Enhanced dendroprovenancing through high-resolution soil- and climate data

Martijn van Sluijs [*], Sytze de Bruin, Peter van der Sleen

Wageningen University, Droevendaalsesteeg 4, Wageningen 6708 PB the Netherlands

A B S T R A C T

Instruments aiming to avoid illegal logging such as certification chains require data-driven solutions to verify timber origin. One approach to timber tracing is dendroprovenancing, which uses the spatial and temporal consistency of tree ring width patterns to match unknown samples to reference samples from known locations. Best matching reference samples indicate the potential source location of the unknown sample. Gaps in temporal and spatial coverage of reference chronologies however currently limit applicability of dendroprovenancing, with additional data acquisition being both time-consuming and expensive. This study presents a novel general dendroprovenancing framework, aiming to overcome this shortcoming. It relies on modelling and spatially exhaustive prediction of reference chronologies using a regression model and gridded high-resolution soil- and climate data with global coverage. The presented framework is explored through a case study on *Quercus robur* using 107 tree-ring chronologies from western and central Europe. We tested three scenarios using leave one out cross-validation: 1) the dating of the chronology is unknown, 2) the source location of the chronology is unknown, and 3) both the dating and source location of the chronology are unknown, with the latter most closely resembling a real-world scenario. We found that tracing accuracy was high, even in the scenario in which both the dating and source location of the chronology were unknown. 82.2% of the chronologies were traced to within a radius of 250 kilometres from the ground truth and correctly dated. The findings highlight newfound potential of dendroprovenancing for timber tracing.

1. Introduction

Illegal logging and trade in illegal timber products are ongoing global issues, causing harm to the environment, local communities, and economies (Khalid et al., 2019). Determining the provenance of timber through forensic diagnostic timber identification is broadly recognised as an important step toward combatting illegal logging (Dormontt et al., 2015). This is also demonstrated by recent policy developments in the European Union (EU), where member states are set to conduct provenance checks to ensure timber products do not contribute to the destruction and degradation of forests (European Commission, 2021).

Instruments aiming to avoid illegal logging include certification chains such as the Forest Stewardship Council (FSC), Programme for the Endorsement of Forest Certification (PEFC), and the EU's Forest Law Enforcement, Governance and Trade Action Plan (FLEGT) (European Commission, 2003). The latter includes bilateral trade agreements known as Voluntary Partnership Agreements (VPAs) (Polo Villanueva et al., 2023). Should such solutions only rely on the accumulated

product of trust and agreements during the multiple stages of the timber chain, they would be susceptible to fraud. Hence, they partly rely on data-driven verification of timber provenance to ensure legality (Forest Stewardship Council, 2017). Data-driven approaches provide independent means of tracking the provenance of timber (Dormontt et al., 2015). They can be used directly on the timber itself at a late- or final stage in the timber chain, circumventing the aforementioned limitations of trust and agreements in certification chains.

Dendroprovenancing uses tree ring growth chronologies to determine the provenance of a wood sample (Bridge, 2012). The premise is that a tree-ring series of unknown origin, or a chronology based on such samples, can be compared to all available chronologies of the same species. The potential regional origin of the unknown wood sample is derived based on the strongest match with a chronology of known origin. A common approach is to assess the correlations between the unknown chronology and reference chronologies and assign the unknown chronology to the location of the reference chronology with which it is most strongly correlated, for example judged by the t-value

---

(Bridge and Fowler, 2019). Further approaches include mapping dots varying in size in proportion to the level of cross-match between the unknown chronology and reference chronologies to indicate provenance (Bridge, 2012; Daly, 2007; Daly and Tyers, 2022), or through employing a network approach (Visser, 2021; Visser and Vorst; 2022).

The availability of reference datasets is currently one of the major challenges in the effective use of dendroprovenancing, as these commonly contain gaps in both spatial and temporal coverage (Pearl et al., 2020). The collection of additional data is often time-consuming and expensive, owing to the effort needed both during in-situ sampling and sample measuring (Pilcher et al., 1990). While other data-driven approaches to enhance timber provenancing have been developed, such as vessel size analysis (Akhmetzyanov et al., 2019), distinguishing between late- and earlywood (Akhmetzyanov et al., 2020b), isotope analysis (Boeschoten et al., 2023a; D'Andrea et al., 2023; Van Ham-Meert and Daly, 2023), multi-element analyses (Boeschoten et al., 2023b) and DNA analysis (Akhmetzyanov et al., 2020a), these face the same fundamental challenges. Reference datasets with full spatial coverage are necessary to determine the provenance of an unlocated sample with confidence, but these currently contain gaps that are difficult and/or expensive to fill.

To address the challenge of incomplete coverage in lacking reference datasets, this study proposes a novel approach that creates a set of modelled reference chronologies at high spatial resolution, based on available climate-growth relationships and gridded meteorological and soil data. Such an approach is possible because the driving factors of tree ring width (TRW) variation have been well-known for many tree species in temperate regions (Schweingruber, 2012) and are used for climate modelling and -reconstruction (Hughes, 2002). The most important environmental factors affecting TRW variation - temperature, precipitation, and soil characteristics - are now available as high-resolution global gridded products (Fick and Hijmans, 2017; Harris et al., 2020; Poggio et al., 2021). As the available gridded meteorological data goes back to the start of the 20th century, this allows modelling tree growth variability over large spatial and temporal extents.

This study serves as a proof of concept for enhanced dendroprovenancing by modelling reference chronologies covering an entire species range using gridded meteorological and soil data. It aims to demonstrate the potential of this novel approach as a tool for timber provenancing. A general framework is presented, of which the effectiveness is judged through a case study using raw TRW data sourced from the International Tree-Ring Data Bank (ITRDB) (Grissino-Mayer and Fritts, 1997) of one tree species; pedunculate oak (*Quercus robur* L.), combined with gridded data from Soilgrids (Poggio et al., 2021), CRU TS (Harris et al., 2020) and Worldclim (Fick and Hijmans, 2017). The aims of this study are to (i) assess to what extent interannual tree growth variability can be explained from globally available gridded data, and (ii) determine the accuracy of dendroprovenancing using the resulting modelled reference chronologies.

2. Material and methods

2.1. General framework

Our framework (Fig. 1) is based on reference TRW time series of individual trees of a particular tree species, which are locally combined into reference chronologies. Thus, TRW measurements of multiple trees are detrended and aggregated into a single chronology at a given location. Traditional approaches would compare a chronology with unknown dating and/or location directly to these reference chronologies. Instead, we select gridded data sources containing variables known to affect this species' TRW variation, such as meteorological- or soil data, and extract data for the locations of each reference chronology. A regression model is then trained using the extracted gridded data as features and individual chronology values of the reference chronologies as targets.

The regression model is used to construct modelled reference chronologies throughout the entire distribution range of the species using a point grid. The undated and/or unlocated chronology is compared to the modelled reference chronologies using the coefficient of determination (R^2) as the evaluation criterion. The greatest R^2 denotes the closest match of the date and/or location of the unknown chronology. Other similarity metrics could also be used but R^2 is easily interpreted and in contrast to parametric hypothesis testing (e.g. the t-test) it does not rely
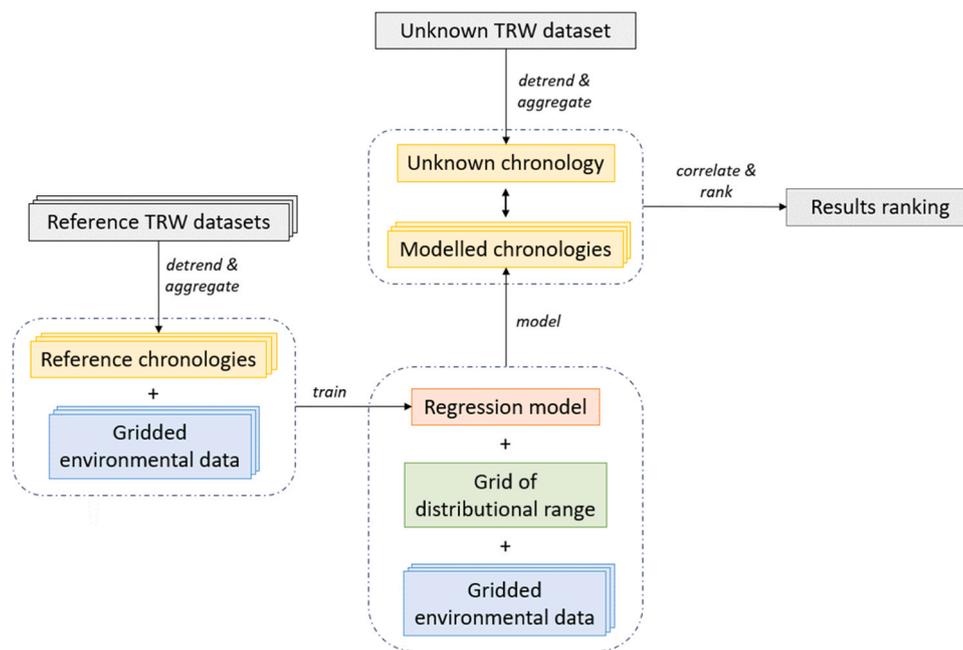

Fig. 1. General flowchart of model framework. Instead of directly comparing and matching an unknown chronology to reference chronologies, a regression model is trained using reference chronologies in combination with relevant gridded environmental data. Using the regression model, modelled reference chronologies are constructed for all points in a user-specified point grid, which encompasses the species distribution with a given sampling density. Next, the unknown chronology is compared to and matched with the modelled reference chronologies.

on sampling and distributional assumptions.

Overall, the approach shows similarities to the work of Babst et al. (2018), Practical example 1), but applied to dendroprovenancing. Note that this framework is broad and requires specific choices to be made regarding the detrending method, the chronology aggregation method, the used gridded data and the regression model, which may all depend on the specific use case.

2.2. Case study

The proposed framework was explored through a case study. The species of interest study was pedunculate oak (*Quercus robur* L.), known for its clear TRW response to environmental conditions (Schweingruber, 1993) and relatively high data availability. The study area encompassed the known species' distribution range (Fig. 2; Caudullo et al., 2017).

2.2.1. Data
TRW data was obtained from the ITRDB, querying for raw TRW data of pedunculate oak (*Quercus robur* L.). The ITRDB is the most comprehensive archive of publicly shared tree-ring data (Zhao et al., 2019). A comprehensive list of ITRDB datasets used in this study can be found in the supplementary reference section (Appendix A).

Gridded meteorological data were obtained by combining Climatic Research Unit Timeseries (CRU TS) v. 4.06 (Harris et al., 2020) and Worldclim (Fick and Hijmans, 2017). Both datasets cover all land areas globally for a set of meteorological variables. CRU TS is a widely used dataset and provides monthly data at a spatial resolution of 0.5°. It has frequently been used in the domain of dendrochronology (Akhmetzyanov et al., 2019; Park et al., 2021; Salehnia and Ahn, 2022). While Worldclim has a finer spatial resolution of 1 km, it only provides long-term (1970–2000), monthly averages of each variable.

