living and relict larch trees (*Larix decidua* Mill.*)* from the Matter and Simplon valleys in the Swiss Alps to develop a maximum latewood density (MXD) chronology calibrating at r = 0.8 (p>0.05, 1901-2017 CE) against May-September temperatures over Western Europe. Machine learning is applied to identify historical wood samples aligning with growth characteristics of sites from elevations above 1900 m asl to extend the modern part of the chronology back to 881 CE. The new Alpine record reveals warmer conditions in the 10th century, followed by an extended cold period during the late Medieval times, a less-pronounced Little Ice Age culminating in the 1810s, and prolonged anthropogenic warming until present. The Samalas eruption likely triggered the coldest reconstructed summer in Western Europe in 1258 CE (-2.32°C), which is in line with a recently published MXD-based reconstruction from the Spanish Pyrenees. Whereas the new Alpine reconstruction is potentially constrained in the lowest frequency, centennial timescale domain, it overcomes variance biases in existing state-of-the-art reconstructions and sets a new standard in site-control of historical samples and calibration/ verification statistics.

3.1 Introduction

There is a rising urgency to improve our understanding of natural and anthropogenic drivers of temperature variations and reduce uncertainties in model predictions (Eyring et al. 2019). Analysing past climate variability, identifying spatial patterns and updating paleoclimate models is indispensable to establish the basic conditions for these models (Esper and Büntgen 2021). Common paleoclimate archives are tree-rings from high-elevation and high-latitude sites, which are used to reconstruct past temperature variability and place anthropogenic warming into historical context (Schweingruber 1988). This is not only true for long-term changes in climate history but for annual fluctuations, for example, exceptionally warm/cold and wet/dry years. With information on extreme events, we can set recent extreme events (and the likelihood of future events) in an historical context regarding their timing, occurrence and strength (e.g., Qin et al. 2011; Lyu et al. 2016; Borkotoky et al. 2021).

The European Alps have been of interest for such studies for decades as the regional warming already exceeded 2°C from 1864 to 2017 CE (Allgaier Leuch et al. 2017). Tree-ring width (TRW) and maximum latewood density (MXD) have been used as proxies to develop eleven Alpine temperature reconstructions (Schweingruber et al. 1987; 1988; Büntgen et al. 2005; Frank et al. 2005; Frank and Esper 2005a; Büntgen et al. 2006; Esper et al. 2007b; Corona et al. 2010; Corona et al. 2011; Trachsel et al. 2012; Coppola et al. 2013; Leonelli et al. 2016, **Tab. 3-1**). Mainly, these publications used larch (*Larix decidua* Mill.*)* tree-ring series with one exception, which is a spruce (*Picea abies* L.*)* reconstruction (Esper et al. 2007b).

Most analyses targeted summer temperatures from June to August (JJA), whereas a few MXD chronologies were used for longer seasons: Schweingruber et al. (1987/1988) and Büntgen et al. (2006) calibrated their MXD chronologies to June to September (JJAS) temperatures, while Frank and Esper (2005) reconstructed April to September temperatures. One exception is Esper et al. (2007b), which found best correlations with temperatures from June-July for TRW and August-September for MXD. Seven of the eleven reconstructions extend back to the first millennium CE (Schweingruber et al. 1987; 1988; Büntgen et al. 2005; Büntgen et al. 2006; Esper et al. 2007b; Corona et al. 2010; Corona et al. 2011; Trachsel et al. 2012). MXD-based chronologies reached correlations up to 0.69 with instrumental JJAS temperatures (Büntgen et al. 2006). The most recent reconstruction reported correlations up to 0.78 with JJA from 1763 to 2014 CE (Leonelli et al. 2016). Out of all mentioned publications, Büntgen et al. (2006) received by far the most citations for their reconstruction from the Swiss Lötschental (hereafter referred to as Lötschental) implying the data often being used in other studies and the findings having a generally high impact on the scientific community. Above all, it is the longest existing MXD-based reconstruction from the Alps based on a single species and serves as comparison for this study.

Tab. 3-1 Existing TRW- or MXD-based temperature reconstructions based on from the European Alps in order of publication year. The table includes the number of included sites, the elevation, the used proxies, the range of years covered, as well as the species (Larix decidua Mill. = LADE, Abies alba Mill. = ABAL, Pinus cembra L. = PICE, Picea abies L. = PIAB), the temperature signal used for reconstructing (letters denote to the first letter of the corresponding months, e.g. JJA June, July, August).

No.	Publication	# of Sites	Elevation [m asl]	Proxy	Period [CE]	Species	Temperature Signal
1	Schweingruber et al. 1987/ 1988[h]	16	100-1700, 1800	MXD	982-1976	PIAB, ABAL	JJAS
2	Büntgen et al. 2005[h]	5	>1500	TRW	951-2002	LADE, PICE	JJA
3	Frank & Esper 2005	53	>1500	TRW / MXD	1600-1988/ 1650-1987	PIAB, ABAL, LADE, PICE	JJA / AMJJAS
4	Frank et al. 2005	53	>1500	TRW	1760-1990	PIAB, ABAL, LADE, PICE	JJA
5	Büntgen et al. 2006[h]	4	>1900	MXD	755-2004	LADE	JJAS
6	Esper et al. 2007b[h]	4	-	TRW /MXD	1028-2003	PIAB	JJ / AS
7	Corona et al. 2010[h]	38	>1725	TRW & MXD	1000-2000	LADE, PICE	JJA
8	Corona et al. 2011[h]	34	>1600	TRW	751-2003	LADE	JJA
9	Coppola et al. 2013	4	>1910	TRW	1610-2008	LADE	JJA
10	Trachsel et al. 2012[h]	4	-	TRW, MXD, BI, Chrionomid/ biogenic Silica	755-2004	LADE, PIAB, PICE	JJA
11	Leonelli et al. 2016	42	> 1800	TRW	1470-2010	LADE, PIAB, PICE	JJA

[h] = used historical series from buildings

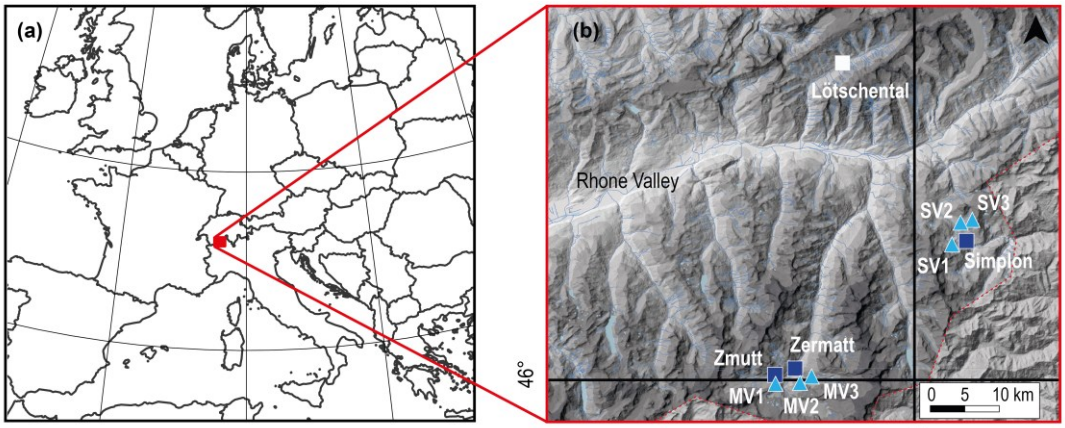
The mentioned millennium-long chronologies all include wood samples from historical buildings (e.g. Corona et al. 2010; Trachsel et al. 2012) as high-elevation forests in the Swiss Alps have been heavily influenced by anthropogenic forest use since the Middle Ages (Conedera et al. 2017). Therefore, the likelihood of living trees spanning earlier ages is low and *in situ* dead wood in the Alps is rare. The long history of human activity, however, enables the use of construction timber of old, high-elevation settlements like Zermatt in Switzerland (Coolidge 1912). The recurring issue with historical tree-ring samples in all of the existing reconstructions spanning a millennium is the limited site control and validation that these woods originate from elevations near treeline and contain strong temperature signals (Riechelmann et al. 2020). Multiple studies showed that the temperature signal in tree-ring parameters weakens with decreasing elevation (Neuwirth et al. 2004; Affolter et al. 2010; King et al. 2013; Hartl-Meier et al. 2014a; Hartl-Meier et al. 2014b; Salzer et al. 2014; Zhang et al. 2015; Hartl-Meier et al. 2017b; Hartl et al. 2022), which can consequently influence the temporal robustness of a reconstruction. Kuhl et al. (2023) recently introduced a method to mitigate these biasing effects by fitting tree-ring parameters and growth characteristics to a classification model to sort historical series to living stands with different distances to treeline. This provenance method was tested on data from the Swiss Simplon Valley in Kuhl et al. (2023) and is now applied to the Swiss Matter Valley data as well to exclude historical series with growth behaviours similar to sites further away from the current treeline. With this approach, we aim to build an improved millennium-length temperature reconstruction by the selection of historical material based on the information we gain from the provenance models.

We present a new MXD-based temperature reconstruction from 352 living and historic larch trees *(Larix decidua* Mill*.)* sampled in the southwestern Swiss Alps. We analyse the influences of detrending and variance stabilisation on reconstruction variability and trends, assess and mitigate the effects of larch budmoth infestation events, and revise Alpine temperature history since 881 CE including severe changes compared to existing reconstructions from the European Alps. The record is compared with other high-resolution warm season temperature reconstructions, and common trends and extremes are discussed to evaluate large-scale climate forcings.

3.2 Methods

3.2.1 Study sites and data distribution

The study sites in the canton Valais in the South-Western Swiss Alps include a total number of 352 series from living trees and historical buildings from valleys south of the Rhone valley (**Fig. 1a-c**). In the Simplon (SV) and Matter (MV) valleys (Hartl et al. 2022; Kunz et al. 2023), 146 living series of larch (Larix *decidua* Mill.) from high elevations (hereafter SV1-3 and MV1-3) were sampled. Likewise, 99 historical series from buildings in Simplon village and 206 historical series from buildings in Zermatt and Zmutt villages (Schmidhalter, M., Riechelmann et al. 2013; Riechelmann et al. 2020) were sampled (**Fig. 3-1 b-d)**). From these samples, X-ray densitometric measurements were conducted using a Walesch2003 (WALESCH, Electronic GmbH, Switzerland) as described in Björklund et al. (2019). The distribution of historical series in time was evaluated to prevent biases arising from an unbalanced dataset (**Fig. 3-2**, Esper et al. 2016).



(d)

Site	Elevation	N
SV1	1900	23
SV2	2000	22
SV3	2200	24
Simplon	1480	36
MV1	2000	30
MV2	2300	24
MV3	2300	23
Zermatt/Zmutt	1600/1900	170

Fig. 3-1 a) Study site in the Swiss Alps close to the Italian border, **b)** and locations of the sampling sites for living trees (turquoise triangles) and historical buildings (dark blue quarters) in the Matter Valley (MV) and the Simplon Valley (SV) **c)** The historical village of Zmutt in the Matter Valley (credits: C. Hartl) **d)** Information about the elevation [m asl] and the number of samples (increment cores and discs) per site.

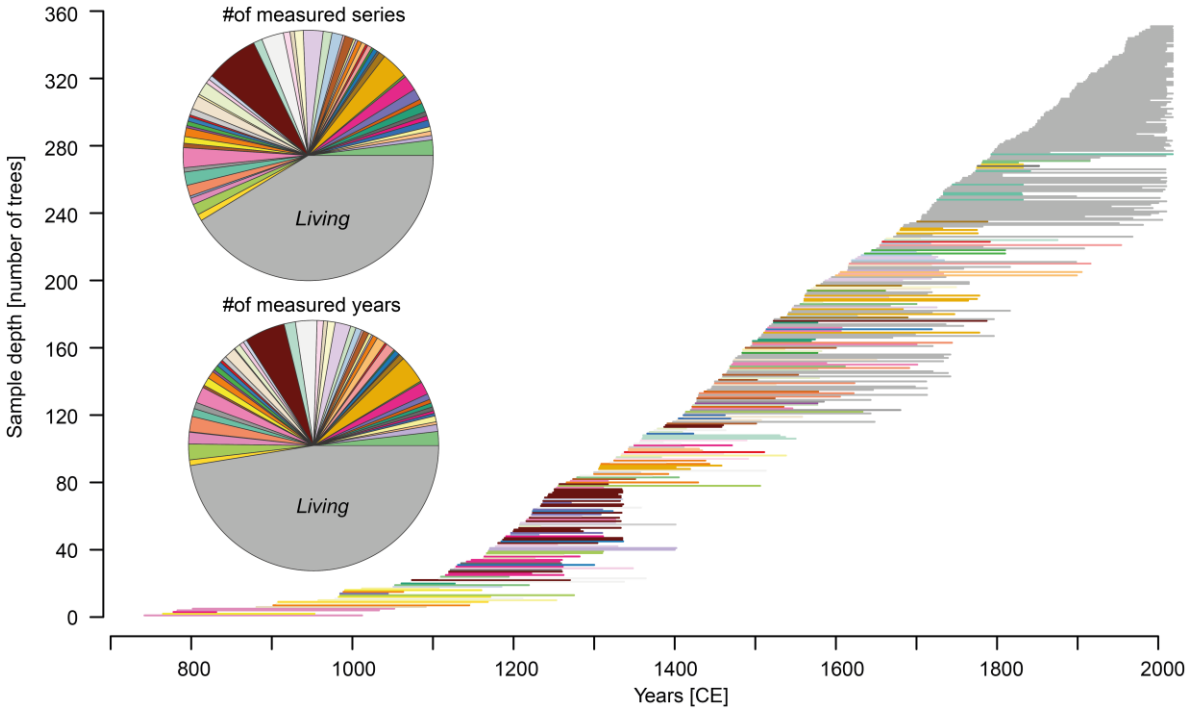


Fig. 3-2 Temporally balanced distribution of the 146 living (grey) and 206 relict (coloured) samples, sorted by their innermost rings in a bar plot. Pie charts show the even distribution of samples per house (top) and the total number of yearly measurement points per building (bottom). Both, pie charts and bar plot support a balanced ratio between the number of historical and living measurements.

3.2.2 Provenancing of the historical material

To determine the provenance of the historical series, two Machine Learning (ML) models were trained based on densitometric measurements and statistical data from transect sites in both valleys. For SV, the procedure is described in Kuhl et al. (2023) and was likewise applied to MV. Compared to SV, MV had only a limited number of available sites from different elevations (MV1 and MV2/MV3). Hence, samples showing a prediction probability lower than 0.8 in the ML model output were excluded from the analysis. In rare cases, when A and B samples were predicted to origin from different elevation classes, these were manually assigned to the most likely elevation. From all measured historical samples, 36 (out of 99) from SV and likewise 170 (out of 206) from MV were used for this study. The final datasets were produced by including the historical series to the corresponding living series collectives (**Fig. 3-1 c)**, **Tab. S 3-1**).

3.2.3 Detrending

After provenancing, the sites were mean-adjusted valley by valley to the MXD level of the highest sites (**Fig. S 3-1**). Differences in mean levels were estimated by calculating the offset between regional curves (RCs) over a period of replication ≥ 15 series (100 cambial years SV, 250 cambial years MV) for each site. The data from the two valleys were then combined and

likewise adjusted in their mean levels over the first 300 years of cambial age meeting > 15 series. Missing rings were in-filled using Arstan (Version 41d, Cook 1985), series were power transformed, pruned (Briffa et al. 2001) beyond 300 years of cambial age (considering pith offsets) and standardized in R 4.2 (R Core Team 2021). To remove age-related density changes (Bräker 1981) the series were detrended as residuals (Cook and Peters 1997) using regional curve standardization (RCS, Esper et al. 2003) with a 67% mean cambial age smoothing filter (frequency response = 0.5). In addition, the signal free (SF) approach (Melvin and Briffa 2008) on RCS and age-dependent spline detrending (Melvin et al. 2007), as well as a 300-year smoothing spline (Cook and Peters 1981) and Hugershoff detrending (Cook et al. 1990) were applied for comparison (see **Fig. S 3-2**). SF detrending was calculated using the program Signal Free (Version 45_v2b, Cook et al. 2017).

3.2.4 Larch budmoth treatment

Larch trees in these elevations and regions of the European Alps are affected by recurring larch budmoth (*Zeiraphera griseana* Hb., LBM) mass outbreaks, leading to defoliation, and consequently to reduced MXD values and radial growth in the event year (Esper et al. 2007a; Baltensweiler et al. 2008; Hartl-Meier et al. 2016; Hartl-Meier et al. 2017b; Kunz et al. 2023). These distinct declines in MXD are independent of weather conditions and need to be removed when reconstructing temperature. Several methods were introduced by Büntgen et al. (2009) or Kunz et al. (2023) for LBM outbreak detection. From these methods, impulse indicator saturation (Pretis et al. 2016, IIS) is the most effective one for our purpose as it not only detects LBM events but also gives a correction factor for these years. The raw MXD series were detrended prior to IIS using a 30-year smoothing spline. This algorithm identifies rapid breaks (**Fig. S 3-3 a)**) as negative outliers and uses these as correction coefficients to be subtracted from the RCS detrended series (**Fig. S 3-3 b)**). The method was initially introduced to accurately detect volcanic events in simulated temperature time series and has proven its potential for LBM detection (Pretis et al. 2017; Schneider et al. 2017; Kunz et al. 2023). When applying IIS, a non-host chronology can be included as a regressor to ensure that the algorithm does not detect outliers that show climate induced declines (e.g. volcanic events) (Kunz et al. 2023). Between 1616 and 2017 CE a non-host Swiss stone pine (*Pinus cembra* L.) site from MV (Kunz et al. 2023) and a non-host larch (*Larix decidua* Mill.) site from the Northern Alps (Hartl-Meier et al. 2014a) were included in the algorithm (**Fig. S 3-3 c)**). As the non-host chronologies only extend back to 1616 CE, prior comparison to the host chronology reaching back to 881 CE was prevented. Consequently, the algorithm detected known volcanic events as LBM events in this period. To avoid the correction of climatic induced growth declines, stratospheric aerosol optical depth (SAOD) years from Sigl et al. (2021) were averaged over

the Westerlies zone (~30-60°N). In years, where SAOD values exceeded 0.03 and matched an IIS detected year, the LBM correction value was not subtracted from the host chronology. The valleys were treated individually to address the lag of outbreak years between the valleys (Johnson et al. 2004; Saulnier et al. 2017; Kunz et al. 2023). Comparison of LBM detections of the IIS algorithm corrected chronology and the non-host chronology (**Fig. S 3-3 c)**) confirms that the algorithm can detect non-climate-related growth declines. The prevention of false detection of volcanic cooling events as LBM outliers by using the SAOD values is necessary to preserve climatic signals in the chronology, but it can lead to overestimated cooling when LBM events and volcanic induced cooling years align.

3.2.5 Variance stabilization

The final chronology was constructed using Tukey's bi-weight robust mean. To account for changes in variance, chronology variance was stabilized using the method detailed in Osborn et al. (1997). On trial, different window lengths were tested during stabilization. Results showed shorter windows (e.g. 31 years) stabilize variance better than larger ones. However, applying a narrow window for variance stabilization increases the risk to eradicate low-frequency trends (Frank et al. 2007). We approached this risk by calculating residuals between the chronology and a 100-year low-pass Butterworth filter. The variance of the high-frequency residuals was then stabilized in a 31-year window, whereas the low-pass chronology was stabilized in a 201-year window (Osborn et al. 1997)). Afterwards, the stabilized sub-chronologies were added back together to the final chronology from 881 to 2017 CE. Calculations were performed in R 4.2 (R Core Team 2021) using the package "dplR" (Bunn 2010).

3.2.6 Climate signals

Pearson's correlations between instrumental temperatures (CRU TS 4.06; Harris et al. 2020) and the MXD chronology were calculated from 1901-2017 CE in a classical bootstrap approach ($boot_n$ = 1000, α = 0.95) using the R package "treeclim" (Zang and Biondi 2015). Correlations were computed for monthly January-September (J-S) and JJA and May-September (MJJAS) seasonal means. For every grid from -10°E to 20°W and 35-65°N, we likewise calculated correlations between MJJAS mean temperature anomalies and the tree-ring chronology. We averaged the temperature data of all grids exceeding r = 0.7. Static and running correlations (31-year window) were performed using this averaged temperature record.

3.2.7 Calibration, verification, and reconstruction

Three different reconstruction approaches (linear regression, polynomial regression, and scaling) were tested using a two-fold cross-validation from 1901-1959 and 1960-2017 CE against MJJAS grid-averaged temperature anomalies. The correlation between predicted and observed data (r), explained variance (R^2) and root mean squared error (RMSE) were used to assess model accuracies. For the final reconstruction, the best performing model was chosen and then calibrated over the full period from 1901-2017 CE.

Uncertainty bands for the reconstruction were estimated by averaging RMSEs of sub-chronology reconstructions for different, stepwise increasing sample depths from n = 5-50 series in five steps. For each of these replication steps, bootstrapped sub-chronologies of n randomly sampled series were built ($boot_n$ = 1000). These individual sub-chronologies were then used to reconstruct MJJAS temperature anomalies on the calibration period 1914-2002 CE. For each reconstruction r, R^2 and RMSE were calculated. The mean RMSE for each replication step ($RMSE_n$) was then used to estimate the error of the reconstruction depending on its sample depth n (error = 2*$RMSE_n$).

3.2.8 Standard reconstruction approach

A reconstruction including all existing historical series was additionally modelled neglecting provenance information. The data were gap-filled, power transformed and pruned (300-years). In contrast to the new approach, no mean adjustment was performed. All series were detrended, LBM corrected, and variance stabilized, just as the main chronology of this paper. A linear model with the same temperature data was used to reconstruct the MJJAS mean temperature variability (averaged grids), and the results compared with the novel provenance considering reconstruction.

3.2.9 Superposed epoch analysis and extreme year analysis

Superposed epoch analysis (Lough and Fritts 1987, SEA) was applied to evaluate the significance of major volcanic eruptions (**Tab. S 3-2**) considering a period of five years prior (reference period) and five years after the event using the package “dplR” (Bunn 2010). For this, a 30-year high-pass filter was run on the new reconstruction and the Lötschental record (Büntgen et al. 2006). From bootstrap resampling ($boot_n$ = 10,000), 99% confidence intervals were calculated to estimate significant deviations after volcanic event year period. In addition, the coldest 20 years of the 50-year high-pass filtered records were compared against volcanic event years.

3.3 Results

3.3.1 Chronology development

The developed chronology covers the period from 881-2017 CE including 1137 years and thus has the potential to revise climate variability back to the early medieval period. The minimum sample depth n = 5 spans the period 881 to 901 CE and the maximum is reached in 1963 CE with n = 93 series. The average correlation between the series is 0.54, with a mean segment length of 146 years due to pruning of the data and a lag-1 autocorrelation of 0.3.

The removal of non-climate related biases in tree growth can help to increase the robustness of a chronology. This improvement can be accomplished by selecting historical series, which have been sorted to high-elevation microclimate growth characteristics by the provenance model and are therefore more likely to have a stronger temperature signal. Considering this information of the series allows for exploration of differences in MXD mean level offsets between the sites (**Fig. S 3-1**). If not addressed, these differences can lead to deflated or inflated values when RCS detrending is applied (Zhang et al. 2015; Römer et al. 2021; Hartl et al. 2022), especially if a period is mainly covered with series from one elevational growth characteristics. RCS mainly requires a great sample depth (here: 352 series), a homogeneous distribution of tree-ring series in time (**Fig. 3-2**), heterogenous tree age and homogeneity in the provenance of historical and living series (Briffa et al. 1992; Düthorn et al. 2015; Esper et al. 2003; Melvin et al. 2013), which is improved by the provenance approach of Kuhl et al. (2023). The RCS detrended chronology shows more variability and a stronger negative trend towards the early 19th century compared to 300-year spline and Hugershoff detrending (**Fig. S 3-2 a)**), indicating a stronger preservation of the low-frequency (Briffa et al. 1992). The standard deviation of the 100-year smoothed chronology is almost twice the size compared to other detrending methods when we apply SF age-dependent spline detrending (**Fig. S 3-2 a)**). The SF RCS chronology, however, does not show significant differences to the classical RCS approach (**Fig. S 3-2 b)**).

Spectrum analysis emphasizes that a smaller window of 31 years can stabilize the variance better than a wider window of 201 years. However, stabilization on narrow window can remove low-frequency signals which are related to temperature (**Fig. S 3-4 a)**). While the 31-year window stabilization has a reduced power spectrum moving towards lower frequencies, greater windows (e.g. 201 years) preserve more low-frequency information but in turn do not stabilize the variance as well (**Fig. S 3-4 b)**). Our new approach detangles low- from high-frequency bands by stabilizing the variance individually on both, the high- and the low-frequency sub-

chronologies. Treating low- and high-frequency on different window lengths reduces the loss of low-frequency signals when applying a short window but at the same time improves stabilization compared to using a wider window for the full chronology.

3.3.2 Climate-growth relationship

Correlations between temperature anomalies and the MXD chronology reveal a positive relationship between tree growth and temperature. Spatial correlations for MJJAS and the chronology between 1901-2017 CE show the significantly high ($p<0.05$) response between Western European temperature variability and MXD indices (**Fig. 3-3 a)**). Grid points over France, Switzerland, Southern Germany, Northern Italy, and Western Austria reached highest correlations of $r \geq 0.7$. Altogether, correlations do not drop below 0.5 in Western Europe. The ability of the Alpine chronology to represent past temperature variability in Western Europe is strengthened by calculating static correlations from 1901-2017 CE between temperatures of the grids exceeding 0.7 for MJJAS in **Fig. 3-3 a)** and the MXD indices (**Fig. 3-3 b)**). From March onwards, significant correlations ($p < 0.05$) with temperature are observed with highest monthly correlations for August ($r = 0.71$). The coherence between the proxy and instrumental temperatures increases when considering seasonal means up to $r = 0.8$ (MJJAS). Moving window correlations indicate a temporally stable relationship over the full instrumental period from 1901-2017 CE (**Fig. 3-3 c)**). Although correlations slightly decline in recent years, r values do not fall below 0.6.

3.3.3 Reconstruction of Alpine summer temperature anomalies

For reconstruction, a linear regression model was chosen as these models perform best on both cross-validation periods (1901-1959 CE, 1960-2017 CE) with correlations of 0.8 on the entire period 1901-2017 CE between observed and predicted time series (**Fig. 3-4 a)**). These cross-validation models explain up to 70% of the variance and are stable in time with the validation periods reaching $r > 0.82$ (**Tab. S 3-2**). Although both models exhibit a high level of performance, they predict underexaggerated values for the latest decades. While the model trained on the early period fails to capture the recent warming from 1995 CE onwards, the model for the later period cannot accurately predict the increase in temperatures since 2010 CE.

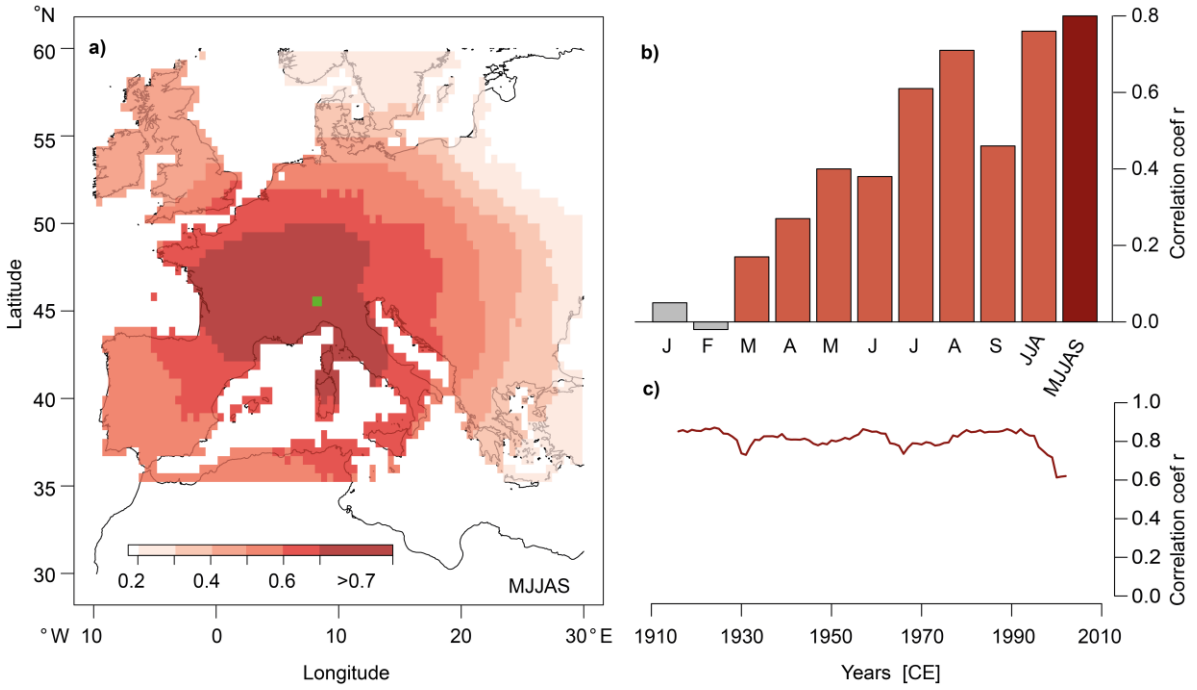


Fig. 3-3 a) Correlations between the final (combined and variance stabilized) MXD chronology and gridded mean May-September (MJJAS) temperature data (CRU TS 4.06 0.5°) from 1901-2017 CE **b)** Temperature data of all grids with spatial MJJAS correlations ≥ 0.7 (a) are averaged to calculate monthly correlations between January and September, as well as using two different summer means (June-August JJA and May-September MJJAS). Grey colouring denotes non-significant correlations ($p<0.05$), orange shows significant correlations and red illustrates the MJJAS window that is chosen for further analysis **c)** Moving 31-year correlations between MJJAS mean temperature and the MXD chronology.

