Chemical timber tracing: combining tree-genera information lowers reference data needs and makes harvest location identification more

Truszkowski *et al. Annals of Forest Science* (2026) 83:18
https://doi.org/10.1186/s13595-026-01341-x

INRAE

Annals of Forest Science

RESEARCH

Open Access

Chemical timber tracing: combining tree-genera information lowers reference data needs and makes harvest location identification more accurate

Jakub Truszkowski[1,2*], Laura Boeschoten[3,4], Thomas Mortier[5], Charlotte Smith[6], Bogdan Buliga[7,8], Caspar Chater[9,10], Steven B Janssens[11,12], Jade Saunders[6], Johann Trischler[13], Pieter Zuidema[14], Alexandre Antonelli[9,15,16,17] and Victor Deklerck[6,9,18]

Abstract

Key Message Chemistry-based tracing techniques are increasingly used for combating illegal timber trade, but they are currently limited by the small and fragmented reference datasets available. We introduce a model that integrates data from multiple tree genera while accounting for statistical differences between them. Our model accurately predicts the harvest location even when relevant data are unavailable in some areas, by leveraging data from other genera. Our approach could lower reference sampling costs and enable tracing in situations where new samples cannot be collected, such as during armed conflict.

Context Chemistry-based techniques for identifying the harvest location of timber are becoming increasingly important for enforcing timber trade regulations. However, their application has been limited by the need for reference samples from all species across all areas of interest.

Aims We investigate whether combining reference data from multiple taxonomic groups can improve timber harvest location determination in regions where reference data is scarce by using the shared natural variability in isotopic composition across species.

Methods We extend the harvest location model of Mortier et al. to jointly model isotope ratios and trace element concentrations in wood from different genera. This is achieved by a new covariance function that accounts for shared patterns of spatial variation between genera. We evaluate our approach on 1020 tree samples from four economically important genera (*Betula*, *Fagus*, *Pinus*, *Quercus*) across 12 Eastern European countries.

Results The multi-genus model substantially outperforms the single-genus model when little or no data for that genus is available in the focus area. When data from all genera are available across the study area, the multi-genus model achieves similar performance to the single-genus model.

Handling editor: John Lhotka.

*Correspondence:
Jakub Truszkowski
jakubt@chalmers.se
Full list of author information is available at the end of the article

Conclusion Our approach strengthens the applicability of timber tracing methods by enabling accurate predictions in areas where sample collection is not currently feasible due to political, logistical and/or security-related challenges, provided that pre-existing samples from other genera are available.

Keywords Timber traceability, Stable isotopes, Trace elements, Gaussian process, Multitask learning

1 Introduction

Persistently high rates of deforestation in recent decades (FAO 2020) have spurred action to tackle illegal logging and trade of illegally harvested forest products, which are key drivers of environmental degradation and biodiversity loss globally (Reboredo 2013). Environmental laws, including the 2008 amendment to the United States Lacey Act, the 2013 European Union Timber Regulation (EUTR), and the 2012 Australian Illegal Logging Prohibition Act, prohibit the placement of illegal timber on the market and mandate operators to exercise due diligence (or due care in the US context) to ensure compliance. Building on lessons from earlier implementation challenges, the European Union (EU) developed the Regulation on deforestation-free products (EUDR), which entered into force in 2023, replacing the EUTR and expanding its scope (EUDR 2023). Due diligence required by the EUDR includes reporting precise geolocation coordinates for harvest location, and enforcement agencies have been empowered to conduct compliance checks using "any technical and scientific means adequate to determine the species or the exact place where the relevant commodity or relevant product was produced, including anatomical, chemical or DNA analysis" (EU Regulation 2023/1115, 2023). This has created a need for robust, accurate, and high-throughput tools to identify the provenance of timber, in order to validate origin claims and ensure regulatory compliance.

There are currently three scientific techniques offering a solution to origin claim validation: genetic, stable isotope, and trace element analysis. Genetic analysis (Capo et al. 2024) uses molecular markers, such as single-nucleotide polymorphisms (SNPs), microsatellite regions, or genome-wide sequence data, to determine the location of timber harvesting through population genetics and phylogeographic methods. Stable isotope ratio analysis (SIRA; West et al. 2009; Bowen et al. 2005; Truszkowski et al. 2025; Aguzzoni et al. 2025) quantifies the ratios of naturally occurring isotopes of common elements, such as oxygen, hydrogen, carbon, nitrogen, and sulfur. Trace element analysis (TEA; Ågren and Weih 2012; Boeschoten et al. 2023a) measures the concentrations of a large set of trace elements in wood tissue, commonly including Al, Si, Li, Fe, Zn, Pb, and others. Genetic methods are based on genetic differences between populations as geographically distant populations tend to be genetically distinct from one another (Rocha Venancio Meyer-Sand et al. 2025). While genetics is a promising technology for timber tracing, its practical application remains limited due to difficulties in extracting DNA from wood and the fact that plantation forests may be genetically distinct from nearby natural forests (Deklerck 2023). TEA and SIRA both measure variation in plants' chemical characteristics caused by climatic and environmental variation across landscapes. Plant uptake of trace elements is unique to each species and location, because the availability of trace elements is spatially variable and different species absorb them from the soil in various quantities (Boeschoten et al. 2022). Similarly, environmental factors such as humidity, precipitation, temperature, and soil composition affect the presence and ratio of stable isotopes in organic matter (Deklerck 2023).

Previous studies have demonstrated the power of combining SIRA and TEA with probabilistic spatial models for timber origin verification (*assessing a harvest location claim*) and determination (*identifying harvest location*) (Mortier et al. 2024). Development of these large-scale spatial models, based on Gaussian process (GP) regression combined with Bayesian inference, has advanced timber origin validation beyond localized, site-specific comparisons (i.e. classification). By capturing the natural variation in stable isotope ratios and trace element concentrations in trees across large continuous areas, GP models can predict the likelihood of a sample originating from any location that is sufficiently close to reference samples, regardless of political borders or other arbitrarily defined regions (Mortier et al. 2024). However, these large-scale models have two key limitations. First, their spatial extent is determined by the location of reference data, thus they are restricted by the small, fragmented timber reference databases currently available (Low et al. 2022). This data availability issue is unlikely to be resolved in a time frame that is useful for regulatory enforcement, due to the considerable time and financial costs associated with collecting and analyzing samples from extensive field campaigns (see also Suarez and Tsutsui 2004). Additionally, collecting physical samples may be impossible in some key regions due to geopolitical factors (e.g., collector safety in war zones) and international agreements that place limitations on the export of scientific samples (e.g., Access and Benefit Sharing rules under the Convention on Biological Diversity's 2010

Nagoya Protocol). Second, current modeling approaches, such as random forest classification (Boeschoten et al. 2023b), clustering (Vlam et al. 2018), discriminant analysis (Paredes-Villanueva et al. 2022), or Gaussian Process-based harvest location models (Truszkowski et al. 2025) are species-specific and do not take into account physiological and ecological similarities between species or genera. This is an inefficient use of data, as variation in stable isotopes and trace elements across space is primarily caused by differences in climate and soil composition (Hartl-Meier et al. 2015). This means that similar patterns are likely to be observed across different species, even if absolute values differ (Hartl-Meier et al. 2015; Boeschoten et al. 2023a). Species-specific analyses, combined with scattered data sources, pose important limitations to the power, performance, and spatial extent of statistical models that are currently used to trace timber.

We introduce a new approach for inferring timber harvest location that harnesses the similarities in stable isotope ratios (SIR) and trace element (TE) measurements between genera, while also addressing the challenge of limited field reference data. Specifically, we investigate the following key question: can a multi-genus model produce accurate origin predictions across all genera considered, even in regions where reference samples from some genera are scarce or entirely absent? To explore the effectiveness of a multi-genus approach, we use a data set of 1,020 timber samples from Eastern Europe. This region provides an important case study for illustrating scenarios in which areas of interest for international sanctions cannot be sampled for political or security reasons. Our model uses a covariance function that accounts for shared patterns of spatial variation between genera, a method closely related to multi-task Gaussian processes (Bonilla et al. 2007; Alvarez et al. 2012). Models using data from multiple species or genera have previously been proposed for tracing timber (Boeschoten et al. 2023b) and marine animals (St. John Glew K, Graham LJ, McGill RA, et al. 2019). However, to our knowledge, this is the first time a multi-genus approach has been developed with the goal of mitigating data scarcity. This approach could eliminate the key barrier of limited reference data to enable the widespread adoption of scientific testing in timber supply chains.

2 Material and Methods

To test whether tracing results improve when multiple genera are combined into one spatial model, we extended the Gaussian Process model developed by Mortier et al. (2024) with a new covariance function designed to accommodate data from multiple genera. The covariance function represents each genus as a vector, such that the angle between two vectors represents the degree of similarity in the spatial pattern of isotope ratios or trace elements between the two genera. The vector coordinates are estimated from training data, allowing the model to learn about the similarities between different genera. We evaluate our approach on a recent data set (Mortier et al. 2024) of four economically important genera (*Betula*, *Pinus*, *Quercus*, *Fagus*) in Eastern Europe.

2.1 Probabilistic models for harvest location inference

Our method has three stages. First, for every SIR or TE variable, a GP regression model is trained on the reference data. Second, we compute the probability density of observing the SIR and TE values for each test sample at each point of the study area. Third, Bayes' theorem is used to compute the posterior probability distribution of possible harvest locations of the test sample. For further information on the model, see Truszkowski et al. (2025).

2.2 Gaussian process regression

Gaussian process (GP) regression is a flexible regression technique that enables the quantification of uncertainty of a model's predictions. A standard GP regression model is defined by a *mean* parameter, a *covariance function*, and a *noise* parameter. The mean parameter specifies the expected value of the responses. The covariance function $f(\mathbf{x}_1, \mathbf{x}_2)$ models the degree of similarity between two locations, usually as a function of distance between them. The standard choice for single-genus models (Truszkowski et al. 2025) is the Matérn covariance function with the shape parameter $\nu = 1.5$ (Abramowitz and Stegun 1972), which can be written as

$$f(\mathbf{x}_1, \mathbf{x}_2) = A(1 + \sqrt{3}d)\exp(-\sqrt{3}d)$$

where $d = \sqrt{(\mathbf{x_1} - \mathbf{x_2})\mathbf{D}^{-1}(\mathbf{x_1} - \mathbf{x_2})^\top}$ is the Euclidean distance between $\mathbf{x_1}$ and $\mathbf{x_2}$ with dimensions scaled by diagonal matrix $\mathbf{D}^{-1}$ (a parameter of the covariance function). A is a parameter governing the magnitude of the differences between locations.