Delta downscaling was used to combine the two, leveraging both the high spatial resolution of Worldclim and the monthly data of CRU TS, following an approach similar to Moreno and Hasenauer (2016). Utilizing inverse distance weighting, we precisely obtained CRU TS data at the coordinates of both ITRDB and point grid sites. A correction was then implemented using the non-interpolated long-term WorldClim data. This involved calculating the difference between the average CRU TS and aligned WorldClim data, applying this discrepancy as either a subtraction or division, depending on the meteorological variable. When evaluating using observed data, downscaling CRU TS using Worldclim has been shown to lower the root-mean-square error (RMSE) when compared to non-downscaled CRU TS (Salvacion et al., 2018). We made use of the meteorological variables mean, minimum, and maximum temperature and precipitation, which are available in both CRU TS and Worldclim.

Gridded soil data were obtained from Soilgrids (Poggio et al., 2021), a collection of soil property maps of the world at 250 m resolution. We used five soil characteristics for this study that were expected to directly or indirectly influence TRW: clay-, sand-, silt-, nitrogen- and organic carbon content (Eckstein et al., 1990; Schweingruber, 2012). The soil data were interpolated for gap filling and consequently masked to Worldclim coverage.

Before analysis, preselection was performed on the TRW data, according to the following ruleset:

- Datasets with (nearly) identical coordinates were checked for duplicate TRW sequences and combined into one if no duplicates were found.
- Datasets exclusively containing late- or earlywood measurements were removed.
- Datasets with known internal quality issues were removed.
- Datasets located at sea were removed.
- Datasets with no temporal overlap with CRU TS were removed.

2.2.2. Data analysis
From here onwards, we refer to the sample of which we aimed to predict the location and/or date as the 'unknown TRW dataset', and the other samples as the 'known TRW datasets'. Similarly, we refer to the chronology for the unknown TRW dataset as the unknown chronology, and the chronologies for the known TRW datasets as the known chronologies.

All TRW sequences were detrended by dividing them by a Gaussian kernel smoother, leaving only year-to-year variability. We opted for a Gaussian Kernel smoother over alternative detrending functions, as it demonstrated superior efficacy in generating robust climate-growth relationships. The standard deviation (sigma) of the Gaussian kernel was selected as follows. For each known TRW dataset we calculated the
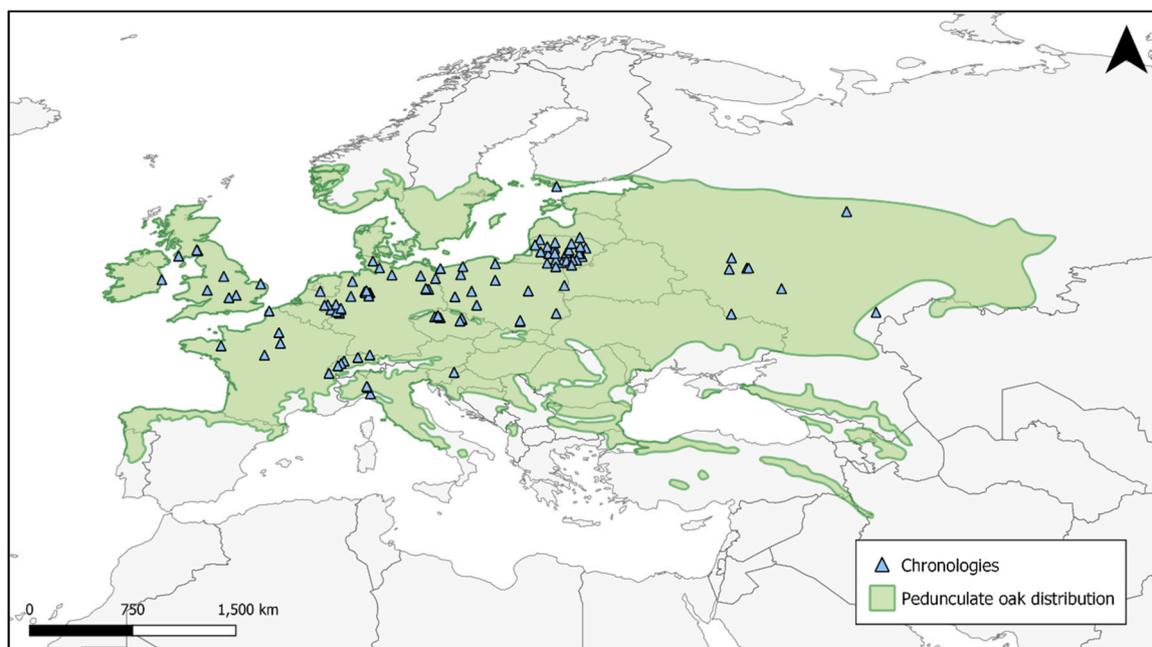

Fig. 2. Overview of the spatial distribution of the 107 chronologies. Overlay on the distribution range of pedunculate oak (Caudullo et al., 2017).

median linear correlation between its TRW sequences and the yearly local temperature and precipitation. This was repeated over a range of possible sigma's in the [0.2 – 10] interval with step size 0.2. The correlation medians were then averaged across all known TRW datasets, resulting in a single performance indicator for each of the possible sigma's. To limit outlier influence, these averaged correlation medians were smoothed by a centred rolling mean with a window width of five sigma interval steps. The sigma with the greatest smoothed mean median correlation value was used for the detrending all TRW sequences in both the known and unknown TRW datasets.

To remove potential outlier TRW sequences before constructing a chronology, the sequence that had the largest p-value when performing linear regression against the mean of all other TRW sequences in its TRW dataset was removed if that p was >0.01. This process was repeated until either the largest p-value was <0.01 or until over half of the sequences in the TRW dataset were removed. If over half of the sequences of a TRW dataset were removed, that TRW dataset was deemed to contain internal quality issues and consequently removed from further analysis.

Next, the remaining detrended TRW sequences were log-transformed and averaged into one chronology for every TRW dataset. Log-transformation was applied to transform relative (multiplicative) effects into additive ones and has an impact on the loss function used in the below-explained machine learning approach. Through initial experimentation (not reported here) the transformation was found to be beneficial.

Gridded data was extracted for the coordinates of each chronology. For modelling chronologies, we opted for random forest (RF) regression as implemented in Python using the scikit-learn package (Pedregosa et al., 2011). A random forest fits a number of regression trees on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting (Breiman, 2001). RF models are known for their high out-of-the-box performance, i.e., without hyperparameter tuning. Default scikit-learn hyperparameter settings were applied except for n_estimators, controlling the number of trees, which was set to 500. This resulted in a bootstrapped RF with 500 trees of maximum depth without pruning.

The RF model was trained using the data from the known chronologies. Features used in the training of the model included the extracted monthly meteorological variables from downscaled CRU TS, the extracted soil variables from Soilgrids, the year belonging to the chronology value, the longitude and latitude belonging to the chronology, and a one-hot encoding for each chronology. In addition to the 'standard' monthly meteorological variables, we added their 15-year monthly average that was centred on the focus month, as well as the deviation of the focus month from this average. Particularly, the deviations from the monthly means were expected to improve fitting the regression model. All meteorological features were also added for a time lag of one year, allowing the model to use the meteorological conditions of the year prior to the tree ring formation.

Finally, a point grid with a 0.2° spatial resolution within the distribution range of pedunculate oak and Worldclim coverage was created. Feature data were extracted at each grid node and modelled reference chronologies were constructed using the random forest models from the previous step. Using the modelled reference chronologies, three scenarios were tested for the unknown chronology:

1. **Year unknown**: The correct location of the chronology is known, dating the chronology is attempted. Here, dating offsets from −5 to +5 years from the ground truth dating were investigated, considering the dating and location of the TRW dataset as indicated by the ITRDB as ground truth. The coefficient of determination (R^2) was used as an evaluation criterion, with the offset with the highest R^2 used as the prediction for the year. Though the focus here is on dating and not provenancing, this scenario helps in the interpretation of the results, especially regarding the results of scenario 3.

2. **Location unknown**: The correct dating of the chronology is known, predicting the location is attempted. Here, all points in the point grid were investigated, considering the dating and location of the TRW dataset as indicated by the ITRDB as ground truth. R^2 was used as an evaluation criterion, with the point with the highest R^2 used as the prediction for the location.

3. **Both unknown**: The scenario likely closest to a real-world case, where the previous two scenarios are combined into one. We attempt to predict both the location of the chronology and the correct dating. To this end, we investigated all points in the point grid for dating offsets from −5 to 5 years, again considering the dating and location of the TRW dataset as indicated by the ITRDB as ground truth. R^2 Was used as the evaluation criterion, with the result with the highest R^2 used as the prediction for the combination of year and location.

We used a leave-one-out cross-validation (LOOCV) approach, each time using one of the chronologies as the unknown chronology and using the remainder of the chronologies for constructing the modelled reference chronologies.

All data analysis was conducted in Python. Essential packages used in the analysis were astropy, Fiona, GDAL, netCDF4, numpy, pandas, rasterio, scipy, shapely, scikit-learn, and statsmodels. For further documentation, methodological details, and non-essential package usage, we refer to the documentation accompanying the publicly available code for this study, available on GitHub (github.com/mvansluijs/Enhanced-Dendroprovenancing) with the DOI: 10.5281/zenodo.10465532.

3. Results

3.1. Data pre-processing

After the preselection of TRW datasets following the ruleset presented in Section 2.2.1, a total of 109 TRW datasets were retained. Two further TRW datasets were removed during chronology formation due to internal quality issues, leaving 107 chronologies (Fig. 2, Appendix B) for the remainder of the analysis. For an overview of all combined and removed TRW datasets, and the reasons for their combination or removal, we refer to Appendix C.

3.2. Dating chronologies

For the 'Year unknown' scenario, in 98 of the 107 test cases the year was correctly predicted, implying a 91.6% prediction accuracy. In 6 of the remaining cases, the results indicated that the TRW datasets were possibly misdated in the ITRDB, shown by a relatively high R^2 value for a small offset (one to three years) compared to the average R^2 value for the other years of that TRW dataset (difference of an order of magnitude 10 or higher).

3.3. Tracing chronologies

For the 'Location unknown' scenario (see example in Fig. 3), we considered a distance under 250 km between the predicted location and ground truth location to be a 'match'. Under that condition, 91 of the 107 test cases matched, i.e., 85.0% accuracy. Note that the 250 km threshold was arbitrarily chosen and does not imply the maximum positional accuracy. Distances between the predicted location and ground truth location were concentrated close to zero, with the histogram showing an asymptotically decreasing frequency at greater distances (Fig. 4).

3.4. Dating and tracing chronologies

Finally, for the 'Both unknown' scenario, we considered a combination of a temporal offset equal to zero and a distance under 250 km between the predicted location and ground truth location to be a match.