Uncertainties of the reconstruction are highly related to changes in replication. With decreased replication further back in time, model uncertainties are likely increasing. To test the reliability of the linear regression model depending on replication, randomly built sub-chronologies of different replications were fitted in linear regression models over a period from 1914-2002 CE, where the replication is constantly over 50 series. This smaller window was chosen to test the replication effect from 5 up to 50 series per sub-chronology. Results show a high correlation of the sub-chronologies to the instrumental data until the replication drops below 15 series ($r_{median} \geq 0.8$) (**Fig. 3-4 b)**). When the sample replication equals ten series or lower, the linear models lose accuracy as the confidence intervals increase towards lower correlation values. Still, r_{median} does not fall below 0.7.

The robust temperature signals and the stability of the linear model with respect to replication fluctuations enable the construction of a millennium-length MJJAS temperature reconstruction. The reconstruction model fitted to 1901-2017 CE explains 64% of the variance with r = 0.8 and is used to predict and revise Alpine temperature anomalies until 881 CE (**Fig. 3-4 c, d)**).

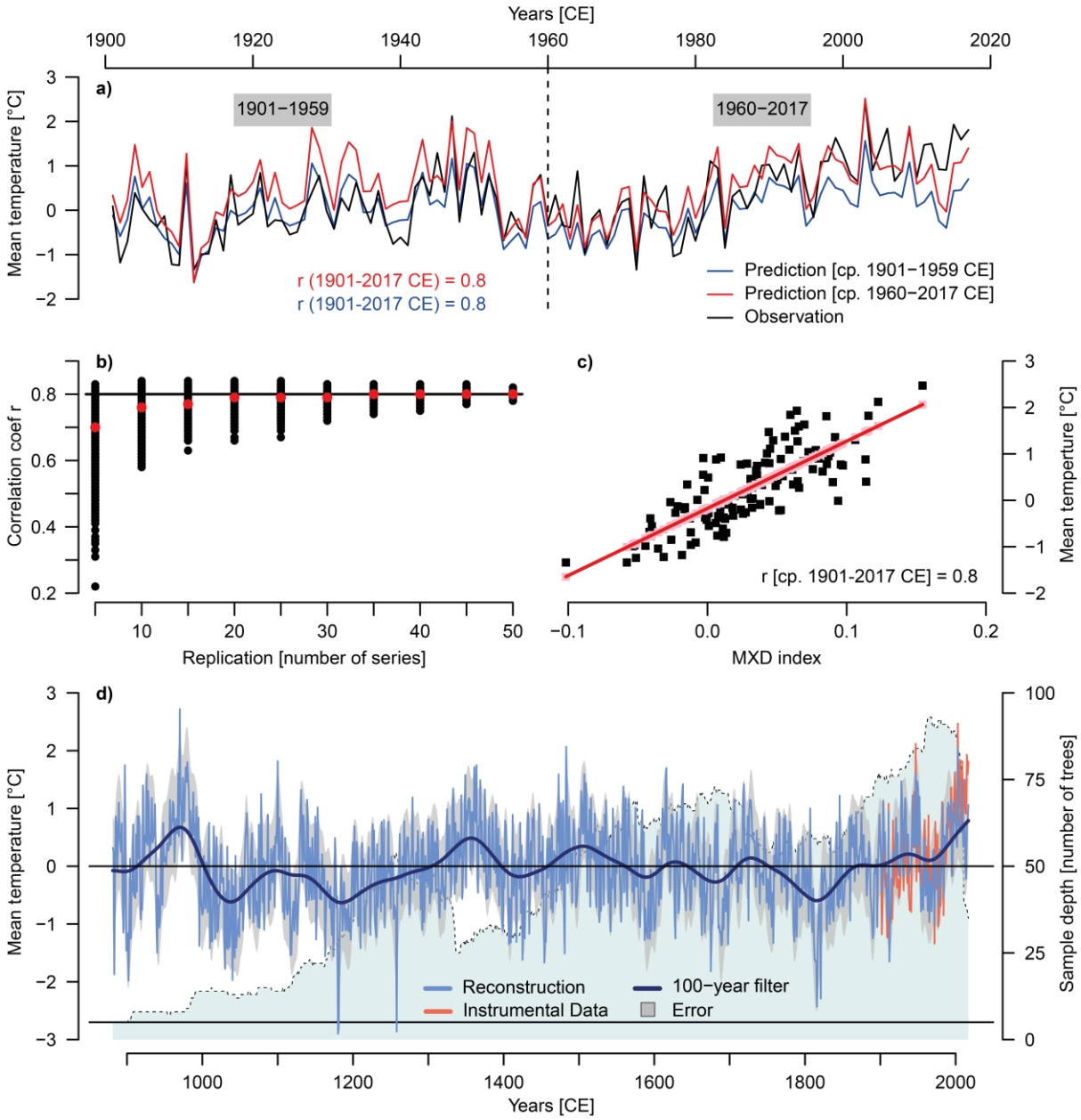


Fig. 3-4 a) Linear Regression Models between the MXD chronology and May-September mean temperature anomalies [w.r.t. 1961-1990 CE] (MJJAS Temp) of calibration periods (cp.) 1901-1959 CE (blue) and 1960-2017 CE (red). Detailed performance measures of the cross-valid cross-validation are listed in Tab. S 3-2 **b)** The effect of replication on regression model accuracy: Correlations r between MJJAS Temp and instrumental data between 1914-2002 CE depending on different sample replications (5-50). Black dots represent correlation coefficient between the individual model runs (1000 per replication step) and instrumental data. Red dots represent the medians of the 1000 runs (rmedian). The black line shows r for the regression performed on the full period 1901-2017 (min. replication 71 samples) **c)** Model fit of the linear regression between the MXD indexes and instrumental data on the full period 1901-2017 CE, which is used for the reconstruction **d)** Final MJJAS Temp reconstruction: the blue line shows the reconstruction with a 100-yr smoothing filter (dark blue). Grey shading presents the error range of the reconstruction and orange the instrumental data. Light blue denotes to the sample depth n.

The temperature reconstruction (**Fig. 3-4 d)**) from the European Alps displays that long-term temperature trends decrease during the early and high medieval times from 881-1200 CE. Temperatures rise again until late medieval ages in the so-called medieval climate anomaly

(MCA) and peak in the 1360s, which are on average 0.76 °C (uncertainty range: -0.08 to 1.6 °C) warmer than the reference period 1961-1990 CE. This maximum is followed by a slow long-term decrease in temperatures during the Little Ice Age (LIA) until the early 19th century, when the Tambora eruption caused the "year without summer" of 1816 CE (Stothers 1984; Oppenheimer 2003). This decade is the coldest decade of the record and marks the peak of the LIA. Here, temperature anomalies drop to -1.28°C (uncertainty range -2.12 to -0.44°C). Between the end of the MCA and the peak of the LIA, temperatures decrease in total by 2.04°C. Afterwards temperatures rise again from 1900 until 2017 CE. This increase is shortly disrupted between 1950 and 1980 CE.

The most prominent long-term periods of cooling are between 1000-1300 and 1750-1870 CE, while warmer phases are found between 910-1000, 1300-1400, 1475-1575 and 1900-2017 CE. The 971-980 CE period represents the warmest decade of the record with a decadal average of +1.17°C (uncertainty range: 0.25 to 2.09 °C), followed by the 960s, 1500s, 980s and 2000s. Besides the 1810s, coldest decades of the record are the 1180s, 1040s, 1030s and 1170s. The absolute warmest and coldest years of the record are not necessarily located within these decades. Warmest years are recorded in 970 (2.72°C), 1483 (2.07 °C), 2003 (2.06°C), 1100 (1.82°C) and 980 CE (1.80°C), while the coldest years are 1180 CE (-2.9°C), 1258 (-2.86°C), 1181 (-2.8°C), 1816 (-2.43°C) and 1821 CE (-2.29°C).

3.4 Discussion

3.4.1 The new approach: Eliminating interference signals from elevation and insects

The here presented new approaches aim to make reconstructions, which depend on historical series, more robust. Elevation and the potential weak temperature signals in historical series have been neglected in the past when building chronologies. Using a ML based provenance model helps to overcome this problem (Kuhl et al. 2023). The provenance of the historical series allows the selection of series, that are likely to be most sensitive to temperature, and therefore improves the temporal stability of the climate sensitivity in periods beyond the lifespans of the living trees.

As MXD increases with decreasing elevations (Zhang et al. 2015; Hartl et al. 2022), mean adjusting the series can reduce elevational biases while detrending with RCS. Lower elevation series have higher MXD levels on cambial age scale compared to higher elevations and can therefore potentially alter the regional curve towards higher mean values. This can be true for elevational differences of 100 meters (**Fig. S 3-1**). Microsite differences like this have already been observed and critically viewed in multiple studies in terms of regional curves biases in RCS detrending (Gunnarson et al. 2011; Düthorn et al. 2013; Hartl et al. 2021; Hartl et al.

2022). In previous reconstructions, these potential offsets between historical series have been neglected, which is mainly due to the lack of information on the elevational origin of historical tree-ring series. Simultaneously, differing MXD values of the series can also result from temperature variability and changes in treeline (Büntgen et al. 2022a) and there is a likelihood that mean adjustment causes a loss of temperature information. However, site control of historical series offers the possibility to improve RCS detrending of living and historical series to reduce visible biases in the data due to elevational non-temperature related MXD offsets as presented in **figure S 3-1**.

The correction of negative outliers is indispensable when chronologies are based on tree species with recurring insect outbreaks like larch (Büntgen et al. 2006; Aryal et al. 2023). As the outliers in a chronology can potentially result from volcanic cooling, the inclusion of a non-host chronology in combination with SAOD-analysis present necessary pre-steps before applying correction coefficients. This can ensure that high-frequency temperature signals are not removed from the data during this process.

The chronology depicts a different course of the temperature history when provenance information of historical series is neglected in the reconstruction process (**Fig. S 3-5**). This reconstruction profits from more available samples to extend the record further until 729 (n=5). Compared to the main reconstruction, this record shows higher temperature estimates in the medieval period until 1000 CE and between the 12th-15th century. A visible decrease in temperatures during the LIA is almost 150 years later than in the main reconstruction but shows a stronger decline due to the increased values from the 12th-15th century. Higher temperature estimates in this period most likely result from the inclusion of historical series from lower elevation growth characteristics with higher MXD levels and from the neglection of potential level offsets (e.g., see Hartl et al. 2022 for SV).

3.4.2 Mitigated low frequency signals

For almost two decades, the Lötschental reconstruction has been the state of the art record for European summer temperatures and was used in several larger network studies to present Western European temperature variability (Esper et al. 2007c; Brázdil et al. 2010; Luterbacher et al. 2016; Wilson et al. 2016; Anchukaitis et al. 2017; Torbenson et al. 2023; Wang et al. 2023). The comparison between the new reconstruction and the Lötschental record highlights some major differences between 1000-1900 CE (**Fig. 3-5 a)**). In contrast to the Lötschental record, the new reconstruction shows less long-term variability. Although a long-term temperature decrease during the LIA is observed, it is not as pronounced as in the Lötschental.

There, the onset of the cold period begins after a peak at 1200 CE whereas the cooling in the new reconstruction starts approximately 160 years later in the mid-14th century. In the Lötschental record, the drop in temperatures (w.r.t. 1901-2000 CE) between the warmest pre-LIA decade in the 1200s (0.66°C) and the 1810s (-2.31°C) is 2.97°C and the increase in temperatures between the coldest decade (1810s) and the warmest following decade of the 1940s is 3.08°C. The reconstructed temperature decrease since the end of the MCA is almost a degree less in the new record and could hint that the European Little Ice Age temperature decrease was not as strong as the Lötschental record estimated and that the reconstruction implies that summer temperatures are not as cool as in the period between 1000 and 1250 CE. However, there is evidence that the LIA must have existed in Europe, for example by historical records or glacier advances in the Swiss Alps (Grove 2001), which challenges the reconstruction's skill in low-frequency variation.

We tested and eliminated various potential sources of error during the development of the reconstruction (see **Tab. S 3-3** for more detailed information). Sources, which can alter the low-frequency signals, include the distribution of tree age in time, the detrending method, the selection of series or an inadequate classification by the ML models (Kuhl et al. 2023). For example, the model for MV uses data of only two different elevations as reference and despite the model performance score (f1-score) of 0.99, series might align more with patterns from lower elevations but will be sorted to either of the two elevations. We tackle this problem by selecting only those series that are sorted to an elevational class with a prediction probability greater than 0.8. As mentioned before, the mean adjustment might alter climate signals, when tree-ring series are adjusted to higher elevation MXD levels. However, excluding these steps (provenancing and mean adjustment) does not increase the low-frequency signal, especially not during the LIA (**Fig. S 3-5**).

Besides the difference in the low-frequency signal in both records, the variance of the new record is more stable in time (**Fig. 3-5 c)**). In the Lötschental record, however, the variance is not as stable and shows stronger temporal fluctuations, especially before the 13th century CE. As variance changes are likely artificial and an artefact of alternating replication or varying ecological habitats, the increase in variance pre 1200 CE might result from a change in replication and the different origin of this historical tail being exclusively composed of the Simplon valley (Büntgen et al. 2006; Esper et al. 2016). Despite these differences, the correlation in the common period (881–2004 CE) between both records is 0.54 and 0.62, when a 500-year high-pass filter is applied to the Lötschental record. It indicates that albeit low-

frequency signals are mitigated in the new reconstruction higher frequency signals are more coherent.

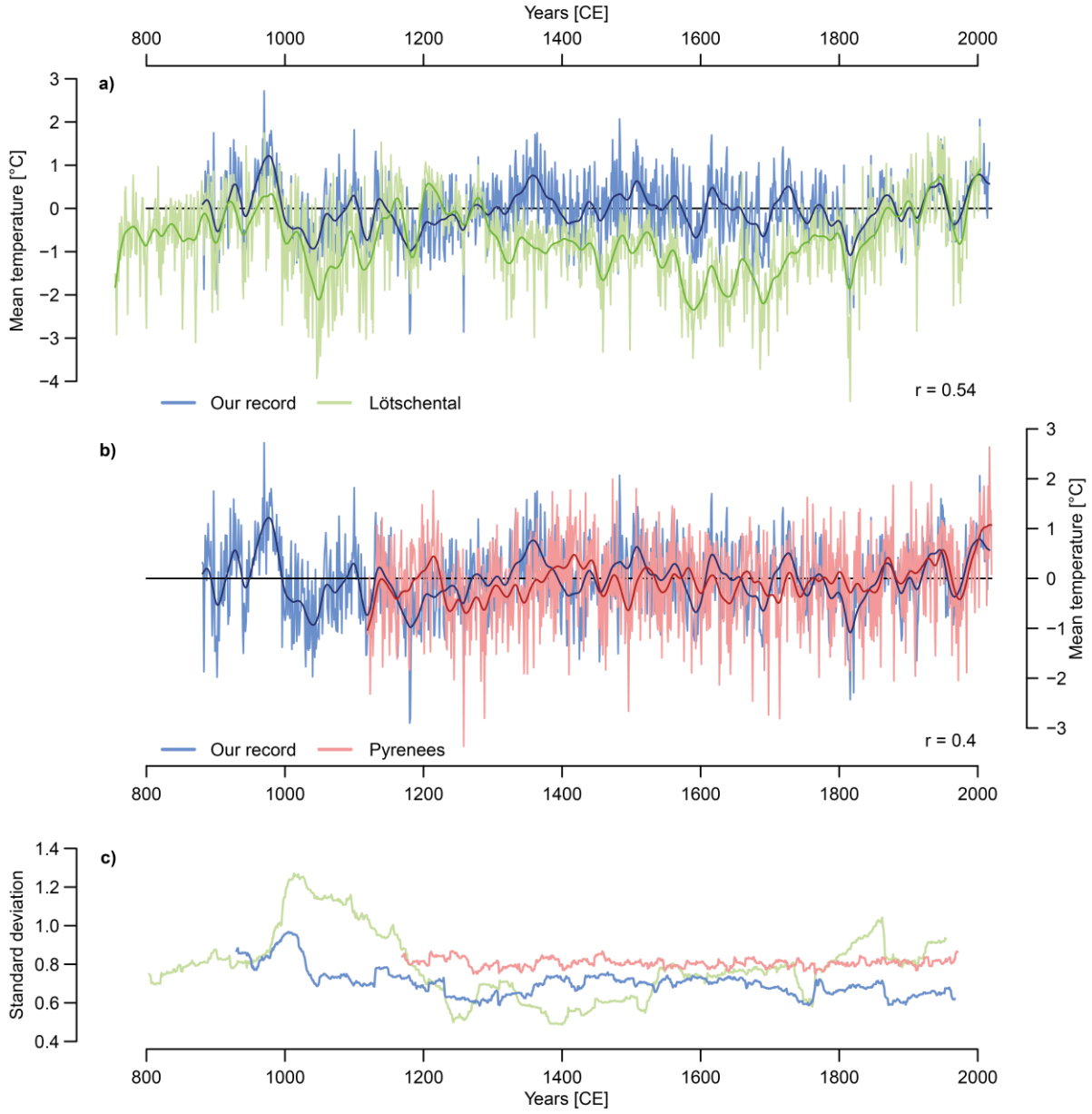


Fig. 3-5 a) Comparison between the Lötschental (green) and our record (blue) and **b)** comparison between the Pyrenees (coral) and our record (blue). Dark lines in a) and b) present the 30-year smoothed data. r equals the correlation between both records in their common period, respectively **c)** Differences in standard deviation of the Pyrenees (coral), the Lötschental (green) and the new Alpine (blue) records (100-year running window).

A recently developed reconstruction from the Pyrenees (Büntgen et al. 2024), however, coincides with the new Alpine record (**Fig. 3-5 b)**). Both reconstructions show a similar low-frequency variability. They correlate with 0.4 but correlations reach 0.74 when both chronologies are 300-year low-pass filtered. Although the LIA period is weakly pronounced in both reconstructions, the Pyrenees record does not record a strong cooling in the 1810s, which

is prominent in most Alpine records (**Tab. 3-1**) (Schweingruber et al. 1987; 1988; Büntgen et al. 2005; Frank and Esper 2005b; Büntgen et al. 2006; Corona et al. 2010; Corona et al. 2011; Trachsel et al. 2012; Coppola et al. 2013; Leonelli et al. 2016). The standard deviation of the Pyrenees record in **figure 3-5 c)** is stable with lower variation than in the Lötschental record or the record of this study.

A focus on the warm and cold decades of the Lötschental and the new record reveals that while both reconstructions agree on the 1810s as the coldest decade, the warmest decade in the new record is not in the last decades but earlier in the 970s. The most recent decades, which are the absolute warmest decades according to the Lötschental record, are not the warmest in the new record anymore. Here, the warmest decades are mainly between 960 – 990 CE. Nonetheless, both records show the 2000s as one of the five warmest decades. Surprisingly, the new record does not estimate the 2010s to be among the warmest decades, which is not in line with the instrumental data.

3.4.3 Revising cold and warm extremes in Alpine temperature history

A stable and unbiased variance is crucial for robust extreme year analysis (Frank et al. 2007). Artificial fluctuations of variance can lead to falsely detected extreme events and consequently alter interpretations of extreme year distributions. This influence becomes visible when 30-year high-pass filter reconstructions and their top 20 extreme years are compared (**Fig. 3-6 a**, **c), Tab. 3-2**). The asterisks in the plots denote the 20 warmest and coldest years of the records, respectively. If variance is not stable in time (**Fig. 3-5 b)**), extreme years like volcanic cooling events are accumulated in periods with increased variance, for example the 11-12th century in the Lötschental record. As these volcanic induced cooling years are recorded by tree-ring proxies like MXD, they are frequently used to detect volcanic events (Esper et al. 2017) and reveal information on their impact on temperatures and on their link to societal developments like mass migrations, famines or cultural heydays (Büntgen et al. 2020). A bias in variance can thus alter volcanic event detection and interpretation. Although the Lötschental record detects some of the strongest 20 volcanic SAOD events during the last millennium, a SEA of these events reveals that a highly significant ($p < 0.01$) cooling in the year of the event and year after the event is not detectable in this record (**Fig. 3-6 d)**). In comparison, the new reconstruction has a temporally more stable variance (**Fig. 3-5 b)**), and the event year as well as following years are found to be significantly cooler than the five years prior to the event (**Fig. 3-6 b)**). As most volcanic events of the SEA fall within periods of lower variance in the Lötschental record, the averaged cooling of these events does not exceed the non-volcanic induced declines in

periods when the variance is higher. Further, the growth decline relative to the years prior to the event is less pronounced than in the new record. With fluctuating variance, the volcanic induced cooling years in the Lötschental record are either no outliers in periods of low variance or surrounded by outliers in periods of higher variance. The residuals between the event year temperatures and the reference period are consequently lower.

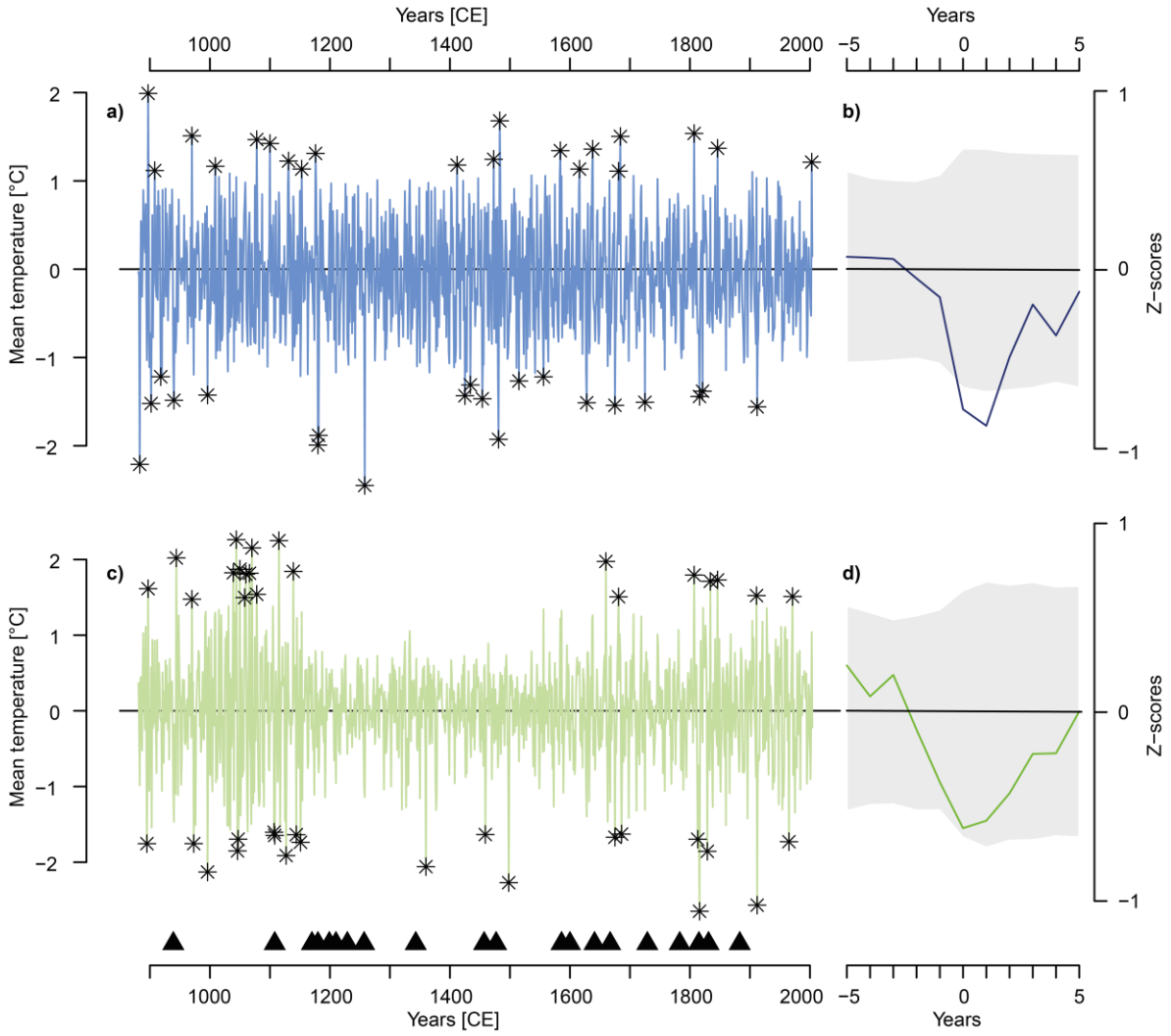


Fig. 3-6 Extreme Year Analysis: 30-year high-pass filtered reconstructions from this study in **a)** and Lötschental in **c)**. Asterisks denote to the 20 warmest and the 20 coldest years, respectively. Triangles show 20 strongest volcanic events between 880-2000 AD (see **Tab. S 3-4**). For these 20 events, panels **b)** and **d)** show the Superposed Epoch Analysis (lag = 5, residuals from the 5 years prior to event) with mean (black line) and 99% confidence intervals (grey) after bootstrap resampling (n = 10,000).

An example of the influence of variance on the interpretation of extreme events is given by focusing on the Samalas eruption in 1257 CE, which is the largest volcanic eruption of the last millennium (Lavigne et al. 2013). The following volcanic cooling is recorded in both reconstructions with -2.32°C in the new record and -1.18°C in the Lötschental record (w.r.t. 1252-1256 CE), respectively. While the Samalas induced cooling in 1258 CE presents the

strongest negative extreme event in the new record, it does not fall under the extreme events detected in the Lötschental record. This might be explained by variance fluctuations in the high-frequency domain. The resulting loss of climate information might explain why this eruption is not imprinted in the top 20 cold events of the record. Guillet et al. (2017) present historical evidence for bad (grape) harvest, increased prices, heavy precipitation and floodings in the regions of Western European region, where temperatures correlate high with the new record (**Fig. 3-3 a)**).

Tab. 3-2 Top 20 minimum and maximum extreme years (descending order) of the new Alpine record between 881-2004 CE. Grey shaded boxed align with an eruption year or the year after using the dates of Tab. S 3-4.

No.	Maximal Extreme Years		Minimal Extreme Years	
	Year [CE]	Temperature Anomaly [°C]	Year [CE]	Temperature Anomaly [°C]
1	897	2.03	1258	-2.42
2	1483	1.71	883	-2.18
3	1807	1.57	1180	-1.96
4	970	1.54	1481	-1.89
5	1684	1.54	1181	-1.85
6	1078	1.50	1912	-1.53
7	1100	1.46	1675	-1.51
8	1846	1.40	902	-1.49
9	1638	1.39	1628	-1.48
10	1584	1.38	1725	-1.47
11	1176	1.35	940	-1.45
12	1473	1.28	1454	-1.43
13	1131	1.26	1816	-1.41
14	2003	1.25	1425	-1.40
15	1412	1.21	996	-1.39
16	1009	1.20	1821	-1.35
17	1153	1.17	1434	-1.28
18	1616	1.17	1515	-1.23
19	908	1.15	1556	-1.19
20	1681	1.14	918	-1.18

A strong cooling in 1258 CE can be observed in the Pyrenees record being -3.26°C colder than the average of the previous five years (see **Fig. 3-5b**, Büntgen et al. 2024). Other NH and Western Europe reconstructions decrease between -0.55 to -1.8°C compared to the ten year mean prior the events (here: -2.52°C) (Esper et al. 2013; Schneider et al. 2015; Stoffel et al. 2015; Wilson et al. 2016; Hartl-Meier et al. 2017a). Less pronounced cooling in these records compared to the Pyrenees or the Alps is likely due to the observed spatially diverse intensity of the post volcanic cooling (Guillet et al. 2017; Büntgen et al. 2022b). For the new Pyrenees and Alpine reconstructions, the 1258 CE cooling is unrivalled in the last millennium (**Figs. 3-5 b), S 3-7**). In both, the cooling following the Samalas eruption is stronger than the cooling observed for 1816 CE after Tambora, which is less prominent in the Pyrenees record compared to the Alpine region. The new record agrees with previous studies that showed that summer temperatures in the 1800/1810s have been relatively cold already before the event compared to the situation in the 1250s (e.g., Frank and Esper 2005b; Corona et al. 2010; Trachsel et al. 2012; Schneider et al. 2015). Although the volcanic induced decrease of 1.01 °C compared to the five years prior to the event appears less severe than the Samalas induced temperature decline, the impact of the 1810s CE cooling on humanity was strong which is proven by historical evidence (e.g., Brázdil et al. 2017; Büntgen et al. 2020). We are aware that the LBM correction can bias the strength of volcanic cooling which is mediated by the reconstruction. In years where LBM events and volcanic events align, LBM induced density decline can amplify the reconstructed compared to the true volcanic cooling. However, as the extreme year 1258 CE is also detected in the new Pyrenees record by Büntgen et al. (2024), which is based on LBM-unaffected pine trees (*Pinus uncinata* Ramond s. str.), it supports our findings that people in Western Europe experienced a year of unusual cold summer temperatures in 1258 CE.

The top 20 summer heat extremes in the new record are relatively balanced throughout the covered period with clusters in the late medieval times and during LIA. Although long-term temperatures during the LIA are decreasing, exceptionally warm summers relative to surrounding decades are observed in a recent study (Wanner et al. 2022). For example, the summer in 1473 CE was extremely hot and dry in Europe causing early (grape) harvests or harvest failures, loss of transportation routes (rivers), water shortages and wildfires (Camenisch et al. 2020). Likely due to a decrease in variance in this period, this year is not observed as an extreme year in the Lötschental record. Similarly, the 2003 CE heatwave appears as an extreme year when long-term trends are included (Büntgen et al. 2006), but not when the record is high-pass filtered. From 2003-2017 CE no cool extremes are observed in the new Alpine record but no heat extremes either which does not agree with instrumental

data. Resuming the fit of the reconstruction model (**Fig. 3-4 a)**) this might result from the models' limitation to fully capture the instrumental warming after 2010 CE.