The GP regression model also includes the noise parameter σ^2, which models the intrinsic noise in the data. Taken together, the model considers the responses as being generated by a multivariate normal distribution with mean μ and covariance matrix $\mathbf{\Sigma}$, where $\mathbf{\Sigma}_{ij} = f(\mathbf{x}_i, \mathbf{x}_j)$ for $i \neq j$ and $\mathbf{\Sigma}_{ii} = f(\mathbf{x}_i, \mathbf{x}_i) + \sigma^2$.

$$\mathbf{y} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$$

We fit a separate model for every SIR or TE variable by finding parameters that maximize the likelihood of the training data. Once the models have been trained, it is possible to compute the mean $\mathbb{E}[y|\mathbf{x}^*]$ and variance $\mathbb{V}(y|\mathbf{x}^*)$ of the response at any test location $\mathbf{x}^*$ using

standard formulae (Truszkowski et al. 2025; Williams and Rasmussen 2006).

2.3 Bayesian inference

Given the mean and variance at location $\mathbf{x}^*$, the likelihood of observing the jth SIR or TE variable at that location is

$$p(y_j|\mathbf{x}^*) = \frac{1}{\sqrt{2\pi\mathbb{V}(y_j|\mathbf{x}^*)}} \exp\left(\frac{-(y_j - \mathbb{E}[y_j|\mathbf{x}^*])^2}{2\mathbb{V}(y_j|\mathbf{x}^*)}\right) \quad (1)$$

For a vector $\mathbf{y}^* = [y_1, \ldots, y_m]$ of SIR/TE values observed in the test sample, the Bayes' theorem gives the posterior distribution of possible harvest locations:

$$p(\mathbf{x}^*|\mathbf{y}^*) = \frac{p(\mathbf{x}^*)\prod_{j=1}^m p(y_j|\mathbf{x}^*)}{\int_{\mathbf{x}\in\mathcal{A}} p(\mathbf{x})\prod_{j=1}^m p(y_j|\mathbf{x})d\mathbf{x}} \quad (2)$$

where the probability densities $p(y_j|\mathbf{x}^*)$ are computed from the GP models for the respective isotopes and elements using Eq. 1 and $\mathcal{A}$ is the study area, represented by a rectangular grid of points. For each genus, we use a uniform prior distribution over the study area – see Section 2.7.

2.4 Verification of harvest location claims

A trained model can also be used to verify a supplier's sourcing claim. We use the statistical testing procedure described in a recent paper (Mortier et al. 2024) to verify harvest location claims. The test takes as input a vector $\mathbf{y}^*$ of SIR/TE values and a set of grid points $\mathcal{X}_c$ corresponding to the territorial unit c that is claimed to be the source of the sample. For each grid point $x \in \mathcal{X}_c$, the residuals of the observed values with respect to the values predicted by the model are computed, normalized by the predicted standard deviation. We then perform the χ^2 test with m degrees of freedom on the sum of squared residuals at each grid point and take the maximum of the p-values to be the p-value of the composite hypothesis. The details of the reasoning can be found in Mortier et al. (2024).

2.5 A multi-genus Gaussian process model

Exact SIR/TE values usually differ between genera. In contrast, patterns of spatial variation are often shared—e.g., the oxygen isotope ratio (δ^{18}O) decreases with increasing latitude for the four genera in our data set (Fig. 1). To extend the single-genus GP regression model to multiple genera, we must take into account (a) different levels of SIR/TE values between genera (Fig. 1a), and (b) varying levels of similarity between spatial variation patterns of SIR/TE values in different genera (Fig. 1c). To address (a), a separate mean parameter θ_g is used for each genus g. To address (b), the standard covariance function is multiplied by a term that models the mutual relevance of data from different genera. For each genus g, a vector $\nu_g \in \mathbf{R}^\ell_{\geq 0}$ is defined (where ℓ is a user-chosen dimension of the model) to represent the genus. The mutual relevance of data from different genera is modeled using the dot product between genus vectors, to create the following covariance function, which acts on sample locations and sample genera:

$$k((\mathbf{x}_1, g_1), (\mathbf{x}_2, g_2)) = \nu_{g_1}^\top \nu_{g_2}(1 + \sqrt{3}d)\exp(-\sqrt{3}d)$$

Importantly, the vectors ν_g are estimated from the data, which allows the model to learn the degree of correlation between different genera. If two vectors ν_{g_1} and ν_{g_2} are orthogonal to each other, the dot product term $\nu_{g_1}^\top \nu_{g_2}$ equals 0, which means that data from the two genera are uncorrelated and the model will treat harvest location inference for g_1 and g_2 as two separate problems. If ν_{g_1} and ν_{g_2} are parallel, this indicates that the responses from g_1 and g_2 at the same location are perfectly correlated and a prediction about g_2 can be obtained from a prediction about g_1 by shifting and rescaling. Appendix Fig. 6 shows an example of estimated genus vectors in two dimensions.

To accommodate different levels of environmental variation across genera, we use a genus-specific noise parameter σ_g^2 for every genus g.

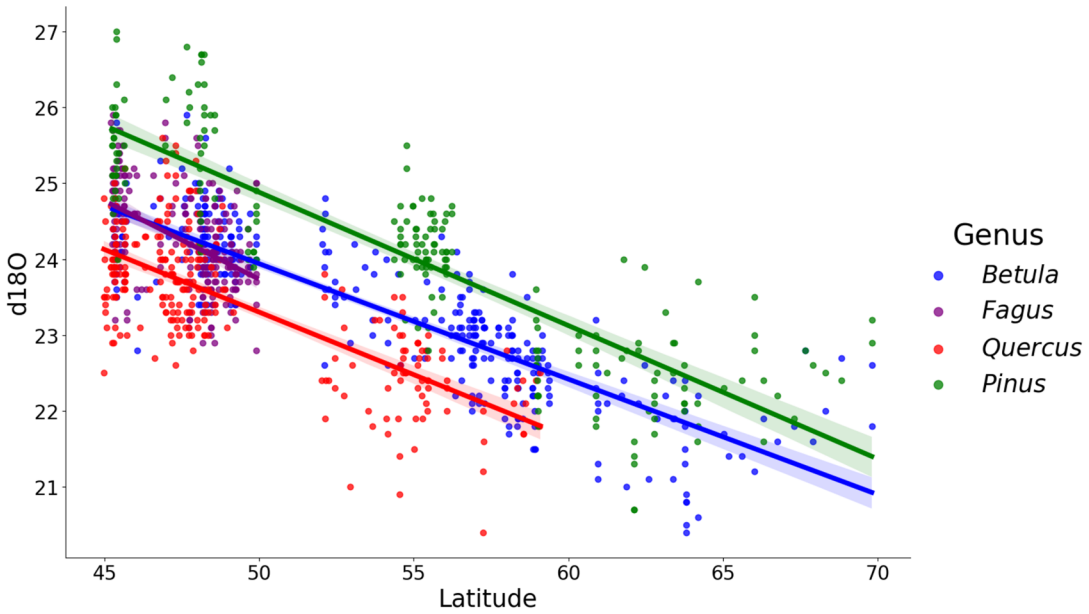
2.6 Data

We evaluate our approach on a data set of 1020 trees from 4 genera (*Betula*, *Fagus*, *Pinus*, *Quercus*) across 12

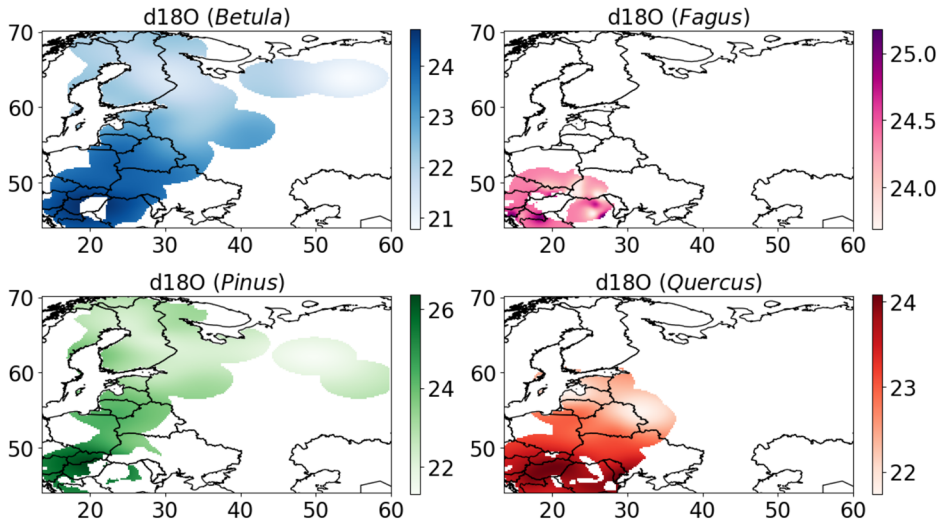
(See figure on next page.)

Fig. 1 Shared patterns of SIR variation between genera in Eastern Europe. **a** $\delta^{18}O$ values plotted against the latitude of the reference samples, colored by genus. $\delta^{18}O$ values decrease with increasing latitude for all genera, even though the average values are distinct for each genus. **b** Isoscapes estimated by the Gaussian process-based single-genus model described by Mortier et al. (2024). High values (dark) dominate in the south, whereas low values (light) dominate in the north of the prediction range. The color scales match the genus colors in the upper figure. Values were computed for points that lie within $d = 300$ km of any sample from that genus and fall within the genus range according to Caudullo et al. (2017). **c** Correlation matrix for single-species isoscape values between genera. Positive correlation can be seen between all genera, with the highest between *Betula* and *Pinus* (0.92). Only the values within the prediction area for each genus were used to compute the correlations

a)



b)



c)

	Betula	Fagus	Pinus	Quercus
Betula	1.00			
Fagus	0.24	1.00		
Pinus	0.92	0.36	1.00	
Quercus	0.88	0.35	0.85	1.00

Fig. 1 (See legend on previous page.)