We present the results for this scenario using the top-N accuracy metric, which we split into a top-1 accuracy and a top-5 accuracy. Top-1 accuracy requires the prediction with the highest R^2 to precisely match
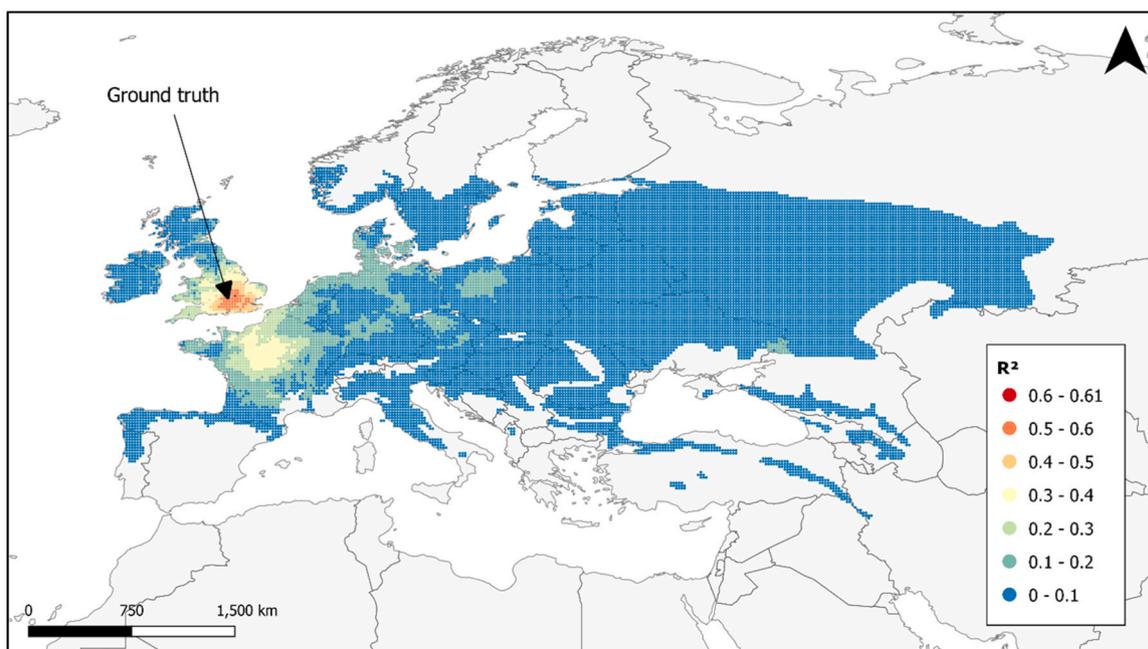
Fig. 3. Example of a 'Location unknown' prediction map (ITRDB code brit4). Every dot represents a modelled reference chronology, with its colour indicating the correlation with the brit4 chronology.
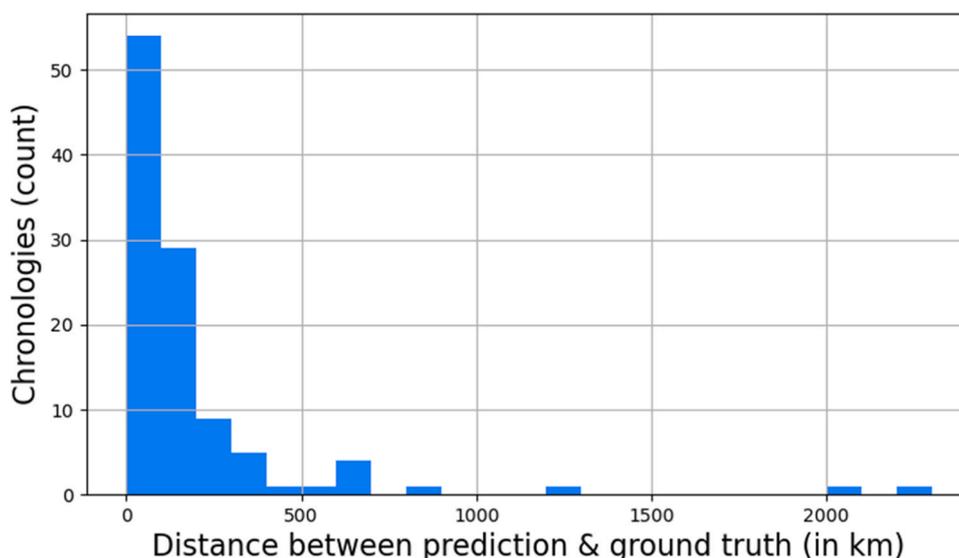

Fig. 4. Histogram of distances between predicted location and ground truth location for the 'Location unknown' scenario.

the ground truth, while top-5 accuracy deems a match successful if the ground truth aligns with any of the five results possessing the highest R^2 values. To avoid matches from the same spatial cluster to populate the entire top-5, the results were filtered such that the top-5 of matches only contained results that were more than 250 kilometres apart and/or dated in a different year.

For 88 of the 107 test cases, the first match was within 250 km of the ground truth location and correctly dated. For 93 of the 107 test cases, these conditions were met by at least one of the top-5 matches. This resulted in a top-1 accuracy of 82.2%, and a top-5 accuracy of 87.0%.

To better understand which minimum chronology length is required to apply this method to timber tracing, the experiments were repeated for reduced chronology lengths. We created centred subsets of the unknown chronologies for a range of lengths and used them for prediction instead of full chronology sequences (Fig. 5). Notably, subsetting the

unknown chronologies to a length of 30 years in this manner still resulted in a top-1 accuracy of 52.3% and a top-5 accuracy of 74.7%.

3.5. Prediction certainty

Fig. 6 shows the mean prediction interval (PI) width according to the method of quantile regression forests (Meinshausen, 2006) to assess spatial patterns of (un)certainty of the modelled reference chronologies. If the predicted location of a TRW dataset falls within a region of high relative uncertainty, the result is expected to be less trustworthy than when it falls in an area of high certainty. Additionally, if there is a claimed location for the TRW dataset which falls in an area of high relative uncertainty, that location is only weakly supported by the reference chronologies, which may limit trust in the claimed provenance.
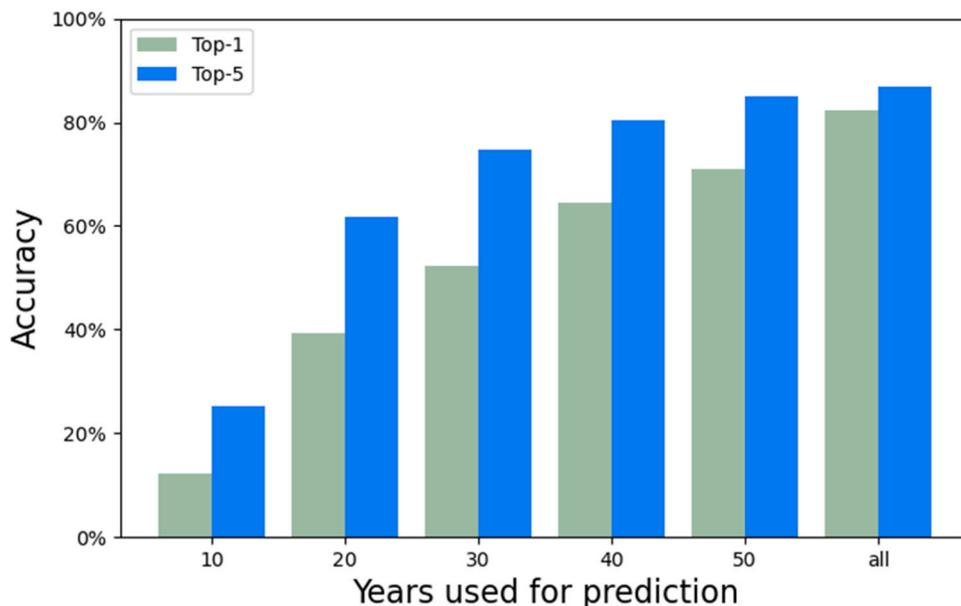
Fig. 5. Accuracies for reduced chronology lengths ('Both unknown scenario').
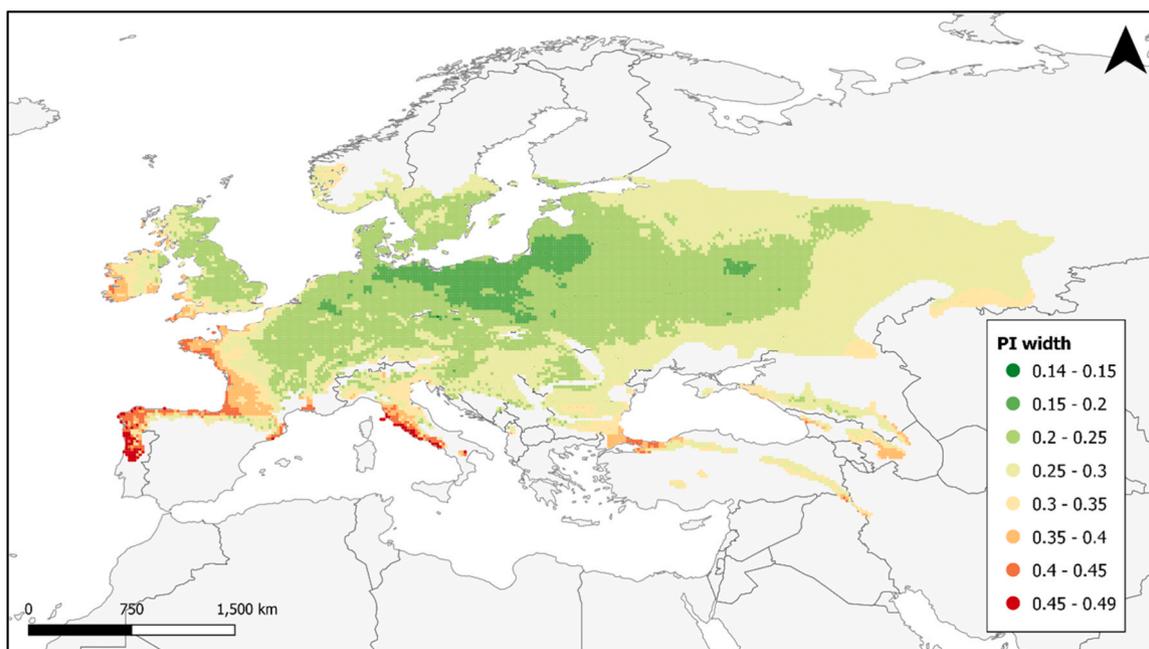

Fig. 6. Spatial distribution of mean PI width (unitless as a result of the detrending) of modelled reference chronologies. Areas with insufficiently represented climate data in the training set exhibit increased uncertainty. Note that this example uses all chronologies for training the RF, while for the LOOCV in the case study subsets are used each iteration.

Furthermore, our findings suggest that the top-1 R^2 value is a predictor for the accuracy of the predicted provenance. Chronologies where the ground truth is in the top-5 of predictions have a significantly higher top-1 R^2 prediction value than chronologies where the ground truth is outside the top-5 of predictions (ANOVA, $F_{1,103}=23.446$, $P<0.001$) (Fig. 7). The value of our ranking metric thus also indicates (un)certainty, and could allow the filtering of results through setting a threshold. For instance, a threshold in this case might involve rejecting any result with an R^2 value below 0.35.

4. Discussion

The results of the case study suggest the proposed framework for dendroprovenancing to be effective in providing a data-driven approach to identify or verify the provenance of timber for 20th and 21st century wood chronologies. Additionally, it provides model-based metrics of prediction quality.