3.4.4 Is there a divergence problem in the Alps?

Although the fit between instrumental and proxy data (r = 0.8) is very high, the dataset shows signs of a beginning divergence from 2010 CE onwards (**Fig. S 3-7**). The divergence phenomenon describes the weakening of the linear relationship between temperature records and tree-ring proxies (TRW and MXD) during the 20th century (D'Arrigo et al. 2008). This phenomenon has been observed in multiple sites all over the Norther Hemisphere (Wilmking et al. 2005; Wilson et al. 2007; D'Arrigo et al. 2008; Andreu-Hayles et al. 2011; Shah and Shah 2015; St. George and Esper 2019) and in low elevation sites in the European Alps (Corona et al. 2010). However, divergence has not been detected in high elevation sites in this region during the 20th century for European larch (*Larix decidua* Mill.) (Büntgen et al. 2008). Here, we find first indications of a developing divergence phenomenon in the Swiss Alps in larch. From 2010 CE onwards MXD does not follow the increased warming observed in instrumental data. This is underlined by a decline in the 30-year running correlations between temperature and MXD indices from 1990 CE until present in **figure 3-3 c)**. The linear relationship is weakened and consequently, the prediction performance of linear regression likely decreases with increasing divergence. Correlations with other climate parameters, for example drought or precipitation, do not show a shift in signal in the Alps yet, but might in the future as it has been observed in other regions, for example Corsica (Römer et al. 2021). There, drought, and consequently summer precipitation become more important for limiting tree growth. Precipitation signals are, however, difficult to assess in network datasets like the new Alpine chronology since precipitation patterns in mountainous regions can vary strongly between valleys and elevations (Sevruk 1997; Auer et al. 2001; Schmidli et al. 2002).

3.5 Conclusion

In this study, we present a new temperature reconstruction based on tree-ring data from the Swiss Alps. We introduce a new approach to improve temperature reconstructions and their temporal reliability by considering elevational differences among data from historical buildings, removing LBM signals with improved methods, and applying novel approaches to stabilize variance. Albeit low-frequency trends are muted in our new reconstruction compared to previous records, high-frequency temperature variability shows strong agreements with large volcanic eruptions which is in line with historical evidence. We show how artificial variance increases may alter extreme year analysis including the 1257 CE Samalas eruption, which led

to stronger regional cooling than the 1815 CE Tambora eruption. Our study thereby emphasizes how decision making, data selection and treatment in dendroclimatology influences the outcome of a reconstruction approach (Büntgen et al. 2021). The different applied methods, in particular the (1) data selection (of proxy and target), (2) mean adjustment, (3) LBM correction, (4) variance stabilization, and (5) the chosen method of reconstruction, alter the resulting record and interpretation of temperature variability towards its extreme events and long-term temperature signals. This study highlights the urgency to continue the development and testing of new methods to improve the temporal stability and reliability of climate reconstructions. It stresses the importance that with new methods revising existing records to improve our knowledge about past climate variability is indispensable.

3.6 Acknowledgements

We thank Markus Kochbeck, Frederick Reinig, Philipp Römer, Christian Gnaneswaran, Benedikt Lang, Philipp Schulz, Lara Meurer, Sophie Spelsberg and Jannes Fischer for help in the laboratory and field. We thank Martin Schmidhalter from Dendrolabor Valais for the provided data.

3.7 References

Affolter P, Büntgen U, Esper J, Rigling A, Weber P, Luterbacher J, Frank D (2010) Inner Alpine conifer response to 20th century drought swings. Eur J Forest Res 129:2889–298. https://doi.org/DOI 10.1007/s10342-009-0327-x

Allgaier Leuch B, Streit K, Brang P (2017) Der Schweizer Wald im Klimawandel: Welche Entwicklungen kommen auf uns zu? Merkblatt für die Praxis (59):1–12

Anchukaitis KJ, Wilson R, Briffa KR, Büntgen U, Cook ER, D'Arrigo R, Davi N, Esper J, Frank D, Gunnarson BE, Hegerl G, Helama S, Klesse S, Krusic PJ, Linderholm HW, Myglan V, Osborn TJ, Zhang P, Rydval M, Schneider L, Schurer A, Wiles G, Zorita E (2017) Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. Quat Sci Rev 163:1–22. https://doi.org/10.1016/j.quascirev.2017.02.020

Andreu-Hayles L, D'Arrigo R, Anchukaitis K, Beck P, Frank D, Goetz S (2011) Varying boreal forest response to Arctic environmental change at the Firth River, Alaska. Environ Res Lett. https://doi.org/10.1088/1748-9326/6/4/045503

Aryal S, Grießinger J, Arsalani M, Meier WJ-H, Fu P-L, Fan Z-X, Bräuning A (2023) Insect infestations have an impact on the quality of climate reconstructions using Larix ring-

width chronologies from the Tibetan plateau. Ecol Indic 148:110124. https://doi.org/10.1016/j.ecolind.2023.110124

Auer I, Böhm R, Maugeri M (2001) A new long-term gridded precipitation data-set for the alps and its application for Map and Alpclim. Phys Chem Earth (B) 26(5–6):421–424. https://doi.org/10.1016/S1464-1909(01)00029-6

Baltensweiler W, Weber UM, Cherubini P (2008) Tracing the influence of larch-bud-moth insect outbreaks and weather conditions on larch tree-ring growth in Engadine (Switzerland). Oikos 117(2):161–172. https://doi.org/10.1111/j.2007.0030-1299.16117.x

Björklund J, von Arx G, Nievergelt D, Wilson RJS, Van den Bulcke J, Günther B, Loader NJ, Rydval M, Fonti P, Scharnweber T, Andreu-Hayles L, Büntgen U, D'Arrigo R, Davi N, De Mil T, Esper J, Gärtner H, Geary J, Gunnarson BE, Hartl C, Hevia A, Song H, Janecka K, Kaczka RJ, Kirdyanov AV, Kochbeck M, Liu Y, Meko M, Mundo I, Nicolussi K, Oelkers R, Pichler T, Sánchez-Salguero R, Schneider L, Schweingruber F, Timonen M, Trouet V, Van Acker J, Verstege A, Villalba R, Wilmking M, Frank D (2019) Scientific Merits and Analytical Challenges of Tree-Ring Densitometry. Rev Geophys 57(4). https://doi.org/10.1029/2019RG000642

Borkotoky SS, Williams AP, Cook ER, Steinschneider S (2021) Reconstructing extreme precipitation in the Sacramento River watershed using tree-ring based proxies of cold-season precipitation. Water Resources Research 57:e2020WR028824. https://doi.org/10.1029/2020WR028824

Bräker OU (1981) Der Alterstrend bei Jahrringdichten und Jahrringbreiten von Nadelhölzern und sein Ausgleich. Mitteilungen der forstlichen Bundesversuchsanstalt Wien 142:75–102

Brázdil R, Dobrovolný P, Luterbacher J, Moberg A, Pfister C, Wheeler D, Zorita E (2010) European climate of the past 500 years: new challenges for historical climatology. Clim Change 101(1–2):7–40. https://doi.org/10.1007/s10584-009-9783-z

Brázdil R, Reznícková L, Valásek H, Dolák L, Oldrich (2017) Climatic and other responses to the Lakagígar 1783 and Tambora 1815 volcanic eruptions in the Czech Lands. Geografie 122:147–168. https://doi.org/DOI: 10.37040/geografie2017122020147

Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber F, Karlén W, Zetterberg P, Eronen M (1992) Fennoscandian summers from ad 500: temperature changes on short and long timescales. Clim Dyn 7(3):111–119. https://doi.org/10.1007/BF00211153

Briffa KR, Osborn TJ, Schweingruber F, Harris IC, Jones PD, Shiyatov SG, Vaganov EA (2001) Low-frequency temperature variations from a northern tree ring density network. J Geophys Res 106(D3):2929–2941. https://doi.org/10.1029/2000JD900617

Bunn AG (2010) Statistical and visual crossdating in R using the dplR library. Dendrochronologia 28(4):251–258. https://doi.org/10.1016/j.dendro.2009.12.001

Büntgen U, Esper J, Frank D, Nicolussi K, Schmidhalter M (2005) A 1052-year tree-ring proxy for Alpine summer temperatures. Clim Dyn 25(2–3):141–153. https://doi.org/10.1007/s00382-005-0028-1

Büntgen U, Frank D, Nievergelt D, Esper J (2006) Summer Temperature Variations in the European Alps, a.d. 755–2004. J Clim 19(21):5606–5623. https://doi.org/10.1175/JCLI3917.1

Büntgen U, Frank D, Wilson RJS, Carrer M, Urbinati C, Esper J (2008) Testing for tree-ring divergence in the European Alps: tree-ring divergence in the European Alps. Glob Change Biol 14(10):2443–2453. https://doi.org/10.1111/j.1365-2486.2008.01640.x

Büntgen U, Frank D, Liebhold A, Johnson D, Carrer M, Urbinati C, Grabner M, Nicolussi K, Levanic T, Esper J (2009) Three centuries of insect outbreaks across the European Alps. New Phytologist 182(4):929–941. https://doi.org/10.1111/j.1469-8137.2009.02825.x

Büntgen U, Arseneault D, Boucher E, Churakova (Sidorova) OV, Gennaretti F, Crivellaro A, Hughes MK, Kirdyanov AV, Klippel L, Krusic PJ, Linderholm HW, Ljungqvist FC, Ludescher J, McCormick M, Myglan VS, Nicolussi K, Piermattei A, Oppenheimer C, Reinig F, Sigl M, Vaganov EA, Esper J (2020) Prominent role of volcanism in Common Era climate variability and human history. Dendrochronologia 64:125757. https://doi.org/10.1016/j.dendro.2020.125757

Büntgen U, Allen K, Anchukaitis KJ, Arseneault D, Boucher E, Bräuning A, Chatterjee S, Cherubini P, Churakova OV, Corona C, Gennaretti F, Grießinger J, Guillet S, Guiot J, Gunnarson B, Helama S, Hochreuther P, Hughes MK, Huybers P, Kirdyanov AV, Krusic PJ, Ludescher J, Meier WJ-H, Myglan VS, Nicolussi K, Oppenheimer C, Reinig F, Salzer MW, Seftigen K, Stine AR, Stoffel M, St. George S, Tejedor E, Trevino A, Trouet V, Wang J, Wilson R, Yang B, Xu G, Esper J (2021) The influence of decision-making in tree ring-based climate reconstructions. Nat Commun 12(1):3411. https://doi.org/10.1038/s41467-021-23627-6

Büntgen U, Piermattei A, Crivellaro A, Reinig F, Krusic PJ, Trnka M, Torbenson M, Esper J (2022a) Common Era treeline fluctuations and their implications for climate reconstructions. Global and Planetary Change 219:103979. https://doi.org/10.1016/j.gloplacha.2022.103979

Büntgen U, Smith SH, Wagner S, Krusic P, Esper J, Piermattei A, Crivellaro A, Reinig F, Tegel W, Kirdyanov A, Trnka M, Oppenheimer C (2022b) Global tree-ring response and inferred climate variation following the mid-thirteenth century Samalas eruption. Clim Dyn 59(1–2):531–546. https://doi.org/10.1007/s00382-022-06141-3

Büntgen U, Reinig F, Verstege A, Piermattei A, Kunz M, Krusic P, Slavin P, Stěpanék P, Torbenson M, Martinez del Castillo E, Arosio T, Kirdyanov A, Oppenheimer C, Trnka

M, Palosse A, Bebchuk T, Camarero JJ, Esper J (2024) Recent summer warming over the western Mediterranean region is unprecedented since medieval times. Glob Planet Change 232:104336. https://doi.org/10.1016/j.gloplacha.2023.104336

Camenisch C, Brázdil R, Kiss A, Pfister C, Wetter O, Rohr C, Contino A, Retsö D (2020) Extreme heat and drought in 1473 and their impacts in Europe in the context of the early 1470s. Reg Environ Change 20(1):19. https://doi.org/10.1007/s10113-020-01601-0

Conedera M, Colombaroli D, Tinner W, Krebs P, Whitlock C (2017) Insights about past forest dynamics as a tool for present and future forest management in Switzerland. Forest Ecol Management 388:100–112. https://doi.org/10.1016/j.foreco.2016.10.027

Cook ER (1985) A time series analysis approach to tree ring standardization. Dissertation, University of Arizona

Cook ER, Briffa KR, Shiyatov SG, Mazepa V (1990) Tree-ring standardization and growth-trend estimation. In: Cook ER, Kairiukstis LA (eds) Methods of Dendrochronology: Applications in the Environmental Sciences, 1st edn. Kluwer Academic Publishers, Dordrecht, pp 104–123

Cook ER, Krusic PJ, Peters K, Melvin TM (2017) Program Signal Free (Version 45_v2b). RCS Signal Free tree-ring standardization program. Tree-Ring Laboratory of Lamont-Doherty Earth Observatory

Cook ER, Peters K (1981) The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. Tree-Ring Bull 41:45–53

Cook ER, Peters K (1997) Calculating unbiased tree-ring indices for the study of climatic and environmental change. The Holocene 7(3):361–370. https://doi.org/10.1177/095968369700700314

Coolidge WAB (1912) The Names of Zermatt. Engl Hist Rev XXVII(CVII):522–530. https://doi.org/10.1093/ehr/XXVII.CVII.522

Coppola A, Leonelli G, Salvatore MC, Pelfini M, Baroni C (2013) Tree-ring–based summer mean temperature variations in the Adamello–Presanella Group (Italian Central Alps), 1610–2008 AD. Clim Past 9(1):211–221. https://doi.org/10.5194/cp-9-211-2013

Corona C, Guiot J, Edouard JL, Chalié F, Büntgen U, Nola P, Urbinati C (2010) Millennium-long summer temperature variations in the European Alps as reconstructed from tree rings. Clim Past 6(3):379–400. https://doi.org/10.5194/cp-6-379-2010

Corona C, Edouard J-L, Guibal F, Guiot J, Bernard S, Thomas A, Denelle N (2011) Long-term summer (AD751–2008) temperature fluctuation in the French Alps based on tree-ring data. BOREAS 40:351–366. https://doi.org/10.1111/j.1502-3885.2010.00185.x.

D'Arrigo R, Wilson RJS, Liepert B, Cherubini P (2008) On the 'Divergence Problem' in Northern Forests: A review of the tree-ring evidence and possible causes. Glob Planet Change 60(3–4):289–305. https://doi.org/10.1016/j.gloplacha.2007.03.004

Düthorn E, Holzkämper S, Timonen M, Esper J (2013) Influence of micro-site conditions on tree-ring climate signals and trends in central and northern Sweden. Trees 27(5):1395–1404. https://doi.org/10.1007/s00468-013-0887-8

Düthorn E, Schneider L, Konter O, Schön P, Timonen M, Esper J (2015) On the hidden significance of differing micro-sites on tree-ring based climate reconstructions. Silva Fenn 49(1). https://doi.org/10.14214/sf.1220

Esper J, Cook ER, Krusic PJ, Peters K, Schweingruber F (2003) Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies. Tree-Ring Res 59(2):81–89

Esper J, Büntgen U, Frank D, Nievergelt D, Liebhold A (2007a) 1200 years of regular outbreaks in alpine insects. Proc R Soc B 274(1610):671–679. https://doi.org/10.1098/rspb.2006.0191

Esper J, Büntgen U, Frank D, Pichler T, Nicolussi K (2007b) Updating the Tyrol tree-ring dataset. In: Haneca K, Verheyden A, Beeckman H, Gärtner H, Helle G, Schleser G (eds) Schirften des Forschungszentrums Jülich. Forschungszentrum Jülich, Tervuren, Belgium, pp 80–84

Esper J, Frank D, Büntgen U, Verstege A, Luterbacher J, Xoplaki E (2007c) Long-term drought severity variations in Morocco. Geophys Res Lett 34(17):L17702. https://doi.org/10.1029/2007GL030844

Esper J, Schneider L, Krusic PJ, Luterbacher J, Büntgen U, Timonen M, Sirocko F, Zorita E (2013) European summer temperature response to annually dated volcanic eruptions over the past nine centuries. Bull Volcanol 75(7):736. https://doi.org/10.1007/s00445-013-0736-z

Esper J, Krusic PJ, Ljungqvist FC, Luterbacher J, Carrer M, Cook ER, Davi N, Hartl-Meier C, Kirdyanov AV, Konter O, Myglan VS, Timonen M, Treydte K, Trouet V, Villalba R, Yang B, Büntgen U (2016) Ranking of tree-ring based temperature reconstructions of the past millennium. Quat Sci Rev 145:134–151. https://doi.org/10.1016/j.quascirev.2016.05.009

Esper J, Büntgen U, Hartl-Meier C, Oppenheimer C, Schneider L (2017) Northern Hemisphere temperature anomalies during the 1450s period of ambiguous volcanic forcing. Bull Volcanol 79(6):41. https://doi.org/10.1007/s00445-017-1125-9

Esper J, Büntgen U (2021) The future of paleoclimate. Clim Res 83:57–59. https://doi.org/10.3354/cr01636

Eyring V, Cox PM, Flato GM, Gleckler PJ, Abramowitz G, Caldwell P, Collins WD, Gier BK, Hall AD, Hoffman FM, Hurtt GC, Jahn A, Jones CD, Klein SA, Krasting JP, Kwiatkowski L, Lorenz R, Maloney E, Meehl GA, Pendergrass AG, Pincus R, Ruane AC, Russell JL, Sanderson BM, Santer BD, Sherwood SC, Simpson IR, Stouffer RJ, Williamson MS (2019) Taking climate model evaluation to the next level. Nat Clim Change 9(2):102–110. https://doi.org/10.1038/s41558-018-0355-y

Frank D, Esper J (2005a) Characterization and climate response patterns of a high-elevation, multi-species tree-ring network in the European Alps. Dendrochronologia 22(2):107–121. https://doi.org/10.1016/j.dendro.2005.02.004

Frank D, Esper J (2005b) Temperature reconstructions and comparisons with instrumental data from a tree-ring network for the European Alps. Int J Climatol 25(11):1437–1454. https://doi.org/10.1002/joc.1210

Frank D, Wilson R, Esper J (2005) Synchronous variability changes in Alpine temperature and tree-ring data over the past two centuries. BOREAS 34:498–505. https://doi.org/10.1080/03009480500231443

Frank D, Esper J, Cook ER (2007) Adjustment for proxy number and coherence in a large-scale temperature reconstruction. Geophys Res Lett 34(16). https://doi.org/10.1029/2007GL030571

Grove JM (2001) The initiation of the" Little Ice Age" in regions round the North Atlantic. Clim Change 48:53–82. https://doi.org/10.1023/A:1005662822136

Guillet S, Corona C, Stoffel M, Khodri M, Lavigne F, Ortega P, Eckert N, Sielenou PD, Daux V, Churakova (Sidorova) OV, Davi N, Edouard J-L, Zhang Y, Luckman BH, Myglan VS, Guiot J, Beniston M, Masson-Delmotte V, Oppenheimer C (2017) Climate response to the Samalas volcanic eruption in 1257 revealed by proxy records. Nature Geosci 10(2):123–128. https://doi.org/10.1038/ngeo2875

Gunnarson BE, Linderholm HW, Moberg A (2011) Improving a tree-ring reconstruction from west-central Scandinavia: 900 years of warm-season temperatures. Clim Dyn 36(1–2):97–108. https://doi.org/10.1007/s00382-010-0783-5

Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci Data 7(1):109. https://doi.org/10.1038/s41597-020-0453-3

Hartl-Meier C, Dittmar C, Zang C, Rothe A (2014a) Mountain forest growth response to climate change in the Northern Limestone Alps. Trees 28(3):819–829. https://doi.org/10.1007/s00468-014-0994-1

Hartl-Meier C, Zang C, Dittmar C, Esper J, Göttlein A, Rothe A (2014b) Vulnerability of Norway spruce to climate change in mountain forests of the European Alps. Clim Res 60(2):119–132. https://doi.org/10.3354/cr01226

Hartl-Meier C, Büntgen U, Esper J (2016) On the occurrence of cyclic larch budmoth outbreaks beyond its geographical hotspots. In: Hevia A, Sánchez-Salguero R, Linares JC, Olano JM, Camarero JJ, Gutiérrez E, Helle G, Gärtner H (eds) Proceedings of the DENDROSYMPOSIUM 2015: May 20th - 23rd, 2015 in Sevilla, Spain, (Scientific Technical Report ; 16/04), 14th TRACE conference (Tree Rings in Archaeology, Climatology and Ecology). Deutsches GeoForschungsZentrum GFZ, (Sevilla, Spain), Potsdam, pp 86–92

Hartl-Meier C, Büntgen U, Smerdon JE, Zorita E, Krusic PJ, Ljungqvist FC, Schneider L, Esper J (2017a) Temperature Covariance in Tree Ring Reconstructions and Model Simulations Over the Past Millennium: Last-Millennium Temperature Covariance. Geophys Res Lett 44(18):9458–9469. https://doi.org/10.1002/2017GL073239

Hartl-Meier C, Esper J, Liebhold A, Konter O, Rothe A, Büntgen U (2017b) Effects of host abundance on larch budmoth outbreaks in the European Alps. Agr For Entomol 19(4):376–387. https://doi.org/10.1111/afe.12216

Hartl C, Düthorn E, Tejedor E, Kirchhefer AJ, Timonen M, Holzkämper S, Büntgen U, Esper J (2021) Micro-site conditions affect Fennoscandian forest growth. Dendrochronologia 65:125787. https://doi.org/10.1016/j.dendro.2020.125787

Hartl C, Schneider L, Riechelmann DFC, Kuhl E, Kochbeck M, Klippel L, Büntgen U, Esper J (2022) The temperature sensitivity along elevational gradients is more stable in maximum latewood density than tree-ring width. Dendrochronologia 73:125958. https://doi.org/10.1016/j.dendro.2022.125958

Johnson DM, Bjørnstad ON, Liebhold AM (2004) Landscape geometry and travelling waves in the larch budmoth: Landscape geometry and travelling waves in the LBM. Ecol Lett 7(10):967–974. https://doi.org/10.1111/j.1461-0248.2004.00659.x

King GM, Gugerli F, Fonti P, Frank D (2013) Tree growth response along an elevational gradient: climate or genetics? Oecologia 173(4):1587–1600. https://doi.org/10.1007/s00442-013-2696-6

Kuhl E, Zang C, Esper J, Riechelmann DFC, Büntgen U, Briesch M, Reinig F, Römer P, Konter O, Schmidhalter M, Hartl C (2023) Using machine learning on tree-ring data to determine the geographical provenance of historical construction timbers. Ecosphere :1–14. https://doi.org/DOI: 10.1002/ecs2.4453

Kunz M, Esper J, Kuhl E, Schneider L, Büntgen U, Hartl C (2023) Combining Tree-ring Width and Density to Separate the Effects of Climate 2 Variation and Insect Defoliation. Forests 14:1478. https://doi.org/10.3390/f14071478

Lavigne F, Degeai J-P, Komorowski J-C, Guillet S, Robert V, Lahitte P, Oppenheimer C, Stoffel M, Vidal CM, Surono, Pratomo I, Wassmer P, Hajdas I, Hadmoko DS, de Belizal E (2013) Source of the great A.D. 1257 mystery eruption unveiled, Samalas volcano,

Rinjani Volcanic Complex, Indonesia. Proc Natl Acad Sci USA 110(42):16742–16747. https://doi.org/10.1073/pnas.1307520110

Leonelli G, Coppola A, Baroni C, Salvatore MC, Maugeri M, Brunetti M, Pelfini M (2016) Multispecies dendroclimatic reconstructions of summer temperature in the European Alps enhanced by trees highly sensitive to temperature. Clim Change 137(1–2):275–291. https://doi.org/10.1007/s10584-016-1658-5

Lough JM, Fritts HC (1987) An assessment of the possible effects of volcanic eruptions on North American climate using tree-ring data, 1602 to 1900 A.D. Clim Change 10(3):219–239. https://doi.org/10.1007/BF00143903

Luterbacher J, Werner JP, Smerdon JE, Fernández-Donado L, González-Rouco FJ, Barriopedro D, Ljungqvist FC, Büntgen U, Zorita E, Wagner S, Esper J, McCarroll D, Toreti A, Frank D, Jungclaus JH, Barriendos M, Bertolin C, Bothe O, Brázdil R, Camuffo D, Dobrovolný P, Gagen M, García-Bustamante E, Ge Q, Gómez-Navarro JJ, Guiot J, Hao Z, Hegerl GC, Holmgren K, Klimenko VV, Martín-Chivelet J, Pfister C, Roberts N, Schindler A, Schurer A, Solomina O, von Gunten L, Wahl E, Wanner H, Wetter O, Xoplaki E, Yuan N, Zanchettin D, Zhang H, Zerefos C (2016) European summer temperatures since Roman times. Environ Res Lett 11(2):024001. https://doi.org/10.1088/1748-9326/11/2/024001

Lyu S, Li Z, Zhang Y, Wang X (2016) A 414-year tree-ring-based April–July minimum temperature reconstruction and its implications for the extreme climate events, northeast China. Clim Past 12:1879–1888. https://doi.org/10.5194/cp-12-1879-2016

Melvin TM, Briffa KR, Nicolussi K, Grabner M (2007) Time-varying-response smoothing. Dendrochronologia 25(1):65–69. https://doi.org/10.1016/j.dendro.2007.01.004

Melvin TM, Briffa KR (2008) A “signal-free” approach to dendroclimatic standardisation. Dendrochronologia 26(2):71–86. https://doi.org/10.1016/j.dendro.2007.12.001

Melvin TM, Grudd H, Briffa KR (2013) Potential bias in ‘updating’ tree-ring chronologies using regional curve standardisation: Re-processing 1500 years of Torneträsk density and ring-width data. The Holocene 23(3):364–373. https://doi.org/10.1177/0959683612460791

Neuwirth B, Esper J, Schweingruber FH, Winiger M (2004) Site ecological differences to the climatic forcing of spruce pointer years from the Loetschental, Switzerland. Dendrochronologia 21(2):69–78. https://doi.org/10.1078/1125-7865-00040

Oppenheimer C (2003) Climatic, environmental and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. Prog Phys Geogr: Earth Environ 27(2):230–259. https://doi.org/10.1191/0309133303pp379ra

Osborn TJ, Briffa KR, Jones PD (1997) Adjusting variance for sample-size in tree-ring chronologies and other regional mean timeseries. Dendrochronologia 15:89–99

Pretis F, Reade J, Sucarrat G (2016) General-to-specific (GETS) modelling and indicator saturation with the R package gets. Department of Economics Discussion Paper Series University of Oxford

Pretis F, Schneider L, Smerdon JE, Hendry DF (2017) Detecting Volcanic Eruptions in Temperature Reconstructions by Designed Break-Indicator Saturation. In: Chi-ang Lin B, Zheng S (eds) Environmental Economics and Sustainability. John Wiley & Sons, Ltd, pp 7–37

Qin C, Yang B, Bräuning A, Sonechkin DM, Huang K (2011) Regional extreme climate events on the northeastern Tibetan Plateau since AD 1450 inferred from tree rings. Glob Planet Change 75:143–154. https://doi.org/10.1016/j.gloplacha.2010.10.013

R Core Team (2021) R: A language and environment for statistical computing. Vienna, Austria

Riechelmann DFC, Schmidhalter M, Büntgen U, Esper J (2013) Extending the high elevation larch ring width chronology from the Simplon region in the Swiss Alps over the past millennium. In: Helle G, Gärtner H, Beck W, Heinrich I, Heussner K-U, Müller A, Sanders T (eds) Proceedings of the DENDROSYMPOSIUM 2012: May 8th - 12th, 2012 in Potsdam and Eberswalde, Germany, (Scientific Technical Report ; 13/05), 11th TRACE conference (Tree Rings in Archaeology, Climatology and Ecology). Deutsches GeoForschungsZentrum GFZ, Potsdam, Eberswalde, pp 103–107

Riechelmann DFC, Hartl C, Esper J (2020) The effect of provenance of historical timber on tree-ring based temperature reconstructions in the Western Central Alps. iForest 13(1):351–359. https://doi.org/10.3832/ifor3412-013

Römer P, Hartl C, Schneider L, Bräuning A, Szymczak S, Huneau F, Lebre S, Reinig F, Büntgen U, Esper J (2021) Reduced Temperature Sensitivity of Maximum Latewood Density Formation in High-Elevation Corsican Pines under Recent Warming. Atmosphere 12(7):804. https://doi.org/10.3390/atmos12070804

Salzer MW, Larson ER, Bunn AG, Hughes MK (2014) Changing climate response in near-treeline bristlecone pine with elevation and aspect. Environ Res Lett 9(11):114007. https://doi.org/10.1088/1748-9326/9/11/114007

Saulnier M, Roques A, Guibal F, Rozenberg P, Saracco G, Corona C, Edouard J-L (2017) Spatiotemporal heterogeneity of larch budmoth outbreaks in the French Alps over the last 500 years. Can J For Res 47(5):667–680. https://doi.org/10.1139/cjfr-2016-0211

Schmidli J, Schmutz C, Frei C, Wanner H, Schär C (2002) Mesoscale precipitation variability in the region of the European Alps during the 20th century. Int J Climatol 22(9):1049–1074. https://doi.org/10.1002/joc.769

Schneider L, Smerdon JE, Büntgen U, Wilson RJS, Myglan VS, Kirdyanov AV, Esper J (2015) Revising midlatitude summer temperatures back to A.D. 600 based on a wood density network. Geophys Res Lett 42(11):4556–4562. https://doi.org/10.1002/2015GL063956

Schneider L, Smerdon JE, Pretis F, Hartl-Meier C, Esper J (2017) A new archive of large volcanic events over the past millennium derived from reconstructed summer temperatures. Environ Res Lett 12(9):094005. https://doi.org/10.1088/1748-9326/aa7a1b

Schweingruber FH, Bräker OU, Schar E (1987) Temperature information from a European dendroclimatological sampling network. Dendrochronologia 5:9–33

Schweingruber F (1988) Tree Rings. Paul Haupt, Bern, Stuttgart

Schweingruber FH, Bartholin T, Schar E, Briffa KR (1988) Radiodensitometric-dendroclimatological conifer chronologies from Lapland (Scandinavia) and the Alps (Switzerland). BOREAS 17:559–566. https://doi.org/10.1111/j.1502-3885.1988.tb00569.x

Sevruk B (1997) Regional Dependency of Precipitation-Altitude Relationship in the Swiss Alps. Clim Change 36:355–369. https://doi.org/10.1023/A:1005302626066

Shah S, Shah C (2015) Tree Rings for the Assessment of the Potential Impact of Climate Change on Forest Growth. Appl Ecol Environ Res 13(1):277–288. https://doi.org/10.15666/aeer/1301_277288

Sigl M, Toohey M, McConnell JR, Cole-Dai J, Severi M (2021) HolVol: Reconstructed volcanic stratospheric sulfur injections and aerosol optical depth for the Holocene (9500 BCE to 1900 CE). PANGAEA. https://doi.org/10.1594/PANGAEA.928646

St. George S, Esper J (2019) Concord and discord among Northern Hemisphere paleotemperature reconstructions from tree rings. Quaternary Science Reviews 203:278–281. https://doi.org/10.1016/j.quascirev.2018.11.013

Stoffel M, Khodri M, Corona C, Guillet S, Poulain V, Bekki S, Guiot J, Luckman BH, Oppenheimer C, Lebas N, Beniston M, Masson-Delmotte V (2015) Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. Nature Geosci 8(10):784–788. https://doi.org/10.1038/ngeo2526

Stothers RB (1984) The Great Tambora Eruption in 1815 and Its Aftermath. Science (224):1191–1198

Torbenson MCA, Büntgen U, Esper J, Urban O, Balek J, Reinig F, Krusic PJ, Martinez del Castillo E, Brázdil R, Semerádová D, Štěpánek P, Pernicová N, Kolář T, Rybníček M, Koňasová E, Arbelaez J, Trnka M (2023) Central European Agroclimate over the Past 2000 Years. J Clim 36(13):4429–4441. https://doi.org/10.1175/JCLI-D-22-0831.1

Trachsel M, Kamenik C, Grosjean M, McCarroll D, Moberg A, Brázdil R, Büntgen U, Dobrovolný P, Esper J, Frank DC, Friedrich M, Glaser R, Larocque-Tobler I, Nicolussi K, Riemann D (2012) Multi-archive summer temperature reconstruction for the European Alps, AD 1053–1996. Quat Sci Rev 46:66–79. https://doi.org/10.1016/j.quascirev.2012.04.021

Wang J, Yang B, Fang M, Wang Z, Liu J, Kang S (2023) Synchronization of summer peak temperatures in the Medieval Climate Anomaly and Little Ice Age across the Northern Hemisphere varies with space and time scales. Clim Dyn 60(11–12):3455–3470. https://doi.org/10.1007/s00382-022-06524-6

Wanner H, Pfister C, Neukom R (2022) The variable European Little Ice Age. Quat Sci Rev 287:107531. https://doi.org/10.1016/j.quascirev.2022.107531

Wilmking M, D'Arrigo RD, Jacoby GC, Juday GP (2005) Increased temperature sensitivity and divergent growth trends in circumpolar boreal forests. Geophys Res Lett 32(15):L15715. https://doi.org/10.1029/2005GL023331

Wilson RJS, D'Arrigo R, Buckley B, Büntgen U, Esper J, Frank D, Luckman B, Payette S, Vose R, Youngblut D (2007) A matter of divergence: Tracking recent warming at hemispheric scales using tree ring data. J Geophys Res 112(D17):D17103. https://doi.org/10.1029/2006JD008318

Wilson RJS, Anchukaitis KJ, Briffa KR, Büntgen U, Cook ER, D'Arrigo R, Davi N, Esper J, Frank D, Gunnarson BE, Hegerl GC, Helama S, Klesse S, Krusic PJ, Linderholm HW, Myglan V, Osborn TJ, Rydval M, Schneider L, Schurer A, Wiles G, Zhang P, Zorita E (2016) Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. Quat Sci Rev 134:1–18. https://doi.org/10.1016/j.quascirev.2015.12.005

Zang C, Biondi F (2015) treeclim: an R package for the numerical calibration of proxy-climate relationships. Ecography 38(4):431–436. https://doi.org/10.1111/ecog.01335

Zhang P, Björklund J, Linderholm HW (2015) The influence of elevational differences in absolute maximum density values on regional climate reconstructions. Trees 29(4):1259–1271. https://doi.org/10.1007/s00468-015-1205-4

3.8 Supplementary material

Tab. S 3-1 Data Selection after applying the Provenance Model to Simplon and Zermatt/ Zmutt historical series (Kuhl et al. 2023). Numbers include living and historical series.