Eastern European countries (Table 1). The data set was first introduced by Mortier et al. (2024), and has since been expanded to include additional samples. Following the method developed by Boner et al. (2007), five SIR values were measured for each sample: $\delta^{13}C$ (ratio between ^{13}C and ^{12}C), hydrogen $\delta^{2}H$ (ratio between ^{2}H and ^{1}H), nitrogen $\delta^{15}N$ (ratio between ^{15}N and ^{14}N), oxygen $\delta^{18}O$ (ratio between ^{18}O and ^{16}O), and sulfur $\delta^{34}S$ (ratio between ^{34}S and ^{32}S). Isotope ratios of covalently bonded hydrogen after nitration ($\delta^{2}H_{nit}$) were also measured. The relative abundance of 17 TEs (Al, Si, P, S, Cl, K, Ca, Mn, Fe, Ni, Cu, Zn, Br, Rb, Sr, Ba, and Pb) was measured using X-Ray Fluorescence Spectroscopy (XRF) (XEPOS, SPECTRO analytical instruments). This resulted in $m = 20$ SIR+TE measurements for every sample (TE measurements of Ba, Mn, and Br were excluded from further analyses due to a large number of missing entries). See Mortier et al. (2024) for a detailed description of sampling protocol and laboratory analytical procedures.

2.7 Study area

For each genus, the study area consisted of all locations that lie within $d = 300$ km of any sample from that genus and fall within the genus range according to Caudullo et al. (2017). We represent the study area as a rectangular grid spanning locations between 13.6° and 60° longitude and 44° and 70.4° latitude. We chose a resolution of 0.2°, in line with previous work (Mortier et al. 2024).

2.8 Experimental setup

All models were implemented in Python (v. 3.10.12) using GPyTorch (v. 1.9.0; Gardner et al. 2018), and PyTorch (v. 1.12.1; Paszke et al. 2017) packages. Training was performed by maximizing the likelihood of the training data using the Adam optimizer (Kingma and Ba 2015) with an initial learning rate of 0.03. We used Stochastic Gradient Descent with Warm Restarts learning rate scheduler (Loshchilov and Hutter 2017) and early stopping with a patience of 100 iterations. The dimension of genus vectors was set to $\ell = 2$ in all experiments unless otherwise specified.

For the determination task (identifying the harvest location for a test sample), we report the *mode distance*, which is defined as the distance between the true harvest location and the most likely harvest location according to the model (Truszkowski et al. 2025).

For the verification task (assessing whether a harvest location claim is true), we report sensitivity (the percentage of incorrect origin claims that are detected by the model) and specificity (the percentage of correct origin claims that are not marked as incorrect by the model). Each test case consists of a data point from the test set and a simulated forest concession of size $0.5 \times 0.5°$, which is the claimed harvest location. Correct claims are simulated by choosing a random forest concession that contains the true location. Incorrect claims are simulated by randomly choosing a location within both the target country (other than the true country of origin) and the prediction area of the determination task. For each test data point, we simulate a single correct claim and 11 incorrect claims (one for each country considered except the true country of origin). We then test the hypothesis that the test sample originates from the claimed forest concession with significance level $\alpha = 0.05$.

We investigate the performance of our model in two settings: the *high-data regime* and the *low-data regime*. The *high-data regime*, refers to a situation where samples are abundant in every part of the study area. To evaluate the performance in this setting, we randomly split all our available samples into 5 folds and perform 5 cross-validation experiments with one fold designated as the test set and the remaining four being used for training. The *low-data regime* mimics a situation where samples from the genus of interest are scarce or missing entirely from part of the study area. To simulate this, we divide our prediction range into 5 testing areas based on country borders (see Fig. 2). Testing areas were chosen such that each contains a sufficient number of samples from each genus of interest, is spatially coherent, and has similar surface area. Each experimental setting involves choosing a focal testing area, setting aside samples for testing (5-fold cross-validation), and down-sampling the training samples within the focal testing area to a required sample density (between 0 and 1 samples per 10,000 km^2). For each combination of focal testing area, sampling density, and cross-validation fold, a model is trained on the available training samples and its accuracy on the test data is evaluated. For the determination task, test samples are

Table 1 Country origins of the samples in the data set

Genus country	*Betula*	*Fagus*	*Pinus*	*Quercus*
Belarus	35	0	0	29
Croatia	20	60	29	57
Estonia	50	0	0	10
Finland	30	0	36	0
Hungary	9	16	13	49
Latvia	56	0	0	4
Lithuania	0	0	64	34
Moldova	0	0	0	51
Romania	22	35	0	16
Russia	34	0	35	0
Slovakia	20	20	0	9
Ukraine	60	88	30	0

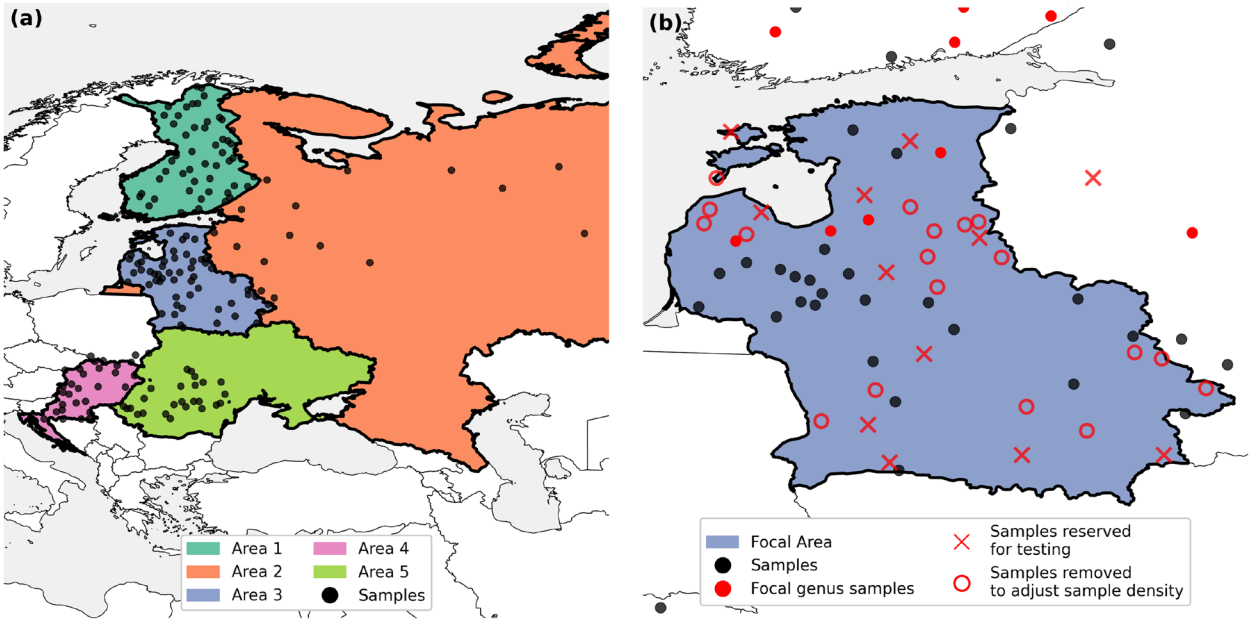


Fig. 2 Defining a test scenario in the low-data regime in Eastern Europe. **a** We divided the study region into five testing areas (colored areas) according to the location of samples (black circles). **b** For each test scenario, a focal genus (e.g., *Betula*; red markers) and a focal testing area (e.g., Area 3; light blue area) was chosen. A subset of the samples for the focal genus (X) was set aside for testing. For determination, we only picked test samples from within the focal testing area. For verification, test samples were picked from across the study region. To create different sample densities for training, samples from the focal genus within the focal testing area were randomly removed (red circle). Note: to protect data confidentiality, and ensure the security of collectors and landowners, not all samples are shown and locations were subjected to spatial obfuscation through random adjustment

chosen from the focal testing area, whereas for the verification task, test samples are chosen from the entire study area, while simulated claim locations are chosen within the focal testing area.

3 Results

3.1 Performance in low-data settings

For the determination task, the multi-genus model substantially outperformed the single-genus model when no data from the target genus were available in the focal testing area (Figs. 3 and 4), with considerably lower average mode distance than the single-genus model. This was most pronounced for *Quercus*, with the average mode distance reduced by 269 km (38%) relative to the single-genus model, and least pronounced for *Pinus* at 68 km (10%). The mode distance for *Fagus* and *Betula* was reduced by 185 km (35%) and 143 km (25%), respectively. See Appendix Figs. 9 and 10 for examples of predictions from both models. For the verification task, the multi-genus model showed higher sensitivity than the single-genus model for all genera, with the largest increase relative to the single-genus model for *Pinus* (49 %) followed by *Betula* (18 %), *Quercus* (17 %), and *Fagus* (9 %)—see Figs. 3 and 4. The specificity of the two models was similar across all genera except *Pinus*, where the single-genus model achieved a substantially higher specificity at low sampling densities (see Appendix Figs. 7 and 8).

Increasing the number of training samples available from the target genus resulted in a gradual performance improvement for both models (single- and multi-genus) in both verification and determination tasks, and the performance gap between single- and multi-genus models diminished with increasing numbers of training samples from the target genus; for the determination task, the difference between the two models became negligible when the sample density reached 0.2 trees per 10000 km^2 for *Betula* and *Pinus*. For the verification task, the multi-genus model retained a sizeable advantage in most experimental settings, even at higher sample densities.