4.1. Case study: limitations of the study

While analysing the 'Year unknown' scenario, 92% of the chronologies were correctly dated by our approach. Out of the nine that were
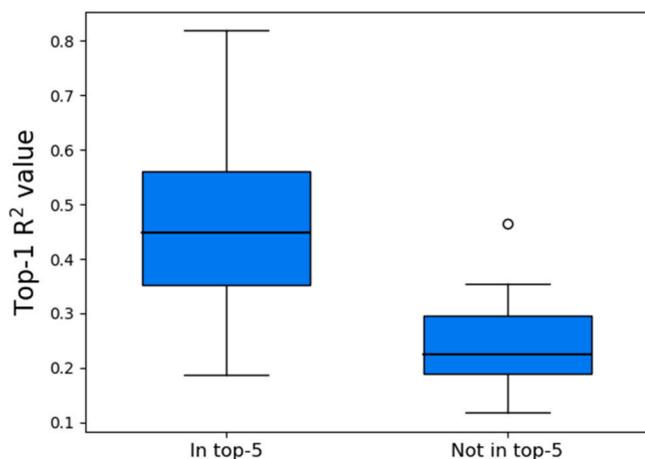
Fig. 7. Boxplot of top-1 R^2 values, grouped by whether the ground truth location was in the top-5 of results or outside of the top-5.

not dated correctly, we suspect that at least some of the TRW datasets were misdated in the ITRDB. More specifically, these suspicions concern lith042, swit264, one or more from the group czec034, czec035 and czec028, and germ011, in descending order of certainty.

Misdated datasets in the ITRDB would lead to an underestimation of the effectiveness of our approach since we assumed the dating and location in the ITRDB to be the ground truth, and not misdated or mislocated. If correct data are used greater accuracy can be achieved for all three scenarios. This is because the same chronologies were, as to be expected, often also mislocated in the 'Location unknown' scenario and mostly fell outside the top-5 of the 'Both unknown' scenario.

We also note the strong spatial clustering of the 107 TRW dataset locations (Fig. 2). An unknown TRW dataset from a location within the species distribution but remote from the known TRWs may have a smaller chance to be correctly dated/located than one that is located closer to the known TRWs. However, this effect depends on the predictive strength of the used features (weather and soil) and on the representativeness of the feature space as sampled by the known TRW dataset (de Bruin et al., 2022).

As stated previously, the chosen species pedunculate oak (*Quercus robur* L.) is an optimal example, both in terms of TRW variation response to the available gridded data and in terms of TRW dataset availability. The method might not work that well for species for which data availability is lower, or climate-growth relationships is weaker, which deserves further research. In addition, the approach outlined and illustrated here may require a chronology of at least >30 years length. This requires the availability of timber samples that have a sufficient number of countable rings. Additionally, it requires the researched timber to not come from a mix of sources; all sampled timber must come from the same location and species. If not, it will not be possible to construct a usable chronology, or to derive a reliable origin.

Furthermore, our approach requires some other conditions to be met: (1) the species in question must have at least some available reference tree-ring chronologies, hence (2) it must form annual growth rings with a limited number of double- or missing rings, (3) its TRW variation must show a sensitivity to environmental variability, and (4) these environmental variables should be available as gridded products with coverage for the region and timespan of interest.

Consequently, our approach may not work well for tropical species due to the lower seasonality in some tropical climates (Zuidema et al., 2022). Here, other provenancing approaches may be more suitable, such as DNA-based provenancing, stable isotope analysis and multi-element analyses. An additional advantage for DNA-based provenancing in the tropics is that the genetic displacement and mixing of specimens are likely to be much lower due to the lower influence of genetic material

displacement by past anthropogenic activities compared to i.e. Europe.

In general, however, a combination of multiple data-driven timber provenancing approaches is likely most effective, as they rely on different, independent data and have different strengths and weaknesses, and the accuracies of combined methods to timber provenancing likely exceed those of the methods independently (Dormontt et al., 2015; Low et al., 2022).

4.2. General framework: potential areas of improvement and further exploration

Two notable areas of potential improvement in the methods became evident during the case study: 1) the use of a finer resolution of the point grid, capitalizing the fine resolution of the Soilgrids- and Worldclim data, and 2) increasing the number of trees in the random forest model. The latter can seem counterintuitive as the number of trees was already set to 500, but even small improvements can have compound effects as the regression results are combined into long chronologies. Both changes are likely to increase regression accuracy and thus improve provenancing and dating accuracy. However, both also come at a cost of increased computing time and storage capacity. For example, increasing the grid spacing by a factor two would quadruple the computation cost and data storage requirements.

The presented framework offers several potential points for further exploration, as there are many different possible approaches to various steps of the process. Important to note is that the case study only highlights one possible approach to the implementation of the general framework. For example, the detrending method of TRW sequences can be adjusted to any other desired detrending method. The same applies to the chosen method for combining individual tree growth time series into population-wide growth chronologies, the evaluation metric for matching between unknown chronology and modelled reference chronologies, and the type of auxiliary gridded data used. For example, alternatives to CRU TS can be considered, but also other variables available in CRU TS and/or Worldclim, or even adding a completely different gridded data source that might influence TRW variation.

In the case study, we chose random forest for its high out-of-the-box performance, and because tuning of hyperparameters was not part of this research. However, in the future it could be replaced by a neural network approach or a gradient-boosting approach like LightGBM (Ke et al., 2017), XGBoost (Chen et al., 2015), or Catboost (Prokhorenkova et al., 2018), which generally outperform RF implementations (Bentéjac et al., 2021) if well-tuned to the problem.

Another particularly interesting area for exploration is the potential to incorporate data from other species during model training. For instance, species within the same family or genus as the unknown sample may help improve prediction accuracy, as such closely related species may have a common signal with the species of interest. Alternatively, it could turn out to be more effective to create separate models for each individual species as we did here in the case study. This requires further research and exploration.

4.3. General framework: advantages and disadvantages

Compared to traditional dendroprovenancing, our framework is less affected by spatial and temporal data gaps. It uses empirical relationships with environmental covariates to achieve spatiotemporal continuity. Other approaches, such as the works of Visser (2021) and Daly (2007), also partly address the challenges of spatial and temporal gaps in dendroprovenancing. Visser's approach, leveraging relations between chronologies in networks, may introduce additional tree ring width data with unknown original provenance into the analysis, which could fill spatiotemporal gaps when combined with archaeological interpretation. Similarly, Daly's work explores the benefits of using different levels, such as site chronologies and single trees, to overcome these gaps. Notably, both previous studies rely on incorporating additional tree ring

width data to address spatiotemporal gaps, which is not the case for our approach.

Another major advantage is the consistent format that the modelled reference chronologies are returned in, all having the same temporal range, eliminating issues with evaluation metrics and different overlap lengths between target and reference chronologies (Bridge and Fowler, 2019). A possible disadvantage could be that our modelled reference chronologies lose information when compared to non-modelled reference chronologies if the RF models fail to explain all TRW variation. On the other hand, one could also argue that the modelled reference chronologies are less prone to erroneous values and data artefacts than non-modelled reference chronologies.

Compared to recently emerging technologies, such as DNA-based provenancing approaches and stable isotope analysis, the main advantages of our approach are (1) the large set of already available reference data in the form of TRW datasets, which can be utilised to model proxy data across the entire species range, and (2) the temporal depth of chronologies, providing a long barcode that can be used to trace origins (compared to single point data in isoscapes; e.g. Watkinson et al., 2020).

For DNA-based provenancing, data is not yet widely available for the large majority of traded timber species (Gasson et al., 2021), nor does it seem feasible to model DNA to get coverage across the species range. Geolocated samples from the entire species range would first need to be gathered and analysed to provenance timber of an unknown source (Gasson et al., 2021). DNA-based provenancing therefore currently mostly seems suitable for confirming a claimed source location, in which case both the claimed source location and the timber can be sampled, but not (yet) for provenancing timber with a completely unknown source.

Furthermore, DNA-based provenancing could be more prone to problems caused by displaced trees and the mixing of genetic material, which is potentially an issue, especially in Europe (Bradshaw, 2004). If trees have been introduced to a particular site in the past while their genetic material originates from a different part of the continent, provenancing based on their DNA may point to that other part of the continent. The higher the fraction of trees of a species that have been displaced in the past, the more this would complicate DNA-based provenancing, as making reliable reference datasets becomes difficult if not impossible due to the genetic mixing. This is not an issue for dendroprovenancing as long as the trees grow their entire lifespan on the location they were cut, since TRW patterns are formed locally and are mostly dependent on meteorological conditions.

4.4. General framework: additional applications

We found that some of the chronologies were likely misdated in the ITRDB. This could mean that the framework may also be used as a means of database quality assurance, finding misdated or mislocated TRW datasets in a database, but also for verifying newly supplied data before adding it to a database.

Additionally, the regression model may also be used for studying general tree growth responses to isolated changes, such as the impact of an increase in winter temperature for a certain location, for example by performing ablation studies or through careful interpretation of RF feature importance (Appendix D).

5. Conclusion

The proposed framework provides an enhanced approach to dendroprovenancing 20th and 21st century wood samples, allowing verification of the indicated provenance of timber. The framework improves spatial and temporal coverage without the need for additional data collection, increasing the potential of dendroprovenancing as a useful tool for forensic timber identification. The framework is flexible, allowing for different approaches in several steps of the process, and also has the potential to provide valuable insights into tree growth patterns and responses to environmental changes. Further research is required to explore the potential application to species beyond the pedunculate oak example demonstrated in this work and geographical limitations to application of the framework. Further work is also needed for optimising components such as data pre-processing, the used similarity metric and machine learning approach. The findings of this study are relevant for the ongoing efforts to combat illegal logging and trade in illegal timber products and contribute to the development of tools that can help to control this global issue.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data sources are described at https://github.com/mvansluijs/Enhanced-Dendroprovenancing.

Appendix A. ITRDB dataset references

Baillie, M.G.L. (2002–04–26): NOAA/WDS Paleoclimatology - Baillie - Raemills - QURO - ITRDB BRIT7. NOAA National Centers for Environmental Information. https://doi.org/10.25921/cys0-hg76. Accessed 2022–10–03.

Bert, D.; Lasnier, J.B.; Capdevielle, X.; Desprez-Loustau, M.L. (2015–09–08): NOAA/WDS Paleoclimatology - Bert - Biscarosse B1 Pontenx P8 Landes de Gascogne - QURO - ITRDB FRAN048. NOAA National Centers for Environmental Information. https://doi.org/10.25921/ac6j-fc45. Accessed 2022–10–03.

Bräker, O.U.; Zingg, A. (2010–06–10): NOAA/WDS Paleoclimatology - Bräker - Boulex - QURO - ITRDB SWIT318. NOAA National Centers for Environmental Information. https://doi.org/10.25921/td0f-mp66. Accessed 2022–10–03.

Bräker, O.U.; Zingg, A. (2010–06–10): NOAA/WDS Paleoclimatology - Bräker - Frischenez - QURO - ITRDB SWIT243. NOAA National Centers for Environmental Information. https://doi.org/10.25921/gkyk-e348. Accessed 2022–10–03.

Bräker, O.U.; Zingg, A. (2010–06–10): NOAA/WDS Paleoclimatology - Bräker - Haslenau - QURO - ITRDB SWIT251. NOAA National Centers for Environmental Information. https://doi.org/10.25921/y06t-pn35. Accessed 2022–10–03.

Bräker, O.U.; Zingg, A. (2010–06–10): NOAA/WDS Paleoclimatology - Bräker - Jussy JU - QURO - ITRDB SWIT264. NOAA National Centers for Environmental Information. https://doi.org/10.25921/p06e-hy56. Accessed 2022–10–03.