	Simplon Valley			Matter Valley			
Site (Elevation [m asl])	**SV1 (1900)**	**SV2 (2000)**	**SV3 (2200)**	**MV1 (2000)**	**MV2 (2300)**	**MV3 (2300)**	**SUM**
Total number of series	56	24	25	180	24	43	**352**

Tab. S 3-2 Performance measures (correlation between predicted and observed (r), explained variance (R^2) and root mean squared error (RMSE)) of the linear models in **Fig. 3-6a)**. Asterisks show the columns for the linear model calibrated on the early period, without asterisks are the columns for the linear model calibrated on the later period.

	Calibration Periods		Validation Periods	
Period [CE]	**1901-1959***	**1960-2017**	**1960-2017***	**1901-1959**
r	0.83	0.84	0.84	0.83
R^2	0.69	0.7	0.7	0.69
RMSE	0.39	0.46	0.7	0.64

Tab. S 3-3 Testing for low frequency signals: Approaches applied to analyse the missing low frequency in our reconstruction.

No	**Approach**	**Visual improvement of low-frequency signal?**
1	No mean adjustment	No
2	Including another high elevation living site.	No
3	Other detrending methods (Spline, Hugershoff, signal-free age-dependent Spline signal-free RCS, see Fig. S 3-2)	No
4	Include low elevation historical series and adjust their mean levels accordingly	No
5	Use all historical series neglecting the new altitudinal approach (see Fig. S 3-5)	No
6	Subsampling (50/50 ratio Simplon/ Zermatt samples)	No
7	Reducing samples of houses with overrepresentation in a time period	No
8	Influences of tree age distribution in time	No
9	Only using only the historical series from Simplon	No (replication too low)

Tab. S 3-4 20 volcanic events between 880-2000 CE linked with the highest volcanic stratospheric aerosol optical depth (SAOD) values in the Westwind Zone (30 - 60°N) from Sigl et al. (2021)[1]. Eruption year and volcanos were taken from Wang et al. (2022)[2] and Büntgen et al. (2020)[3]. Greyed event no. 20 is taken from Wang et al. (2022) to extend the record into the 19th century (SAOD Peak was calculated over 30-90°N[4]).

No.	Eruption Year [CE]	SAOD Peak Year[1] [CE]	SAOD Peak 30-60N[1]	Volcano
1	1257[2,3]	1258	0.46	Samalas[2,3]
2	939[3]	939	0.43	Katla[3]
3	1180[2]	1181	0.34	UE[2] /Katla[3]
4	1783[2,3]	1783	0.32	Laki[3]
5	1457[2]	1458	0.26	UE[2,3]
6	1815[2,3]	1815	0.26	Tambora[2,3]
7	1229[2]	1229	0.24	UE[2,3]
8	1641[2,3]	1642	0.23	Parker[2,3]
9	1831[2,3]	1832	0.2	UE[3]
10	1108[3]	1109	0.16	UE[3]
11	1477[2,3]	1477	0.16	Veidivötn[2,3]
12	1199	1199	0.15	UE
13	1170[2]	1171	0.14	UE[2,3]
14	1600[2,3]	1600	0.13	Huaynaputina[2,3]
15	1210[2,3]	1209	0.12	Katla[2]
16	1343[2]	1344	0.12	UE[2]
17	1586[2]	1587	0.12	Kelud[2]
18	1667[2]	1666	0.12	Shikotsu[2]
19	1729[3]	1729	0.12	UE[3]
20	1883[2,3]	1883	0.12[4]	Krakatau[2,3]

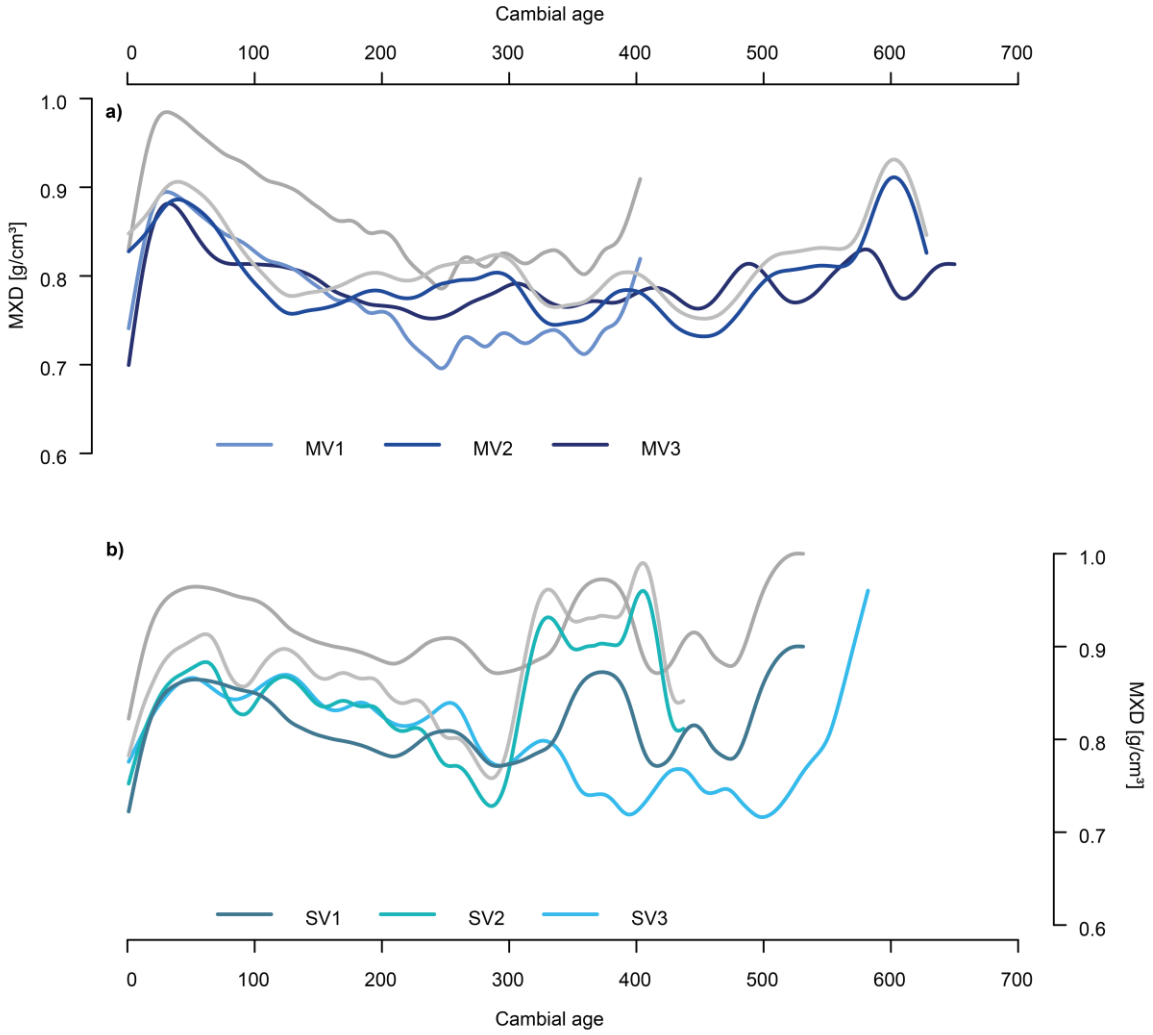


Fig. S 3-1 Mean adjustment of the Matter valley in **a)** and Simplon valley in **b)** depending on the regional curves. Lower elevation sites were adjusted to the highest one. Differences were calculated over the period of replication ≥ 15. Grey lines show the original curves of the sites.

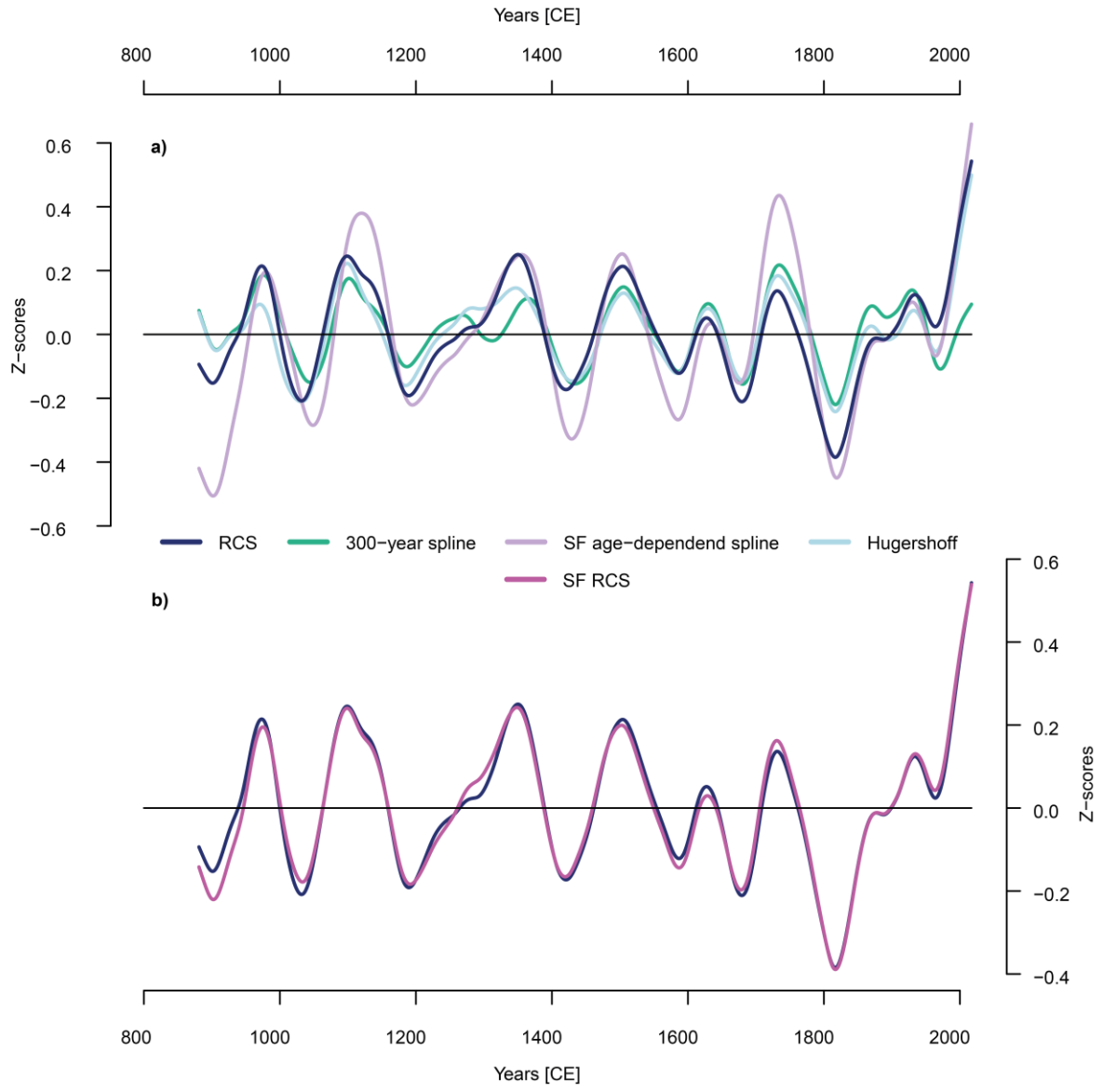


Fig. S 3-2 Different detrending methods compared with each other, smoothed with a 100-year smoothing spline. **a)** Regional Curve Standardization (RCS, Briffa et al. 1992), 300-year Spline detrending (Cook and Peters 1981), Hugershoff detrending (Cook et al. 1990) and a signal-free (SF) age-dependent spline detrending (Melvin and Briffa 2008) **b)** Comparison of classical RCS with the SF RCS (Melvin and Briffa 2014) detrending.

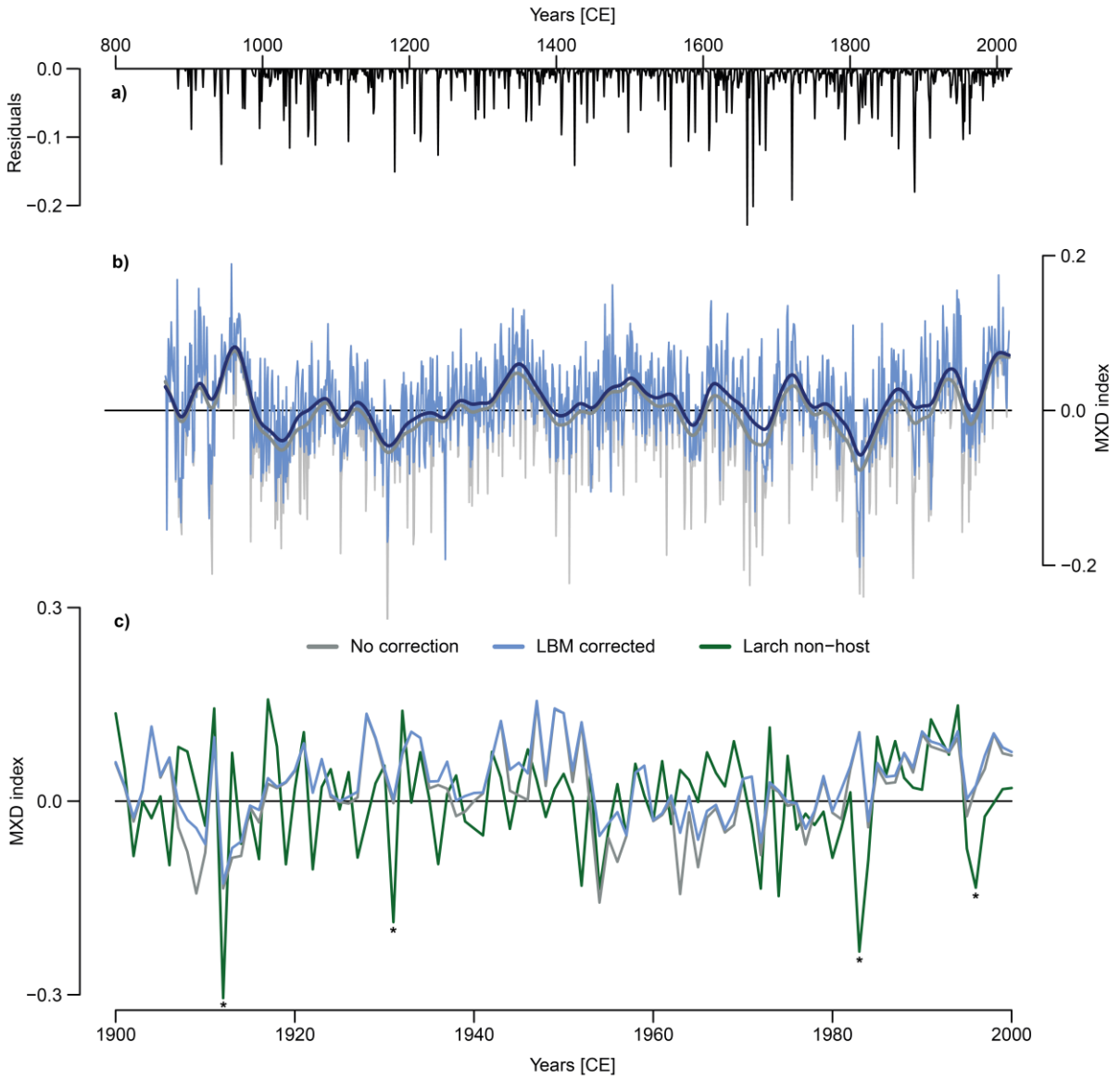


Fig. S 3-3 Larch budmoth (LBM) detection and correction using Impulse Indicator Saturation (IIS) after Pretis et al. (2016). Residuals between the original and the corrected chronology in **a)** present a frequent detection of LBM events throughout the timeseries **b)** The resulting corrected chronology **c)** A zoom into the 20th century in strengthens how a larch non-host as addition indicator is used by the algorithm to exclude potential climate related declines (asterisks) from the correction.

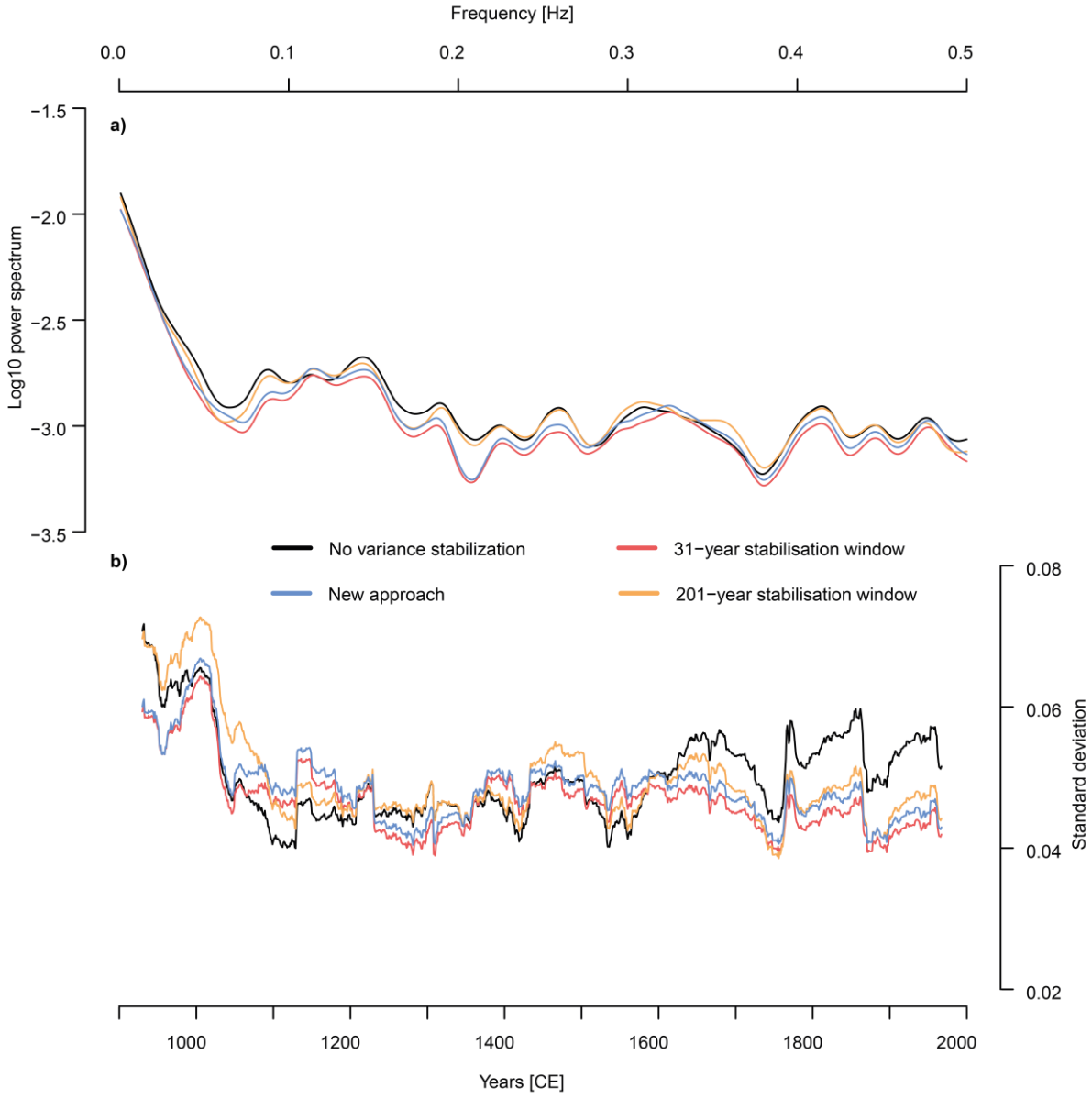


Fig. S 3-4 Variance stabilization effects of window-size and the new presented split-window stabilization using the method of Osborn et al. (1997): Panel **a)** shows the spectrum analysis (log10 power spectrum) calculated using Fast Fourier Transformation **b)** Resulting standard deviations of the chronologies were calculated over a 100-year running window with a 1-year lag.

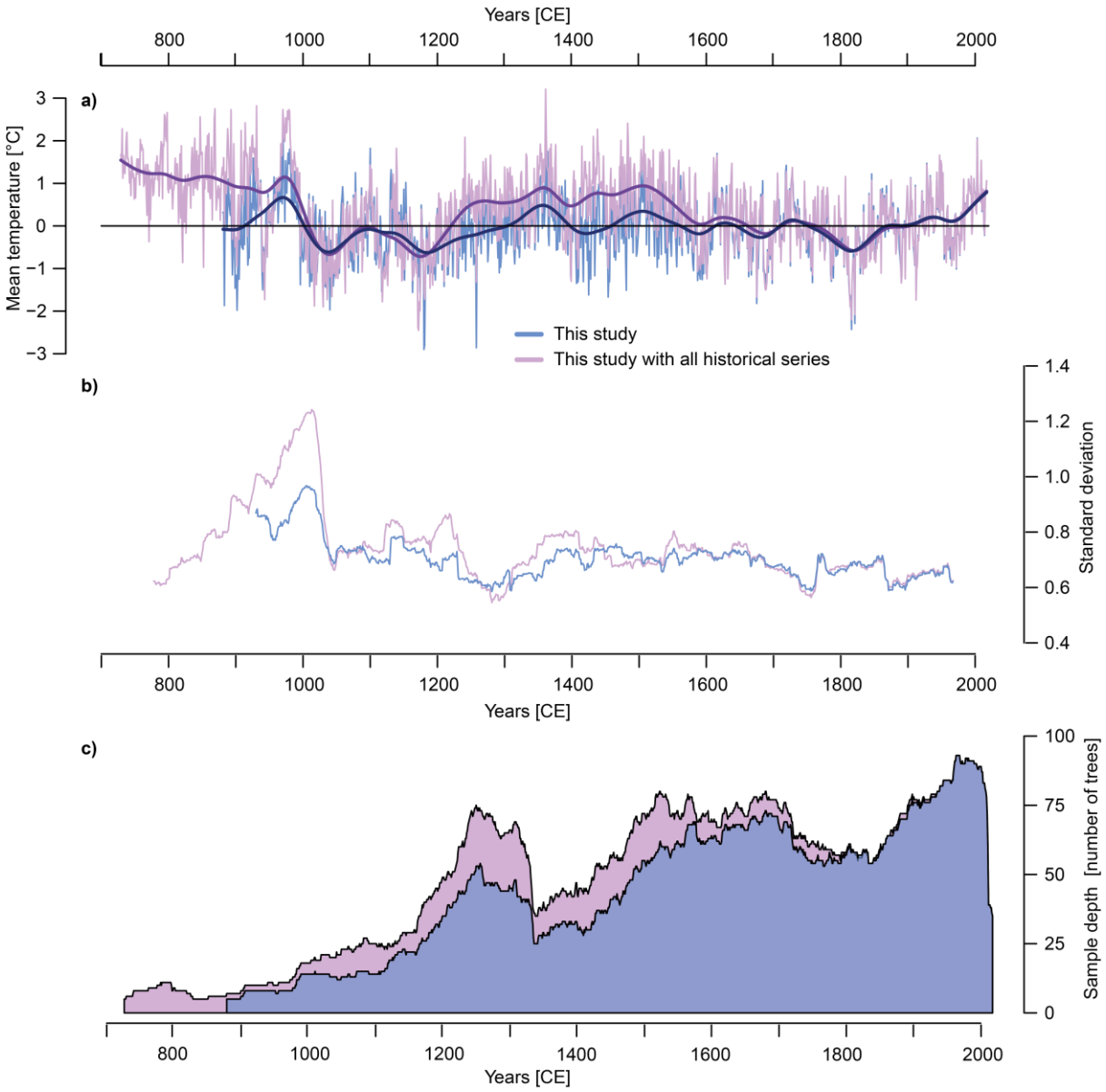


Fig. S 3-5 Comparison between a classical reconstruction (see methods chapter for more details on procedure) and the new altitude considering approach in **a)**. Panel **b)** shows the standard deviations of the reconstructions over a 100-year running window (lag 1) **c)** presents the difference in sample depth between an altitude adjusted and a classical dataset.

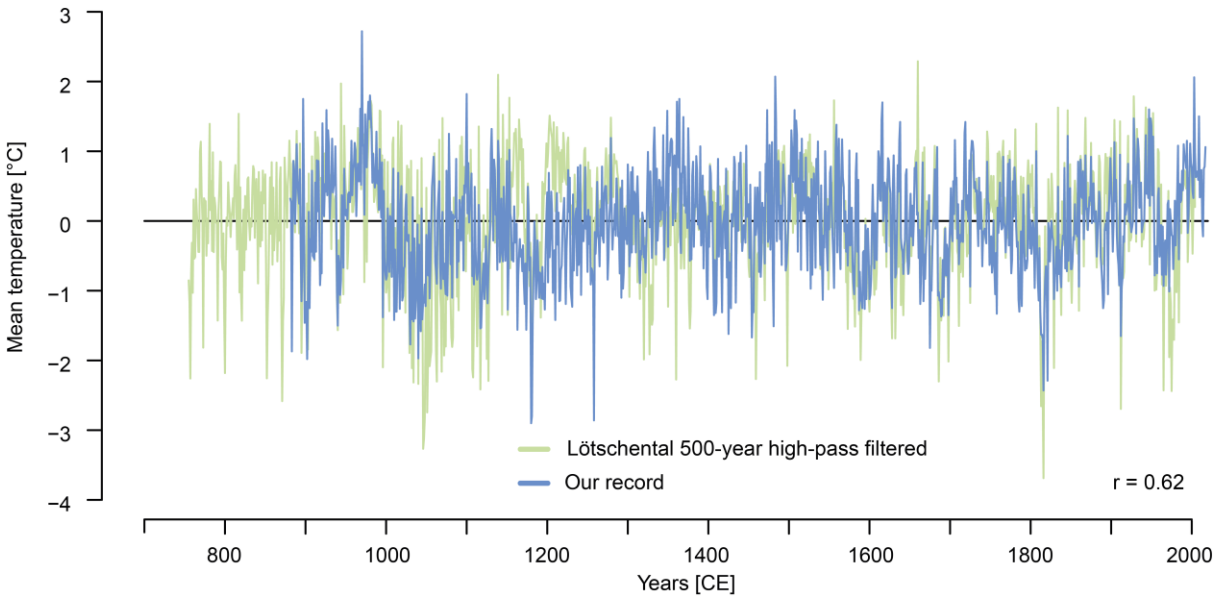


Fig. S 3-6 Comparison between here presented reconstruction and the 500-year high-pass filtered Lötschental reconstruction (Büntgen et al. 2006).