3.2 Performance on all available data

We investigated the performance of the single-genus model and the multi-genus model on a determination task when provided with all available data (Fig. 5). This

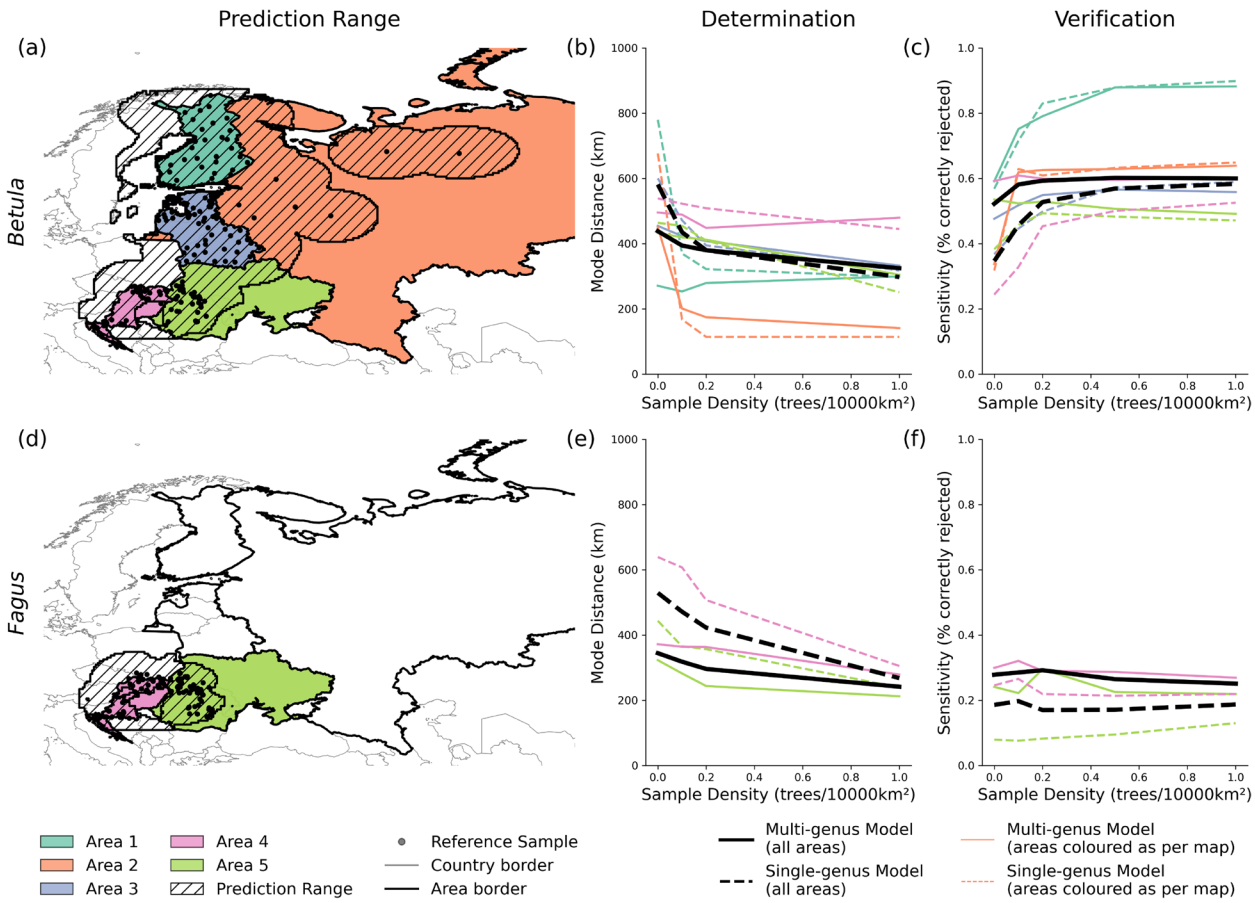


Fig. 3 Performance of single- and multi-genus models for determination and verification tasks for *Betula* and *Fagus* in the low-data regime in Eastern Europe. **a**, **d** A map of sample locations (black circle), prediction ranges (lines), and testing areas (colored areas) used to generate test scenarios for each genus. **b**, **e** The average mode distance values for the determination task and (**c**, **f**) the average sensitivity values for the verification task, as a function of sample density in the low-data region using a single-genus model (dashed) and the multi-genus model (solid). Colored lines show the mean values per focal testing area (line colors correspond to map), while black lines show the mean value across all test scenarios. For determination (**b**, **e**), lower values correspond to more accurate predictions, while for verification (**c**, **f**), higher values correspond to more accurate predictions

corresponds to an average sampling density of 2.0 trees per 10,000 km^2 across the four genera. The accuracy of the multi-genus model depended on the chosen dimension of the genus vectors – increasing the dimension from 1 to 2 decreased mode-distance by between 12 km (*Fagus*) and 64 km (*Quercus*). Increasing the dimension further yielded only modest accuracy improvements (at most 13 km compared to the 2-dimensional model for all genera and dimensions analyzed), and at worst lead to a slight decrease in performance. The best multi-genus model yielded lower error than the single-genus model for *Fagus* and *Pinus*, while the single-genus model was more accurate for *Betula* and *Quercus*. For all genera, these differences were less than 14 km, indicating broadly similar performance between the models in the simulation with high-data availability.

For the the verification task, the multi-genus model gave higher sensitivity for *Betula*, *Fagus*, and *Quercus*, but was outperformed by the single-genus model on *Pinus* samples (Table 2). The biggest difference in performance was seen for *Fagus*, where the multi-genus model had 6.7% higher sensitivity. Both models showed similar specificity for all genera except *Fagus*, where the single-genus model achieved a 4.1% higher specificity than the multi-genus model.

4 Discussion

Building upon the single-genus approach developed for Eastern Europe by Mortier et al. (2024), we present the first large-scale study on multi-genus modeling for timber traceability. Our results show that leveraging information from every available sample, across all the genera,

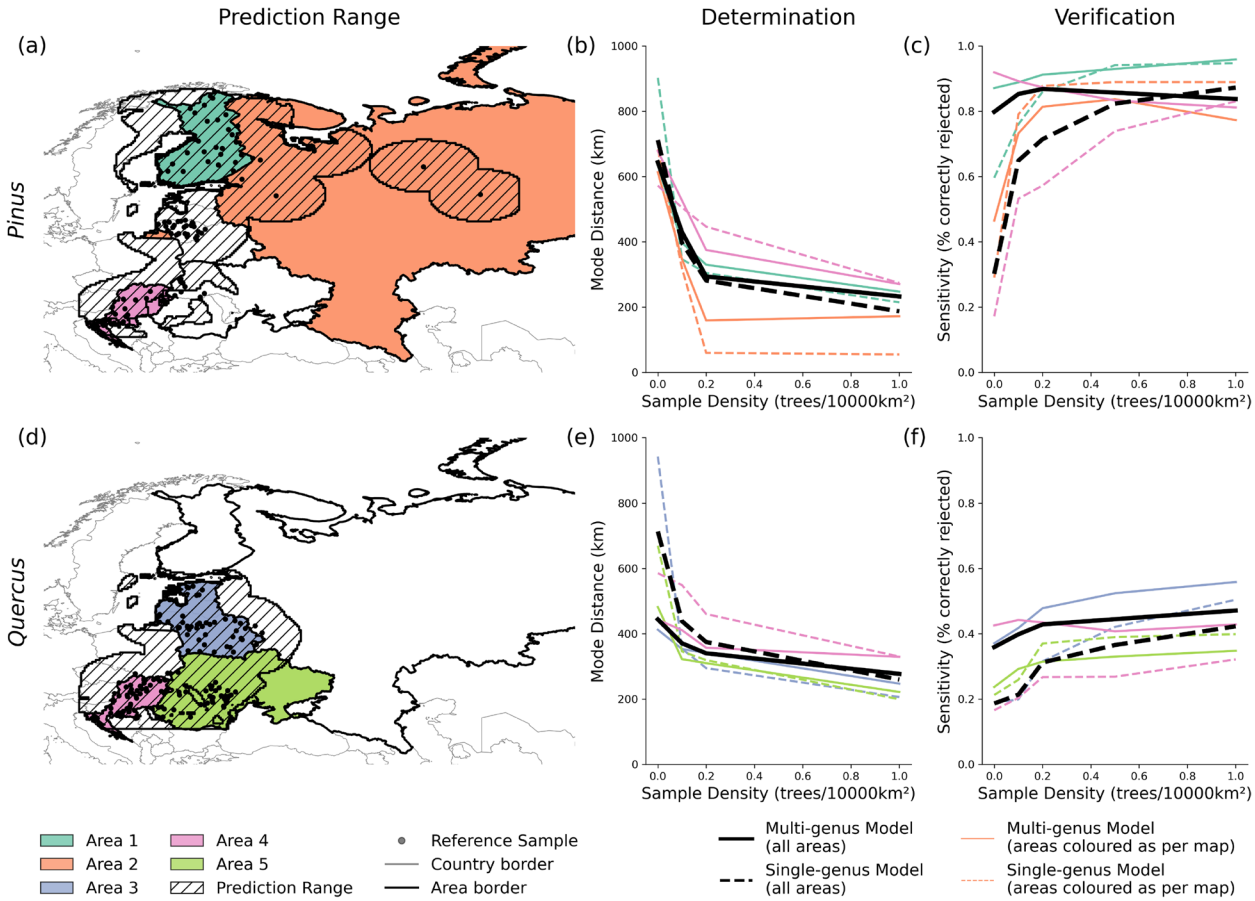


Fig. 4 Performance of single- and multi-genus models for determination and verification tasks for *Pinus* and *Quercus* in the low-data regime in Eastern Europe. **a**, **d** A map of sample locations (black circle), prediction ranges (lines), and testing areas (colored areas) used to generate test scenarios for each genus. **b**, **e** The average mode distance values for the determination task and (**c**, **f**) the average sensitivity values for the verification task, as a function of sample density in the low-data region using a single-genus model (dashed) and the multi-genus model (solid). Colored lines show the mean values per focal testing area (line colors correspond to map), while black lines show the mean value across all test scenarios. For determination (**b**, **e**), lower values correspond to more accurate predictions, while for verification (**c**, **f**), higher values correspond to more accurate predictions

improves tracing results when data from the relevant genus are scarce or unavailable. This demonstrates the advantage of our approach over traditional methods. We expect that multi-genus approaches to tracing will be particularly useful for regions where collecting new reference samples is infeasible due to political or security reasons.