Bräker, O.U.; Zingg, A. (2010–06–10): NOAA/WDS Paleoclimatology - Bräker - Saueinschlag - QURO - ITRDB SWIT216. NOAA National Centers for Environmental Information. https://doi.org/10.25921/hj50-mn06. Accessed 2022–10–03.

Bräker, O.U.; Zingg, A. (2010–06–10): NOAA/WDS Paleoclimatology - Bräker - unt.Bremgartenstrasse - QURO - ITRDB SWIT233. NOAA National Centers for Environmental Information. https://doi.org/10.25921/jh7p-2j57. Accessed 2022–10–03.

Eckstein, D. (2002–04–26): NOAA/WDS Paleoclimatology - Eckstein - Hamburg - QURO - ITRDB GERM011. NOAA National Centers for

Environmental Information. https://doi.org/10.25921/zjmb-jv68. Accessed 2022–10–03.

Eckstein, D. (2002–04–26): NOAA/WDS Paleoclimatology - Eckstein - Oldenburg - QURO - ITRDB GERM6. NOAA National Centers for Environmental Information. https://doi.org/10.25921/tntt-q680. Accessed 2022–10–03.

Eckstein, D. (2002–04–26): NOAA/WDS Paleoclimatology - Eckstein - Schleswig - QURO - ITRDB GERM012. NOAA National Centers for Environmental Information. https://doi.org/10.25921/az18-vr59. Accessed 2022–10–03.

Fischer, S. (2014–02–27): NOAA/WDS Paleoclimatology - Fischer - Niederrheinisches Tiefland drn01 - QURO - ITRDB GERM204. NOAA National Centers for Environmental Information. https://doi.org/10.25921/gx77-vt29. Accessed 2022–10–03.

Fischer, S. (2014–02–27): NOAA/WDS Paleoclimatology - Fischer - Nordrhein-Westfalen/Kottenforst-Ville drb12 - QURO - ITRDB GERM168. NOAA National Centers for Environmental Information. https://doi.org/10.25921/jry7-tw08. Accessed 2022–10–03.

Fischer, S. (2014–02–27): NOAA/WDS Paleoclimatology - Fischer - Nordrhein-Westfalen/Kottenforst-Ville drb13 - QURO - ITRDB GERM169. NOAA National Centers for Environmental Information. https://doi.org/10.25921/wy4d-zk41. Accessed 2022–10–03.

Fischer, S. (2014–02–27): NOAA/WDS Paleoclimatology - Fischer - Nordrhein-Westfalen/Kottenforst-Ville drb31 - QURO - ITRDB GERM170. NOAA National Centers for Environmental Information. https://doi.org/10.25921/0319–0d56. Accessed 2022–10–03.

Fischer, S. (2014–02–27): NOAA/WDS Paleoclimatology - Fischer - Nordrhein-Westfalen/Münsterland drm04 - QURO - ITRDB GERM203. NOAA National Centers for Environmental Information. https://doi.org/10.25921/tn2s-8e71. Accessed 2022–10–03.

Gutiérrez, E. (2019–08–19): NOAA/WDS Paleoclimatology - Gutiérrez - Bromarv - QURO - ITRDB FINL075. NOAA National Centers for Environmental Information. https://doi.org/10.25921/6s3g-zj44. Accessed 2022–10–03.

Gutiérrez, E. (2019–08–19): NOAA/WDS Paleoclimatology - Gutiérrez - Lochwood - QURO - ITRDB BRIT072. NOAA National Centers for Environmental Information. https://doi.org/10.25921/j2wq-p403. Accessed 2022–10–03.

Gutiérrez, E. (2019–08–19): NOAA/WDS Paleoclimatology - Gutiérrez - Niepolomice, Gibiel - QURO - ITRDB POLA039. NOAA National Centers for Environmental Information. https://doi.org/10.25921/86af-7n32. Accessed 2022–10–03.

Gutiérrez, E. (2019–08–19): NOAA/WDS Paleoclimatology - Gutiérrez - Rennes - QURO - ITRDB FRAN050. NOAA National Centers for Environmental Information. https://doi.org/10.25921/ddps-se70. Accessed 2022–10–03.

Gutiérrez, E. (2019–08–19): NOAA/WDS Paleoclimatology - Gutiérrez - Woburn Abbey - QURO - ITRDB BRIT071. NOAA National Centers for Environmental Information. https://doi.org/10.25921/hgxa-mb39. Accessed 2022–10–03.

Hogholt, I.; Passini, E.; Vacchiano, G.; Vaglieri, T. (2019–06–18): NOAA/WDS Paleoclimatology - Hogholt - Morimondo - QURO - ITRDB ITAL052. NOAA National Centers for Environmental Information. https://doi.org/10.25921/kdbr-wp41. Accessed 2022–10–03.

Jevsenak, J.; Levanic, T. (2021–01–09): NOAA/WDS Paleoclimatology - Jevsenak - Mlace - QURO - ITRDB SLOV004. NOAA National Centers for Environmental Information. https://doi.org/10.25921/wecx-zt88. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Babtai - QURO - ITRDB LITH026. NOAA National Centers for Environmental Information. https://doi.org/10.25921/0xtn-4n41. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Birzai - QURO - ITRDB LITH027. NOAA National Centers for Environmental Information. https://doi.org/10.25921/mk5t-s809. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Gelvonai - QURO - ITRDB LITH030. NOAA National Centers for Environmental Information. https://doi.org/10.25921/amx0-jt43. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Kaltinenai - QURO - ITRDB LITH031. NOAA National Centers for Environmental Information. https://doi.org/10.25921/vqn5-0k16. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Vidukle - QURO - ITRDB LITH033. NOAA National Centers for Environmental Information. https://doi.org/10.25921/c7m8-p723. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Zalioji - QURO - ITRDB LITH034. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rvxg-m981. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Kurtuvenai - QURO - ITRDB LITH036. NOAA National Centers for Environmental Information. https://doi.org/10.25921/jrwn-ht54. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Naukaimis - QURO - ITRDB LITH038. NOAA National Centers for Environmental Information. https://doi.org/10.25921/ma38-r286. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Pagegiai - QURO - ITRDB LITH039. NOAA National Centers for Environmental Information. https://doi.org/10.25921/vqzs-5v13. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Plateliai - QURO - ITRDB LITH040. NOAA National Centers for Environmental Information. https://doi.org/10.25921/wywy-jb49. Accessed 2022–10–03.

Kairaitis, J.; Sohar, K.; Vitas, A. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Vezaiciai - QURO - ITRDB LITH045. NOAA National Centers for Environmental Information. https://doi.org/10.25921/dfqp-gr50. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–06): NOAA/WDS Paleoclimatology - Kairaitis - Anyksciai - QURO - ITRDB LITH025. NOAA National Centers for Environmental Information. https://doi.org/10.25921/f617-da50. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Bukta - QURO - ITRDB LITH028. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rjkf-gz86. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Dukstos - QURO - ITRDB LITH029. NOAA National Centers for Environmental Information. https://doi.org/10.25921/01jq-wp04. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–13): NOAA/WDS Paleoclimatology - Kairaitis - Kedainiai - QURO - ITRDB LITH032. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rag2-y729. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Aukstadvaris - QURO - ITRDB LITH035. NOAA National Centers for Environmental Information. https://doi.org/10.25921/bcc3-5r43. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Naujoji Uta - QURO - ITRDB LITH037. NOAA National Centers for Environmental Information. https://doi.org/10.25921/6902–4d98. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Spirakiai - QURO - ITRDB LITH041. NOAA National Centers for Environmental Information. https://doi.org/10.25921/kq5c-th39. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Stakliskes - QURO - ITRDB LITH042. NOAA National

Centers for Environmental Information. https://doi.org/10.25921/d3xh-rq53. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Subartonys - QURO - ITRDB LITH043. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rpwr-9s23. Accessed 2022–10–03.

Kairaitis, J.; Vitas, A.; Sohar, K. (2017–02–17): NOAA/WDS Paleoclimatology - Kairaitis - Troskunai - QURO - ITRDB LITH044. NOAA National Centers for Environmental Information. https://doi.org/10.25921/j4c2-4a44. Accessed 2022–10–03.

Khasanov, B. (2020–05–04): NOAA/WDS Paleoclimatology - Khasanov - Kozelsk - QURO - ITRDB RUSS282. NOAA National Centers for Environmental Information. https://doi.org/10.25921/76em-c470. Accessed 2022–10–03.

Khasanov, B. (2022–04–20): NOAA/WDS Paleoclimatology - Khasanov - Dyakovskiy Forest - DL - QURO - ITRDB RUS328. NOAA National Centers for Environmental Information. https://doi.org/10.25921/sym9-vb03. Accessed 2022–10–03.

Khasanov, B. (2022–04–20): NOAA/WDS Paleoclimatology - Khasanov - Dzerzhinskiy - QURO - ITRDB RUS353. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rnr6-8v18. Accessed 2022–10–03.

Khasanov, B. (2022–04–20): NOAA/WDS Paleoclimatology - Khasanov - Krapivna - TZ - QURO - ITRDB RUS377. NOAA National Centers for Environmental Information. https://doi.org/10.25921/2yp0-gr23. Accessed 2022–10–03.

Khasanov, B. (2022–04–20): NOAA/WDS Paleoclimatology - Khasanov - Leontievo - KS - QURO - ITRDB RUS341. NOAA National Centers for Environmental Information. https://doi.org/10.25921/d4vp-6332. Accessed 2022–10–03.

Maessen, P.P.Th.M. (1997–01–06): NOAA/WDS Paleoclimatology - Maessen - Bijvanck All Trees - QURO - ITRDB NETH023. NOAA National Centers for Environmental Information. https://doi.org/10.25921/drwa-yf09. Accessed 2022–10–03.

Maessen, P.P.Th.M. (1997–01–06): NOAA/WDS Paleoclimatology - Maessen - Leudal - QURO - ITRDB NETH026. NOAA National Centers for Environmental Information. https://doi.org/10.25921/bphw-pn86. Accessed 2022–10–03.

Maessen, P.P.Th.M. (1997–01–06): NOAA/WDS Paleoclimatology - Maessen - Mattemburgh All Trees - QURO - ITRDB NETH028. NOAA National Centers for Environmental Information. https://doi.org/10.25921/c1nz-hr89. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Arcen - QURO - ITRDB NETH020. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rt6g-xz61. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Bijvanck Dominant Trees - QURO - ITRDB NETH024. NOAA National Centers for Environmental Information. https://doi.org/10.25921/yemn-td19. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Bijvanck Sub-dominant Trees - QURO - ITRDB NETH022. NOAA National Centers for Environmental Information. https://doi.org/10.25921/wdat-0240. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Landgoed Hillenraad - QURO - ITRDB NETH025. NOAA National Centers for Environmental Information. https://doi.org/10.25921/yp8x-w009. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Mattemburgh Dominant Trees - QURO - ITRDB NETH029. NOAA National Centers for Environmental Information. https://doi.org/10.25921/v12s-c677. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Mattemburgh Sub-dominant Trees - QURO - ITRDB NETH030. NOAA National Centers for Environmental Information. https://doi.org/10.25921/1g1q-n627. Accessed 2022–10–03.

Maessen, P.P.Th.M. (2002–04–26): NOAA/WDS Paleoclimatology - Maessen - Oud Groevenbeek - QURO - ITRDB NETH032. NOAA National Centers for Environmental Information. https://doi.org/10.25921/752e-ca06. Accessed 2022–10–03.