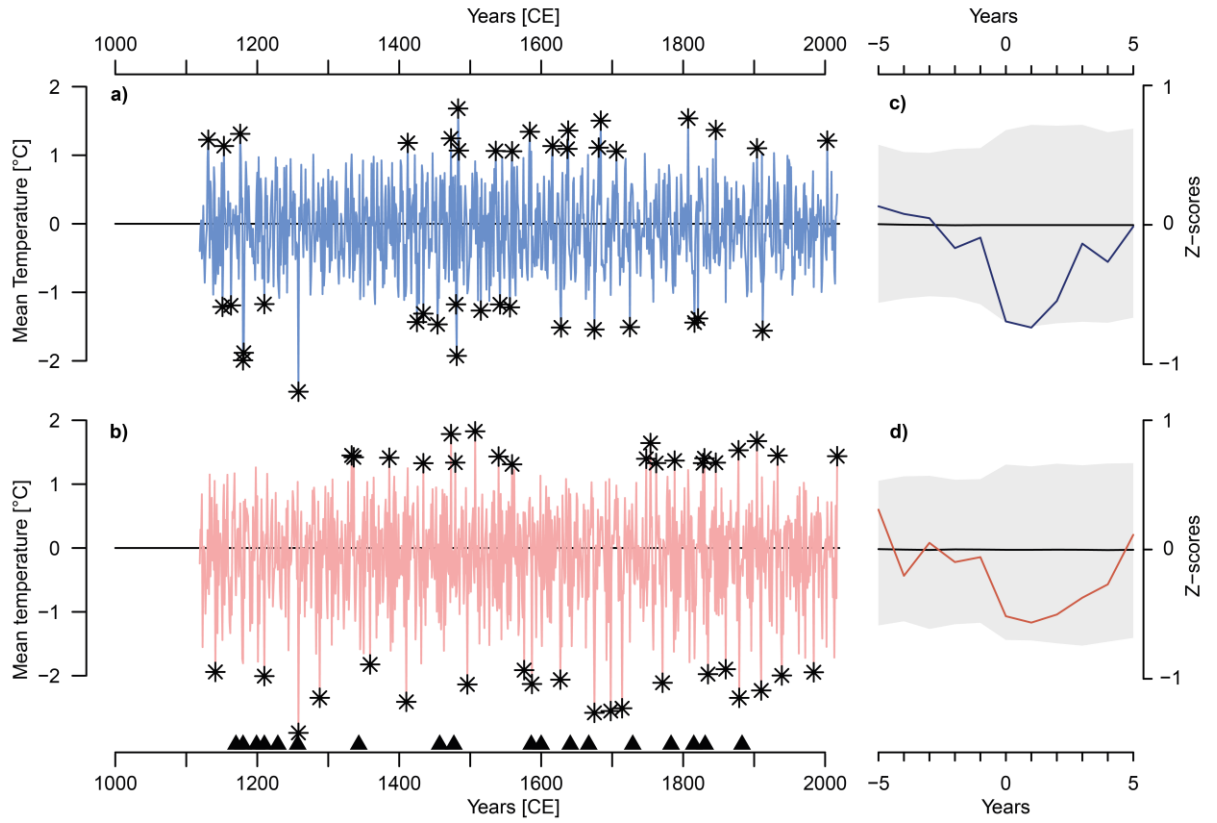


Fig. S 3-7 a) 30-year high-pass filtered reconstruction from this study and **b)** 30-year high-pass filtered Pyrenees record. Asterisks denote to the 20 warmest and the 20 coldest years, respectively. Triangles show 18 strongest volcanic events between 1119-2017 AD (see **Tab. S 3-4**). **c)** and **d)** Superposed epoch analysis for these 18 events (lag = 5, residuals from the 5 years prior to event) with mean (black line) and 99% confidence intervals (grey) after bootstrap resampling (n = 10,000).

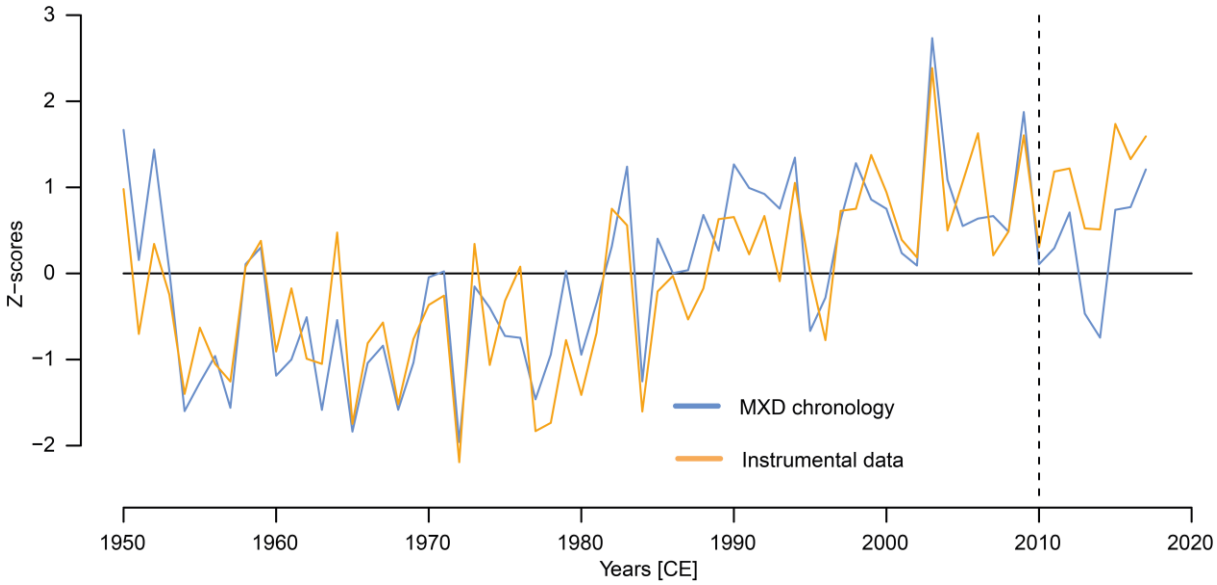


Fig. S 3-8 Z-scores of instrumental data versus the detrended MXD chronology.

Supplementary references

Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber F, Karlén W, Zetterberg P, Eronen M (1992) Fennoscandian summers from ad 500: temperature changes on short and long timescales. Clim Dyn 7(3):111–119. https://doi.org/10.1007/BF00211153

Büntgen U, Arseneault D, Boucher E, Churakova (Sidorova) OV, Gennaretti F, Crivellaro A, Hughes MK, Kirdyanov AV, Klippel L, Krusic PJ, Linderholm HW, Ljungqvist FC, Ludescher J, McCormick M, Myglan VS, Nicolussi K, Piermattei A, Oppenheimer C, Reinig F, Sigl M, Vaganov EA, Esper J (2020) Prominent role of volcanism in Common Era climate variability and human history. Dendrochronologia 64:125757. https://doi.org/10.1016/j.dendro.2020.125757

Büntgen U, Frank D, Nievergelt D, Esper J (2006) Summer Temperature Variations in the European Alps, a.d. 755–2004. J Clim 19(21):5606–5623. https://doi.org/10.1175/JCLI3917.1

Cook ER, Briffa KR, Shiyatov SG, Mazepa V (1990) Tree-ring standardization and growth-trend estimation. In: Cook ER, Kairiukstis LA (eds) Methods of Dendrochronology: Applications in the Environmental Sciences. Kluwer Academic Publishers, Dordrecht, pp 104–123

Cook ER, Peters K (1981) The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. Tree-Ring Bull 41:45–53

Kuhl E, Zang C, Esper J, Riechelmann DFC, Büntgen U, Briesch M, Reinig F, Römer P, Konter O, Schmidhalter M, Hartl C (2023) Using machine learning on tree-ring data to

determine the geographical provenance of historical construction timbers. Ecosphere :1–14. https://doi.org/DOI: 10.1002/ecs2.4453

Melvin TM, Briffa KR (2008) A “signal-free” approach to dendroclimatic standardisation. Dendrochronologia 26(2):71–86. https://doi.org/10.1016/j.dendro.2007.12.001

Melvin TM, Briffa KR (2014) CRUST: Software for the implementation of Regional Chronology Standardisation: Part 1. Signal-Free RCS. Dendrochronologia 32(1):7–20. https://doi.org/10.1016/j.dendro.2013.06.002

Osborn TJ, Briffa KR, Jones PD (1997) Adjusting variance for sample-size in tree-ring chronologies and other regional mean timeseries. Dendrochronologia 15:89–99

Pretis F, Reade J, Sucarrat G (2016) General-to-specific (GETS) modelling and indicator saturation with the R package gets. Department of Economics Discussion Paper Series University of Oxford

Sigl M, Toohey M, McConnell JR, Cole-Dai J, Severi M (2021) HolVol: Reconstructed volcanic stratospheric sulfur injections and aerosol optical depth for the Holocene (9500 BCE to 1900 CE). PANGAEA. https://doi.org/10.1594/PANGAEA.928646

Wang F, Arseneault D, Boucher É, Gennaretti F, Yu S, Zhang T (2022) Tropical volcanoes synchronize eastern Canada with Northern Hemisphere millennial temperature variability. Nat Commun 13(1):5042. https://doi.org/10.1038/s41467-022-32682-6

4 A machine learning approach to fill gaps in dendrometer data

Eileen Kuhl[1], Emanuele Ziaco[1], Jan Esper[1,2], Oliver Konter[1], Edurne Martinez del Castillo[1]

[1] Department of Geography, Johannes Gutenberg University, Mainz, Germany

[2]Global Change Research Centre (CzechGlobe), Brno, Czech Republic

Published in October 2024: Trees. DOI:10.1007/s00468-024-02573-y

Summary

The susceptibility of dendrometer devices to technical failures often makes time-series analyses challenging. Resulting data gaps decrease sample size and complicate time-series comparison and integration. Existing methods either focus on bridging smaller gaps, are dependent on data from other trees or rely on climate parameters. In this study, we test eight machine learning (ML) algorithms to fill gaps in dendrometer data of individual trees in urban and non-urban environments. Among these algorithms, extreme gradient boosting (XGB) demonstrates the best skill to bridge artificially created gaps throughout the growing seasons of individual trees. The individual tree models are suited to fill gaps up to 30 consecutive days and perform particularly well at the start and end of the growing season. The method is independent of climate input variables or dendrometer data from neighbouring trees. The varying limitations among existing approaches call for cross-comparison of multiple methods and visual control. Our findings indicate that ML is a valid approach to fill gaps in individual trees, which can be of particular importance in situations of limited inter-tree co-variance, such as in urban environments.

4.1 Introduction

The growth of trees on intra-annual level has been the subject of numerous studies ranging from urban tree growth (Lindén et al. 2016; Moser-Reischl et al. 2019) over experimental orchard settings (Corell et al. 2014) to forest tree analyses (King et al. 2013; Ziaco and Biondi 2018; Salomón et al. 2022; Zhang et al. 2024). Whilst undisturbed time series of dendrometer data over multiple years are desirable, many datasets contain longer periods of missing data. The primary cause of data loss is irregular physical monitoring due to the accessibility of the sites (e.g. remoteness of sites, time/cost minimization or travel restrictions during pandemics), which can result in battery power or logger failure, full data storage capacity or dendrometers at maximum. Furthermore, other technical damages like moisture intrusion, animal bites, extreme weather events (e.g. storm damage) or vandalism can result in missing values over multiple days to months. The presence of prolonged phases of missing data can impede the ability to conduct a comprehensive analysis on a given dataset, particularly when these periods coincide with the growing season, and can reduce sample size (e.g. in King et al. 2013; Corell et al. 2014; Dulamsuren et al. 2023).

To date, the most common approaches for addressing gaps in dendrometer data have been incorporated into R packages like treenetproc (Haeni et al. 2020; Knüsel et al. 2021), or dendRoAnalyst (Aryal et al. 2020). The imputation approaches are primarily based on linear or spline interpolation and are constrained to a short period of consecutive missing values (e.g. 24 measuring points) in order to achieve acceptable results (Aryal et al. 2020; Knüsel et al. 2021). In addition to these methods, Aryal et al. (2020) introduced a linear-regression-based network interpolation approach, which assumes that all individuals of a tree species at one location share similar stem growth variability. For the successful gap filling this approach requires neighbouring trees with no missing values and high co-variance. Luković et al. (2022) tested different deep neural network architectures for gap filling and found that the combination of long short memory (LSTM) and convolutional neural networks (CNNs) performed better when an input of stem radius data and multiple climatic parameters, including temperature, relative humidity, solar radiation, and vapour pressure deficit was given. While the results of this study were promising, the authors suggested further tests of other machine learning (ML) methods and on other data.

In this study, we present a ML approach to test multiple supervised algorithms, datasets and feature combinations to reconstruct missing growth data from dendrometers. The novelty of this approach lies in its ability to fill data gaps exceeding 12 hours (> 24 measuring points) of

individual trees, when no supplementary data from other trees or climatic parameters are available. Furthermore, we include an evaluation of the multicollinearity of input variables and provide a straightforward scheme to be reproduced. The method was not only tested on stem growth data, but also on raw dendrometer data to expand the method to broader research applications.

4.2 Methods

4.2.1 Study location and data collection

The city of Mainz is located in western Germany (50.0° N, 8.3° E) and is defined by a temperate climate with warm summers and without a dry season (cfb, Beck et al. 2018b). The average yearly mean temperature and the average precipitation sum between 1991 and 2020 were 10.8°C and 579.3 mm, respectively (Mainz-Lerchenberg Station: Deutscher Wetterdienst 2024).

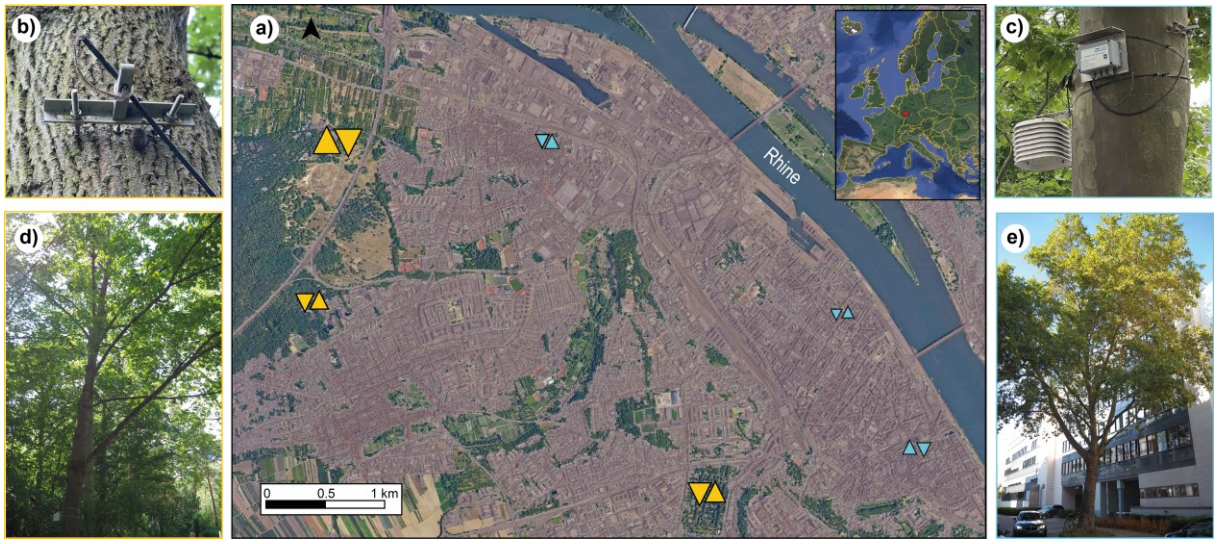


Fig. 4-1 a) Map of Mainz showing the locations in the urban (blue triangles) and non-urban (yellow triangles) areas. Triangles pointing upwards show maple trees and triangles pointing downwards show plane trees. Size of the triangles equal the size of the unsealed area around the trees relative to minimum and maximum (This figure has been prepared using European Union's Copernicus Land Monitoring Service information; https://doi.org/10.2909/3bf542bd-eebd-4d73-b53c-a0243f2ed862 and Google Satellite Image (2024)). **b)** Example of a point dendrometer. **c)** Example of a full set up including a Stevenson screen. **d)** Example of a non-urban maple tree. **e)** Example of an urban plane tree (Photo Credits: D. Thimm (b, c); S. Schöfl (d, e))

Norway maple (*Acer platanoides* L.) and London plane trees (*Platanus x hispanica* Münchh.) are common urban tree species in Europe. In Mainz, these maple and plane species make up for 32.5% and 8% of the total urban tree population (Landeshauptstadt Mainz 2024 (status of 2023)). In February 2019, six maple and six plane trees were selected at different locations throughout the urban and surrounding non-urban areas of the city. On each location, point dendrometers (Ecomatik GmbH DR2), temperature and relative humidity (RH) sensors were

installed (BMC Solutions GmbH HOBO U23-001 Pro v2 data loggers) (**Fig. 4-1**). All sensors measured in a 30-minute interval. For this study, all measurements in the common period from April 2019 to October 2023 were used.

4.2.2 Preparation of sensor and dendrometer data

First, the temperature data was subjected to quality control procedures based on the methods described in Eischeid et al. (1995), Beck et al. (2018a) and Barraro et al. (2022). The raw dendrometer data were examined using the R package treenetproc (Haeni et al. 2020; Knüsel et al. 2021) to identify and remove any erroneous measurements such as shifts in the data (here called jumps) related to technical failures. Here, the temperature data were used to assure no removal of frost indicated jumps. Consecutive missing values due to erroneous data in the temperature or dendrometer data, which did not exceed 24 measuring points (i.e. 12 hours), were gap filled using standard linear interpolation (Haeni et al. 2020; Knüsel et al. 2021). To extract tree growth from the cleaned dendrometer data, the zero-growth model (Zweifel et al. 2016) was applied. Additional information on the trees included metadata such as the area of unsealed soil [m²], the diameter at breast height [cm] or the tree height [mm] (Table 1). In areas where trees were growing in a completely unsealed environment, a maximum value of unsealed area was set at 144 m².

To find the optimal approach for gap filling, a series of tests was conducted utilizing a combination of different datasets (**Tab. 4-1**) and methods (**Tab. S 4-1**). Two hourly-resolved datasets for each species, maple, and plane, were built including measurements from all locations (datasets #1-2). Four additional datasets were constructed by splitting the species-specific ones into urban and non-urban locations per species (datasets #3-6). Furthermore, these datasets were utilized to investigate the effects of data size and hyperparameter tuning on the model performances. Hyperparameter tuning enables data scientists to adjust model parameters for optimal performance, thereby avoiding overfitting to the training data (Géron 2019). Additionally, 12 datasets representing the individual trees were subject to testing (summarized in #7-8). The presence of multicollinearity among predictive variables was evaluated using a variance inflation factor threshold of 5 (Dormann et al. 2013). Correlated variables were excluded, leaving a specific set of predictors for each dataset (hereafter called features Xi): day of the year (DOY), year, hour, and area (set as constant value per tree). In the non-urban datasets, tree height was included as predictor, as VIF values were below 5. All mentioned data processing steps were computed in R 4.2.2 (R Core Team 2021).

Tab. 4-1 Tested datasets for gap filling. X_i present the used features for prediction (DOY = day of year, area = unsealed area around the trees). Average diameter at breast height (DBH) and average unsealed area are given. VIF is the maximum variance inflation factor after multicollinearity tests and feature selection

#	Dataset	X_i	Average DBH [cm]	Average Unsealed area [m²]	VIF
1	Maple	DOY, year, hour, area	35.00	69.20	≤ 1.06
2	Plane	DOY, year, hour, area	55.22	52.14	≤ 1.06
3	Urban Maple	DOY, year, hour, area	30.37	15.24	≤ 1.06
4	Urban Plane	DOY, year, hour, area	58.67	13.22	≤ 1.06
5	Non-urban Maple	DOY, year, hour, area, tree height	39.63	123.474	≤ 1.08
6	Non-urban Plane	DOY, year, hour, area, tree height	51.77	90.75	≤ 1.08
7	6 ind. Maple datasets	DOY, year, hour	24.5 – 42.8	1.82 – 144	≤ 1.07
8	6 ind. Plane datasets	DOY, year, hour	41.5 – 88.7	2.77 – 144	≤ 1.07

4.2.3 Machine learning implementation

First, the outputs from the zero-growth models were controlled for incorrect values, namely negative values, or growth at the end (DOY > 304) or start (DOY < 60) of each year. Manual data quality control is required to detect incorrect growth values or artefacts in winter, as bark cell degradation is not taken into account in the zero-growth model (Zweifel et al. 2016). Matrices comprising the features Xi, and the growth labels y were built and rows with missing dendrometer data were excluded during model training process. Subsequently, X and y were split into training (80%) and test (20%) subsets using stratified sampling of the features ‘year’ and ‘hour’ to ensure that the distribution of data in the subsets is balanced. Afterwards, all features were normalized using z-transformation. The parameters of the z-transformation of the training subset were employed to normalize the test subset data and the input data in periods of missing values.

To find the best model, datasets #1-6 were fitted to eight supervised ML algorithms for regression problems and tested though repeated 10-fold cross validation (10 repeats) on the training subsets before hyperparameter tuning (see **Tab. S 4-1** for a list of the algorithms). To evaluate the performance of these ML regression models, the root mean squared error

(RMSE) and the adjusted R^2 were calculated. Significant differences in the performances of the algorithms were checked with the Friedmans test (Rainio et al. 2024). RMSE results between two models were analyzed using the Mann-Whitney-U test. The best performing algorithms were hyperparameter tuned using a 10-fold cross-validation and Bayesian Optimization Search (iterations= 70) (Bischl et al. 2023). Afterwards, tuned models were evaluated on the test subsets. The effects of the hyperparameter tuning and its necessity for gap filling were analyzed. Permutation feature importance (PFI) was used to analyze the features and their predictive power by calibrating and validating the model for each permutation (boot = 50). The PFI value is here defined as the mean difference between the original and the permuted R^2 of the model (Breiman 2001; Schwarz et al. 2024). A 'random' feature was included to provide a statistical baseline for random performance decline.

A process scheme was developed following the necessary steps and functions for gap filling dendrometer data (**Fig. 4-2**). The functions created for this project, along with detailed explanations and data to replicate the process, are available at GitHub (https://github.com/ESKuhl/DM_GF_XGB). The functions are based on the scikit-learn package (Pedregosa et al. 2011) and are intended to simplify the application. The initial stage of the process is the data preparation, during which the zero-growth model is applied (Zweifel et al. 2016; Haeni et al. 2020; Knüsel et al. 2021) and corrected for erroneous values in the winter months. A dataset associated with the growth data must be constructed with continuous values in the features 'DOY', 'year', and 'hour'. For the functions to work properly, a variable called 'Label' must be added, which includes the output of the zero-growth model (i.e., growth labels y) with data gaps. Once the data have been prepared, they are fitted to an ML algorithm in the second step using the function testxgb(), which returns the performance results and the model. If the model performance on the test subset indicates a high degree of accuracy, the gaps in the dendrometer data can be filled in the third step using the model.predict() function of scikit-learn.

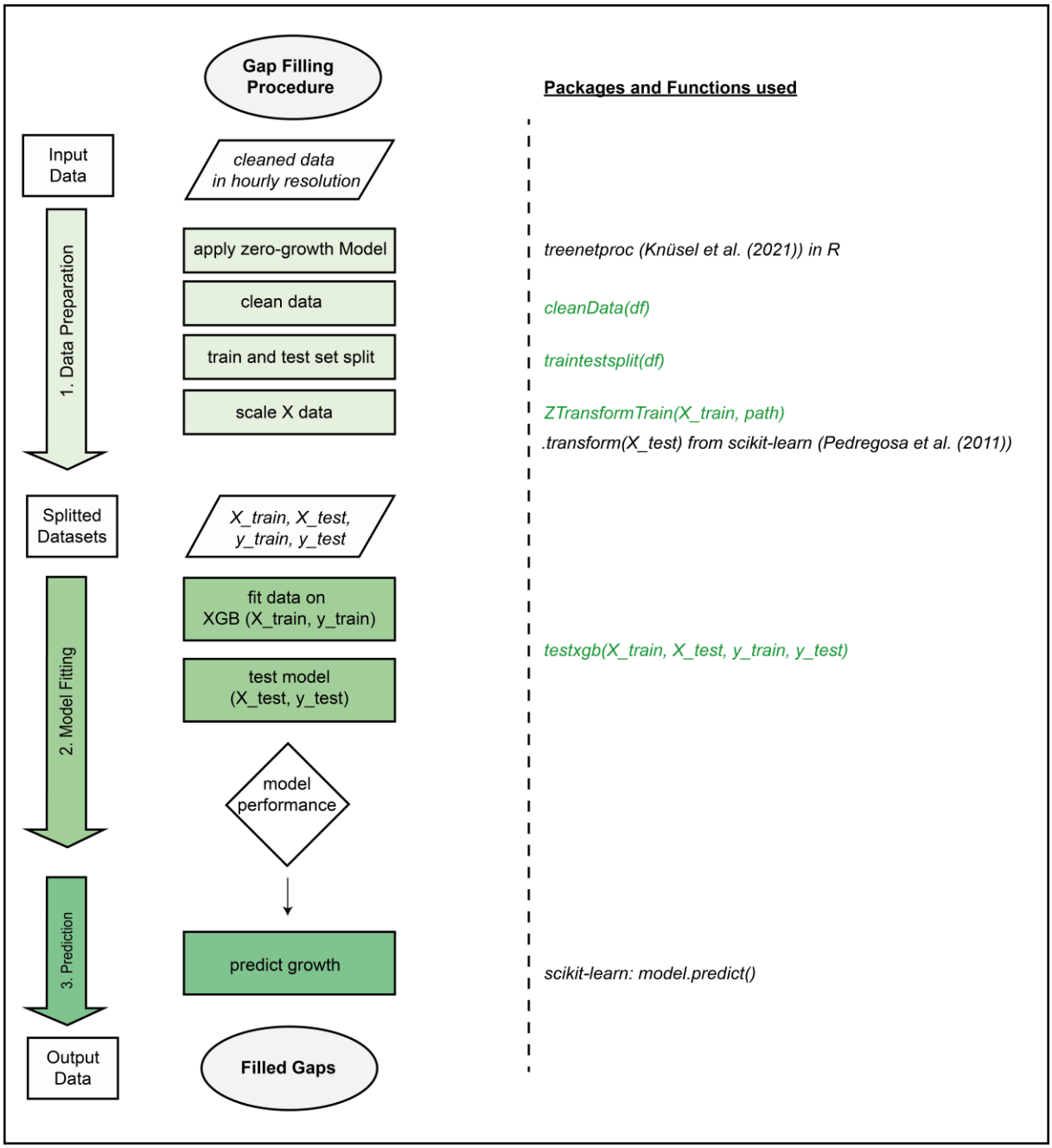


Fig. 4-2 Gap filling process scheme using Extreme Gradient Boosting to fill long data gaps and needed functions and packages. Green functions are implemented for this study and are available on GitHub

4.2.4 Gap filling model performance

To assess the functionality of this approach across different phases of the growing season, artificial gaps of 30 consecutive days were created on datasets #3-8 for the start (April 16th to May 15th), middle (June 1st to 30th) and end (September 1st to 30th) of the growing season for each year of observation (2019-2023). The resulting 15 subsets, comprising the three-period/ five-years combinations for each original dataset, were individually fitted to the ML algorithms and tested, yielding 240 models for comparison (**Fig. 4-2**). Moreover, to benchmark

the new approach against existing methods, network interpolation from Aryal et al. (2020) and spline interpolation were applied to the same artificial gaps, if the data allowed for it.

All model fittings, runs and analyses were conducted with Python 3.9.19 and the packages scikit-learn (Pedregosa et al. 2011), scikit-optimize (Head et al. 2018), and xgboost (Chen and Guestrin 2016). The analysis can be replicated in R by utilizing the package reticulate (Ushey et al. 2024), which enables the execution of Python functions in R studio when Python and the necessary packages are installed on the device (for further details see: https://rstudio.github.io/reticulate/).

4.3 Results and discussion

4.3.1 Algorithm selection and model evaluation

The eight algorithms significantly differed in their performances across all datasets (**Tab. S 4-2** for datasets #3-6). In all runs, nonlinear algorithms outperformed linear models and exhibited higher adjusted R^2 and lower RMSE mean values. The three best-performing untuned algorithms were identified as random forest (RF, Breiman 2001), extreme gradient boosting (XGB, Chen and Guestrin 2016) and k-nearest-neighbour (kNN, Fix and Hodges 1951). Although the linear models performed significantly worse relative to the nonlinear ones, RMSE values did not differ significantly between each other. We selected the best three algorithms (RF, XGB and kNN) for further hyperparameter tuning. Ridge regression (Hoerl and Kennard 2000) was additionally included as fourth algorithm to assess the impact of hyperparameter tuning on a linear model.

After hyperparameter tuning, the RMSE values for kNN and XGB showed either minimal improvement or no change (see **Tab. S 4-3** for datasets #3-6). The RMSE values for RF increased marginally after tuning from 0.00 to a maximum RMSE of 0.07. Ridge regression RMSE values ranged between 0.56-0.96 and could not be improved by tuning. Model performances before and after the tunings revealed no significant differences. The species-specific models showed similar results. The results of the test subsets of these models (**Tab. S 4-4**) demonstrated that the models exhibit comparable performances on the test subset data and on the training subset data after tuning. When the algorithms were fitted to individual tree data without hyperparameter tuning, the results showed similarly low RMSE values.

The computation time required for hyperparameter tuning is substantially growing with the dataset size. While the time to fit the data without hyperparameter tuning has a maximum of a few minutes, hyperparameter tuning can take from 20 minutes (individual small datasets, approx. 25,000 datapoints) to 8 hours (model #3-6, approx. 90,500 data points) and exceeds

24 hours when the species-specific datasets are tuned (180,000 data points). Due to the increased computation time, which did not significantly improve model performances, the suggested procedure (**Fig. 4-2**) did not include hyperparameter tuning.