While elemental and isotopic compositions are known to reflect environmental baselines (Boeschoten et al. 2022; Raju and Raju 2000; Heineman et al. 2016; Allen et al. 2022; Baker et al. 2015), earlier studies reported notable interspecific differences within similar regions, especially for stable isotopes (Boeschoten et al. 2023b; Watkinson et al. 2022a). However, our model successfully combines data from a variety of genera, resulting in improved harvest location predictions across both conifers and broad-leaved taxa. These findings demonstrate that wood chemistry shares common spatial patterns across diverse lineages. The innovation of our GP approach lies in its flexibility: genus-specific mean parameters allow for deviation from overall signals, while data-derived genus vectors allow for varying levels of similarity between spatial variation in different genera. This structure allows the model to account for statistical differences between genera, which enables the use of data from different genera for making predictions. More traditional approaches, such as cluster analysis or discriminant analysis, cannot readily distinguish between spatial and inter-genus variation. This framework reflects

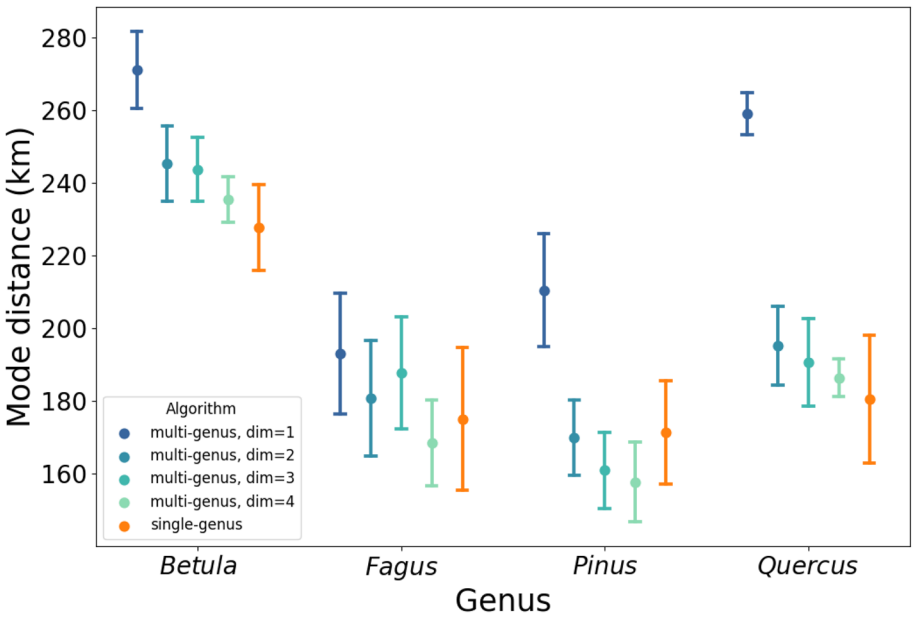


Fig. 5 Determination accuracy in high-data regime. Average mode distance (distance between the predicted and true harvest location) for the single-genus model and multi-genus models with different genus vector dimensions for the determination task when all available data are used. Error bars denote standard errors of the mean

underlying biological processes such as interspecific differences in element uptake and isotope fractionation (Lehmann et al. 2022). Our successful application across four distantly related genera highlights the potential of this method to be applied more broadly, particularly for improving timber tracing in poorly sampled regions such as Central Africa.

A key barrier to routine use of harvest location identification is the availability of reference data covering a genus's harvestable range (Low et al. 2022; Dormontt et al. 2015). Sampling to the required extent is both time- and cost-intensive (Suarez and Tsutsui 2004); and often made impossible by logistical challenges in the geopolitically unstable and war-torn regions primarily targeted by economic sanctions (Deb et al. 2022). For example, following the invasion of Ukraine, sanctions on Russian timber disrupted legal trade channels, leading to concerns about increased illegal logging and misdeclaration of origin to bypass restrictions. This led to the creation of the dataset and model presented in Mortier et al. (2024). However, the applicability of that approach was limited by the difficulties in obtaining reference samples in some countries. Our multi-genus model could alleviate this problem by making use of historical samples collected before the conflict began, even if they do not belong to the genera of interest.

Table 2 Sensitivity and specificity for the single- and multi-genus models for the verification task in the high-data regime

Genus	Model	Sensitivity	Specificity
Betula	Multi-genus	61.5 ± 2.2	92.9 ± 2.1
Betula	Single-genus	59.5 ± 2.6	93.2 ± 1.9
Fagus	Multi-genus	24.6 ± 2.5	91.8 ± 1.8
Fagus	Single-genus	17.9 ± 1.4	95.9 ± 1.3
Pinus	Multi-genus	82.7 ± 0.9	90.8 ± 1.6
Pinus	Single-genus	83.8 ± 1.6	91.2 ± 1.7
Quercus	Multi-genus	50.7 ± 1.9	89.6 ± 1.6
Quercus	Single-genus	47.6 ± 2.2	89.6 ± 1.0

Values after ± denote the standard error of the mean

The cost of reference data generation and testing operations continues to slow down the adoption of chemical tracing technologies. TEA currently costs 20–120 USD per sample, while the cost of SIRA is 200–500 USD per sample, depending on the number of measurements and the type of technology used. While samples are costly to obtain and analyze, our model can lower reference data requirements by interpolating between reference data points and utilizing data from different genera, which effectively replaces physical samples with statistical inference. We expect that the cost of laboratory analyses will gradually decrease given the increasing popularity of chemical tracing (Gori et al. 2025; Deklerck 2023). Moreover, one could argue that the testing cost is minimal relative to the economic impacts of the illegal timber trade.

The accuracy of the multi-genus model generally increases with the dimension of the vectors used to

represent each genus (see Fig. 5). We observe a large difference in accuracy between models using 1- and 2-dimensional vectors, with the 1-dimensional model yielding substantially worse accuracy than all the other models tested, including the single-genus model, for all genera. The 1-dimensional model assumes that the exact measurements for one genus can be obtained by shifting and rescaling the corresponding measurements for another genus, which is similar to the model applied to different species of marine animals by St. John Glew K, Graham LJ, McGill RA, et al. (2019). Our results suggest that this approach is insufficient for datasets composed of distantly related timber species, such as the ones used in this study. Representing genera with higher-dimensional vectors allows the model to better represent the noisy relationship between measurements from different genera.

In addition to reducing sample collection costs, our model could also be used to identify strategic collection locations. In the low-data regime, the advantage of the multi-genus model is substantially smaller for *Pinus* than for other genera. This is likely due to the relatively discontinuous distribution of *Pinus* sample locations, with most of the locations being close to the boundaries of the study area, making it difficult to estimate the expected values of variables in areas where few *Pinus* samples are available. Knowing this allows future sampling campaigns to be targeted more specifically for improving model outputs. Alternatively, future research could explore employing models with more expressive mean functions to further alleviate the need for additional samples.

This study focuses on an existing set of samples for genera and species that are harvested commercially in temperate forests. Future work could apply the model to tropical species from areas where illegal logging and deforestation pose immediate threats to species and ecosystems, such as in Central Africa, where over 50% of annual timber production is estimated to come from illegal logging (African Natural Resources Centre 2021). Previous studies on Central Africa have described similarities in spatial variation of TEs, and an assignment-based timber tracing approach using multi-element analysis showed promising results using two different timber species (Boeschoten et al. 2023a). Incorporating our multi-genus approach could be beneficial for any tracing study that involves diverse timber species (Watkinson et al. 2022a; Aguzzoni et al. 2025; Watkinson et al. 2022b).

Investigating the usefulness of our multi-genus approach beyond timber offers another interesting direction for future research. Several commodities relevant to EUDR, such as soybean and cocoa, are harvested annually, which means that their chemical characteristics could vary between years due to climate variability. Our approach could be used to flexibly model samples collected in different years as distinct but related spatial patterns, analogously to the multi-genus model presented in this study. This could considerably reduce the amount of sampling that needs to be done every year to maintain sufficient model accuracy, making traceability more practical and cost-effective in the long run.

5 Conclusion

In this study, we demonstrate that it is possible to create accurate and cost-effective tools for identifying the harvest location of traded timber using a multi-genus approach. This requires substantially fewer ground-truthed physical samples than previous, single-genus models, thereby increasing the applicability of harvest location identification methods. This is achieved by increasing the amount of information derived from each physical sample through an integrated model that leverages cross-generic data to make predictions for any sample of interest.

Our modeling approach enables practitioners to leverage data sets that were previously of limited use due to their small size or limited geographic range by combining them with other data sets from the same area to obtain more robust harvest location predictions. Such models will also offer more flexibility to sample collectors when planning future collections or considering sample collection possibilities in difficult geopolitical contexts. Our work highlights the potential for efficient and cost-effective scientific testing techniques and tools to support both regulated companies and enforcement agencies to ensure compliance with import regulations and promote legal and sustainable forest products trade.

Appendix: Supplementary Figures

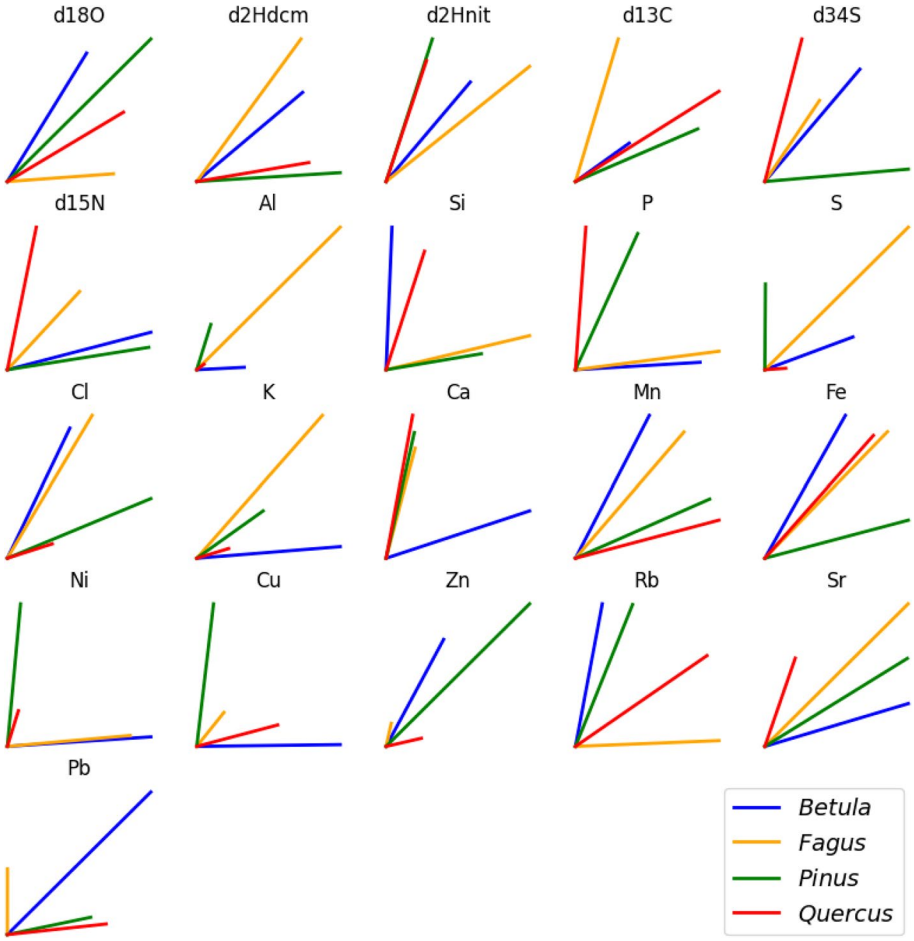


Fig. 6 Learned genus vectors for all SIR and TE variables with latent dimension set to $\ell = 2$. The vectors were estimated from the first cross-validation fold in the high data regime