Morgan, R. (1996–05–08): NOAA/WDS Paleoclimatology - Morgan - Towy Valley - QURO - ITRDB BRIT10. NOAA National Centers for Environmental Information. https://doi.org/10.25921/te0g-w736. Accessed 2022–10–03.

Neuwirth, B. (2014–02–27): NOAA/WDS Paleoclimatology - Neuwirth - Nordrhein-Westfalen/Jülich drj01 - QURO - ITRDB GERM195. NOAA National Centers for Environmental Information. https://doi.org/10.25921/1ns8-5561. Accessed 2022–10–03.

Neuwirth, B. (2014–02–27): NOAA/WDS Paleoclimatology - Neuwirth - Nordrhein-Westfalen/Jülich drj03 - QURO - ITRDB GERM196. NOAA National Centers for Environmental Information. https://doi.org/10.25921/1n8z-4087. Accessed 2022–10–03.

Neuwirth, B. (2014–02–27): NOAA/WDS Paleoclimatology - Neuwirth - Nordrhein-Westfalen/Köln-Bonner Bucht drk03 - QURO - ITRDB GERM199. NOAA National Centers for Environmental Information. https://doi.org/10.25921/s1qq-ax94. Accessed 2022–10–03.

Nola, P. (2002–04–26): NOAA/WDS Paleoclimatology - Nola - Corte Brugnatella (Piacenza) - QURO - ITRDB ITAL019. NOAA National Centers for Environmental Information. https://doi.org/10.25921/vchc-vn49. Accessed 2022–10–03.

Nola, P. (2002–04–26): NOAA/WDS Paleoclimatology - Nola - Gambolo (Pavia) - QURO - ITRDB ITAL021. NOAA National Centers for Environmental Information. https://doi.org/10.25921/x13j-bt52. Accessed 2022–10–03.

Nola, P. (2002–04–26): NOAA/WDS Paleoclimatology - Nola - Pavia - QURO - ITRDB ITAL020. NOAA National Centers for Environmental Information. https://doi.org/10.25921/wjpg-gq37. Accessed 2022–10–03.

Pilcher, J.R. (1996–05–08): NOAA/WDS Paleoclimatology - Pilcher - Glen of the Downs - QURO - ITRDB BRIT006. NOAA National Centers for Environmental Information. https://doi.org/10.25921/zxb7-4972. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Blickling - QURO - ITRDB BRIT011. NOAA National Centers for Environmental Information. https://doi.org/10.25921/bk59-v407. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Chambord - QURO - ITRDB FRAN001. NOAA National Centers for Environmental Information. https://doi.org/10.25921/866c-yr25. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Fontainebleau - QURO - ITRDB FRAN003. NOAA National Centers for Environmental Information. https://doi.org/10.25921/g06k-g948. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Foret de Guines - QURO - ITRDB FRAN004. NOAA National Centers for Environmental Information. https://doi.org/10.25921/6tqa-xv12. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Foret d'Halatte (Both) - QURO - ITRDB FRAN005. NOAA National Centers for Environmental Information. https://doi.org/10.25921/27wk-aj23. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Foret d'Halatte (Wet) - QURO - ITRDB FRAN007. NOAA National Centers for Environmental Information. https://doi.org/10.25921/3vhg-hv26. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Glenluce - QURO - ITRDB BRIT005. NOAA National Centers for Environmental Information. https://doi.org/10.25921/24bv-2a13. Accessed 2022–10–03.

Pilcher, J.R. (2002–04–26): NOAA/WDS Paleoclimatology - Pilcher - Oakley - QURO - ITRDB BRIT4. NOAA National Centers for Environmental

Information. https://doi.org/10.25921/kwvd-rw67. Accessed 2022–10–03.

Pilcher, J.R. (2015–05–06): NOAA/WDS Paleoclimatology - Pilcher - Ludlow - QURO - ITRDB BRIT010. NOAA National Centers for Environmental Information. https://doi.org/10.25921/nmpj-xx18. Accessed 2022–10–03.

Scharnweber, T. (2015–09–23): NOAA/WDS Paleoclimatology - Scharnweber - Blumberg - QURO - ITRDB GERM218. NOAA National Centers for Environmental Information. https://doi.org/10.25921/k6tv-b540. Accessed 2022–10–03.

Scharnweber, T. (2015–09–23): NOAA/WDS Paleoclimatology - Scharnweber - Carlshof - QURO - ITRDB GERM220. NOAA National Centers for Environmental Information. https://doi.org/10.25921/a6fb-d546. Accessed 2022–10–03.

Scharnweber, T. (2015–09–23): NOAA/WDS Paleoclimatology - Scharnweber - Holzkrug - QURO - ITRDB GERM222. NOAA National Centers for Environmental Information. https://doi.org/10.25921/0hpg-8433. Accessed 2022–10–03.

Schmidt, B. (1996–05–08): NOAA/WDS Paleoclimatology - Schmidt - Odenthal Abt. 17 - QURO - ITRDB GERM300. NOAA National Centers for Environmental Information. https://doi.org/10.25921/a6ky-9978. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Barkhausen - QURO - ITRDB GERM004. NOAA National Centers for Environmental Information. https://doi.org/10.25921/99ve-2f10. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Hessisch-Oldendorf - QURO - ITRDB GERM005. NOAA National Centers for Environmental Information. https://doi.org/10.25921/z20y-vw84. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Hesslingen - QURO - ITRDB GERM006. NOAA National Centers for Environmental Information. https://doi.org/10.25921/n3ya-x296. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Koln Abt. 48 - QURO - ITRDB GERM003. NOAA National Centers for Environmental Information. https://doi.org/10.25921/zjz7-kf86. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Odenthal Abt. 13 - QURO - ITRDB GERM001. NOAA National Centers for Environmental Information. https://doi.org/10.25921/2fgr-vp75. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Schaumburg - QURO - ITRDB GERM009. NOAA National Centers for Environmental Information. https://doi.org/10.25921/3j68-hp32. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Schwalenberger Wald - QURO - ITRDB GERM008. NOAA National Centers for Environmental Information. https://doi.org/10.25921/z2rv-9380. Accessed 2022–10–03.

Schmidt, B. (2002–04–26): NOAA/WDS Paleoclimatology - Schmidt - Varenholz - QURO - ITRDB GERM007. NOAA National Centers for Environmental Information. https://doi.org/10.25921/9g8j-gh61. Accessed 2022–10–03.

Schneider, C.; Neuwirth, B.; Schneider, S.; Balanzategui, D.; Elsholz, S.; Fenner, D.; Meier, F.; Heinrich, I. (2022–05–11): NOAA/WDS Paleoclimatology - Schneider - Berlin Alpenrose den14 - QURO - ITRDB DEU314. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rqte-jb16. Accessed 2022–10–03.

Schneider, C.; Neuwirth, B.; Schneider, S.; Balanzategui, D.; Elsholz, S.; Fenner, D.; Meier, F.; Heinrich, I. (2022–05–11): NOAA/WDS Paleoclimatology - Schneider - Berlin Grünanlage Britz-Süd den09 - QURO - ITRDB DEU312. NOAA National Centers for Environmental Information. https://doi.org/10.25921/jkqj-hb17. Accessed 2022–10–03.

Schneider, C.; Neuwirth, B.; Schneider, S.; Balanzategui, D.; Elsholz, S.; Fenner, D.; Meier, F.; Heinrich, I. (2022–05–11): NOAA/WDS Paleoclimatology - Schneider - Berlin Teufelssee dek01 - QURO - ITRDB DEU303. NOAA National Centers for Environmental Information. https://doi.org/10.25921/m4ap-w933. Accessed 2022–10–03.

Schneider, C.; Neuwirth, B.; Schneider, S.; Balanzategui, D.; Elsholz, S.; Fenner, D.; Meier, F.; Heinrich, I. (2022–05–11): NOAA/WDS Paleoclimatology - Schneider - Berlin Volkspark Hasenheide den15 - QURO - ITRDB DEU315. NOAA National Centers for Environmental Information. https://doi.org/10.25921/vhfm-rg22. Accessed 2022–10–03.

Schneider, C.; Neuwirth, B.; Schneider, S.; Balanzategui, D.; Elsholz, S.; Fenner, D.; Meier, F.; Heinrich, I. (2022–05–11): NOAA/WDS Paleoclimatology - Schneider - Berlin Volkspark Hasenheide Northeast den16 - QURO - ITRDB DEU316. NOAA National Centers for Environmental Information. https://doi.org/10.25921/x5sc-1947. Accessed 2022–10–03.

Solomina, O. (2022–04–20): NOAA/WDS Paleoclimatology - Solomina - Selivanovo Krapivnenskiy Dendrariy - W01D - QURO - ITRDB RUS382. NOAA National Centers for Environmental Information. https://doi.org/10.25921/hk4f-mm29. Accessed 2022–10–03.

Solomina, O.; Matskovsky, V.V. (2022–04–20): NOAA/WDS Paleoclimatology - Solomina - Belogoriye Preserve Les Na Vorskle - V01D - QURO - ITRDB RUS378. NOAA National Centers for Environmental Information. https://doi.org/10.25921/xyjb-bf52. Accessed 2022–10–03.

Solomina, O.; Mikhalenko, V.N. (2022–04–20): NOAA/WDS Paleoclimatology - Solomina - Faschovka - T12D - QURO - ITRDB RUS368. NOAA National Centers for Environmental Information. https://doi.org/10.25921/xnh7-7d34. Accessed 2022–10–03.

Tumajer, J. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Roudnice 1 - QURO - ITRDB CZEC030. NOAA National Centers for Environmental Information. https://doi.org/10.25921/rgbp-he63. Accessed 2022–10–03.

Tumajer, J. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Roudnice 2 - QURO - ITRDB CZEC031. NOAA National Centers for Environmental Information. https://doi.org/10.25921/825c-hn39. Accessed 2022–10–03.

Tumajer, J. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Zaluzi - QURO - ITRDB CZEC033. NOAA National Centers for Environmental Information. https://doi.org/10.25921/ja5m-df31. Accessed 2022–10–03.

Tumajer, J.; Treml, V. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Myslivna - QURO - ITRDB CZEC029. NOAA National Centers for Environmental Information. https://doi.org/10.25921/b86b-ht53. Accessed 2022–10–03.

Tumajer, J.; Treml, V. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - U Houkvice - QURO - ITRDB CZEC028. NOAA National Centers for Environmental Information. https://doi.org/10.25921/vq9a-s483. Accessed 2022–10–03.

Tumajer, J.; Treml, V. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Upor - QURO - ITRDB CZEC032. NOAA National Centers for Environmental Information. https://doi.org/10.25921/m083-p380. Accessed 2022–10–03.

Tumajer, J.; Treml, V. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Zbytka - QURO - ITRDB CZEC034. NOAA National Centers for Environmental Information. https://doi.org/10.25921/weyj-8q15. Accessed 2022–10–03.

Tumajer, J.; Treml, V. (2021–10–15): NOAA/WDS Paleoclimatology - Tumajer - Zernov - QURO - ITRDB CZEC035. NOAA National Centers for Environmental Information. https://doi.org/10.25921/ry38-w148. Accessed 2022–10–03.