Among the predictor variables included in the datasets, the day of the year ('DOY') was the most important feature of the grouped tree models (**Fig. 4-3**). The PFI values exceeded 1 for the 'DOY' variable in all grouped tree models, with the highest values observed in the plane tree models. This indicates that the permuted 'DOY' models are poorly fitting with negative R^2 values. In the maple tree-growth models, the feature 'area' achieved a higher ranking than in the plane-tree-based models. The feature 'year' also had a significant predictive value for all four algorithms. Despite this, an increase in model error was not significantly observed for the features 'tree height' or 'hour', as their PFI values were not exceeding the PFI values from the introduced random feature. Additional tests including various climatic parameters revealed no significant PFI values for these parameters, which is example wise shown for VPD in figure 3 and demonstrates that these variables are not essential for gap reconstruction using this ML approach.

All four algorithms were trained as individual and grouped tree models on the datasets with the consecutive gaps at the start, middle and end of the growing seasons. No hyperparameter tuning was applied. The average performance of these models on their test subsets highlights the superiority of the individual tree models (**Tab. S 4-5**). The comparison between the four algorithms clearly showed that in most of the cases, the decision-tree-based algorithms outperformed kNN and ridge regression. The performance of ridge regression and kNN was found to be significantly different to XGB and RF, respectively ($p < 0.01$), being similar in their performance ($p > 0.1$). All individual tree models performed on average similarly good than on their training data. The grouped tree models, on the other hand, had higher average RMSE values on the test subset data compared to the training subset data. This is an indicator for generalization errors in the grouped tree models. Although these models performed well on the test data with hyperparameter tuning (**Tab. S 4-4**), it is probable that the grouped tree models are dependent on the tuning to prevent overfitting on the training subset data. The diverse growth patterns of multiple trees might have forced the models to learn detailed patterns, which are characteristically for the training data, but are too specific for a general application to unseen data (Géron 2019).

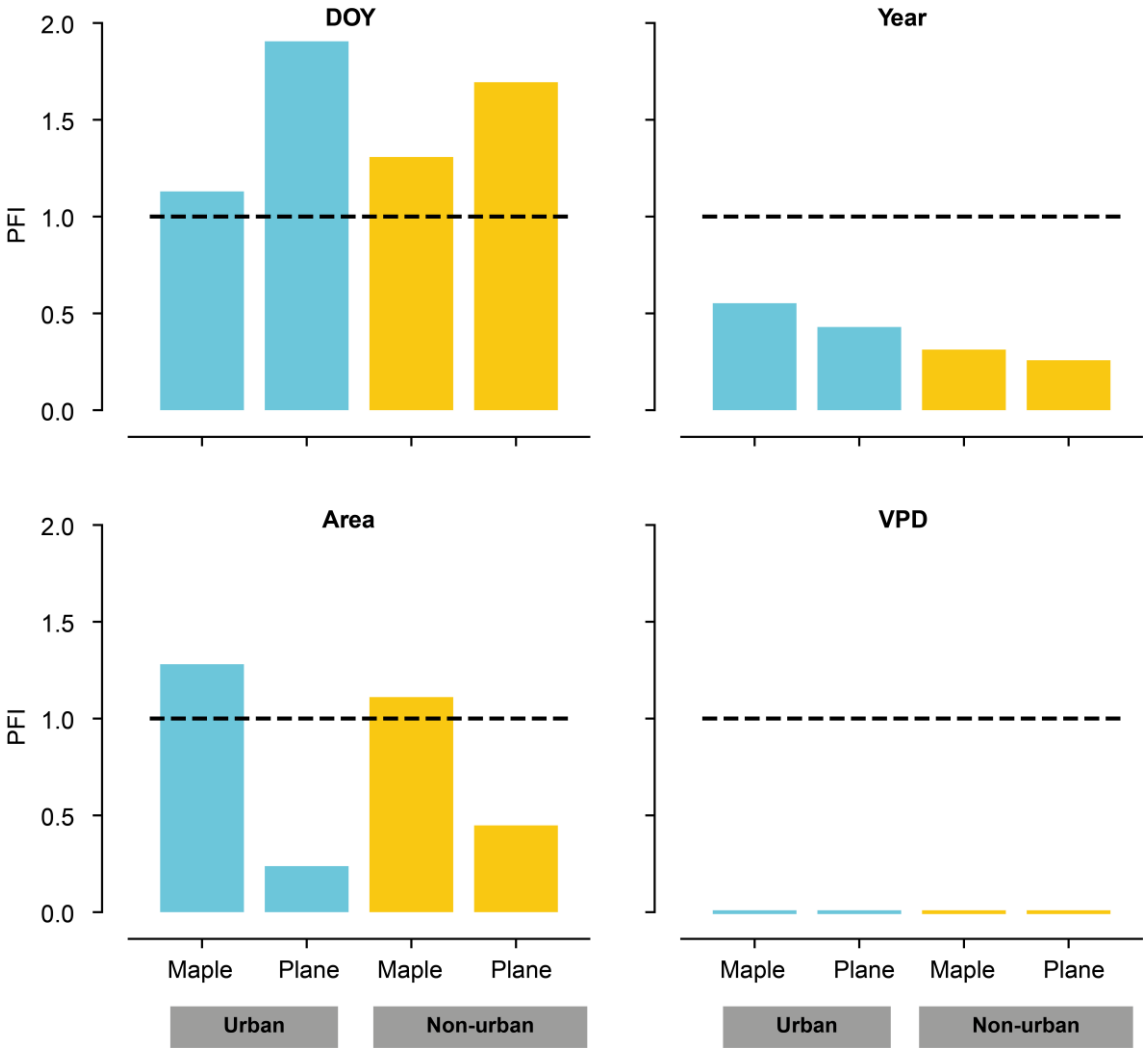


Fig. 4-3 Example of permutation feature importance (PFI) results for the tree most important features DOY (day of year), year and are of the models derived from the urban (blue) and non-urban (yellow) multi tree datasets #3-6 (**Tab. 4-1**). The figure contains an additional example of the PFI values from climatic features like the vapour pressure deficit (VPD), when these are included in the models.

As RF and XGB showed similar performance on the test subsets, a Mann-Whitney-U test validated that the RMSE values of the artificial gap predictions were not significant different between the algorithms ($p > 0.01$) (**Fig. S1**). These results indicate that XGB is the optimal algorithm for this gap-filling approach. The algorithm is known for its high scalability, short computation time, the ability to handle unbalanced datasets and its iterative learning process (Chen and Guestrin 2016; Fatima et al. 2023). Compared to randomly build decision trees in RF, XGB grounds on a gradient boosting algorithm (Friedman 2001) and iteratively builds an ensemble model of decision trees. The objective of each iteration is to minimize the loss function. The model optimization stops, when the performance of the training or the validation subset, a 20% subset from the training subset, no longer improves. Consequently, overfitting can be mitigated without the need for hyper-parameter tuning (Ying 2019).

4.3.2 Evaluation on artificial gaps

Both, individual and grouped tree models predicted data for the artificially created growth gaps to evaluate the performances on consecutive missing values. RMSE values for the various gaps were calculated using the excluded observations (**Figs. 4-4** and **4-5**). The values for the grouped tree models in **figures 4-4d)** and **4-5d)** were generally higher than for the individual tree models in **a-c)**. Despite the performances of the grouped models, gap filling in the middle of the growing season showed higher variance in RMSE values than at the start and middle for the individual tree models. Best predictions of the individual tree models were observed for the gaps at the end of the growing season. In this period, the performances in the grouped models were worse compared to all other periods. It is likely, that XGB was unable to predict growth with the same degree of accuracy when multiple trees with differing growth behaviors and diverse growing season lengths were given as training data. From the grouped tree models, the urban models had higher RMSE values than the non-urban models, which could indicate more coherent tree-growth patterns between non-urban trees. Urban environments have shown high heterogeneity of tree-growth influencing conditions between urban locations and compared to non-urban environments (Iakovoglou et al. 2001; Cedro and Nowak 2006; Moser-Reischl et al. 2019; Lv et al. 2024).

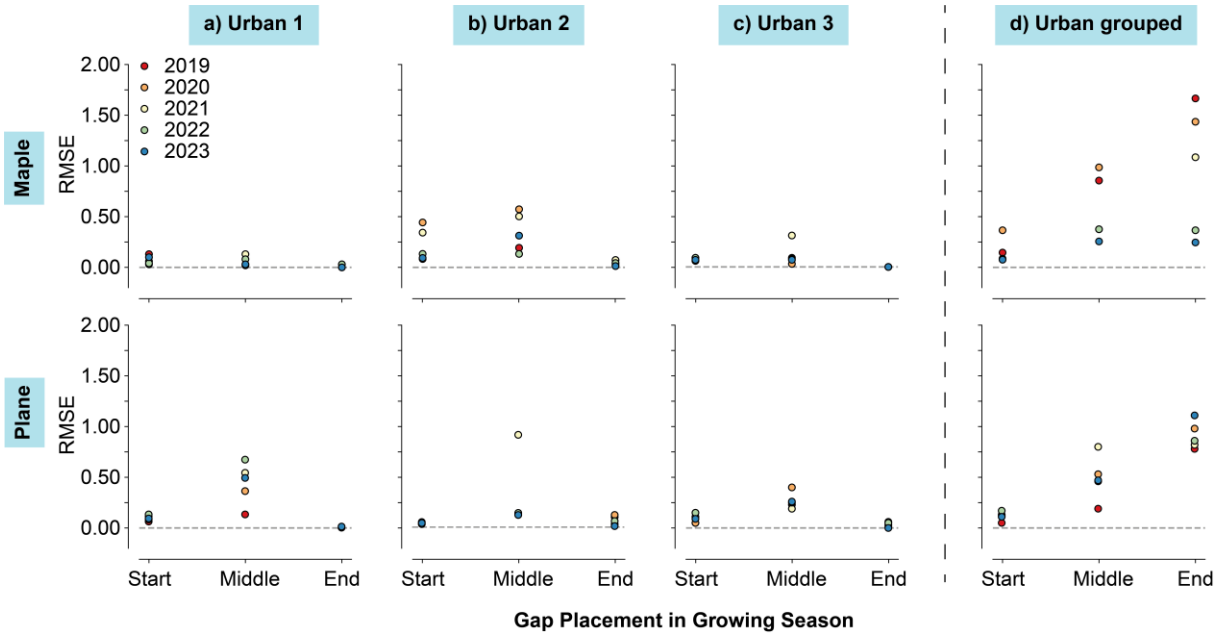


Fig. 4-4 RMSE results for the predictions on all different phases (Start, Middle and End of growing season) and all years (2019-2023) for the urban individual (a-c)) grouped tree models (d))

The distribution of RMSE values for each seasonal gap demonstrated that network interpolation exhibited similar performance to the individual tree models during the middle of

the growing season but had significantly higher RMSE values during the beginning of the growing season (**Fig. 4-6**). At the end of the growing season, the variance in RMSEs was considerably smaller for the network interpolation, although the number of gaps, where network interpolation could be applied, was limited (22.4%). For the start and middle of the season, it was possible to apply network interpolation to 16 and 26% of the artificially created gaps, respectively. Spline interpolation showed higher variance in RMSEs for the middle and end of the growing season. Significant differences ($p<0.01$) were found between the individual tree and spline approach for the start of the growing season. For all approaches, the mid of the growing season means were higher than for the start and end.

In **figure S 4-2**, the plane growth reconstruction for gaps at an urban example location for the year 2022 was highlighted in comparison to existing methods. The location and year were chosen as examples for the purpose of facilitating a more comprehensive comparison between methods, given that network interpolation could not be applied to all examples. Applied to the plane trees, the ML approach worked for all phases of the growing season with RMSE values ≤ 0.15 and the other approaches had higher RMSEs in most cases. The most striking visual observations were those when spline interpolation predicted negative growth for the start of the season in 2022 and when the network interpolation reconstructed reduced growth for mid of growing season with an RMSE of 0.55.

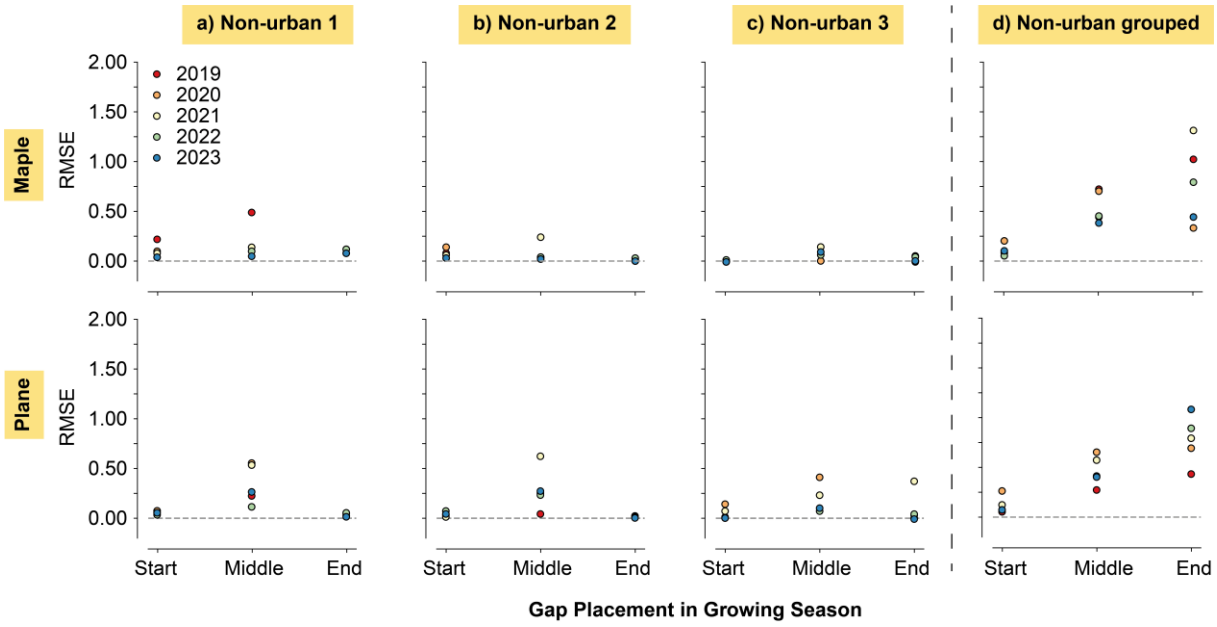


Fig. 4-5 RMSE results for the predictions on all different phases (Start, Middle and End of growing season) and all years (2019-2023) for the non-urban individual (a-c)) grouped tree models (d))

The predictions for the maple tree at the same urban location (**Fig. S3**) revealed some shortcomings of the ML approach. The network interpolation approach made the most

accurate predictions for the middle and end of the growing season. However, at the start of the growing season network interpolation did not capture the onset on the growing season, while the spline interpolation captured it best. Due to the way spline interpolation operates, this approach was unable to capture the characteristic stepwise growth in any gaps compared to the other approaches although RMSE values were generally low.

When the scheme is applied to true gaps of the datasets, the results clearly demonstrate the strength of the ML models to reconstruct long gaps in dendrometer data (**Fig. 4-7**). At visual inspection, the individual tree models were able to comprehensibly fill these gaps. It should be noted that the predictions for the non-urban plane tree (**Fig. 4-7 d)**) did not display reasonable growth values for the year 2019. The model predicted a reduction in growth values with increasing time. In such instances we recommend that these erroneous predictions should be deleted and alternative methods like spline interpolation should be considered for these values. Users are encouraged to test various method on their capability to reconstruct data gaps. We recommend that any method used should be accompanied by a visual control of the reconstructed data.

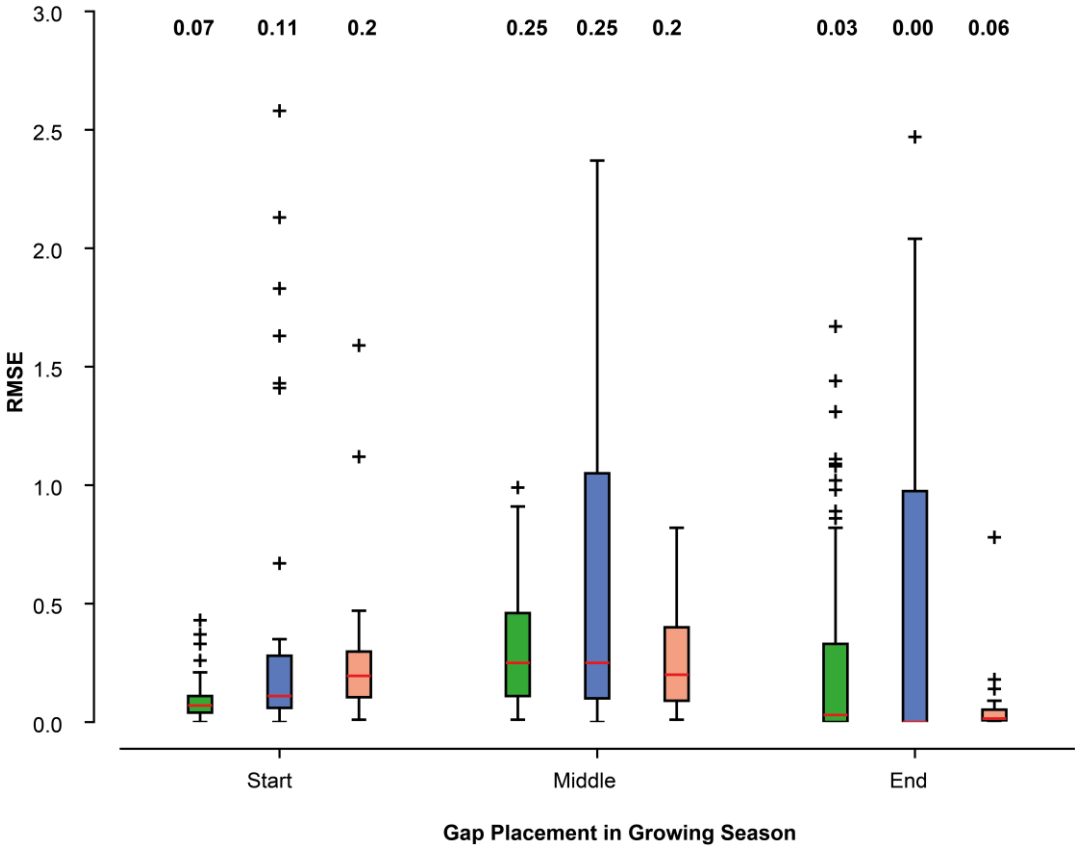


Fig. 4-6 Comparison of RMSE values of the artificial gaps, when the gaps are filed with the individual tree models (green), spline interpolation (blue) or network interpolation (orange). On each boxplot, the red bar indicates the median, bottom and top edges indicate the 25th and 75th percentiles; the whiskers extend to all data points except outliers (drawn as "+"). The median RMSE values are given at the top of the figure. Highly significant ($p > 0.01$, Mann-Whitney-U test) are the differences for the start of the growing season

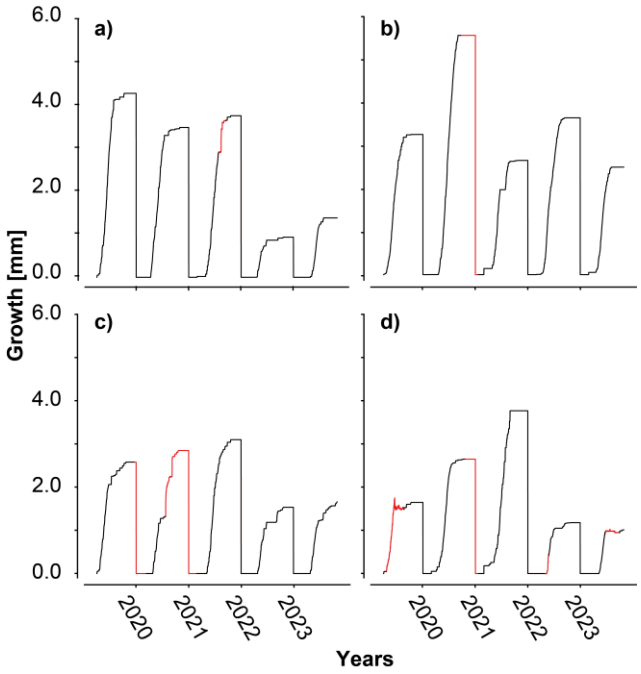


Fig. 4-7 Reconstructed growth in actual existing gaps in the dendrometer data of this study: a) urban maple, b) urban plane, c) non-urban maple and d) non-urban plane

The proposed ML method represents a complementary tool to existing methods as it enables the gap filling for individuals of a tree species. Consequently, the necessity for neighbouring trees with no gaps in the dendrometer data is negated. The findings indicate that individual tree models are capable of filling large data gaps. The application to actual gaps revealed that the models were capable of predicting reasonable values of growth even in gaps exceeding 30 days that span the transition between growth periods (start, middle, end) (**Fig. 4-7**). The low RMSE values (**Fig. 4-6**) suggest that the individual tree models can predict growth at the start and end of the growing season with high confidence. This allows, for example, a gapless analysis of the onset and cessation of tree growth in consecutive years. The ML method is not dependent on any climatic features. All features induced into the model are based on temporal variables and are hence available at any time. Features were controlled for multicollinearity before the application of any ML regression algorithm and dendrometer data has been carefully checked for quality. When the quality is poor, it is reflected in the generalization error and predictions made with these models are correspondingly unsatisfactory (Geiger et al. 2021; Briesch et al. 2022).

The application of this method, however, did not work for predicting raw dendrometer data and the models did not perform as good as the existing methods **(Fig. S 4-4**). This is likely due to cyclical fluctuations of raw data due to stem growth, shrinkage and extension being more complex than continuously increasing stem growth data. Still, the introduced method offers a key advantage in handling long data gaps, especially when methods relying on sufficient data are not suitable (i.e. due to a lack of data from neighbouring trees). Future work should be

undertaken to develop and test machine learning methods, which not only work for individual trees and independent from climatic variables but can also address positive and negative fluctuations in raw dendrometer data.

4.4 Conclusion

We here introduced a novel method based on ML to fill dendrometer data gaps of individual trees exceeding 24h. Unlike previous approaches, our method is not dependent on climatic features or additional tree-growth data from neighbouring trees. Furthermore, we provided a comprehensive scheme that enables the replication of the method for other dendrometer data. The results showed that XGB-based models are capable of reconstructing tree-growth derived from dendrometer data at the start, middle and end of the growing season. Nonetheless, the comparison between methods revealed that no single universal method exists to fill long data gaps perfectly. The process pf gap filling is complex and must be applied with caution. A visual evaluation of the predicted data is essential for any given case. Further research should explore the combination of methods and the functionality of the introduced ML approach on other tree species and environments.

4.5 Acknowledgements

We thank Markus Kochbeck, Benedikt Lang, Nora Jubelius, Daniel Thimm, Daniel Bonn and Sophie Spelsberg for their help with the logger and sensor monitoring as well as the data collection and pre-processing.

4.6 References

Aryal S, Häusser M, Grießinger J, Fan Z, Bräuning A (2020) “dendRoAnalyst”: A tool for processing and analysing dendrometer data. Dendrochronologia 64:125772. https://doi.org/10.1016/j.dendro.2020.125772

Barrao S, Serrano-Notivoli R, Cuadrat JM, Tejedor E, Saz Sánchez MA (2022) Characterization of the UHI in Zaragoza (Spain) using a quality-controlled hourly sensor-based urban climate network. Urban Climate 44:101207. https://doi.org/10.1016/j.uclim.2022.101207

Beck C, Straub A, Breitner S, Cyrys J, Philipp A, Rathmann J, Schneider A, Wolf K, Jacobeit J (2018a) Air temperature characteristics of local climate zones in the Augsburg urban area (Bavaria, southern Germany) under varying synoptic conditions. Urban Clim 25:152.166. https://doi.org/10.1016/j.uclim.2018.04.007

Beck HE, Zimmermann NE, McVicar TR, Vergopolan, N., Berg A, Wood EF (2018b) Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data 5:180214. https://doi.org/10.1038/sdata.2018.214

Bischl B, Richter J, Becker M, Binder M, Lang M, Pielok T, Coors S, Thomas J, Ullmann T, Boulesteix A-L, Deng D, Lindauer M (2023) Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. WIREs Data Mining Knowl Discov 13:1484. https://doi.org/10.1002/widm.1484

Breiman L (2001) Random Forests. Machine Learning 45:5–32. https://doi.org/10.1023/A:1010933404324

Briesch M, Sobania D, Rothlauf F (2022) The Randomness of Input Data Spaces is an A Priori Predictor for Generalization. In: Bergmann R, Malburg L, Rodermund SC, Timm IJ (eds) KI 2022: Advances in Artificial Intelligence. Springer International Publishing, Cham, pp 17–30

Cedro A, Nowak G (2006) Effects of climatic conditions on annual tree ring growth of the Platanus × hispanica 'Acerifolia' under urban conditions of Szczecin. Dendrobiology 55:11–17

Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, San Francisco California USA, pp 785–794

Corell M, Girón IF, Galindo A, Torrecillas A, Torres-Sánchez R, Pérez-Pastor A, Moreno F, Moriana A (2014) Using band dendrometers in irrigation scheduling: Influence of the location inside the tree and comparison with point dendrometer. Agric Water Manag 142:29–37. http://dx.doi.org/10.1016/j.agwat.2014.04.005

Deutscher Wetterdienst (2024) Mainz-Lerchenberg Station data. https://www.dwd.de/DE/leistungen/klimadatendeutschland/vielj_mittelwerte.html

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:027–046. https://doi.org/doi: 10.1111/j.1600-0587.2012.07348.x

Dulamsuren C, Coners H, Leuschner C, Hauck M (2023) Climatic control of high-resolution stem radius changes in a drought-limited southern boreal forest. Trees 37:797–810. https://doi.org/10.1007/s00468-022-02384-z

Eischeid JK, Baker CB, Karl TR, Diaz HF (1995) The quality control of long-term climatological data using objective data analysis. J Appl Meteorol 34(12):2787–2795. https://doi.org/10.1175/1520-0450(1995)034<2787:TQCOLT>2.0.CO;2

Fatima S, Hussain A, Amir SB, Ahmed SH, Aslam SMH (2023) XGBoost and Random Forest Algorithms: An in Depth Analysis. Pak J Sci Res 3(1):26–31. https://doi.org/10.57041/pjosr.v3i1.946

Fix E, Hodges JL (1951) Discriminatory analysis. nonparametric discrimination: consistency properties. USA Air Force School of Aviation Medicine, Randolph Field, Texas, USA

Friedman JH (2001) Greedy function approximation: A gradient boosting machine. Ann Statist 29(5):1189–1232. https://doi.org/10.1214/aos/1013203451

Geiger RS, Cope D, Ip J, Lotosh M, Shah A, Weng J, Tang R (2021) "Garbage In, Garbage Out" Revisited: What Do Machine Learning Application Papers Report About Human-Labeled Training Data? Quant Sci Stud 2(2). https://doi.org/doi.org/10.1162/qss_a_00144

Géron A (2019) Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd edn. O'Reilly, Sebastopol

Google Satellite Image (2024) Google Maps. In: AeroWest Airbus, CNES/Airbus, GeoBasis-DE/BKH, GeoContent, Landsat/Copernicus, Maxar Technologies, Map data: GeoBasis-DE/BKG, Google

Haeni M, Knüsel S, Wilhelm M, Peters RL, Zweifel R (2020) treenetproc - Clean, process and visualise dendrometer data. R package

Head T, MechCoder GL, Shcherbatyi I (2018) scikit-optimize/scikit-optimize: v0. 5.2. Version v0 5

Hoerl AE, Kennard RW (2000) Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 42(1):80–86. https://doi.org/10.1080/00401706.2000.10485983

Iakovoglou V, Thompson J, Burras L, Kipper R (2001) Factors related to tree growth across urban-rural gradients in the Midwest, USA. Urban Ecosyst 5:71–85. https://doi.org/10.1023/A:1021829702654

King G, Fonti P, Nievergelt D, Büntgen U, Frank D (2013) Climatic drivers of hourly to yearly tree radius variations along a 6°C natural warming gradient. Agric For Meteorol 168:36–46. http://dx.doi.org/10.1016/j.agrformet.2012.08.002

Knüsel S, Peters RL, Haeni M, Wilhelm M, Zweifel R (2021) Processing and Extraction of Seasonal Tree Physiological Parameters from Stem Radius Time Series. Forests 12:765. https://doi.org/10.3390/f12060765

Landeshauptstadt Mainz (2024) Jahresbericht Juli 2022- Juni 2023: Hauptamt Abteilung Öffentlichkeitsarbeit und Protokoll sowie städtische Ämter, Mainz

Lindén J, Fonti P, Esper J (2016) Temporal variations in microclimate cooling induced by urban trees in Mainz, Germany. Urban For Urban Green 20:198–209. https://doi.org/10.1016/j.ufug.2016.09.001

Lukovic M, Zweifel R, Thiry G, Zhang C, Schubert M (2022) Reconstructing radial stem size changes of trees with machine learning. J R Soc Interface 19:20220349. https://doi.org/10.1098/rsif.2022.0349

Lv H, Dermann A, Dermann F, Petridis Z, Köhler M, Saha S (2024) Comparable diameter resulted in larger leaf area and denser foliage in the park trees than in street trees: A study on Norway maples of Karlsruhe city, Germany. Heliyon 10:e23647. https://doi.org/10.1016/j.heliyon.2023.e23647

Moser-Reischl A, Rahman MA, Pauleit S, Pretsch H, Rötzer T (2019) Growth patterns and effects of urban micro-climate on two physiologically contrasting urban tree species. Landsc Urban Plan 183:88–99. https://doi.org/10.1016/j.landurbplan.2018.11.004

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12:2825–2830. https://doi.org/10.48550/arXiv.1201.0490

R Core Team (2021) R: A language and environment for statistical computing. Vienna, Austria

Rainio O, Teuho J, Klén R (2024) Evaluation metrics and statistical tests for machine learning. Sci Rep 14:6086. https://doi.org/10.1038/s41598-024-56706-x