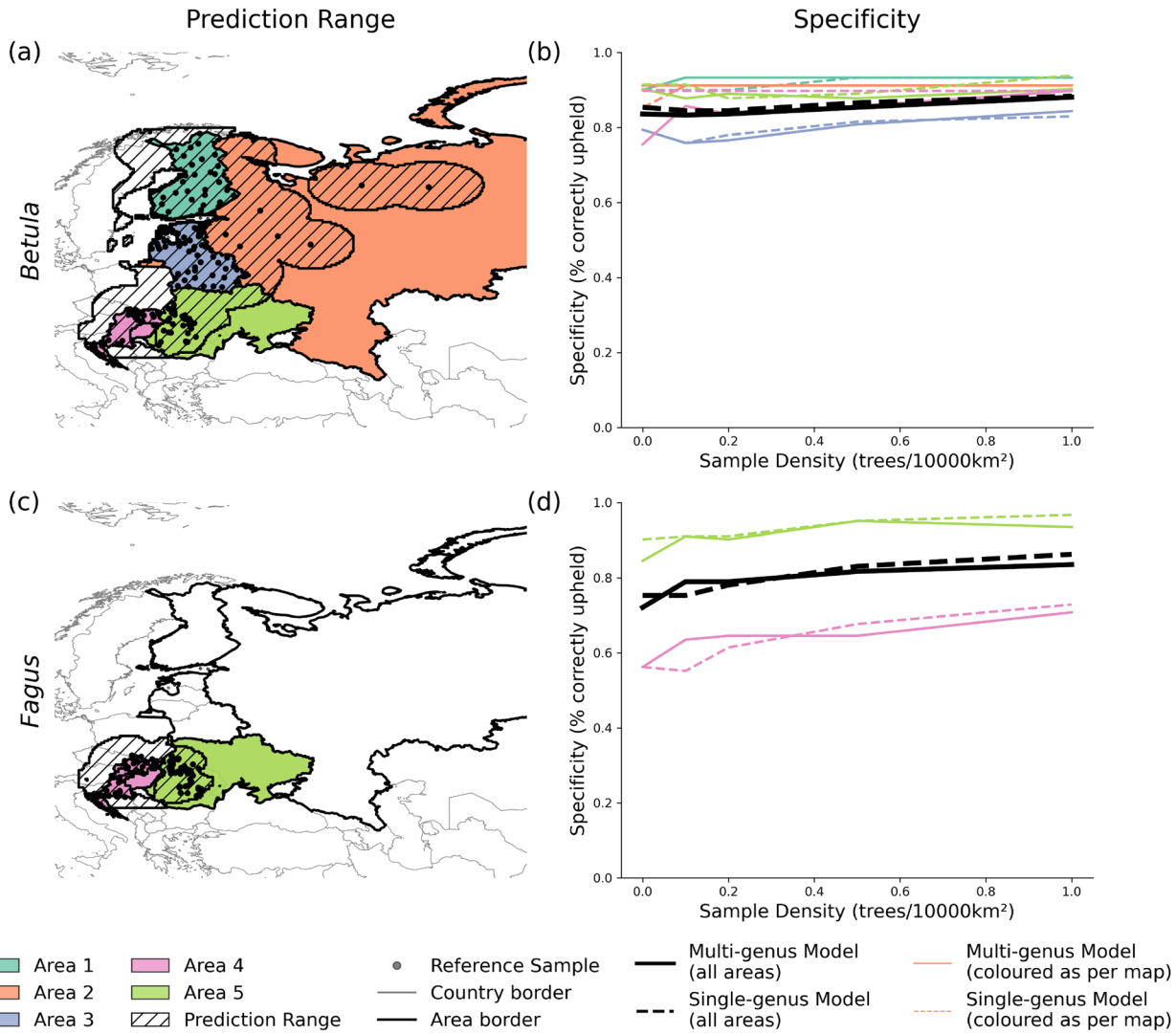


Fig. 7 Specificity of single- and multi-genus models for the verification task for *Betula* and *Fagus* in the low-data regime. **a**, **c** A map of sample locations (black circle), prediction ranges (lines), and testing areas (colored areas) used to generate test scenarios for each genus. **b**, **d** the average specificity values for the verification task, as a function of sample density in the low-data region using a single-genus model (dashed) and the multi-genus model (solid). Colored lines show the mean values per focal testing area (line colors correspond to map), while black lines show the mean value across all test scenarios. Higher values correspond to more accurate predictions

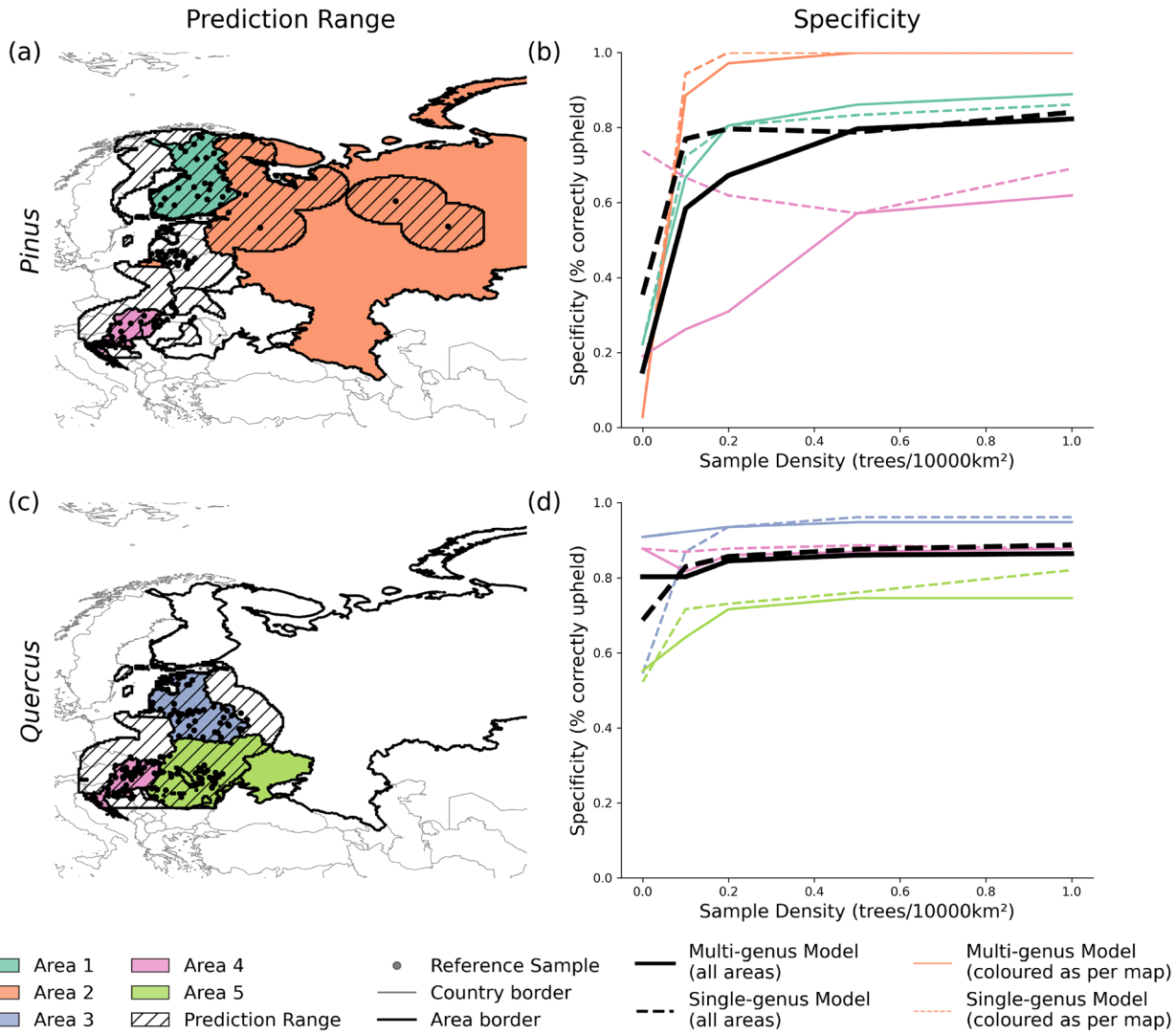


Fig. 8 Specificity of single- and multi-genus models for the verification task for *Pinus* and *Quercus* in the low-data regime. **a**, **c** A map of sample locations (black circle), prediction ranges (lines), and testing areas (colored areas) used to generate test scenarios for each genus. **b**, **d** the average specificity values for the verification task, as a function of sample density in the low-data region using a single-genus model (dashed) and the multi-genus model (solid). Colored lines show the mean values per focal testing area (line colors correspond to map), while black lines show the mean value across all test scenarios. Higher values correspond to more accurate predictions