Tyers, I. (2004–04–30): NOAA/WDS Paleoclimatology - Tyers - Sheffield, Bingham Park Wood and Endcliffe Wood - QURO - ITRDB BRIT053. NOAA National Centers for Environmental Information. https://doi.org/10.25921/kx85-gn41. Accessed 2022–10–03.

(*continued*)

ITRDB code	Latitude	Longitude	First year	Last year	Sigma	Usable chronology length	Year known offset	Location unknown distance (km)	Both unknown result	Both unknown top-1 R^2
germ008	51.92	9.2	1830	1972	4	69	0	42.0	Top-1	0.399
germ009	52.32	9.03	1819	1969	4	69	0	107.6	Top-1	0.566
germ011	54	10	1340	1967	4	64	-3	320.0	Not in top-5	0.297
germ012	54.5	9.5	1800	1969	4	66	0	193.6	Top-1	0.358
germ168+germ169	50.6667	7.0333	1829	2005	4	102	0	7.8	Top-1	0.819
germ170	50.6833	7.0333	1846	2006	4	103	0	15.0	Top-1	0.776
germ195+germ196	50.9167	6.4167	1749	2005	4	102	0	41.9	Top-1	0.626
germ199	50.7833	6.8333	1848	2005	4	102	0	25.8	Top-1	0.453
germ203	51.9	7.8833	1867	2007	4	104	0	211.2	Top-1	0.260
germ204	51.3	6.7833	1845	2005	4	102	0	35.7	Top-1	0.457
germ218	53.223	14.13	1850	2009	4	106	0	84.7	Top-1	0.550
germ220	53.417	13.033	1872	2009	4	106	0	121.4	Top-1	0.277
germ222	53.485	10.915	1766	2009	4	106	0	100.1	Top-1	0.402
germ300	51.02	7.15	1806	1972	4	69	0	70.9	Top-1	0.635
germ6	53	8	1850	1973	3.8	69	0	195.4	Top-1	0.542
ital019	44.72	9.32	1779	1989	3.8	86	-3	660.7	Not in top-5	0.197
ital020	45.18	9.12	1875	1989	4	86	0	40.8	Top-1	0.205
ital021	45.27	9.05	1888	1989	4	86	0	55.8	Top-1	0.269
lith011	55.07	22.48	1878	2002	4	99	0	49.3	Top-1	0.485
lith025	55.46	25.23	1838	1994	4	91	0	39.7	Top-1	0.543
lith026	55.105	23.795	1803	1971	4	68	0	200.6	Top-1	0.576
lith027	56.245	24.79	1809	1974	4.2	71	0	127.8	Top-1	0.526
lith028	54.4	23.45	1852	1971	4	68	0	119.0	Top-1	0.365
lith029	54.84	24.955	1800	1994	4	91	0	133.5	Top-1	0.492
lith030	55.05	24.745	1800	1971	4	68	0	63.2	Top-1	0.722
lith031	55.565	22.395	1819	1971	4	68	0	45.5	Top-1	0.662
lith032	55.32	23.99	1815	1994	4	91	0	29.2	Top-1	0.498
lith033	55.435	22.93	1833	1971	4	68	0	384.8	Top-5	0.638
lith034	54.6	23.02	1840	1971	3.8	68	0	147.9	Top-1	0.472
lith035	54.585	24.55	1749	1969	4	66	0	125.6	Top-1	0.648
lith036	55.875	22.965	1837	1992	4	69	0	60.1	Top-1	0.560
lith037	54.535	23.8	1821	1997	4	94	0	29.9	Top-1	0.426
lith038	55.09	22.95	1841	1971	4	68	0	98.0	Top-1	0.486
lith039	55.19	21.89	1827	1971	4	68	0	71.4	Top-1	0.651
lith040	56.08	21.835	1851	1972	4	69	0	90.1	Top-1	0.614
lith041	55.8015	24.1715	1823	1994	4	91	0	103.2	Top-1	0.423
lith042	54.565	24.265	1731	1972	4	69	-3	2040.7	Not in top-5	0.465
lith043	54.195	24.185	1862	1996	4	93	0	58.8	Top-1	0.487
lith044	55.565	24.8195	1834	1994	4	91	0	64.1	Top-1	0.578
lith045	55.697	21.4815	1812	1971	4	68	0	38.6	Top-1	0.537
neth020	51.28	6.18	1883	1986	4	83	0	247.5	Top-1	0.465
neth026	51.25	5.93	1861	1986	4	83	0	69.2	Top-1	0.604
neth032	52.27	5.62	1879	1986	4	83	0	338.3	Top-5	0.266
pola005	54.3	18.55	1762	1986	4	82	0	194.7	Top-1	0.794
pola006	53.5	16	996	1986	4	82	0	187.3	Top-1	0.785
pola007	54.35	22.38	1871	1987	4	83	0	120.7	Top-1	0.207
pola008	52.7	23.65	1720	1985	4	81	0	104.7	Not in top-5	0.192
pola009	50.65	23.05	1782	1989	4	85	0	165.7	Top-1	0.348
pola010	54.1	16.15	1782	1987	4	83	0	532.7	Not in top-5	0.277
pola011	50.05	20.37	1792	1986	4	82	0	34.9	Top-1	0.512
pola012	52.27	16.8	1836	1987	4	83	0	39.1	Top-1	0.272
pola013	54.08	23.02	1861	1987	4	83	0	152.6	Top-1	0.409
pola014	53.08	18.55	1713	1987	4	83	-3	58.6	Top-1	0.302
pola015	52.3	20.98	1690	1985	4	81	0	299.7	Top-5	0.299
pola016	53.95	14.5	1554	1987	4	83	0	137.3	Top-1	0.300
pola017	51.25	17.17	1727	1987	4	83	0	156.1	Top-1	0.321
pola018	51.87	15.57	1774	1987	4	83	0	336.7	Top-5	0.227
pola039	50.12	20.38	1663	2003	4	100	0	18.0	Top-1	0.373
rus328	50.7299	46.6695	1908	2008	4	97	0	459.4	Top-5	0.229
rus341	58.1505	44.4952	1923	2012	4	86	0	345.9	Not in top-5	0.189
rus353	54.7323	36.0002	1806	2014	4	111	0	232.1	Top-1	0.301
rus368	52.46	39.69	1816	2014	4	111	0	174.4	Top-1	0.367
rus377	53.9781	37.1162	1770	2014	4	111	0	88.4	Top-1	0.353
rus378	50.604055	35.981444	1732	2014	4	111	0	650.4	Not in top-5	0.117
rus382	53.98837	37.25537	1809	2014	4	111	0	176.7	Top-1	0.407
russ282	53.91	35.825	1823	2015	4	112	0	64.5	Top-1	0.448
slov004	46.307071	15.509417	1848	2012	4	109	0	223.6	Not in top-5	0.147
swit216	47.1333	7.3833	1886	1995	4	92	0	47.1	Top-1	0.510
swit233	47.3925	8.3986	1927	1995	4	65	0	75.1	Top-1	0.423
swit243	46.9497	7.1628	1905	1991	4	83	0	33.1	Top-1	0.430
swit251	47.5828	9.2742	1932	1991	4	56	0	112.8	Top-1	0.386
swit264	46.2358	6.2669	1916	1998	4	79	-1	2233.1	Not in top-5	0.240
swit318	46.7844	6.9247	1938	1991	4	50	0	70.6	Top-1	0.437

Appendix C. Combined and removed TRW datasets

ITRDB code	Action	Reason
brit10	Removed	Not located on land
fran005 & fran007	Combined	Geographically very close
fran048	Removed	Temporally too short
germ012l	Removed	Contains only early- and/or latewood measurements
germ168 & germ169	Combined	Geographically very close
germ195 & germ196	Combined	Geographically very close
ital052	Removed	Internal quality issues detected during chronology formation
lith011e	Removed	Contains only early- and/or latewood measurements
lith011l	Removed	Contains only early- and/or latewood measurements
neth022	Removed	Stated in correlation stats that they are not useful; too many problems/flags/misdated samples
neth023	Removed	Stated in correlation stats that they are not useful; too many problems/flags/misdated samples
neth024	Removed	Stated in correlation stats that they are not useful; too many problems/flags/misdated samples
neth025	Removed	Internal quality issues detected during chronology formation
neth028	Removed	Stated in correlation stats that they are not useful; too many problems/flags/misdated samples
neth029	Removed	Stated in correlation stats that they are not useful; too many problems/flags/misdated samples
neth030	Removed	Stated in correlation stats that they are not useful; too many problems/flags/misdated samples

Appendix D. Feature importances

Feature importance analysis should be conducted with care as most features are strongly correlated. As such, their importance may be spread across multiple correlated features, complicating individual feature importance analysis. Here, we present the total feature importance based on the mean decrease in impurity for the large feature groups:

	Total feature importance
Meteorological data (temperature, precipitation)	0.955
Soil data	0.021
Coordinates (latitude, longitude)	0.004
One-hot encoding	0.018
Year	0.002
Total	**1**

Similarly, the meteorological data can be grouped into one importance value for each month:

Previous year	January	0.036
	February	0.031
	March	0.037
	April	0.036
	May	0.043
	June	0.039
	July	0.044
	August	0.050
	September	0.037
	October	0.044
	November	0.039
	December	0.032
Growing year	January	0.032
	February	0.039
	March	0.034
	April	0.040
	May	0.048
	June	0.062
	July	0.046
	August	0.049
	September	0.036
	October	0.038
	November	0.032
	December	0.032
	Total	**0.955**

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, 1163–1177.

Akhmetzyanov, L., Copini, P., Sass-Klaassen, U., Schroeder, H., de Groot, G.A., Laros, I., Daly, A., 2020a. DNA of centuries-old timber can reveal its origin. Sci. Rep. 10, 20316.

Akhmetzyanov, L., Sánchez-Salguero, R., García-González, I., Buras, A., Dominguez-Delmás, M., Mohren, F., den Ouden, J., Sass-Klaassen, U., 2020b. Towards a new approach for dendroprovenancing pines in the Mediterranean Iberian Peninsula. Dendrochronologia 60, 125688.

Babst, F., Bodesheim, P., Charney, N., Friend, A.D., Girardin, M.P., Klesse, S., Moore, D. J., Seftigen, K., Björklund, J., Bouriaud, O., 2018. When tree rings go global: challenges and opportunities for retro-and prospective insight. Quat. Sci. Rev. 197, 1–20.

Bentéjac, C., Csörgő, A., Martínez-Muñoz, G., 2021. A comparative analysis of gradient boosting algorithms. Artif. Intell. Rev. 54, 1937–1967.

Boeschoten, L.E., Vlam, M., Sass-Klaassen, U., Meyer-Sand, B.R.V., Boom, A., Bouka, G.U. D., Ciliane-Madikou, J.C.U., Obiang, N.L.E., Guieshon-Engongoro, M., Loumeto, J.J., Mbika, D.-mM.F., Moundounga, C.G., Ndangani, R.M.D., Bourobou, D.N., van der Sleen, P., Tassiamba, S.N., Tchamba, M.T., Toumba-Paka, B.B.L., Zanguim, H.T., Zemtsa, P.T., Zuidema, P.A., 2023a. Stable isotope ratios in wood show little potential for sub-country origin verification in Central Africa. For. Ecol. Manag. 544, 121231.