Salomón RL, Peters RL, Zweifel R, Sass-Klaassen UGW, Stegehuis AI, Smiljanic M, Poyatos R, Babst F, Cienciala E, Fonti P, Lerink BJW, Lindner M, Martinez-Vilalta J, Mencuccini M, Nabuurs G-J, van der Maaten E, von Arx G, Bär A, Akhmetzyanov L, Balanzategui D, Bellan M, Bendix J, Berveiller D, Blaženec M, Čada V, Carraro V, Cecchini S, Chan T, Conedera M, Delpierre N, Delzon S, Ditmarová Ľ, Dolezal J, Dufrêne E, Edvardsson J, Ehekircher S, Forner A, Frouz J, Ganthaler A, Gryc V, Güney A, Heinrich I, Hentschel R, Janda P, Ježík M, Kahle H-P, Knüsel S, Krejza J, Kuberski Ł, Kučera J, Lebourgeois F, Mikoláš M, Matula R, Mayr S, Oberhuber W, Obojes N, Osborne B, Paljakka T, Plichta R, Rabbel I, Rathgeber CBK, Salmon Y, Saunders M, Scharnweber T, Sitková Z, Stangler DF, Stereńczak K, Stojanović M, Střelcová K, Světlík J, Svoboda M, Tobin B, Trotsiuk V, Urban J, Valladares F, Vavrčík H, Vejpustková M, Walthert L, Wilmking M, Zin E, Zou J, Steppe K (2022) The 2018 European heatwave led to stem dehydration but not to consistent growth reductions in forests. Nature Communications 13(1):28. https://doi.org/10.1038/s41467-021-27579-9

Schwarz L, Sobania D, Rothlauf F (2024) On relevant features for the recurrence prediction of urothelial carcinoma of the bladder. Int J of Med Inform 186:105414. https://doi.org/10.1016/j.ijmedinf.2024.105414

Ushey K, Allaire JJ, Tang Y (2024) reticulate: Interface to “Python”

Ying X (2019) An Overview of Overfitting and its Solutions. J Phys: Conf Ser 1168:022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Zhang Y, Gao L, Deng Y, Huang Q, Yuan Y, Shi X (2024) Seasonal aridity regulates drivers and temporal variability of wood phenology: A meta-analysis of dendrometer monitoring data across the Northern Hemisphere. Dendrochronologia 85:126201. https://doi.org/10.1016/j.dendro.2024.126201

Ziaco E, Biondi F (2018) Stem Circadian Phenology of Four Pine Species in Naturally Contrasting Climates from Sky-Island Forests of the Western USA. Forests 9(7):396. https://doi.org/10.3390/f9070396

Zweifel R, Haeni M, Buchmann N, Eugster W (2016) Are trees able to grow in periods of stem shrinkage? New Phytol 211:839–849. https://doi.org/10.1111/nph.13995

4.7 Supplementary material

Tab. S 4-1 Regression algorithms tested in this approach

#	Regression Algorithm	Type
1	Linear	Linear
2	Ridge	Linear
3	Lasso	Linear
4	Random Forest (RF)	Nonlinear
5	k-Nearest-Neighbor (kNN)	Nonlinear
6	Support Vector Machine (SVM)	Linear
7	Extreme Gradient Boosting (XGB)	Nonlinear
8	Partial Least Square (PLS)	Linear

Tab. S 4-2 Performance results (RMSE and adjusted R^2) of eight algorithms trained on datasets #3-6 (Tabl. 1) to predicting growth. Performance measures were estimated using repeated k-fold cross validation (k = 10, repeats = 10). Performance values shown here equal the mean of the 100 runs

Urban Maple			**Urban Plane**			**Non-urban Maple**			**Non-urban Plane**		
Model	**RMSE**	**adjR²**	**Model**	**RMSE**	**adjR²**	**Model**	**RMSE**	**adjR²**	**Model**	**RMSE**	**adjR²**
RF	0.00	1.00	RF	0.00	1.00	RF	0.00	1.00	RF	0.00	1.00
XGB	0.01	1.00	XGB	0.02	1.00	XGB	0.01	1.00	XGB	0.01	1.00
kNN	0.10	0.99	kNN	0.13	0.99	kNN	0.04	1.00	kNN	0.06	1.00
SVM	0.49	0.8	SVM	0.92	0.71	SVM	0.33	0.82	SVM	0.51	0.82
Ridge	0.74	0.54	Ridge	0.96	0.69	Ridge	0.47	0.63	Ridge	0.64	0.73
Linear	0.74	0.54	Linear	0.96	0.69	Linear	0.47	0.63	Linear	0.64	0.73
Lasso	0.74	0.54	Lasso	0.96	0.69	Lasso	0.47	0.63	Lasso	0.64	0.73
PLS	0.74	0.54	PLS	0.96	0.69	PLS	0.52	0.56	PLS	0.64	0.72

Tab. S 4-3 Ranked performance results from hyperparameter tuning to fit different algorithms to the datasets #3-6 (Tabl. 1). Best RMSE scores for all four datasets are shown in reference to growth in millimetres. Param. denotes to the best hyperparameters found via Bayesian Optimization Search and 10-fold cross validation

Position	Urbane Maple			Urban Plane			Non-urban Maple			Non-urban Plane		
	Model	**RMSE**	**Param.**	**Model**	**RMSE**	**Param.**	**Model**	**RMSE**	**Param.**	**Model**	**RMSE**	**Param.**
1	**XGB**	0.01	colsample_bytree= 1, eta= 0.16, gamma= 10, max_depth= 15, min_child_weight= 7, subsample= 0.9, early_stopping_rounds= 10	**XGB**	0.01	colsample_bytree= 1, eta= 0.1, gamma= 30, max_depth= 13, min_child_weight= 8, subsample= 0.9, early_stopping_rounds= 10	**XGB**	0.00	colsample_bytree= 1, eta= 0.16, gamma= 50, max_depth= 13, min_child_weight= 6, subsample= 0.9, early_stopping_rounds= 10	**XGB**	0.01	colsample_bytree= 1, eta= 0.1, gamma= 30, max_depth= 15, min_child_weight= 3, subsample= 0.6, early_stopping_rounds= 10
2	**RF**	0.02	bootstrap= True, max_depth= 86, max_features= 1.0, max_leaf_nodes= 95, n_estimators= 116, random_state= 42	**RF**	0.07	bootstrap= True, max_depth= 66, max_features= 1.0, max_leaf_nodes= 95, n_estimators= 186, random_state= 42	**RF**	0.02	bootstrap= True, max_depth= 75, max_features= 1.0, max_leaf_nodes= 95, n_estimators= 170, random_state= 42	**RF**	0.05	bootstrap= True, max_depth= 46, max_features= 1.0, max_leaf_nodes= 98, n_estimators= 196, random_state= 42
3	**kNN**	0.08	algorithm= 'ball_tree', leaf_size= 31, metric= 'manhattan', n_neighbors= 5, p= 2, weights= 'distance'	**kNN**	0.11	algorithm= 'brute', leaf_size= 11, metric= 'manhattan', n_neighbors= 5, p= 1, weights= 'distance'	**kNN**	0.04	algorithm= 'kd_tree', leaf_size= 66, metric= 'euclidean', n_neighbors= 5, p= 2, weights= 'distance'	**kNN**	0.06	algorithm= 'brute', leaf_size= 11, metric= 'euclidean', n_neighbors= 7, p= 2, weights= 'distance'
4	**Ridge**	0.74	alpha= 1.0	**Ridge**	0.96	alpha= 1.0	**Ridge**	0.47	alpha= 0.4	**Ridge**	0.64	alpha= 1.0

Tab. S 4-4 Test set RMSE values of the models based on the datasets #1-6 and the four algorithms extreme gradient boosting (XGB), k-nearest neighbor (kNN), random forest (RF) and ridge regression after hyperparameter tuning

#	X	XGB	RF	kNN	Ridge
1	Maple	0.01	0.05	0.07	0.78
2	Plane	0.01	0.10	0.10	0.94
3	Urban Maple	0.01	0.02	0.09	0.74
4	Urban Plane	0.01	0.07	0.12	0.95
5	Non-Urban Maple	0.00	0.02	0.04	0.47
6	Non-Urban Plane	0.01	0.05	0.06	0.64

Tab. S 4-5 Test set RMSE mean values and standard deviation (in brackets) of the models based on the datasets #3-8 and the four algorithms extreme gradient boosting (XGB), k-nearest neighbor (kNN), random forest (RF) and ridge regression without hyperparameter tuning and VPD

#	X	XGB	RF	kNN	Ridge
3	Urban Maple	1.03 (0.01)	1.36 (0.01)	1.15 (0.01)	1.14 (0.01)
4	Urban Plane	0.69 (0.01)	0.92 (0.01)	0.78 (0.01)	1.55 (0.01)
5	Non-Urban Maple	0.77 (0.01)	1.00 (0.01)	0.88 (0.01)	1.02 (0.01)
6	Non-Urban Plane	0.61 (0.01)	0.80 (0.01)	0.70 (0.01)	1.05 (0.00)
7	Urban 1 Maple	0.02 (0.01)	0.00 (0.00)	0.03 (0.00)	0.27 (0.00)
8	Urban 1 Plane	0.07 (0.08)	0.00 (0.00)	0.18 (0.01)	1.20 (0.01)
9	Urban 2 Maple	0.13 (0.10)	0.00 (0.00)	0.19 (0.01)	1.29 (0.01)
10	Urban 2 Plane	0.01 (0.00)	0.00 (0.00)	0.25 (0.01)	2.00 (0.01)
11	Urban 3 Maple	0.01 (0.00)	0.00 (0.00)	0.06 (0.01)	0.40 (0.00)
12	Urban 3 Plane	0.04 (0.07)	0.00 (0.00)	0.16 (0.01)	1.25 (0.02)
13	Non-urban 1 Maple	0.05 (0.09)	0.00 (0.00)	0.10 (0.00)	0.87 (0.01)
14	Non-urban 1 Plane	0.11 (0.09)	0.00 (0.00)	0.17 (0.00)	1.16 (0.01)
15	Non-urban 2 Maple	0.03 (0.03)	0.00 (0.00)	0.05 (0.00)	0.36 (0.00)
16	Non-urban 2 Plane	0.06 (0.05)	0.00 (0.00)	0.10 (0.00)	0.66 (0.01)
17	Non-urban 3 Maple	0.01 (0.01)	0.00 (0.00)	0.03 (0.00)	0.25 (0.00)
18	Non-urban 3 Plane	0.01 (0.00)	0.00 (0.00)	0.14 (0.01)	1.16 (0.02)

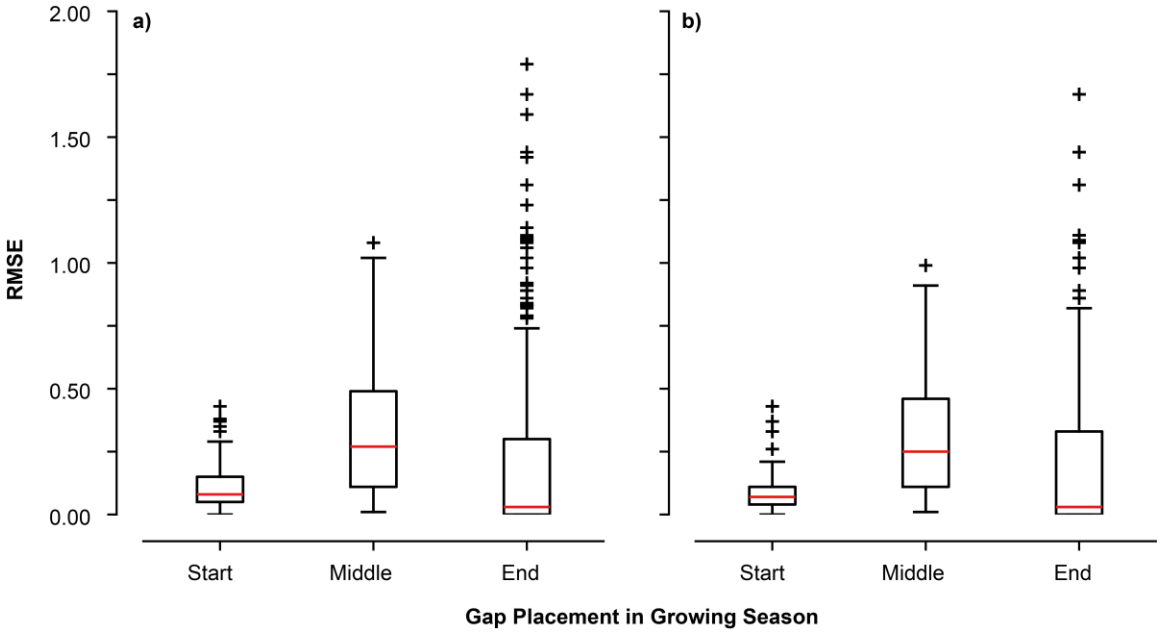


Fig. S 4-1 Comparison of RMSE values of the artificial gaps, when the models are based on a) random forest and b) extreme gradient boosting. On each boxplot, the red bar indicates the median, bottom and top edges indicate the 25th and 75th percentiles; the whiskers extend to all data points except outliers (drawn as "+"). No significant differences could be found for the middle and end of the growing seasons ($p > 0.01$, Mann-Whitney-U test)

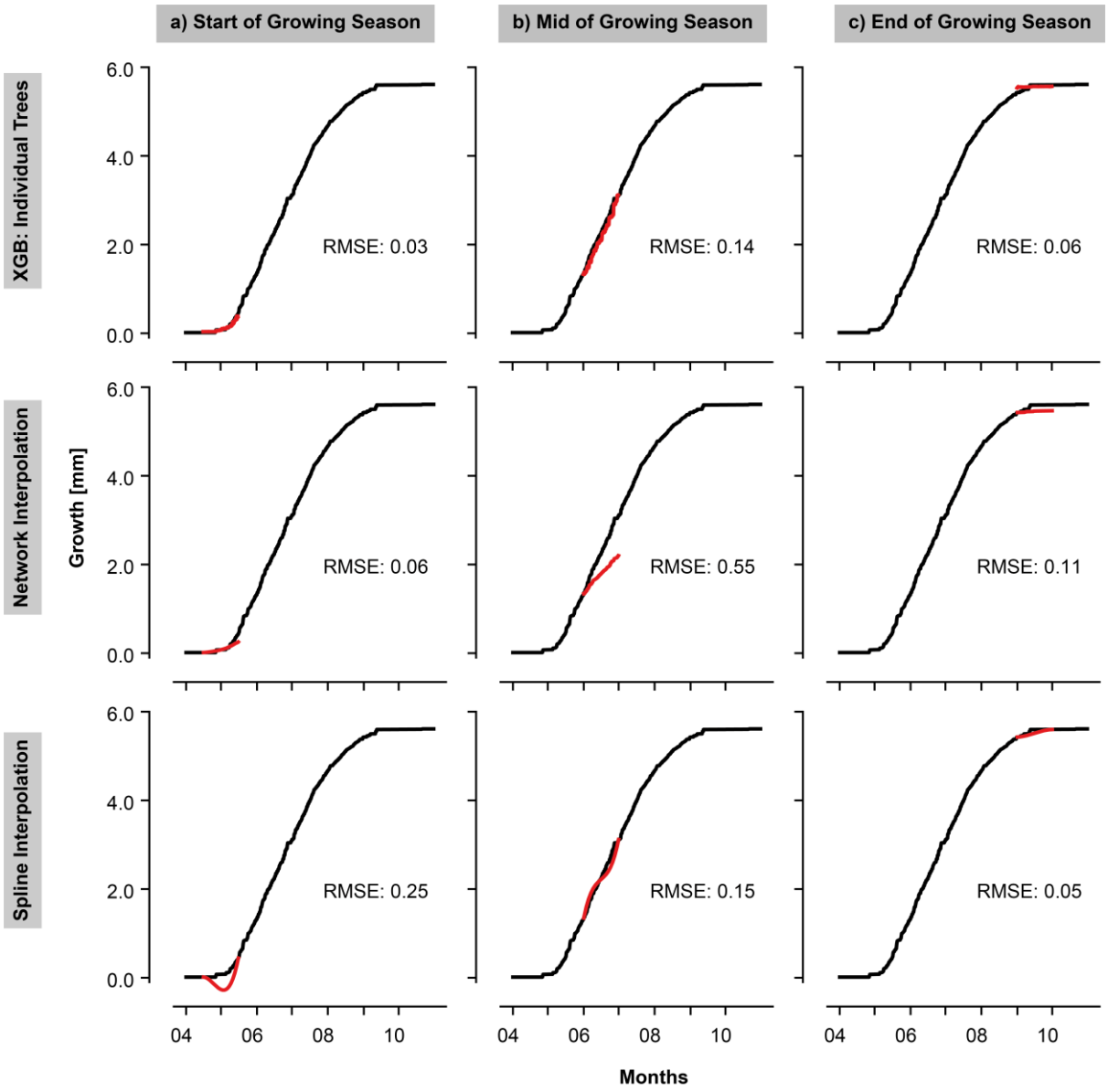


Fig. S 4-2 Results of different approaches for predicting growth of the plane tree at an urban location of 2022 CE fitted to the algorithm Extreme Gradient Boosting (XGB) for different seasons (a-c)). Top row shows the results of the individual Urban Plane models. Middle and bottom row present the comparison to the network interpolation by Aryal et al. (2020) and the classic spline interpolation

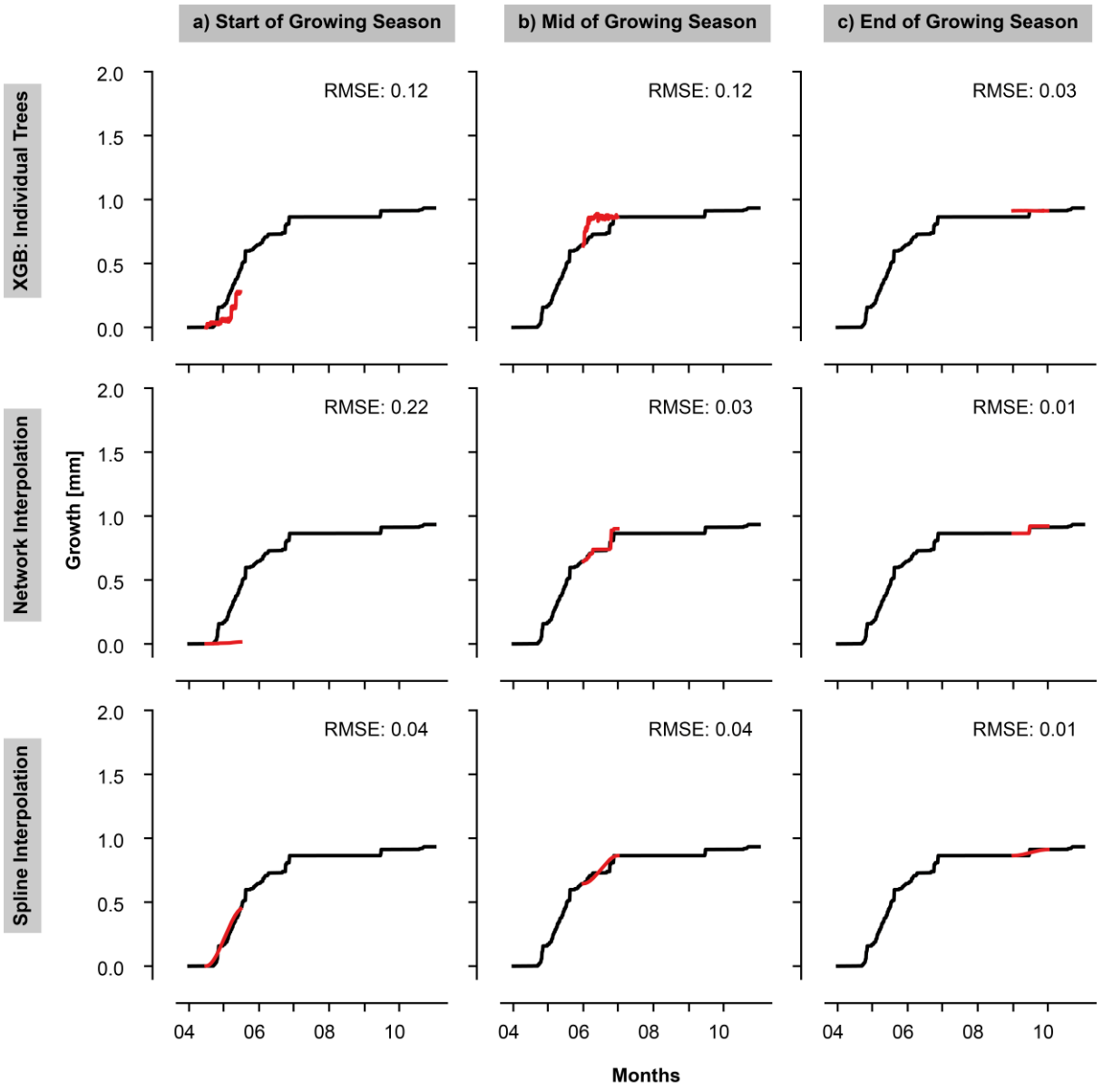


Fig. S 4-3 Results of different approaches for predicting growth of the maple tree at an urban location in 2022 CE fitted to the algorithm Extreme Gradient Boosting (XGB) for different seasons (a-c)). Top row shows the results of the individual Urban Maple models. Middle and bottom row present the comparison to the network interpolation by Aryal et al. (2020) and the classic spline interpolation

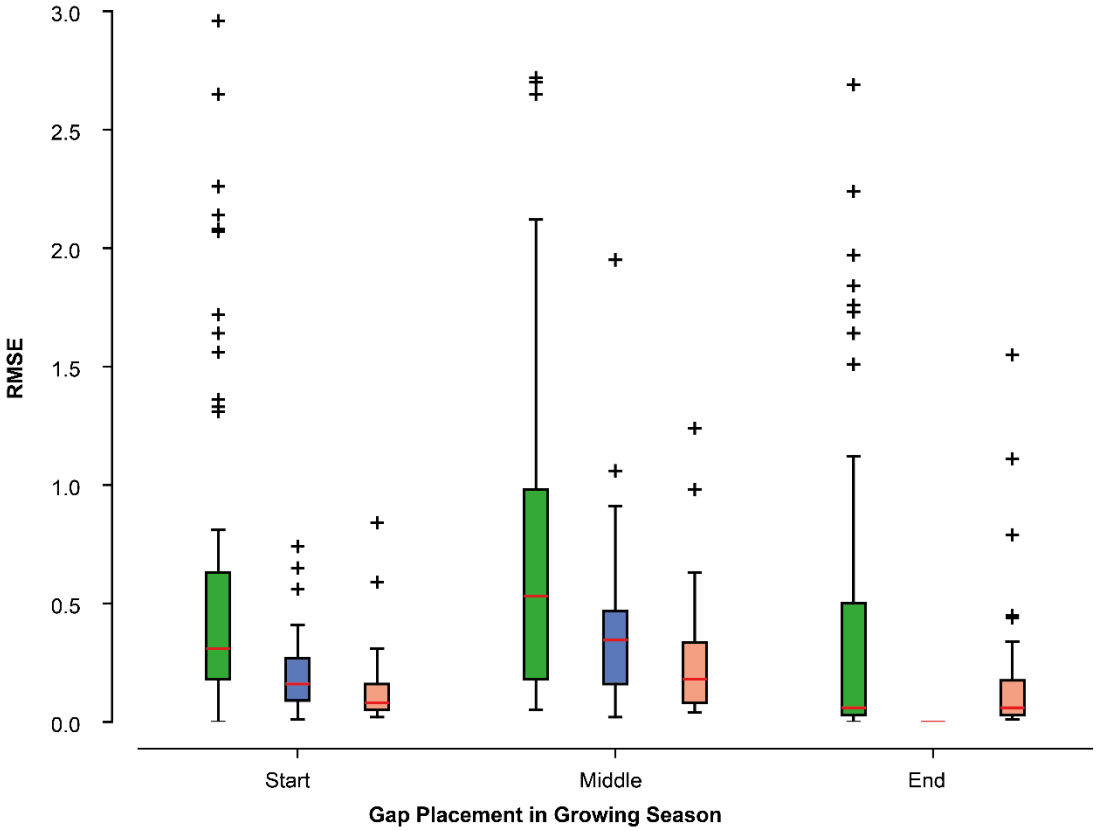


Fig. S 4-4 Comparison of RMSE values of the artificial gaps from raw dendrometer data, when the gaps are filed with the individual tree models (green), spline interpolation (blue) or network interpolation (orange). On each boxplot, the red bar indicates the median, bottom and top edges indicate the 25th and 75th percentiles; the whiskers extend to all data points except outliers (drawn as "+").

5 Conclusions and perspectives

This dissertation aims to advance tools for assessing climate-growth relationships using ML algorithms, highlighting the benefits and limitations in two major applications in dendroclimatology. It demonstrates how ML methods can improve the development of millennium-length temperature reconstructions and enhance our understanding of seasonal tree growth using dendrometer data. Analyzing the relationship between tree- and forest-growth responses to climate variations enable scientists to refine their understanding of the complex spatial and temporal dynamics of the Earth's system. Beside ML approaches, numerous techniques have been devised to analyse past climate variability and the different ways in which these can be used and combined can influence the results of a study. For example, Büntgen et al. (2021) highlighted how different techniques throughout the reconstruction process from series selection and detrending to the proxy calibration can influence a reconstruction's outcome. While uncertainties in paleoclimate modelling can never be entirely eliminated due to the proxies being indirect indicators of climatic changes, a reduction of uncertainty is feasible through continual refinement, development, and testing of methods and approaches.

In this thesis, a method to find the provenance of historical timbers was established and then applied to develop a new millennium-length summer temperature reconstruction using larch trees from the European Alps. Additionally, a new approach to fill gaps in dendrometer data was established and extended by a process scheme to apply the new method with a few, simple steps. The findings of these study not only allow for an evaluation of ML as methodological prospect, but also enable a revision of past climate variability in the European Alps.

First, an XGB-based classification model successfully predicted the origin of historical larch trees, particularly when incorporating wood density parameters data along with traditional TRW data (f1 score = 0.8). A total of 99 historical samples was successfully sorted to six different elevations, but the model performance was significantly lower, when the model was only trained with TRW data. This ML approach improves temperature reconstructions and advances dendroarchaeological studies by reducing dependence on shared temporal periods when trained on width and density parameters.

In the second step, the XGB algorithm is applied to tree-ring series from the Matter valley. Here, the thesis identifies data availability and quality as critical factors for ML model performance. The provenance model requires adequate data from the specific original locations. A model trained on data from the Simplon valley cannot accurately predict the origins

of historical trees from the Matter valley, and vice versa. Moreover, the Matter valley dataset includes only two elevational transect steps compared to six in the Simplon valley, which enhances the likelihood of accurately classifying tree-ring series using the Simplon model versus the Zermatt model. Additionally, the data quality significantly impacts model performance. Erroneous, noisy, or inadequately prepared input data generally lead to increased generalization errors and diminished model performance (Geiger et al. 2021; Briesch et al. 2022).

Based on the provenance method, the development of the MXD reconstruction enabled a revision of temperature variability in the European Alps. Techniques to mitigate signals from LBM defoliation and variance stabilization refined the results of the reconstruction. While low-frequency trends were muted, the reconstruction exhibits robustness in analysing high-frequency variability, especially volcanic cooling events. These findings underscore the methodological influence on outcomes (Büntgen et al. 2021). Previous studies have highlighted potential risks of losing low-frequency signals due to methodological factors such as standardization biases, as well as changes in site ecology or climate-growth relationships (Briffa et al. 1996). Regional curve standardization (Briffa et al. 1992; Esper et al. 2003) has shown to mitigate the loss of low-frequency information compared to traditional standardization approaches (Esper 2002; Bunn et al. 2004; Peters et al. 2015). However, the potential loss of low frequency due to changing climate-growth relationships remains unresolved. Studies have found that instrumental and proxy data can diverge with rising temperatures, biasing climate reconstructions (D'Arrigo et al. 2008). This thesis supports these findings by identifying signs of a developing divergence in the European Alps.

The new XGB-based gap-filling approach is able to fill gaps of up to 30 consecutive days of missing values. It has its advantage in being independent from neighbouring trees or any climatic input parameters and performs with high grades of accuracy, especially when predicting data for gaps at the start and end of the growing season. Still, the XGB models do not consistently outperform network or spline interpolation. The findings indicate that there is no single universally effective method for all scenarios yet and emphasize the need for visual evaluation of any gap filling results. Further limitations arise in predicting future growth under shifting climatic conditions, as the algorithms fail to establish models with significant predictive power of critical climatic features. Initially, we attempted to incorporate additional climatic variables while excluding non-climatic variables like DOY to increase the predictive power of the former and forecast future tree growth. This, however, led to severe overfitting on the

training data and poor generalization on the test data. Consequently, forward modelling using this approach is not feasible.

The findings of this thesis demonstrate the significant potential of applying ML algorithms in dendroclimatology to enhance provenance methods, refine temperature reconstructions and thereby improve our understanding of past climate variability. However, these results also pose challenges and thus opportunities for further testing of the introduced approaches. For instance, future research should explore the combination and performance of different tree-growth proxies such as blue intensity (BI), SI, or wood anatomical features with ML algorithms. Especially BI could lower the hurdle for the provenance method. BI, which assesses light reflectance from scanned images of tree-ring series to estimate density values, offers a promising alternative to the time- and cost-intensive wood densitometry measurements (Björklund et al. 2019). If replacing features of the provenance model with BI-based parameters proves equally effective, it could broaden the accessibility of this method to more researchers. A recently introduced method called Binary Surface Intensity (Rydval et al. 2024) could add to less cost- and labour-intensive methods to measure tree-growth parameters.

Additional investigation is also necessary to address questions about low-frequency trends in reconstructions and why they are occasionally absent. Nonlinear ML algorithms present an opportunity to mitigate these biases in climate reconstructions (Zhang et al. 2000; Gholami et al. 2017) and warrant further exploration. For example, Jevšenak et al. (2018) considered non-linear algorithms and a multi-proxy input reconstruct mean temperatures for April-May and June-August.

To successfully establish ML as integral part of dendroclimatology, access to the application via code sharing and data availability must be ensured. Studies, which utilize ML should usually follow standard procedures to find the best model for their purpose, but research showed that a large number of publications have not gone through the individual steps in their entirety (Zhu et al. 2023). The process of testing and tuning models can be computationally consuming and often, the user must not necessarily go through all these steps, when the extensive tests have been carried out beforehand. It is therefore important to enable the reproduction of the research by developing easy to follow schemes and providing code, which can be made public on platforms like GitHub. As many ML packages are produced and optimized for Python but the dendroclimatology community is mainly working with R packages, an interface between Python and R based codes should be available for the users. Existing R packages for dendroclimatology should be made available for Python to allow for an uncomplicated

integration of ML in basic tree-ring analyses. For example, the package ‘dplR’ has recently been translated to Python and has been published as ‘dplPy’ (Bunn 2010).