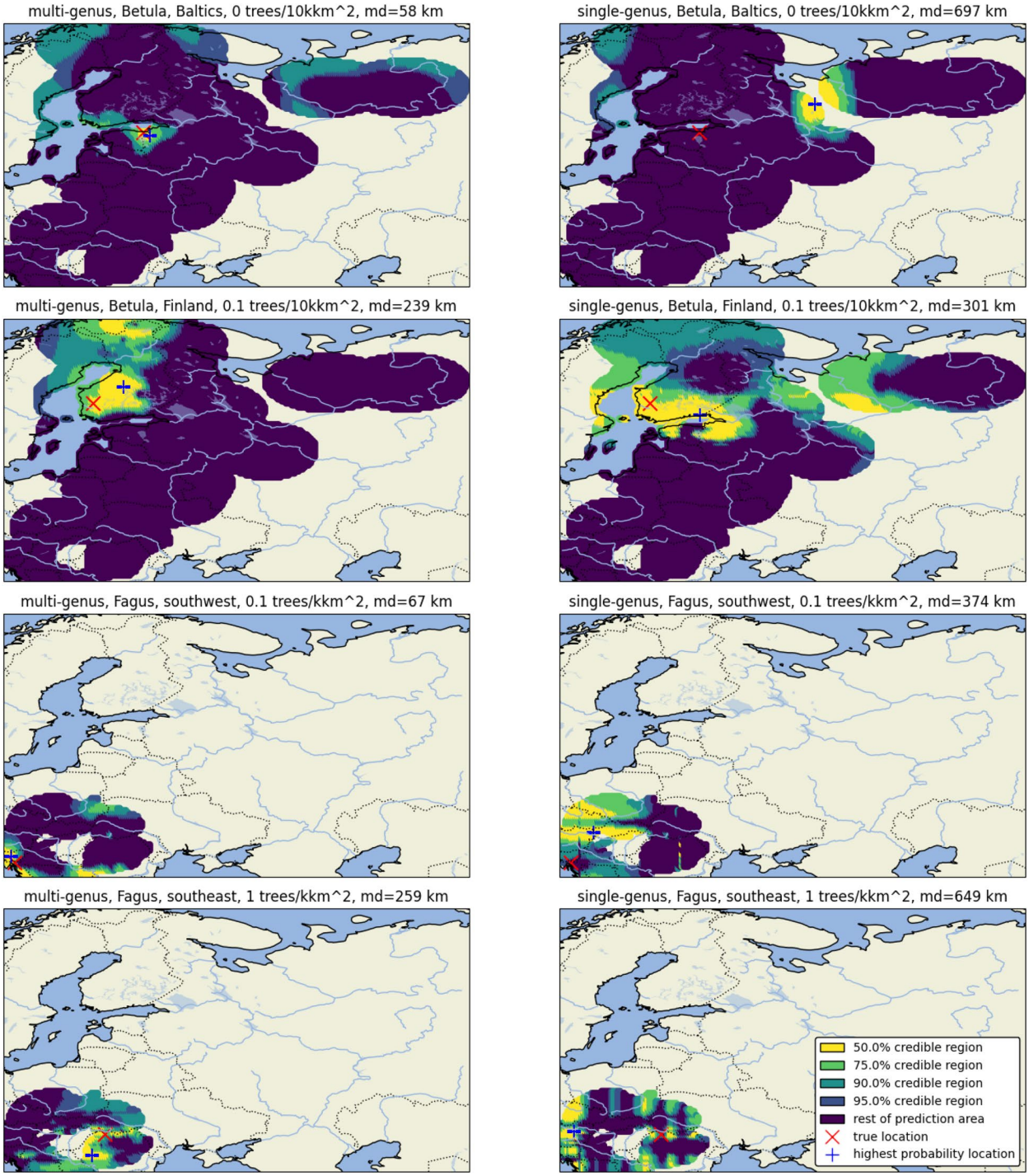


Fig. 9 Examples of predicted harvest location probabilities for *Betula* and *Fagus* in the low-data regime. The left column shows predictions from the multi-genus model, whereas the right column shows predictions from the single-genus model for the same sample. The choice of the focal testing area, the sampling density in the focal testing area and the mode distance (distance between the highest-probability location and the true location) are indicated above each figure. Different colors denote credible regions for different probability thresholds

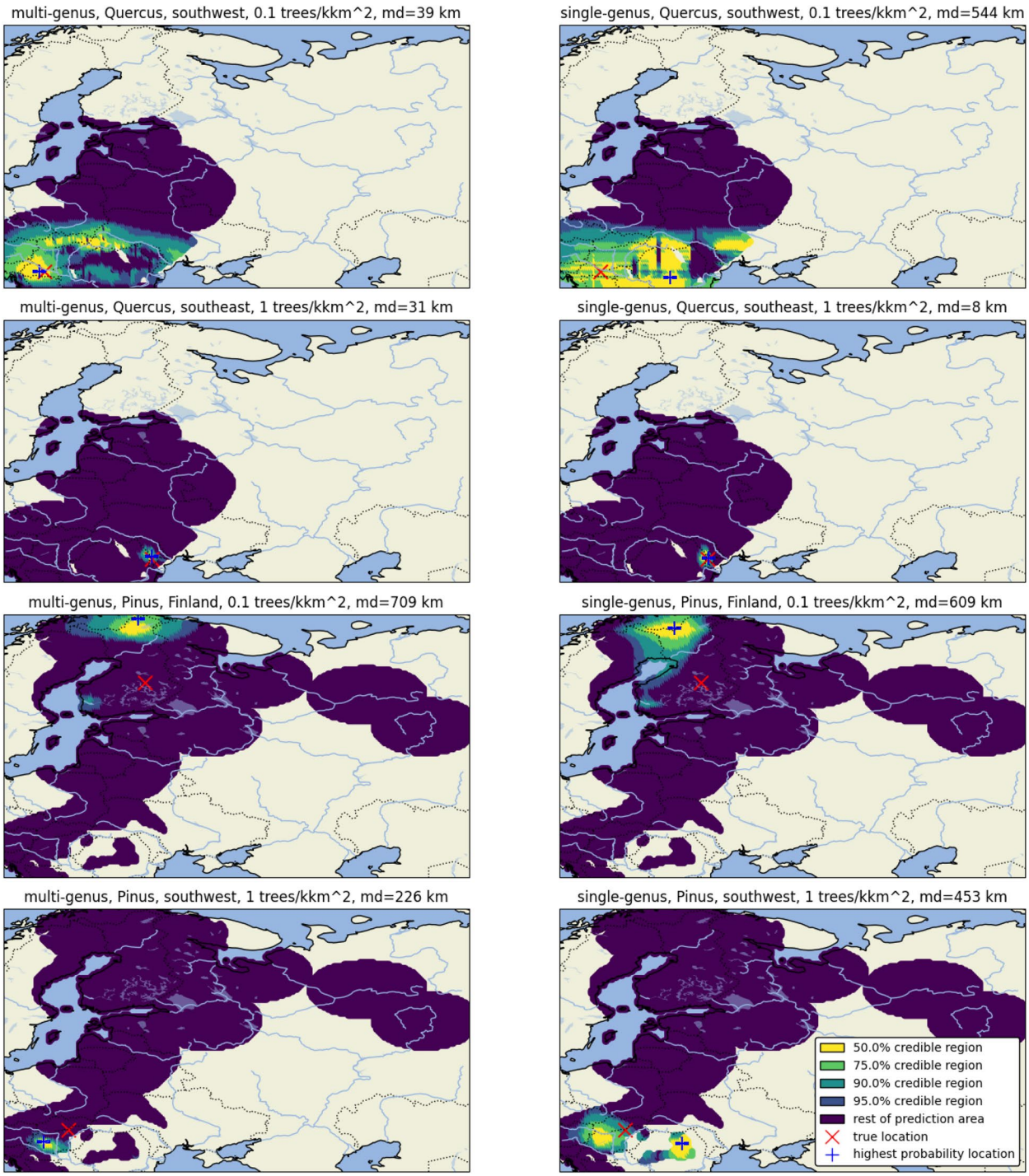


Fig. 10 Examples of predicted harvest location probabilities for *Quercus* and *Pinus* in the low-data regime. The left column shows predictions from the multi-genus model, whereas the right column shows predictions from the single-genus model for the same sample. The choice of the focal testing area, the sampling density in the focal testing area and the mode distance (distance between the highest-probability location and the true location) are indicated above each figure. Different colors denote credible regions for different probability thresholds

Authors' contributions
J.T., J.S., and V.D.: conceptualization; J.T., T.M., and C.S.: data curation; J.T., T.M., and V.D: formal analysis, investigation and methodology; J.T., J.S., and A.A.: funding acquisition; J.S. and V.D.: project administration; B.B.: resources; J.T., C.S., and T.M.: software; J.S. and V.D.: supervision; C.S., L.B., and V.D.: validation; C.S and J.T.: visualization; J.T., V.D., and J.S.: writing–original draft; L.B., B.B., C.C., P.Z., J.T., S.B.J., and A.A.: writing—review and editing.

Funding
Open access funding provided by Chalmers University of Technology.

Data availability
The datasets analyzed in the current study are available from the corresponding author on reasonable request. Source code can be accessed upon request at https://zenodo.org/records/16423363.

Code availability
The source code can be accessed from this reference: Truszkowski (2025).

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Chalmers University of Technology, Gothenburg, Sweden. [2]World Forest ID, Washington DC, United States. [3]Columbia University, New York, United States. [4]Ghent University, Ghent, Belgium. [5]Ghent University, Ghent, Belgium. [6]World Forest ID, Washington DC, United States. [7]Preferred by Nature, Vilnius, Lithuania. [8]Ştefan cel Mare University of Suceava, Suceava, Romania. [9]Royal Botanic Gardens, London, United Kingdom. [10]University of Sheffield, Sheffield, United Kingdom. [11]Meise Botanic Garden, Meise, Belgium. [12]KU Leuven, Leuven, Belgium. [13]IKEA, Älmhult, Sweden. [14]Wageningen University & Research, Wageningen, Netherlands. [15]University of Gothenburg, Gothenburg, Sweden. [16]Department of Biology, University of Oxford, South Parks Road, Oxford, United Kingdom. [17]Gothenburg Global Biodiversity Centre, Gothenburg, Sweden. [18]Meise Botanic Garden, Meise, Belgium.