Boeschoten, L.E., Vlam, M., Sass-Klaassen, U., Rocha Venâncio Meyer-Sand, B., Adzkia, U., Bouka, G.D.U., Ciliane-Madikou, J.C.U., Engone Obiang, N.L., Guieshon-Engongoro, M., Loumeto, J.J., Mbika, D.-mM.F., Moundounga, C.G., Ndangani, R.M. D., Bourobou, D.N., Rahman, M.M., Siregar, I.Z., Tassiamba, S.N., Tchamba, M.T., Toumba-Paka, B.B.L., Zanguim, H.T., Zemtsa, P.T., Zuidema, P.A., 2023b. A new method for the timber tracing toolbox: applying multi-element analysis to determine wood origin. Environ. Res. Lett. 18, 054001.

Bradshaw, R.H.W., 2004. Past anthropogenic influence on European forests and some possible genetic consequences. For. Ecol. Manag. 197, 203–212.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Bridge, M., 2012. Locating the origins of wood resources: a review of dendroprovenancing. J. Archaeol. Sci. 39, 2828–2834.

Bridge, M.C., Fowler, A.M., 2019. A new way of looking at dendroprovenancing: spatial field correlations of residuals. Dendrochronologia 57, 125627.

Caudullo, G., Welk, E., San-Miguel-Ayanz, J., 2017. Chorological maps for the main European woody species. Data Brief. 12, 662–666.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., 2015. Xgboost: extreme gradient boosting. R package version 0.4-2 1, 1-4.

Daly, A., 2007. Timber, Trade and Tree-rings: A Dendrochronological Analysis of Strucural Oak Timber in Northern Europe, C. AD 1000 to C. AD 1650. Syddansk Universitet.

Daly, A., Tyers, I., 2022. The sources of Baltic oak. J. Archaeol. Sci. 139, 105550.

D'Andrea, R., Corona, C., Poszwa, A., Belingard, C., Domínguez-Delmás, M., Stoffel, M., Crivellaro, A., Crouzevialle, R., Cerbelaud, F., Costa, G., Paradis-Grenouillet, S., 2023. Combining conventional tree-ring measurements with wood anatomy and strontium isotope analyses enables dendroprovenancing at the local scale. Sci. Total Environ. 858, 159887.

de Bruin, S., Brus, D.J., Heuvelink, G.B.M., van Ebbenhorst Tengbergen, T., Wadoux, A. M.J.C., 2022. Dealing with clustered samples for assessing map accuracy by cross-validation. Ecol. Inform. 69, 101665.

Dormontt, E.E., Boner, M., Braun, B., Breulmann, G., Degen, B., Espinoza, E., Gardner, S., Guillery, P., Hermanson, J.C., Koch, G., Lee, S.L., Kanashiro, M., Rimbawanto, A., Thomas, D., Wiedenhoeft, A.C., Yin, Y., Zahnen, J., Lowe, A.J., 2015. Forensic timber identification: it's time to integrate disciplines to combat illegal logging. Biol. Conserv. 191, 790–798.

Eckstein, D., Innes, J., Kairiukstis, L., Kocharov, G.E., Nash, T.H., Kincaid, W.B., Briffa, K., Cook, E., Serre-Bachet, F., Tessier, L., Downing, D.J., McLaughlin, S.B., Visser, H., Molenaar, J., Schweingruber, F.H., Norton, D.A., Ogden, J., 1990. Tree-Ring/Environment Interactions and Their Assessment. In: Cook, E.R., Kairiukstis, L. A. (Eds.), Methods of Dendrochronology: Applications in the Environmental Sciences. Springer Netherlands, Dordrecht, pp. 219–288.

European Commission, 2003. Forest Law Enforcement, Governance and Trade (FLEGT) Proposal for an EU Action Plan.

European Commission, 2021. European Green Deal: Commission adopts new proposals to stop deforestation, innovate sustainable waste management and make soils healthy for people, nature and climate.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315.

Forest Stewardship Council, 2017. FSC and Corruption.

Gasson, P.E., Lancaster, C.A., Young, R., Redstone, S., Miles-Bunch, I.A., Rees, G., Guillery, R.P., Parker-Forney, M., Lebow, E.T., 2021. WorldForestID: Addressing the need for standardized wood reference collections to support authentication analysis technologies; a way forward for checking the origin and identity of traded timber. Plants People Planet 3, 130–141.

Grissino-Mayer, H.D., Fritts, H.C., 1997. The International Tree-Ring Data Bank: an enhanced global database serving the global scientific community. Holocene 7, 235–238.

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7, 109.

Hughes, M.K., 2002. Dendrochronology in climatology – the state of the art. Dendrochronologia 20, 95–116.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm: a highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 30.

Khalid, F., Babar Taj, M., Jamil, A., Kamal, H., Afzal, T., Jamshed Iqbal, M., Khan, T., Ashiq, M., Raheel, A., Sharif, M., 2019. Multiple impacts of illegal logging: A key to deforestation over the globe.

Low, M.C., Schmitz, N., Boeschoten, L.E., Cabezas, J.A., Cramm, M., Haag, V., Koch, G., Meyer-Sand, B.R.V., Paredes-Villanueva, K., Price, E., Thornhill, A.H., Van Brusselen, J., Zuidema, P.A., Deklerck, V., Dormontt, E.E., Shapcott, A., Lowe, A.J., 2022. Tracing the world's timber: the status of scientific verification technologies for species and origin identification. IAWA J. 44, 63–84.

Meinshausen, N., 2006. Quantile regression forests. J. Mach. Learn. Res. 7.

Moreno, A., Hasenauer, H., 2016. Spatial downscaling of European climate data. Int. J. Climatol. 36, 1444–1458.

Park, C.-H., Lee, U.-C., Kim, S.-C., Lee, K.-H., 2021. The relationship between tree-ring growths of pinus densiflora and climate from three mountains in central region, the Republic of Korea. Atmosphere 12, 878.

Pearl, J.K., Keck, J.R., Tintor, W., Siekacz, L., Herrick, H.M., Meko, M.D., Pearson, C.L., 2020. New frontiers in tree-ring research. Holocene 30, 923–941.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Pilcher, J.R., Schweingruber, F.H., Kairiukstis, L., Shiyatov, S., Worbes, M., Kolishchuk, V.G., Vaganov, E.A., Jagels, R., Telewski, F.W., 1990. Primary Data. In: Cook, E.R., Kairiukstis, L.A. (Eds.), Methods of Dendrochronology: Applications in the Environmental Sciences. Springer Netherlands, Dordrecht, pp. 23–96.

Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. Soil 7, 217–240.

Polo Villanueva, F.D., Tegegne, Y.T., Winkel, G., Cerutti, P.O., Ramcilovic-Suominen, S., McDermott, C.L., Zeitlin, J., Sotirov, M., Cashore, B., Wardell, D.A., Haywood, A., Giessen, L., 2023. Effects of EU illegal logging policy on timber-supplying countries: a systematic review. J. Environ. Manag. 327, 116874.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A., 2018. CatBoost: unbiased boosting with categorical features. Adv. Neural Inf. Process. Syst. 31.

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.

Salvacion, A.R., Magcale-Macandog, D.B., Cruz, P.C.S., Saludes, R.B., Pangga, I.B., Cumagun, C.J.R., 2018. Evaluation and spatial downscaling of CRU TS precipitation data in the Philippines. Model. Earth Syst. Environ. 4, 891–898.

Schweingruber, F.H., 1993. Tree Species in Eurasia and Japan. In: Schweingruber, F.H. (Ed.), Trees and Wood in Dendrochronology: Morphological, Anatomical, and Tree-Ring Analytical Characteristics of Trees Frequently Used in Dendrochronology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 28–207.

Schweingruber, F.H., 2012. Tree rings: basics and applications of dendrochronology. Springer Science & Business Media.

Van Ham-Meert, A., Daly, A., 2023. Provenancing 16th and 17th century CE building timbers in Denmark–combining dendroprovenance and Sr isotopic analysis. Plos One 18, e0278513.

Visser, R.M., 2021. Dendrochronological provenance patterns. Network analysis of tree-ring material reveals spatial and economic relations of roman timber in the continental north-western provinces. J. Comput. Appl. Archaeol. 4.

Visser, R.M., Vorst, Y., 2022. Connecting ships: using dendrochronological network analysis to determine the wood provenance of Roman-period river barges found in the Lower Rhine region and visualise wood use patterns. Int. J. Wood Cult. 3, 123–151.

Watkinson, C.J., Gasson, P., Rees, G.O., Boner, M., 2020. The development and use of isoscapes to determine the geographical origin of Quercus spp. in the United States. Forests 11, 862.

Zhao, S., Pederson, N., D'Orangeville, L., HilleRisLambers, J., Boose, E., Penone, C., Bauer, B., Jiang, Y., Manzanedo, Rubén D., 2019. The International Tree-Ring Data Bank (ITRDB) revisited: data availability and global ecological representativity. J. Biogeogr. 46, 355–368.

Zuidema, P.A., Babst, F., Groenendijk, P., Trouet, V., Abiyu, A., Acuña-Soto, R., Adenesky-Filho, E., Alfaro-Sánchez, R., Aragão, J.R.V., Assis-Pereira, G., Bai, X., Barbosa, A.C., Battipaglia, G., Beeckman, H., Botosso, P.C., Bradley, T., Bräuning, A., Brienen, R., Buckley, B.M., Camarero, J.J., Carvalho, A., Ceccantini, G., Centeno-Erguera, L.R., Cerano-Paredes, J., Chávez-Durán, Á.A., Cintra, B.B.L., Cleaveland, M. K., Couralet, C., D'Arrigo, R., del Valle, J.I., Dünisch, O., Enquist, B.J., Esemann-Quadros, K., Eshetu, Z., Fan, Z.-X., Ferrero, M.E., Fichtler, E., Fontana, C., Francisco, K.S., Gebrekirstos, A., Gloor, E., Granato-Souza, D., Haneca, K., Harley, G. L., Heinrich, I., Helle, G., Inga, J.G., Islam, M., Jiang, Y.-m, Kaib, M., Khamisi, Z.H., Koprowski, M., Kruijt, B., Layme, E., Leemans, R., Leffler, A.J., Lisi, C.S., Loader, N. J., Locosselli, G.M., Lopez, L., López-Hernández, M.I., Lousada, J.L.P.C., Mendivelso, H.A., Mokria, M., Montóia, V.R., Moors, E., Nabais, C., Ngoma, J., Nogueira Júnior, Fd.C., Oliveira, J.M., Olmedo, G.M., Pagotto, M.A., Panthi, S., Pérez-De-Lis, G., Pucha-Cofrep, D., Pumijumnong, N., Rahman, M., Ramirez, J.A., Requena-Rojas, E.J., Ribeiro, Ad.S., Robertson, I., Roig, F.A., Rubio-Camacho, E.A., Sass-Klaassen, U., Schöngart, J., Sheppard, P.R., Slotta, F., Speer, J.H., Therrell, M. D., Toirambe, B., Tomazello-Filho, M., Torbenson, M.C.A., Touchan, R., Venegas-González, A., Villalba, R., Villanueva-Diaz, J., Vinya, R., Vlam, M., Wils, T., Zhou, Z.-K., 2022. Tropical tree growth driven by dry-season climate variability. Nat. Geosci. 15, 269–276.