Future studies can highly benefit from the use of ML in assessing non-linear climate-growth relationships to reconstruct past climate variability or to predict future growth of trees under various climate scenarios. New ML algorithms should be tested for their potential to reconstruct climate in different regions and from different species. Additionally, the potential to predict future tree growth with should be further explored, which is especially of interest for urban planners and foresters. The number of publications working with ML to address these topics is increasing and shows that ML is no longer ‘a cutting edge’ in dendroclimatology. While Popa et al. (2024) used ML to explore the non-linear relationship between climate and growth of Norway spruce, Keret et al. (2024) developed a ML approach to detect xylem cell types, which is easily applicable though a software and open access. Shindo et al. (2024) investigated the use of ML to improve predictions of cambial age, based on the number of sapwood rings for European larch. The variability and response of Scots pine growth in the Kazakh steppe to changes in climate was analysed using RF and linear mixed effect models by Zhao et al. (2024). Cho et al. (2024) finds that integrating ML algorithms into existing process-based biogeochemical models enables to monitor and simulate tree mortality caused by droughts. The introduced approaches of this thesis integrate well in these upcoming publications and not only explore benefits and challenges occurring from ML application but also help the community to make steps towards improved assessments of climate and tree-growth relationships.

References

Akhmetzyanov L, Sánchez-Salguero R, García-González I, Buras A, Dominguez-Delmás M, Mohren F, den Ouden J, Sass-Klaassen U (2020) Towards a new approach for dendroprovenancing pines in the Mediterranean Iberian Peninsula. Dendrochronologia 60:125688. https://doi.org/10.1016/j.dendro.2020.125688

Anchukaitis KJ (2017) Tree Rings Reveal Climate Change Past, Present, and Future. Proc Am Philos Soc 161(3):244–263. https://www.jstor.org/stable/45211559

Aryal S, Grießinger J, Dyola N, Gaire NP, Bhattarai T, Bräuning A (2023) INTRAGRO: A machine learning approach to predict future growth of trees under climate change. Ecol Evol 13:e10626. https://doi.org/DOI: 10.1002/ece3.10626

Aryal S, Häusser M, Grießinger J, Fan Z, Bräuning A (2020) “dendRoAnalyst”: A tool for processing and analysing dendrometer data. Dendrochronologia 64:125772. https://doi.org/10.1016/j.dendro.2020.125772

Babst F, Poulter B, Trouet V, Tan K, Neuwirth B, Wilson RJS, Carrer M, Grabner M, Tegel W, Levanic T, Panayotov M, Urbinati C, Bouriaud O, Ciais P, Frank D (2013) Site- and species-specific responses of forest growth to climate across the European continent: Climate sensitivity of forest growth across Europe. GEB 22(6):706–717. https://doi.org/10.1111/geb.12023

Björklund J, von Arx G, Nievergelt D, Wilson RJS, Van den Bulcke J, Günther B, Loader NJ, Rydval M, Fonti P, Scharnweber T, Andreu-Hayles L, Büntgen U, D’Arrigo R, Davi N, De Mil T, Esper J, Gärtner H, Geary J, Gunnarson BE, Hartl C, Hevia A, Song H, Janecka K, Kaczka RJ, Kirdyanov AV, Kochbeck M, Liu Y, Meko M, Mundo I, Nicolussi K, Oelkers R, Pichler T, Sánchez-Salguero R, Schneider L, Schweingruber F, Timonen M, Trouet V, Van Acker J, Verstege A, Villalba R, Wilmking M, Frank D (2019) Scientific Merits and Analytical Challenges of Tree-Ring Densitometry. Rev Geophys 57(4). https://doi.org/10.1029/2019RG000642

Bonde N, Tyers I, Wazny T (1997) Where Does The Timer Come From? Dendrochronological Evidence of the Timber Trade in Northern Europe. In: Sinclair A, Slater E, Gowlett J (eds) Archaeological Sciences 1995. Proceedings of a conference on the application of scientific techiques to the study of archaeology. Oxbow Books, Liverpool, pp 201–204

Bridge M (2012) Locating the origins of wood resources: a review of dendroprovenancing. J Archaeol Sci 39(8):2828–2834. https://doi.org/10.1016/j.jas.2012.04.028

Briesch M, Sobania D, Rothlauf F (2022) The Randomness of Input Data Spaces is an A Priori Predictor for Generalization. In: Bergmann R, Malburg L, Rodermund SC, Timm IJ (eds) KI 2022: Advances in Artificial Intelligence. Springer International Publishing, Cham, pp 17–30

Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber F, Karlén W, Zetterberg P, Eronen M (1992) Fennoscandian summers from ad 500: temperature changes on short and long timescales. Clim Dyn 7(3):111–119. https://doi.org/10.1007/BF00211153

Briffa KR, Jones PD, Schweingruber F, Osborn TJ (1998) Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. Nature 393(6684):450–455. https://doi.org/10.1038/30943

Briffa KR, Jones PD, Schweingruber FH, Karlén W, Shiyatov SG (1996) Tree-ring variables as proxy-climate indicators: Problems with low-frequency signals. In: Jones PD, Bradley RS, Jouzel J (eds) Climatic Variations and Forcing Mechanisms of the Last 2000 Years. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 9–41

Briffa KR, Melvin TM, Osborn TJ, Hantemirov RM, Kirdyanov AV, Mazepa VS, Shiyatov SG, Esper J (2013) Reassessing the evidence for tree-growth and inferred temperature change during the Common Era in Yamalia, northwest Siberia. Quat Sci Rev 72:83–107. https://doi.org/10.1016/j.quascirev.2013.04.008

Bunn AG (2010) Statistical and visual crossdating in R using the dplR library. Dendrochronologia 28(4):251–258. https://doi.org/10.1016/j.dendro.2009.12.001

Bunn AG, Sharac TJ, Graumlich LJ (2004) Using a Simulation Model to Compare Methods of Tree-Ring Detrending and to Investigate the Detectability of Low-Frequency Signals. Tree-Ring Res 60(2):77–90. https://doi.org/10.3959/1536-1098-60.2.77

Büntgen U, Allen K, Anchukaitis KJ, Arseneault D, Boucher E, Bräuning A, Chatterjee S, Cherubini P, Churakova OV, Corona C, Gennaretti F, Grießinger J, Guillet S, Guiot J, Gunnarson B, Helama S, Hochreuther P, Hughes MK, Huybers P, Kirdyanov AV, Krusic PJ, Ludescher J, Meier WJ-H, Myglan VS, Nicolussi K, Oppenheimer C, Reinig F, Salzer MW, Seftigen K, Stine AR, Stoffel M, St. George S, Tejedor E, Trevino A, Trouet V, Wang J, Wilson R, Yang B, Xu G, Esper J (2021) The influence of decision-making in tree ring-based climate reconstructions. Nat Commun 12(1):3411. https://doi.org/10.1038/s41467-021-23627-6

Büntgen U, Esper J, Frank D, Nicolussi K, Schmidhalter M (2005) A 1052-year tree-ring proxy for Alpine summer temperatures. Clim Dyn 25(2–3):141–153. https://doi.org/10.1007/s00382-005-0028-1

Büntgen U, Reinig F, Verstege A, Piermattei A, Kunz M, Krusic P, Slavin P, Stěpanék P, Torbenson M, Martinez del Castillo E, Arosio T, Kirdyanov A, Oppenheimer C, Trnka M, Palosse A, Bebchuk T, Camarero JJ, Esper J (2024) Recent summer warming over the western Mediterranean region is unprecedented since medieval times. Glob Planet Change 232:104336. https://doi.org/10.1016/j.gloplacha.2023.104336

Buras A, van der Maaten-Theunissen M, van der Maaten E, Ahlgrimm S, Hermann P, Simard S, Heinrich I, Helle G, Unterseher M, Schnittler M, Eusemann P, Wilmking M (2016) Tuning the Voices of a Choir: Detecting Ecological Gradients in Time-Series Populations. PLoS ONE 11(7):e0158346. https://doi.org/10.1371/journal.pone.0158346

Cherubini P, Gärtner H, Esper J, Dobbertin MK, Kaiser KF, Rigling A, Treydte K, Zimmermann NE, Bräker OU (2004) Jahrringe als Archive für interdisziplinäre Umweltforschung. Schweiz Z Forstwes 155(6):162–168

Cho N, Agossou C, Kim E, Lim J-H, Hwang T, Kang S (2024) Evaluating key climatic and ecophysiological parameters of worldwide tree mortality with a process-based BGC model and machine learning algorithms. Ecol Model 491:110688. https://doi.org/10.1016/j.ecolmodel.2024.110688.

D'Arrigo R, Wilson RJS, Liepert B, Cherubini P (2008) On the 'Divergence Problem' in Northern Forests: A review of the tree-ring evidence and possible causes. Glob Planet Change 60(3–4):289–305. https://doi.org/10.1016/j.gloplacha.2007.03.004

Deslauriers A, Rossi S, Anfodillo T (2007) Dendrometer and intra-annual tree growth: What kind of information can be inferred? Dendrochronologia 25:113–124. https://doi.org/doi:10.1016/j.dendro.2007.05.003

Drew DM, Downes GM (2009) The use of precision dendrometers in research on daily stem size and wood property variation A review. Dendrochronologia 27:159–172. https://doi.org/doi:10.1016/j.dendro.2009.06.008

Esper J (2002) Low-Frequency Signals in Long Tree-Ring Chronologies for Reconstructing Past Temperature Variability. Science 295(5563):2250–2253. https://doi.org/10.1126/science.1066208

Esper J, Büntgen U (2021) The future of paleoclimate. Clim Res 83: 57-59. https://doi.org/10.3354/cr01636

Esper J, Cook ER, Krusic PJ, Peters K, Schweingruber F (2003) Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies. Tree-Ring Res 59(2):81–89

Esper J, Düthorn E, Krusic PJ, Timonen M, Büntgen U (2014) Northern European summer temperature variations over the Common Era from integrated tree-ring density records. J Quat Sci 29(5):487–494. https://doi.org/10.1002/jqs.2726

Esper J, Gärtner H (2001) Interpretation of tree-ring chronologies. Erdkunde 55:277–288

Esper J, Torbenson M, Büntgen U (2024) 2023 summer warmth unparalleled over the past 2,000 years. Nature. https://doi.org/10.1038/ s41586-024-07512-y

Eyring V, Cox PM, Flato GM, Gleckler PJ, Abramowitz G, Caldwell P, Collins WD, Gier BK, Hall AD, Hoffman FM, Hurtt GC, Jahn A, Jones CD, Klein SA, Krasting JP, Kwiatkowski L, Lorenz R, Maloney E, Meehl GA, Pendergrass AG, Pincus R, Ruane AC, Russell JL, Sanderson BM, Santer BD, Sherwood SC, Simpson IR, Stouffer RJ, Williamson MS (2019) Taking climate model evaluation to the next level. Nat Clim Change 9(2):102–110. https://doi.org/10.1038/s41558-018-0355-y

Fritts HC (1976) Tree Rings and Climate. Academic Press, Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona, U.S.A.

Geiger RS, Cope D, Ip J, Lotosh M, Shah A, Weng J, Tang R (2021) “Garbage In, Garbage Out” Revisited: What Do Machine Learning Application Papers Report About Human-Labeled Training Data? Quant Sci Stud 2(2). https://doi.org/doi.org/10.1162/qss_a_00144

Géron A (2019) Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd edn. O’Reilly, Sebastopol

Gholami V, Ahmadi Jolandan M, Torkaman J (2017) Evaluation of climate change in northern Iran during the last four centuries by using dendroclimatology. Nat Hazards 85(3):1835–1850. https://doi.org/10.1007/s11069-016-2667-4

Gut U (2018) Evaluating the key assumptions underlying dendro-provenancing: How to spruce it up with a scissor plot. Dendrochronologia 52:131–145. https://doi.org/10.1016/j.dendro.2018.09.008

Hartl C, Düthorn E, Tejedor E, Kirchhefer AJ, Timonen M, Holzkämper S, Büntgen U, Esper J (2021) Micro-site conditions affect Fennoscandian forest growth. Dendrochronologia 65:125787. https://doi.org/10.1016/j.dendro.2020.125787

Hartl C, Schneider L, Riechelmann DFC, Kuhl E, Kochbeck M, Klippel L, Büntgen U, Esper J (2022) The temperature sensitivity along elevational gradients is more stable in maximum latewood density than tree-ring width. Dendrochronologia 73:125958. https://doi.org/10.1016/j.dendro.2022.125958

Hartl-Meier C, Dittmar C, Zang C, Rothe A (2014) Mountain forest growth response to climate change in the Northern Limestone Alps. Trees 28(3):819–829. https://doi.org/10.1007/s00468-014-0994-1

He C, Zhou L, Yao Y, Ma W, Kinney PL (2021) Cooling effect of urban trees and its spatiotemporal characteristics: A comparative study. Build Environ 204:108103. https://doi.org/10.1016/j.buildenv.2021.108103

Jevšenak J, Džeroski S, Zavadlav S, Levanic T (2018) A Machine Learning Approach to Analyzing the Relationship Between Temperatures and Multi-Proxy Tree-Ring Records. Tree-Ring Res 74(2):210–224. https://doi.org/10.3959/1536-1098-74.2.210

Jiang C, Xu H, Tong Y, Li J (2024) Anomalous Warm Temperatures Recorded using Tree Rings in the Headwater of the Jinsha River during the Little Ice Age. Forests 15(6). https://doi.org/10.3390/f15060972

Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. Science 349(6245):255–260. https://doi.org/10.1126/science.aaa8415

Keitt TH, Abelson ES (2021) Ecology in the age of automation. Science 373(6557):858–859. https://doi.org/10.1126/science.abi4692

Keret, R, Schliephack P M, Stangler D F, Seifert T, Kahle H-P, Drew D M, Hills P N (2024) An

open-source machine-learning approach for obtaining high-quality quantitative wood anatomy data from E. grandis and P. radiata xylem. Plant Sci 340:111971. https://doi.org/10.1016/j.plantsci.2023.111970.

Klippel L, Büntgen U, Konter O, Kyncl T, Esper J (2020) Climate sensitivity of high- and low-elevation Larix decidua MXD chronologies from the Tatra Mountains. Dendrochronologia 60:125674. https://doi.org/10.1016/j.dendro.2020.125674

Klippel L, Krusic PJ, Brandes R, Hartl C, Belmecheri S, Dienst M, Esper J (2018) A 1286-year hydro-climate reconstruction for the Balkan Peninsula. Boreas 47(4):1218–1229. https://doi.org/10.1111/bor.12320

Knüsel S, Peters RL, Haeni M, Wilhelm M, Zweifel R (2021) Processing and Extraction of Seasonal Tree Physiological Parameters from Stem Radius Time Series. Forests 12:765. https://doi.org/10.3390/f12060765

Kuhl E, Esper J, Schneider L, Trouet V, Kunz M, Klippel L, Büntgen U, Hartl C (2024) Revising Alpine summer temperatures since 881 CE. Clim Dyn. https://doi.org/10.1007/s00382-024-07195-1

Kuhl E, Zang C, Esper J, Riechelmann DFC, Büntgen U, Briesch M, Reinig F, Römer P, Konter O, Schmidhalter M, Hartl C (2023) Using machine learning on tree-ring data to determine the geographical provenance of historical construction timbers. Ecosphere :1–14. https://doi.org/10.1002/ecs2.4453

Kuhl E, Ziaco E, Esper E, Konter O, Martinez de Castillo E (in review) A machine learning approach to fill gaps in dendrometer data. Trees

Lara A, Villalba R, Urrutia-Jalabert R, González-Reyes A, Aravena JC, Luckman BH, Cuq E, Rodríguez C, Wolodarsky-Franke A (2020) +A 5680-year tree-ring temperature record for southern South America. Quat Sci Rev 228:106087. https://doi.org/10.1016/j.quascirev.2019.106087

Linderholm HW, Gunnarson BE, Fuentes M, Büntgen U, Hormes A (2021) The origin of driftwood on eastern and south-western Svalbard. Polar Sci 29:100658. https://doi.org/10.1016/j.polar.2021.100658

Liu Y, An ZS, Linderholm HW, Chen DL, Song HM, Cai QF, Sun JY, Tian H (2009) Annual temperatures during the last 2485 years in the mid-eastern Tibetan Plateau inferred from tree rings. Sci China Ser D-Earth Sci 52(3):348–359. https://doi.org/10.1007/s11430-009-0025-z

Luckman BH, Wilson RJS (2005) Summer temperatures in the Canadian Rockies during the last millennium: a revised record. Clim Dyn 24:131–144. https://doi.org/10.1007/s00382-004-0511-0

Lukovic M, Zweifel R, Thiry G, Zhang C, Schubert M (2022) Reconstructing radial stem size changes of trees with machine learning. J R Soc Interface 19:20220349. https://doi.org/10.1098/rsif.2022.0349

Martinez del Castillo E, Prislan P, Gričar J, Gryc V, Merela M, Giagli K, de Luis M, Vavrčík H, Čufar K (2018) Challenges for growth of beech and co-occurring conifers in a changing climate context. Dendrochronologia 52:1–10. https://doi.org/10.1016/j.dendro.2018.09.001

Martinez del Castillo E, Zang CS, Buras A, Hacket-Pain A, Esper J, Serrano-Notivoli R, Hartl C, Weigel R, Klesse S, Resco de Dios V, Scharnweber T, Dorado-Liñán I, van der Maaten-Theunissen M, van der Maaten E, Jump A, Mikac S, Banzragch B-E, Beck W, Cavin L, Claessens H, Čada V, Čufar K, Dulamsuren C, Gričar J, Gil-Pelegrín E, Janda P, Kazimirovic M, Kreyling J, Latte N, Leuschner C, Longares LA, Menzel A, Merela M, Motta R, Muffler L, Nola P, Petritan AM, Petritan IC, Prislan P, Rubio-Cuadrado A, Rydval M, Stajić B, Svoboda M, Toromani E, Trotsiuk V, Wilmking M, Zlatanov T, de Luis M (2022) Climate-change-driven growth decline of European beech forests. Commun Biol 5(1):163. https://doi.org/10.1038/s42003-022-03107-3

Mendes MP, Cherubini P, Plieninger T, Ribeiro L, Costa A (2019) Climate effects on stem radial growth of Quercus suber L.: does tree size matter? Forestry 92:73–84. https://doi.org/doi:10.1093/forestry/cpy034

Mickler RA, Earnhardt TS, Moore JA (2002) Regional estimation of current and future forest biomass. Environ Pollut 116:S7–S16. https://doi.org/10.1016/S0269-7491(01)00241-X

Miller TW, Stangler DF, Larysch E, Honer H, Seifert T, Kahle H-P (2022) A methodological framework to optimize models predicting critical dates of xylem phenology based on dendrometer data. Dendrochronologia 72:125940. https://doi.org/10.1016/j.dendro.2022.125940

Netherer S, Schopf A (2010) Potential effects of climate change on insect herbivores in European forests—General aspects and the pine processionary moth as specific example. For Ecol Manag 259(4):831–838. https://doi.org/10.1016/j.foreco.2009.07.034

Oberhuber W, Gruber A (2010) Climatic influences on intra-annual stem radial increment of Pinus sylvestris (L.) exposed to drought. Trees 24:887–898. https://doi.org/10.1007/s00468-010-0458-1

Pandey S (2021) Climatic influence on tree wood anatomy: a review. J Wood Sci 67(1):24. https://doi.org/10.1186/s10086-021-01956-w

Pattnaik N, Honold M, Franceschi E, Moser-Reischl A, Rötzer T, Pretzsch H, Pauleit S, Rahman MA (2024) Growth and cooling potential of urban trees across different levels

of imperviousness. J Environ Manag 361:121242. https://doi.org/10.1016/j.jenvman.2024.121242

Peters RL, Groenendijk P, Vlam M, Zuidema PA (2015) Detecting long-term growth trends using tree rings: a critical evaluation of methods. Glob Change Biol 21:2040–2054. https://doi.org/doi: 10.1111/gcb.12826

Phillips ET, Höll M, Kantz H, Zhou Y (2023) Trend analysis in the presence of short- and long-range correlations with application to regional warming. Phys Rev E 108:034301. https://doi.org/10.1103/PhysRevE.108.034301

Andrei P, Popa I, Badea O, Bosela M (2024): Non-linear response of Norway spruce to climate variation along elevational and age gradients in the Carpathians. Environ Res 252 (4): 119073. https://doi.org/10.1016/j.envres.2024.119073.

Rahmani AM, Yousefpoor E, Yousefpoor MS, Mehmood Z, Haider A, Hosseinzadeh M, Ali Naqvi R (2021) Machine Learning (ML) in Medicine: Review, Applications, and Challenges. Mathematics 9(22). https://doi.org/10.3390/math9222970

Rennenberg H, Loreto F, Polle A, Brilli F, Fares S, Beniwal RS, Gessler A (2006) Physiological Responses of Forest Trees to Heat and Drought. Plant Biol 8(5):556–571. https://doi.org/10.1055/s-2006-924084

Rocha E, Holzkämper S (2023) Assessing urban climate effects on Pinus sylvestris with point dendrometers: a case study from Stockholm, Sweden. Trees 37:31–40. https://doi.org/10.1007/s00468-020-02082-8

Rötzer T, Moser-Reischl A, Rahman MA, Grote R, Pauleit S, Pretzsch H (2021) Modelling Urban Tree Growth and Ecosystem Services: Review and Perspectives. In: Cánovas FM, Lüttge U, Risueño M-C, Pretzsch H (eds) Progress in Botany Vol. 82. Springer International Publishing, Cham, pp 405–464

Rydval M, Björklund J, von Arx G, Begović K, Lexa M, Nogueira J, Schurman JS, Jiang Y (2024) Ultra-high-resolution reflected-light imaging for dendrochronology. Dendrochronologia 83:126160. https://doi.org/10.1016/j.dendro.2023.126160

Samuel AL (1959) Some Studies in Machine Learning Using the Game of Checkers. IBM J Res Dev 3(3):210–229. https://doi.org/10.1147/rd.33.0210

Shindo L, Saulnier M, Raese H, Guibal F, Edouard J-L, Bolka M, Carrer M, Corona C, Gassmann P, Grabner M, Guillet S, Nicolussi K, Nola P, Pignatelli O, Stoffel M (2024) European larch sapwood: A model for predicting the cambial age and for a more accurate dating. Dendrochronologia 83: 126150. https://doi.org/10.1016/j.dendro.2023.126150.

Schwaab J, Meier R, Mussetti G, Seneviratne S, Bürgi C, Davin EL (2021) The role of urban trees in reducing land surface temperatures in European cities. Nat Commun 12(1):6763. https://doi.org/10.1038/s41467-021-26768-w

Speer JH (2010) Fundamentals of tree-ring research. Tucson

Torbenson MCA, Büntgen U, Esper J, Urban O, Balek J, Reinig F, Krusic PJ, Martinez del Castillo E, Brázdil R, Semerádová D, Štěpánek P, Pernicová N, Kolář T, Rybníček M, Koňasová E, Arbelaez J, Trnka M (2023) Central European Agroclimate over the Past 2000 Years. J Clim 36(13):4429–4441. https://doi.org/10.1175/JCLI-D-22-0831.1

Treydte K, Liu L, Padrón RS, Martínez-Sancho E, Babst F, Frank DC, Gessler A, Kahmen A, Poulter B, Seneviratne SI, Stegehuis AI, Wilson R, Andreu-Hayles L, Bale R, Bednarz Z, Boettger T, Berninger F, Büntgen U, Daux V, Dorado-Liñán I, Esper J, Friedrich M, Gagen M, Grabner M, Grudd H, Gunnarsson BE, Gutiérrez E, Hafner P, Haupt M, Hilasvuori E, Heinrich I, Helle G, Jalkanen R, Jungner H, Kalela-Brundin M, Kessler A, Kirchhefer A, Klesse S, Krapiec M, Levanič T, Leuenberger M, Linderholm HW, McCarroll D, Masson-Delmotte V, Pawelczyk A, Pazdur A, Planells O, Pukiene R, Rinne-Garmston KT, Robertson I, Saracino A, Saurer M, Schleser GH, Seftigen K, Siegwolf RTW, Sonninen E, Stievenard M, Szychowska-Krapiec E, Szymaszek M, Todaro L, Waterhouse JS, Weigl-Kuska M, Weigt RB, Wimmer R, Woodley EJ, Vitas A, Young G, Loader NJ (2024) Recent human-induced atmospheric drying across Europe unprecedented in the last 400 years. Nat Geosci 17(1):58–65. https://doi.org/10.1038/s41561-023-01335-8

Vilenski E, Bak P, Rosenblatt JD (2019) Multivariate anomaly detection for ensuring data quality of dendrometer sensor networks. Comput Electron Agric 162:412–421. https://doi.org/10.1016/j.compag.2019.04.018

Wilson RJS, Anchukaitis KJ, Briffa KR, Büntgen U, Cook ER, D'Arrigo R, Davi N, Esper J, Frank D, Gunnarson BE, Hegerl GC, Helama S, Klesse S, Krusic PJ, Linderholm HW, Myglan V, Osborn TJ, Rydval M, Schneider L, Schurer A, Wiles G, Zhang P, Zorita E (2016) Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. Quat Sci Rev 134:1–18. https://doi.org/10.1016/j.quascirev.2015.12.005

Wilson RJS, Esper J, Luckman BH (2004) Utilising historical tree-ring data for dendroclimatology: A case study from the Bavarian Forest, Germany. Dendrochronologia 21(2):53–68. https://doi.org/10.1078/1125-7865-00041

Wilson RJS, Luckman BH, Esper J (2005) A 500 year dendroclimatic reconstruction of spring-summer precipitation from the lower Bavarian Forest region, Germany. Int J Climatol 25(5):611–630. https://doi.org/10.1002/joc.1150

Ziaco E, Biondi F (2018) Stem Circadian Phenology of Four Pine Species in Naturally

Contrasting Climates from Sky-Island Forests of the Western USA. Forests 9: 396. doi:10.3390/f9070396

Zhang P, Björklund J, Linderholm HW (2015) The influence of elevational differences in absolute maximum density values on regional climate reconstructions. Trees 29(4):1259–1271. https://doi.org/10.1007/s00468-015-1205-4

Zhang Q-B, Hebda RJ, Zhang Q-J, Alfaro RI (2000) Modeling Tree-Ring Growth Responses to Climatic Variables Using Artificial Neural Networks. Forest Science 46(2):229–239. https://doi.org/10.1093/forestscience/46.2.229

Zhang Y, Gao L, Deng Y, Huang Q, Yuan Y, Shi X (2024) Seasonal aridity regulates drivers and temporal variability of wood phenology: A meta-analysis of dendrometer monitoring data across the Northern Hemisphere. Dendrochronologia 85:126201. https://doi.org/10.1016/j.dendro.2024.126201

Zhao X, Chen F, Seim A, Hu M, Akkemik Ü, Kopabayeva A, Mazarzhanova K, Zhang R, Maisupova B, Kirillov V, Mambetov B, Yu S, He Q, Dosmanbetov D, Kelgenbayev N (2024) Global warming leads to growth increase in Pinus sylvestris in the Kazakh steppe. For Ecol Manag 553:121635. https://doi.org/10.1016/j.foreco.2023.121635.

Zuh J-J, Yang M, Ren Z J (2023) Machine Learning in Environmental Research: Common Pitfalls and Best Practices. Environ Sci Technol 57:17671-17689. 10.1021/acs.est.3c00026

List of figures

List of tables

Index of abbreviation

A1 TRW	– first order autocorrelation of TRW
BI	– blue intensity
DM_{gen}	– general density model
DM_{sp}	– species-specific density model
ED/LD ratio	– earlywood/ latewood density ratio
EWD	– earlywood density
EWW	– earlywood width
JJA(S)/MJJAS	– June, July, August, (September)/ May- September
kNN	– k-nearest neighbor
LBM	– larch budmoth
LIA	– little ice age
LWD	– latewood density
LWW	– latewood width
ML	– machine learning
MV	– Matter Valley
MXD	– maximum latewood density
PCA	– principal component analysis
PCGA	– principal component gradient analysis
RC(S)	– regional curve (standardization)
RF	– random forest
RWM_{gen}	– general ring-width model
RWM_{sp}	– species-specific ring-width model
SAOD	– stratospheric aerosol optical depth
SEA	– superposed epoch analysis
SI	– stable isotopes
SF	– signal-free RCS
SV	– Simplon Valley
TRW	– tree-ring width
XGBoost/ XGB	– extreme gradient boosting

List of publications

Kuhl E, Ziaco E, Esper E, Konter O, Martinez de Castillo E (2024) A machine learning approach to fill gaps in dendrometer data. Trees. https://doi.org/10.1007/s00468-024-02573-y

Kuhl E, Esper J, Schneider L, Trouet V, Kunz M, Klippel L, Büntgen U, Hartl C (2024) Revising Alpine summer temperatures since 881 CE. Climate Dynamics. https://doi.org/10.1007/s00382-024-07195-1

Kunz M, Esper J, **Kuhl E**, Schneider L, Büntgen U, Hartl C. (2023) Combining Tree-RingWidth and Density to Separate the Effects of Climate Variation and Insect Defoliation. Forests 14(7). https://doi.org/10.3390/f14071478

Kuhl E, Zang C, Esper J, Riechelmann D, Büntgen U, Briesch M, Reinig F, Römer P, Konter O, Schmidhalter M, Hartl C (2023) Using Machine Learning on tree-ring data to determine the geographical provenance of historical construction timbers. Ecosphere 14(3). https://doi.org/10.1002/ecs2.4453

Hartl C, Schneider L, Riechelmann DFC, **Kuhl E**, Kochbeck M, Klippel L, Büntgen U, Esper J (2022) The temperature sensitivity along elevational gradients is more stable in maximum latewood density than tree-ring width. Dendrochronologia 73. https://doi.org/10.1016/j.dendro.2022.125958

Curriculum Vitae

Not available online.