Received: 5 January 2026 Accepted: 6 May 2026
Published online: 19 May 2026

References

Abramowitz M, Stegun IA (1972) Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series 55. Tenth Printing. ERIC. https://doi.org/10.1115/1.3625776

African Natural Resources Centre (2021) Illicit trading in Africa's forest products: Focus on timber. Tech. rep., African Development Bank, Abidjan, Côte d'Ivoire, pp. 1–20. https://www.afdb.org/sites/default/files/documents/publications/illicit_timber_trade_report.pdf. Accessed 12 May 2026

Ågren GI, Weih M (2012) Plant stoichiometry at different scales: element concentration patterns reflect environment more than genotype. New Phytol 194(4):944–952. https://doi.org/10.1111/j.1469-8137.2012.04114.x

Aguzzoni A, Giammarchi F, Mundo IA et al (2025) Tracing timber origin: geographic provenancing at regional scale with multielement and strontium isotope ratio analyses. For Ecol Manage 579:122494. https://doi.org/10.1016/j.foreco.2025.122494

Allen ST, Sprenger M, Bowen GJ et al (2022) Spatial and temporal variations in plant source water: O and h isotope ratios from precipitation to xylem water. In: Stable isotopes in tree rings: Inferring physiological, climatic and environmental responses. Springer International Publishing Cham, pp 501–535. https://doi.org/10.1007/978-3-030-92698-4_18

Alvarez MA, Rosasco L, Lawrence ND et al (2012) Kernels for vector-valued functions: A review. Found Trends® Mach Learn 4(3):195–266. https://doi.org/10.1561/9781601985590

Baker JC, Hunt SF, Clerici SJ et al (2015) Oxygen isotopes in tree rings show good coherence between species and sites in Bolivia. Glob Planet Change 133:298–308. https://doi.org/10.1016/j.gloplacha.2015.09.008

Boeschoten LE, Sass-Klaassen U, Vlam M et al (2022) Clay and soil organic matter drive wood multi-elemental composition of a tropical tree species: implications for timber tracing. Sci Total Environ 849:157877. https://doi.org/10.2139/ssrn.4151661

Boeschoten LE, Vlam M, Sass-Klaassen U et al (2023) A new method for the timber tracing toolbox: applying multi-element analysis to determine wood origin. Environ Res Lett 18(5):054001. https://doi.org/10.1088/1748-9326/acc81b

Boeschoten LE, Vlam M, Sass-Klaassen U et al (2023) Stable isotope ratios in wood show little potential for sub-country origin verification in central Africa. For Ecol Manage 544:121231. https://doi.org/10.1016/j.foreco.2023.121231

Boner M, Sommer T, Erven C et al (2007) Stable isotopes as a tool to trace back the origin of wood. In: Proceedings of the international workshop, Fingerprinting methods for the identification of timber origins, pp 47–57. https://literatur.thuenen.de/digbib_extern/dk040646.pdf. Retrieved on 28 Feb 2026

Bonilla EV, Agakov FV, Williams CK (2007) Kernel multi-task learning using task-specific features. In: Artificial Intelligence and Statistics. PMLR, pp 43–50. https://doi.org/10.1145/1143844.1143860

Bowen GJ, Wassenaar LI, Hobson KA (2005) Global application of stable hydrogen and oxygen isotopes to wildlife forensics. Oecologia 143(3):337–348. https://doi.org/10.1007/s00442-004-1813-y

Capo LFM, Degen B, Blanc-Jolivet C et al (2024) Timber tracking of *Jacaranda copaia* from the Amazon Forest using DNA fingerprinting. Forests. https://doi.org/10.3390/f15081478

Caudullo G, Welk E, San-Miguel-Ayanz J (2017) Chorological maps for the main European woody species. Data Brief 12:662–666. https://doi.org/10.1016/j.dib.2017.05.007

Deb S, Baudais V et al (2022) The challenges of data collection in conflict-affected areas: A case study in the liptako-gourma region. SIPRI Insights Peace Secur 7:1–20. https://www.sipri.org/sites/default/files/2022-10/sipriinsight2207_data_collection_1.pdf. Accessed 12 May 2026

Deklerck V (2023) Timber origin verification using mass spectrometry: challenges, opportunities, and way forward. Forensic Sci Int Anim Environ 3:100057. https://doi.org/10.1016/j.fsiae.2022.100057

Dormontt EE, Boner M, Braun B et al (2015) Forensic timber identification: it's time to integrate disciplines to combat illegal logging. Biol Conserv 191:790–798. https://doi.org/10.1016/j.biocon.2015.06.038

EUDR (2023) Regulation on Deforestation-free Products. https://environment.ec.europa.eu/topics/forests/deforestation/regulation-deforestation-free-products_en. https://doi.org/10.5040/9781509977581.ch-005. Accessed 10 Jan 2025

FAO (2020) Global forest resources assessment 2020: Main report. https://doi.org/10.4060/ca9825en

Gardner JR, Pleiss G, Bindel D et al (2018) Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. In: Advances in Neural Information Processing Systems, 31. pp 7576–7586. https://proceedings.neurips.cc/paper/2018/hash/27e8e17134dd7083b050476733207ea1-Abstract.html. Accessed 12 May 2026

Gori Y, Deklerck V, Camin F (2025) Stable isotope ratio analysis to determine the geographical origin of timber: a review. TrAC Trends Anal Chem 193:118448. https://doi.org/10.1016/j.trac.2025.118448

Hartl-Meier C, Zang C, Büntgen U et al (2015) Uniform climate sensitivity in tree-ring stable isotopes across species and sites in a mid-latitude temperate forest. Tree Physiol 35(1):4–15. https://doi.org/10.1093/treephys/tpu096

Heineman KD, Turner BL, Dalling JW (2016) Variation in wood nutrients along a tropical soil fertility gradient. New Phytol 211(2):440–454. https://doi.org/10.1111/nph.13904

Kingma DP, Ba J (2015) Adam: A method for stochastic optimization. In: Bengio Y, LeCun Y (eds) 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. https://arxiv.org/abs/1412.6980. Accessed 12 May 2026

Lehmann MM, Schuler P, Cormier MA et al (2022) The stable hydrogen isotopic signature: from source water to tree rings. In: Stable isotopes in tree rings: inferring physiological, climatic and environmental responses. Springer International Publishing Cham, pp 331–359. https://doi.org/10.1007/978-3-030-92698-4_11

Loshchilov I, Hutter F (2017) SGDR: stochastic gradient descent with warm restarts. In: 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings. OpenReview.net. https://openreview.net/forum?id=Skq89Scxx. Accessed 12 May 2026

Low MC, Schmitz N, Boeschoten LE et al (2022) Tracing the world's timber: the status of scientific verification technologies for species and origin identification. IAWA J 44(1):63–84. https://doi.org/10.1163/22941932-bja10097

Meyer-Sand RVB, Boeschoten LE, Bouka GU et al (2025) Unlocking the geography of azobé timber (*Lophira alata*): revealing spatial genetic structure beyond species boundaries. BMC Plant Biol 25(1):315. https://doi.org/10.1186/s12870-025-06287-2

Mortier T, Truszkowski J, Norman M et al (2024) A framework for tracing timber following the Ukraine invasion. Nat Plants 10(3):390–401. https://doi.org/10.1038/s41477-024-01648-5

Paredes-Villanueva K, Boom A, Ottenburghs J et al (2022) Isotopic characterization of *Cedrela* to verify species and regional provenance of Bolivian timber. Tree-Ring Res 78(2):73–89. https://doi.org/10.3959/2021-17

Paszke A, Gross S, Chintala S et al (2017) Automatic differentiation in PyTorch. In: NIPS 2017 Workshop on Autodiff, Long Beach, California, USA. https://openreview.net/forum?id=BJJsrmfCZ. Accessed 12 May 2026

Raju K, Raju A (2000) Biogeochemical investigation in South Eastern Andhra Pradesh: the distribution of rare earths, thorium and uranium in plants and soils. Environ Geol 39(10):1102–1106. https://doi.org/10.1007/s002540000111

Reboredo F (2013) Socio-economic, environmental, and governance impacts of illegal logging. Environ Syst Decis 33:295–304. https://doi.org/10.1007/s10669-013-9444-7

St. John Glew K, Graham LJ, McGill RA et al (2019) Spatial models of carbon, nitrogen and sulphur stable isotope distributions (isoscapes) across a shelf sea: an inla approach. Methods Ecol Evol 10(4):518–531. https://doi.org/10.1111/2041-210x.13138

Suarez AV, Tsutsui ND (2004) The value of museum collections for research and society. Bioscience 54(1):66–74. https://doi.org/10.1641/0006-3568(2004)054[0066:tvomcf]2.0.co;2

Truszkowski J (2025) Code for manuscript: Combining forces: Multi-genus modeling for stable isotope-based timber traceability. https://doi.org/10.5281/zenodo.16423362

Truszkowski J, Maor R, Yousuf RB et al (2025) A probabilistic approach to estimating timber harvest location. Ecol Appl 35(1):e3077. https://doi.org/10.1002/eap.3077

Vlam M, de Groot GA, Boom A et al (2018) Developing forensic tools for an African timber: regional origin is revealed by genetic characteristics, but not by isotopic signature. Biol Conserv 220:262–271. https://doi.org/10.1016/j.biocon.2018.01.031

Watkinson CJ, Rees GO, Gwenael MC et al (2022) Stable isotope ratio analysis for the comparison of timber from two forest concessions in Gabon. Front For Glob Change 4:650257. https://doi.org/10.3389/ffgc.2021.650257

Watkinson CJ, Rees GO, Hofem S et al (2022) A case study to establish a basis for evaluating geographic origin claims of timber from the Solomon Islands using stable isotope ratio analysis. Front For Glob Change 4:645222. https://doi.org/10.3389/ffgc.2021.645222

West JB, Bowen GJ, Dawson TE, et al (2009) Isoscapes: understanding movement, pattern, and process on Earth through isotope mapping. Springer. https://doi.org/10.1007/978-90-481-3354-3

Williams CK, Rasmussen CE (2006) Gaussian processes for machine learning, vol 2. MIT press, Cambridge. https://doi.org/10.7551/mitpress/3206.001.0001

Publisher's Note